Derivatives Pricing In a Markov Chain Jump-Diffusion Setting

Shaoul Nathan

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

To The

Department of Statistics
London School of Economics
University of London

February 2005

בס״ד

Dedicated to my loving parents.

Acknowledgements

I would like to offer my deep gratitude to *Dr. Angelos Dassios* for his expert supervision and guidance over the course of my Ph.D, as well as for his inspiration during my undergraduate studies, which motivated me to continue and perform this research. Many thanks to all the Department of Statistics for their help and advice.

990925

Abstract

In this work we develop a Markov Chain Jump-Diffusion (MCJD) model, where we have a financial market in which there are several possible states. Asset prices in the market follow a generalised geometric Brownian motion, with drift and volatility depending on the state of the market. So for example, one state may represent a *bull market* where drifts are high, whilst another state may represent a *bear market* where where drifts are low. The state the market is in is governed by a continuous time Markov chain. We add to this diffusion process jumps in the asset price which occur when the market changes state, and the jump sizes are dependent on the states the market is transiting to and transiting from. We also allow the market to transit to the same state, which corresponds to a jump in the asset price with no change to the drift or volatility.

We will develop conditions of no arbitrage in such a market, and methods for pricing derivatives of assets whose prices follow MCJD processes. We will also consider Term-Structure models where the short rate (or forward rate) follows an MCJD process.

Table of Contents

Notation

During the course of this work there will be the need to define many terms and symbols in order to develop all the theory. In order to assist the reader we have listed below some of the most commonly used terms, as well as the page number where they are first defined and a brief description of what they are used to represent. We have also included a list of matrices.

Term	Page no.	Description
Y_t	5	Markov chain representing the state of the market at time t.
$\mathcal{S}$	5	The set of possible states the market can be in which can take values $1, \ldots, n$.
n	5	Total number of states in the market .
j	5	Subscript denoting the state of the market. $j \in \mathcal{S}$.
k	5	Subscript denoting the state of the market. $k \in \mathcal{S}$.
λ^{jk}	5	The transition intensity from state j to state k.

Term	Page no.	Description
N_t^{jk}	6	The total number of times the market has transited from state j to state k up until time t.
N_t	6	The total number of transitions of the state up to time t.
I_t^j	7	Indicator variable taking the value 1 if the market is in state j at time t and 0 otherwise .
p_t^{jk}	8	The probability at time 0 that the market is in state k at time t, given that at time 0 it is state j .
P_t^{jk}	8	The total expected time spent in state k in the interval $[0, t]$, given that at time 0 the market is in state j.
$\mathcal{J}$	9	An $x+1$ dimensional row vector representing the jump sequence of the first x jumps.
D_t	9	Markov chain used for phase-type distributions.
w	11	Subscript denoting the state of the Markov chain D_t. $w = 1, \ldots, x+1$.
$p^{j_1 \cdots j_{x+1}}(T)$	12	The probability of observing jump sequence $\mathcal{J}$ in time T.
p^{jk}	13	The probability of transiting from state j to state k conditional on a transition occurring.
r	15	Total number of Brownian motions.
b	15	Subscript denoting number of Brownian motion. $b = 1, \ldots, r$.
W_t^b	15	Value of Brownian motion number b at time t .
γ_i^{jk}	17	The jump size of asset i when transiting from state j to state k. The i may be suppressed when there is only one asset.
B_t	18	Bank-account process.

Term	Page no.	Description
r_j	19	The rate of interest whilst in state j.
$S_{i,t}$	20	The price of asset i at time t. The i may be suppressed when there is only one asset .
$\tilde{S}_{i,t}$	20	The discounted price of asset i.
$\mu_{i,j}$	20	The drift of asset i whilst in state j. The i may be suppressed when there is only one asset.
$\sigma_{i,b,j}$	20	The volatility of asset i due to Brownian motion b whilst in state j. The i may be suppressed when there is only one asset. The b may be suppressed when there is only one Brownian motion.
$\hat{\mu}_{i,j}$	22	Drift of asset i in state j under transformed (risk-neutral) measure.
$\theta_{b,j}$	22	Addition to the drift for Brownian motion b in state j.
ψ^{jk}	25	Transformation to transition intensity from state j to state k under transformed (risk-neutral) measure.
r_t	108	The value of the short rate at time t.
$\mu(t, Y_{t-})$	108	The drift of the short-rate process at time t.
$\sigma(t, Y_{t-})$	108	The volatility of the short-rate process at time t.
$\gamma^r(t, Y_{t-}, Y_t)$	108	The size of the jump in the short-rate process at time t.
$p(t, T)$	108	The value of a zero-coupon bond at time t expiring at time T.
$m(t, T, Y_{t-})$	108	The drift of the zero-coupon price process.
$v(t, T, Y_{t-})$	108	The volatility of the zero-coupon price process.
$\gamma^p(t, T, Y_{t-}, Y_t)$	108	The size of the jump in the value of the zero-coupon price process.

Term	Page no.	Description
$f(t,T)$	108	The forward rate between times t and T.
$a(t,T,Y_{t-})$	108	The drift of the forward-rate process.
$b(t,T,Y_{t-})$	108	The volatility of the forward-rate process.
$\gamma^f(t,T,Y_{t-},Y_t)$	108	The size of the jump in the forward-rate process.
$A(t,T,Y_{t-})$	113	Minus the integral of $a(t,T,Y_{t-})$.
$B(t,T,Y_{t-})$	113	Minus the integral of $b(t,T,Y_{t-})$.
$\Gamma^f(t,T,Y_{t-},Y_t)$	113	Minus the integral of $\gamma^f(t,T,Y_{t-})$.
$\hat{p}(0,T)$	124	Empirical bond prices.
$\hat{f}(0,T)$	124	Empirical forward rates.

We will now include a table of matrices.

Matrix	Page no.	Description
$\mathbf{D(a)}$	21	$m \times m$ matrix with elements of $\mathbf{a}$ down the principal diagonal where $\mathbf{a} \in \mathbf{R}^m$.
$\tilde{\mathbf{S}}_t^{m\times 1}$	21	Vector of discounted stock prices.
$\mathbf{U}_j^{m\times 1}$	21	Vector of drifts.
$\mathbf{\Sigma}_j^{m\times r}$	21	Matrix of volatilities.
$\mathbf{W}_t^{r\times 1}$	21	Matrix of Brownian motions.
$\mathbf{\Gamma}_j^{m\times n}$	21	Matrix of jump sizes.
$\mathbf{N}_t^{j\ n\times 1}$	21	Vector of counting processes.
$\hat{\mathbf{U}}_j^{m\times 1}$	22	Vector of drifts under transformed (risk-neutral) measure.
$\mathbf{\Theta}_j^{r\times 1}$	22	Vector of additions to the drifts for each Brownian motion.
$\tilde{\mathbf{W}}_t^{r\times 1}$	22	Vector of Brownian motions under the changed (risk-neutral) measure.
$\mathbf{\Lambda}_j^{n\times 1}$	27	Vector of transition intensities.
$\mathbf{\Psi}_j^{n\times 1}$	27	Vector of transformations to transition intensities under changed (risk-neutral) measure.

Chapter 1

Introduction

1.1 General Introduction

Over 30 years ago Black and Scholes produced their seminal paper Black and Scholes [1973], which together with Merton [1973] paved the way for the development of mathematical finance as we know it. Their papers were based on the assumption that the price of the underlying asset follows the behaviour of a *diffusion process*, most notably a geometric Brownian motion. Another watershed was the development of the Arbitrage Pricing Technique in Ross [1976] and Ross [1978], and the martingale approach to arbitrage pricing developed in Harrison and Kreps [1979] and Harrison and Pliska [1981].

The Black-Scholes model has become very popular due to its simplicity in that it quantifies risk through a single constant volatility parameter. However, it is clear that this assumption of constant volatility will not be totally adequate when attempting model the behaviour of today's complex financial markets, and this accusation has been supported by various empirical studies (for example see Bakshi et al. [1997]). There have subsequently been many attempts to modify this model and relax its over-simplistic assumptions. Stochastic volatility models have been extensively studied where the volatility is allowed to evolve over time (see Hull and White [1987], Stein and Stein [1991] and Heston [1993], or alternatively for a synopsis see Fouque et al. [2000]).

On an alternate front, there have been attempts to develop models which incorporate jumps into the asset price behaviour. Such jumping behaviour in asset prices has been supported by empirical evidence such as in Ball and Torous [1985] and Jorion [1988]. Pure jump processes were developed in papers such as Merton [1976] and Björk et al. [1997]. A natural extension to these models are processes that include both a diffusion part and a jump part, known as *jump-diffusion processes*, such as in Andersen and Andersen [2000] and Madan [2001]. Many other varying models have been developed to try to improve on the Black-Scholes model, although in their increased sophistication they sacrifice a lot in terms of ease of calculation, as well as intuitiveness of the models. This last factor is fairly important, as any model which requires a so-called 'rocket scientist' to understand is unlikely to be used widely by practitioners. They prefer to employ more simplistic models they can understand even though they may not be as accurate.

In this work we shall develop a different type of Jump-diffusion model which we shall call a *Markov Chain Jump-Diffusion Model* (MCJD). The motivation for this lies behind two papers where totally different models have been developed. Firstly, in Norberg [2003] a pure-jump process is considered where the market is driven by a continuous-time homogenous financial market. Alternatively, in Runggaldier [2003] a jump-diffusion process is developed where the jumps are modelled by a marked point process.

In our MCJD model we consider a market in which there are several states of the market. Asset prices in the market follow a generalised geometric Brownian motion, with drift and volatility depending on the state of the market. So for example, one state may represent a *bull market* where drifts are high, whilst another state may represent a *bear market* where drifts are low. The state the market is in is governed by a continuous-time Markov chain. We add to this diffusion process jumps in the asset price which occur when the market changes state, and the jump sizes are dependent on the states the market is transiting to and transiting from. We also allow the market to transit to the same state, which corresponds to a jump in the asset price with no change to the drift or volatility.

This model constitutes a stochastic drift and volatility model as these parameters are allowed to change, as well as being a jump process. It is very intuitive to see how this model may represent the behaviour of financial assets, as it is widely recognized that there are trends in the market, and periods where asset prices behave in different ways. This model provides the advantages of stochastic volatility models and jumps processes in that it should describe more accurately the behaviour of financial assets, and at the same time regulates these features in a restricted sense so as to facilitate pricing, and hopefully making intuitive sense.

1.2 Order of Work

In the conclusion of this chapter we will describe the market characteristics common to all the subsequent models, and develop several results concerning Markov chains which will prove useful in our subsequent investigations.

In chapter 2 we develop the *Equity model* where asset prices follow our MCJD model. We will deal with issues of completeness, replicating contingent claims and finally pricing derivatives.

In chapter 3 we shall develop and compare various numerical methods for pricing derivatives on the assets described in chapter 2. We will look at a particular example and see how all the methodologies had priced call options on this asset.

Chapter 4 sees us turning our attention to *Term-Structure models* where we will use our MCJD model to describe the behaviour of the short rate for *short-rate models*, or of the forward rate for *HJM models.* We again discuss issues regarding completeness and derivative pricing for these models, as well as parameter estimation.

Finally in Chapter 5 we will develop numerical methods for pricing the interest-rate derivatives developed in chapter 4.

To try to make reading this thesis as easy as possible for the reader, we have also included a notation page which includes many of the terms and symbols that are used repeatedly throughout this work.

1.3 Preliminaries

1.3.1 The Markov Chain Market

As mentioned above, in this work we shall be considering a market in which there are n states represented by the continuous-time Markov chain $(Y_t)_{t\geq 0}$ with finite state space $\mathcal{S} = \{1 \ldots n\}$. The process Y_t transits between states j and k where $j, k \in \mathcal{S}$ with intensity λ^{jk}, so that the generator of this process is given by

$$\mathbf{G} = \begin{pmatrix} -\hat{\lambda}^1 & \lambda^{12} & . & . & \lambda^{1n} \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \lambda^{n1} & . & . & . & -\hat{\lambda}^n \end{pmatrix}, \qquad (1.3.1)$$

where $\hat{\lambda}^j = \sum_{j \neq k}^n \lambda^{jk}$. This process is time-homogeneous, so that for $j \neq k$ we have that

$$\Pr[Y_{t+dt} = k | Y_t = j] = \lambda^{jk} dt + o(dt). \tag{1.3.2}$$

We also have that

$$\lambda^{jk} \geq 0,$$

and we are assuming there can be at most one transition for any small length of time dt.

Let us add to the above Markov chain the ability to **transit** to the same state, the probability of which is given by $\lambda^{jj} dt$. This should not be confused with the probability of **remaining** in the same state which has probability equal to $1 - \sum_{k=1}^n \lambda^{jk} dt + o(dt)$. The motivation for doing this will become clear in section 2.2. We can therefore regard this extended process as being a *multivariate point process*, with state-dependent intensity vector Λ_j given by

$$\Lambda_j = \{\lambda^{j1}, \ldots, \lambda^{jn}\}^T \quad \forall\, j \in \mathcal{S}.$$

Define N_t^{jk} as being the number of transitions from state j to state k up to time t, so that for $j \neq k$

$$N_t^{jk} = |\{s;\ 0 < s < t\,,\ Y_s = k\,,\ Y_{s-} = j\}|,$$

whilst N_t^{jj} is the number of times the process was in state j and **transited** to the same state. N_t denotes the total number of transitions up to time t:

$$N_t = \sum_{j \in \mathcal{S}} \sum_{k \in \mathcal{S}} N_t^{jk}$$

where we set $N_0 = 0$. We shall call this point process the *jump process* and its intensities the *jump intensities*, as opposed to the *transition process* given in 1.3.1.

We make the assumption of non-explosion, so that $N_t < \infty$ for $t \geq 0$ (similarly $\lambda^{jk} < \infty \ \forall \ j, k$), and assume N_t is defined on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with filtration $\mathcal{F}_t$ to which N_t is adapted. This process can be characterised as a *doubly stochastic Poisson process* with state-dependent intensity λ_t, where

$$\lambda_t = \lambda(Y_t) = \sum_{j \in \mathcal{S}} \mathbf{1}\Lambda_j I_t^j, \tag{1.3.3}$$

$\mathbf{1}$ is a 1x4 row vector with all entries equal to 1, and I_t^j is the indicator variable which takes values

$$I_t^j = \begin{cases} 1 & \text{if } Y_t = j \\ 0 & \text{otherwise.} \end{cases} \tag{1.3.4}$$

The expected time spent in state j before transiting out is therefore exponentially distributed with parameter $(\mathbf{1}\Lambda_j - \lambda^{jj})$. We shall denote the times at which each of these N_t jumps occur by $t_1, \ldots, t_{N_t}$. Finally, we shall set the process Y_t to be *left continuous* and hence it will also be *previsible.* In the remainder of this work we shall denote previsible state-dependent processes as functions of Y_{t-}.

1.3.2 Properties of the Model

We will now derive some properties of this model which we will make use of in the forthcoming chapters. In order to do this, let us first write the following definition (for example see Grimmett and Stirzaker [2001]):

Definition 1.3.1. Given a Markov chain setting described in section 1.3.1, the probability of being in state k at time t given that at time 0 we were in state j for $j, k \in \mathcal{S}$ is given by

$$p_t^{jk} = P[Y_t = k|\, Y_0 = j].$$

The probability p_t^{jk} is then given by

$$\begin{aligned} p_t^{jk} &= \mathbf{1}^{j\prime} exp\{t\mathbf{G}^y\}\, \mathbf{1}^k \\ &= \sum_{y=0}^{\infty} \frac{t^y}{y!} \mathbf{1}^{j\prime} \mathbf{G}^y \mathbf{1}^k, \end{aligned} \tag{1.3.5}$$

where $\mathbf{1}^j$ is the n-dimensional column vector with the j^{th} entry equal to 1 and all other entries equal to 0, and $\mathbf{G}$ is the generator defined in (1.3.1). We denote the transpose by $'$. Let us also define the expected total time spent over the interval $[0, t]$ in state k, given at time 0 we are in state j, by

$$P_t^{jk} = \int_{s=0}^{t} p_s^{jk} ds.$$

We therefore have

$$P_t^{jk} = \sum_{y=0}^{\infty} \frac{t^{y+1}}{(y+1)!} \mathbf{1}^{j\prime} \mathbf{G}^y \mathbf{1}^k. \tag{1.3.6}$$

In order to perform many of the calculations in subsequent chapters, we will need to condition on the path taken by the Markov chain Y_t. We will now derive some results when conditioning on this path, which we will make use of when pricing derivatives later on.

Let us condition on the Markov chain Y_t following a given path. Suppose the Markov chain starts in state j_1, and that the first $x+1$ transitions of the Markov chain (where a jump to the same state is considered a transition) are at times $t_1, \ldots, t_{x+1}$. Let the jump sequence (and hence state of the Markov chain) of the first x of these jumps be represented by the $x+1$-dimensional row vector $\mathcal{J} = \{j_1, \ldots, j_{x+1}\}$. Under this setup, we shall now calculate the probability the Markov chain is in each state at any time $t \in [0, T]$. In order to do this we shall set up our model as a *phase-type distribution* (see Asmussen [2000] or O'Cinneade [1990]).

Constructing a phase-type distribution involves representing this conditional Markov chain as a different Markov chain D_t which has generator $\mathbf{Q}_{\mathcal{J}}$, where the subscript shows dependence on the path we are conditioning on. We shall use the subscripts j and k to denote states of the original Markov chain Y_t, whilst the subscript w is reserved for the new Markov chain D_t.

Whilst Y_t is in state j_1, it transits to state j_2 with intensity $\lambda^{j_1 j_2}$. Similarly, whilst in state j_2 it transits to state j_3 with intensity $\lambda^{j_2 j_3}$ and so on, until Y_t arrives at state j_{x+1}. Since we are not conditioning on any subsequent transitions, the transition intensity out of state j_{x+1} is the total intensity for transiting out of state j_{x+1} given by $\hat{\lambda}^{j_{x+1}}$ where $\hat{\lambda}^{j_{x+1}} = \sum_{l=1}^{n} \lambda^{j_{x+1} l}$. This conditional Y_t process can be represented by the continuous-time Markov chain D_t where $D_t \in \{1, \dots, x+2\}$, which has generator $\mathbf{Q}_{\mathcal{J}}$ given by the $x+2$-dimensional square matrix below:

$$\mathbf{Q}_{\mathcal{J}} = \begin{bmatrix} -\lambda^{j_1 j_2} & \lambda^{j_1 j_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\lambda^{j_2 j_3} & \lambda^{j_2 j_3} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda^{j_3 j_4} & \lambda^{j_3 j_4} & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & -\lambda^{j_x j_{x+1}} & \lambda^{j_x j_{x+1}} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\hat{\lambda}^{j_{x+1}} & \hat{\lambda}^{j_{x+1}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

If all transition intensities are different then $\mathbf{Q}_{\mathcal{J}}$ can be diagonalised, which would simplify many of the calculations below. Let us now write the following lemmas:

Lemma 1.3.2. *We will now calculate the value of p_t^{jk} conditional on the jump sequence $\mathcal{J}$. Suppose we have $D_0 = 1$. Using (1.3.5) the probability that $D_T = x + 1$, i.e. that we have had exactly x transitions up until time T (and so are in state $x+1$ as every transition increases the value of D_t by 1),* *is given by*

$$P[D_T = x+1] = \sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1}, \tag{1.3.7}$$

where once again $\mathbf{1}^j$ *is the* $x+1$*-dimensional column vector with all entries equal to* 0 *and the* j^{th} *entry equal to* 1*. Conditioning on* $D_T = x+1$ *being true, we can further see that the probability of being in state* w *for* $w = 1, \ldots, x+1$ *at time* t *where* $t \in [0, T]$ *is given by*

$$P[D_t = w | D_T = x+1] = \frac{(\sum_{y=0}^{\infty} \frac{t^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{w})(\sum_{y=0}^{\infty} \frac{(T-t)^y}{y!} \mathbf{1}^{w\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1})}{\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1}}. \tag{1.3.8}$$

We can re-write (1.3.8) as

$$P[D_t = w | D_T = x+1] = \frac{\sum_{y_1=0}^{\infty} \sum_{y_2=0}^{\infty} V (\mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y_1} \mathbf{1}^{w})(\mathbf{1}^{w\prime} \mathbf{Q}_{\mathcal{J}}^{y_2} \mathbf{1}^{x+1})}{\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1}} \tag{1.3.9}$$

where

$$V = \frac{t^{y_1} (T-t)^{y_2}}{y_1! y_2!}.$$

We now have that the conditional probability at time t *of being in state* k *of our Markov chain* Y_t*, given that we start in state* j *at time* 0 *is given by*

$$p_t^{jk} | \mathcal{J} = \sum_{\{w : j_w = k\}} P[D_t = w | D_T = x+1]. \tag{1.3.10}$$

Using equation (1.3.9) this becomes

$$p_t^{jk}|\mathcal{J} = \sum_{\{w:j_w=k\}} \frac{\sum_{y_1=0}^{\infty}\sum_{y_2=0}^{\infty} V(\mathbf{1}^{1\prime}\mathbf{Q}_{\mathcal{J}}^{y_1}\mathbf{1}^{w})(\mathbf{1}^{w\prime}\mathbf{Q}_{\mathcal{J}}^{y_2}\mathbf{1}^{x+1})}{\sum_{y=0}^{\infty}\frac{T^y}{y!}\mathbf{1}^{1\prime}\mathbf{Q}_{\mathcal{J}}^{y}\mathbf{1}^{x+1}}. \tag{1.3.11}$$

We now also define P_T^{jk} to be the expected total time spent in state k so that

$$P_T^{jk}|\mathcal{J} = \int_{t=0}^{T} [p_t^{jk}|\mathcal{J}]dt. \tag{1.3.12}$$

We can integrate (1.3.11) to get

$$P_T^{jk}|\mathcal{J} = \sum_{\{w:j_w=k\}} \frac{\sum_{y_1=0}^{\infty}\sum_{y_2=0}^{\infty} Z(\mathbf{1}^{1\prime}\mathbf{Q}_{\mathcal{J}}^{y_1}\mathbf{1}^{w})(\mathbf{1}^{w\prime}\mathbf{Q}_{\mathcal{J}}^{y_2}\mathbf{1}^{x+1})}{\sum_{y=0}^{\infty}\frac{T^y}{y!}\mathbf{1}^{1\prime}\mathbf{Q}_{\mathcal{J}}^{y}\mathbf{1}^{x+1}} \tag{1.3.13}$$

where

$$Z = \frac{T^{(y_1+y_2+1)}}{(y_1+y_2+1)!}.$$

□

Lemma 1.3.3. *We will now calculate the probability of observing jump sequence $\mathcal{J} = (j_1, \ldots, j_{x+1})$ in an interval $[0, T]$, which we will denote by $p^{j_1 \cdots j_{x+1}}(T)$, by setting up a phase-type distribution as follows. We set*

$$
\bar{\mathbf{Q}}_{\mathcal{J}} = \begin{bmatrix}
-\hat{\lambda}_1^j & \lambda^{j_1 j_2} & 0 & 0 & . & . & 0 & 0 & \bar{\lambda}^{j_1 j_2} \\
0 & -\hat{\lambda}_2^j & \lambda^{j_2 j_3} & 0 & . & . & 0 & 0 & \bar{\lambda}^{j_2 j_3} \\
0 & 0 & -\hat{\lambda}_3^j & \lambda^{j_3 j_4} & . & . & 0 & 0 & \bar{\lambda}^{j_3 j_4} \\
. & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . \\
0 & 0 & 0 & 0 & . & . & -\hat{\lambda}_x^j & \lambda^{j_x j_{x+1}} & \bar{\lambda}^{j_x j_{x+1}} \\
0 & 0 & 0 & 0 & . & . & 0 & -\hat{\lambda}^{j_{x+1}} & \hat{\lambda}^{j_{x+1}} \\
0 & 0 & 0 & 0 & . & . & 0 & 0 & 0
\end{bmatrix},
$$

where as previously we have $\hat{\lambda}^j = \sum_{k=1}^n \lambda^{jk}$ *as well as* $\bar{\lambda}^{jk} = \hat{\lambda}^j - \lambda^{jk}$. *We therefore have*

$$
p^{j_1 \cdots j_{x+1}}(T) = \sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \bar{\mathbf{Q}}_{\mathcal{J}}^y \mathbf{1}^{x+1}. \tag{1.3.14}
$$

□

Finally, we have the following definition:

Definition 1.3.4. When in state j and given that a transition will occur, the probability that the process will transit to state k is given by

$$
\begin{aligned}
p^{jk} &= P[Y_{t+dt} = k | Y_t = j, \ \textit{Transition has occurred}] \\
&= \frac{\lambda^{jk}}{\lambda^{j1} + \ldots + \lambda^{jn}}
\end{aligned} \tag{1.3.15}
$$

for $k = 1, \ldots, n$, so that we also have that $\sum_{k=1}^n p^{jk} = 1$.

The usefulness of these lemmas will soon become apparent. We shall now begin to develop the *Equity Model.*

Chapter 2

The Equity Model

2.1 Introduction

In this section we will begin by introducing the model, and then use the Black-Scholes methodology to price derivatives whose underlying is represented by this model. We will obtain a risk-neutral measure under which our model will be a martingale, find a self-financing replicating strategy, and then finally develop an equation to price the derivatives.

2.2 The Model

Our market contains assets whose price processes are dependent on the state of the Markov chain market described in the previous chapter. Suppose we have an asset whose price process, denoted by S_t (where $S_t \geq 0 \;\; \forall t$), follows a generalised geometric Brownian motion so that

$$dS_t = \mu(Y_{t-})S_t dt + \sum_{b=1}^{r} \sigma_b(Y_{t-})S_t dW_t^b, \tag{2.2.1}$$

where the W_t^b for $b = 1 \dots r$ are independent Brownian motions under the probability measure $\mathbf{P}$, and the drift function $\mu(\cdot)$ and volatility functions $\sigma_b(\cdot)$ are deterministic functions of the state variable Y_{t-}. Note that the mean and drift functions are dependent on $t-$ as they are predictable. The price process of an asset behaving according to this model will therefore have constant drift and volatilities until a transition occurs in our Markov chain.

We so far have the *Markov chain* and the *diffusion* part of our model, and we will now add the *jump* part. We shall add to equation (2.2.1) a jump process as follows. Suppose the Markov chain Y_t has transited from state j to state k. Let γ_t^{jk} be a random variable representing the size of the jump in the asset price due to this transition, such that $\gamma_t^{jk} \geq -1 \ \forall t$. Being dependent on t and not $t-$ means that γ_t will not be predictable. Adding this random variable to (2.2.1) we get

$$dS_t = S_t \left[\mu(Y_{t-})dt + \sum_{b=1}^{r} \sigma_b(Y_{t-})dW_t^b + \sum_{j=1}^{n}\sum_{k=1}^{n} \gamma_t^{jk} dN_t^{jk} \right], \tag{2.2.2}$$

where the counting process N_t^{jk} is defined as in section 1.3.1. This implies that the jumps occur only when N_t^{jk} has increased for any j and k, i.e. a transition of the Markov chain has occurred (which includes a transition to the same state). We shall assume that the jumps and the Brownian motions are independent. Note that forcing γ_t^{jk} to be greater than or equal to -1 ensures that the asset price will never jump to a negative value. We also have that

$$\Pr[S_{t_2} = 0 | S_{t_1} = 0] = 1 \qquad \forall t_2 > t_1,$$

so that once the asset has lost all of its value it can never regain it.

We are now left with the task of assigning a distribution to γ_t^{jk}. We will confine ourselves to using a distribution with a finite event space, because should the event space be infinite we would then need an infinite number of assets to obtain a risk-neutral measure, as shall be seen later on.

Consider a model under which for each transition from state j to k there are l possible jump sizes given by $(\beta_1^{jk}, \dots, \beta_l^{jk})$, where all the jump sizes are finite and greater than -1. So given that a jump from state j to state k occurs, the jump size is represented by γ_t^{jk} which has distribution

$$\gamma_t^{jk} = \begin{cases} \beta_1^{jk} & \text{with probability } p_1^{jk} \\ \vdots & \qquad\qquad \vdots \\ \beta_l^{jk} & \text{with probability } p_l^{jk} \end{cases}$$

for all $j, k \in \mathcal{S}$ and where $\sum_{z=1}^{l} p_z^{jk} = 1 \quad \forall j, k$. With this setup we can replicate practically any distribution for γ_t^{jk} with a suitable choice of l.

We are also able to represent this model in a different way. Our model thus far consists of n states, and for every transition between states there are

l possible jump sizes in the asset price. We could however represent each of these jump sizes as their own state in the market, all of which have the same drift and volatility but only with differing jump sizes. We would then be left with a model where there are $n \times l$ states, with only one possible jump size between any two given states. For the rest of this work, we will therefore consider models where there is only one possible jump size when transiting from state j to state k so that $\gamma_t^{jk} = \gamma^{jk}$. All the results which will be developed will therefore also hold true for models where there are more possible jump sizes between any two states, since this would simply correspond to a model with more states in the market.

Now that we have all the components of our model we can write the following proposition:

Proposition 2.2.1. *Assume that between times $[0, t]$ jumps occur at times $t_1, \ldots, t_{N_t}$. The solution to (2.2.2) is given by the following exponential formula:*

$$\begin{aligned} S_t &= S_0 exp\left\{\int_{s=0}^{t}\left(\mu(Y_{s-}) - \frac{1}{2}\sum_{b=1}^{r}\sigma_b^2(Y_{s-})\right)ds + \int_{s=0}^{t}\sum_{b=1}^{r}\sigma_b(Y_{s-})dW_s^b\right. \\ &\qquad \left. + \int_{s=0}^{t}\sum_{j=1}^{n}\sum_{k=1}^{n} log[1+\gamma^{jk}]dN_s^{jk}\right\}. \end{aligned} \tag{2.2.3}$$

Proof. The result can be obtained by applying the standard Itô formula for the diffusion part, as well as the *exponential formula* of Stieltjes-Lebesgue

Calculus (see Appendix A4 in Brémaud [1981]). Alternatively, it can be obtained from the generalised Itô formula as in Runggaldier [2003] or Applebaum [2004].

□

This model has a lot of appeal because it is describes the manner in which many of the world's financial markets behave: a stable period of fluctuation and drift, followed by a sudden change in the market conditions. This shock to the system causes asset prices to jump and there to be new levels of drift and volatility. It is particularly apt to fit such a model to less liquid markets, where price behavior is generally stable until external stimuli cause temporary shocks to the system, whereby a new equilibrium is reached.

We shall now introduce into our market a numéraire in the form of a bank account process B_t, which grows by a predictable state-dependent interest rate $r(Y_{t-})$. We will assume $r(Y_{t-}) \geq 0 \ \forall t$. This bank-account process has dynamics

$$dB_t = r(Y_{t-})B_t dt,$$

which has solution

$$B_t = exp\left\{\int_{s=0}^{t} r(Y_{s-})ds\right\}$$

with condition $B_0 = 1$.

We can now define the discounted asset price process $\tilde{S}_t$ as

$$\tilde{S}_t = S_t/B_t,$$

which has dynamics

$$d\tilde{S}_t = S_t \left[[\mu(Y_{t-}) - r(Y_{t-})]\, dt + \sum_{b=1}^{r} \sigma_b(Y_{t-}) dW_t^b + \sum_{j=1}^{n}\sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right]. \tag{2.2.4}$$

Since the drift, volatility and interest-rate functions are only dependent of the value of Y_{t-}, we can therefore denote them whilst in state j as being μ_j, $\sigma_{b,j}$ and r_j respectively for all $j \in \mathcal{S}$. We can therefore write equation (2.2.4) more simply as

$$d\tilde{S}_t = S_t \sum_{j=1}^{n} I_t^j \left[(\mu_j - r_j)dt + \sum_{b=1}^{r} \sigma_{b,j} dW_t^b + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right], \tag{2.2.5}$$

where I_t^j is the indicator variable that the market is in state j at time t, and N_t^{jk} is the number of times that the Markov chain has transited from state j to state k as described in section 1.3.1.

We can now write the following corollary:

Corollary 2.2.2. *The dynamics of the discounted asset-price process which* *is the solution of (2.2.5), can be obtained using the same manner as for the undiscounted price process in proposition 2.2.1 to give*

$$\tilde{S}_t = S_0 exp\left\{\int_{s=0}^{t}\sum_{j=1}^{n} I_t^j\left[\left(\mu_j - r_j - \frac{1}{2}\sum_{b=1}^{r}\sigma_{b,j}^2\right)dt + \sum_{b=1}^{r}\sigma_{b,j}dW_t^b + \sum_{k=1}^{n} log[1+\gamma^{jk}]dN_t^{jk}\right]\right\}. \qquad (2.2.6)$$

□

Finally, our market consists of a set of m assets $\mathcal{M} = \{1,\dots,m\}$, whose discounted price processes are all described by equations similar to (2.2.5), although with different drift, volatility and jump sizes. For asset $i \in \mathcal{M}$, let the price process be denoted by $S_{i,t}$ and the drift and volatility functions by $u_{i,j}$ and $\sigma_{i,b,j}$ respectively, as well as the jump sizes by γ_i^{jk} for and all $j,k \in \mathcal{S}$. We can now re-write the discounted asset-price dynamics equation (2.2.5) for all assets $i \in \mathcal{M}$ as

$$d\tilde{S}_{i,t} = \tilde{S}_{i,t}\sum_{j=1}^{n} I_t^j\left[(\mu_{i,j} - r_j)dt + \sum_{b=1}^{r}\sigma_{i,b,j}dW_t^b + \sum_{k=1}^{n}\gamma_i^{jk}dN_t^{jk}\right], \qquad (2.2.7)$$

or alternatively in matrix form

$$d\tilde{\mathbf{S}}_t = \sum_{j=1}^{n} I_t^j \mathbf{D}(\tilde{\mathbf{S}}_t)\left[(\mathbf{U}_j - \mathbf{1}^m r_j)dt + \mathbf{\Sigma}_j d\mathbf{W}_t + \mathbf{\Gamma}_j d\mathbf{N}_t^j\right], \qquad (2.2.8)$$

where $\mathbf{D}(\mathbf{a})$ is an $m \times m$ diagonal matrix when $\mathbf{a} \in \mathbf{R}^m$, with the elements of $\mathbf{a}$ down the principal diagonal and 0 elsewhere. $\mathbf{1}^m$ is an $m \times 1$ column vector of 1's, and we also define the following matrices:

$$
\begin{aligned}
\tilde{\mathbf{S}}_t^{m\times 1} &= \{\tilde{S}_{i,t}\}_{i=1\ldots m} \\
\mathbf{U}_j^{m\times 1} &= \{\mu_{i,j}\}_{i=1\ldots m} \\
\mathbf{\Sigma}_j^{m\times r} &= \{\sigma_{i,b,j}\}_{i=1\ldots m\; b=1\ldots r} \\
\mathbf{W}_t^{r\times 1} &= \{W_t^b\}_{b=1\ldots r} \\
\mathbf{\Gamma}_j^{m\times n} &= \{\gamma_i^{j,k}\}_{i=1\ldots m\; k=1\ldots n} \\
\mathbf{N}_t^{j\ n\times 1} &= \{N_t^{jk}\}_{k=1\ldots n}.
\end{aligned}
$$

2.3 Risk-Neutral Measure

We will now establish the set of price processes which do not permit any arbitrage opportunities. In order to do this, we will first develop a Girsanov-type change of measure. We will then proceed to finding the necessary conditions under which this change of measure is a *martingale measure*, that is a measure under which the discounted asset-price processes in equation (2.2.7) are martingales. It was shown by Dybvig and Huang [1988], as well as Harrison and Kreps [1979] and Harrison and Pliska [1981], that the existence of such a measure is equivalent to a lack of any arbitrage opportunities in a finite-state finite-time economy like ours, by the fundamental theorem of asset pricing.

2.3.1 Change of Measure

The main tool for transforming processes into martingales is Girsanov's theorem (or Cameron-Martin-Girsanov theorem), which is discussed in the context of stochastic differential equations in Øksendal [2000], or for its use in mathematical finance see Bingham and Kiesel [2004] . This is done by inducing a change in the drift of a Wiener process by choosing a suitably different probability measure. We will need to adapt the standard version of this theorem for use in our model, but let us begin by stating this classic theorem for when S_t follows the process defined by equation (2.2.1), that is without any jumps.

Theorem 2.3.1 (Girsanov). *Suppose we have a financial market as described in section 1.3.1 where there are no jumps. The prices of the m assets in this market $\mathbf{S}(t) \in \mathbf{R}^m$ follow an Itô process defined on the probability space (Ω, $\mathcal{F}$,P) of the form*

$$dS_t = \sum_{j=1}^{n} I_t^j \mathbf{D}(\mathbf{S}_t) \left[\mathbf{U}_j dt + \mathbf{\Sigma}_j d\mathbf{W}_t\right], \quad t \leq T, \tag{2.3.1}$$

where the notation is as defined above. Suppose predictable processes $\hat{\mathbf{U}}_j^{m\times 1}$ and $\mathbf{\Theta}_j^{r\times 1}$ exist where

$$\begin{aligned} \hat{\mathbf{U}}_j^{m\times 1} &= \{\hat{\mu}_{i,j}\}_{b=1\ldots m}, \\ \mathbf{\Theta}_j^{r\times 1} &= \{\theta_{b,j}\}_{b=1\ldots r}, \end{aligned}$$

such that

$$\mathbf{\Sigma}_j \mathbf{\Theta}_j = \mathbf{U}_j - \hat{\mathbf{U}}_j, \tag{2.3.2}$$

and where $\mathbf{\Theta}_j$ *satisfies Novikov's condition*

$$E\left[exp\left(\frac{1}{2}\int_0^T \sum_{j=1}^n I_s^j \|\mathbf{\Theta}_j\|^2 ds\right)\right] < \infty.$$

Let

$$dL_t = \sum_{j=1}^n I_t^j L_t \mathbf{\Theta}_j' d\mathbf{W}(t), \quad L_0 = 0, \tag{2.3.3}$$

where

$$E^P[L_t] = 1 \tag{2.3.4}$$

and

$$dQ_t = L_t dP_t$$

on $\mathcal{F}$. *We can now define*

$$\tilde{\mathbf{W}}_t := \int_0^t \sum_{j=1}^n I_s^j \mathbf{\Theta}_j ds + \mathbf{W}_t; \quad t \leq T,$$

where $\tilde{\mathbf{W}}_t$ *is a b-dimensional Brownian motion w.r.t. the equivalent probability measure* $\mathbf{Q}$. *We can now represent the process* $\mathbf{S}_t$ *in terms of this new Brownian motion under* $\mathbf{Q}$ *as follows:*

$$d\mathbf{S}_t = \sum_{j=1}^n I_t^j \mathbf{D}(\mathbf{S}_t)\left[\hat{\mathbf{U}}_j dt + \mathbf{\Sigma}_j d\tilde{\mathbf{W}}(t)\right].$$

□

Note that if $m = b$ and $\boldsymbol{\Sigma}_j^{m \times m}$ is invertible, then the process $\boldsymbol{\Theta}_j$ defined in equation (2.3.2) is uniquely given by

$$\boldsymbol{\Theta}_j = \boldsymbol{\Sigma}_j^{-1} \left[\mathbf{U}_j - \hat{\mathbf{U}}_j \right].$$

We can now set the process $\mathbf{S}_t$ to be a martingale by letting $\hat{\mathbf{U}}_j = \mathbf{0}$ for all $j \in \mathcal{S}$. For the above model to be *complete* we require the existence of a unique martingale measure, and hence a unique $\boldsymbol{\Theta}_j$ for all $j \in \mathcal{S}$. For this to be achieved in the no-jumps case above, we require that the number of assets to be equal to the number of Brownian motions, as well as the invertibility condition above to be fulfilled. Having fewer assets than this means that there will be an infinite number of such measures, and more assets than this could mean no risk-neutral measures. Either case will leave us with a market which is either *incomplete*, or where there potentially exist arbitrage opportunities.

It is interesting to note that the number of states of the Markov chain does not affect the number of assets required for the market to be complete. This is because since there are no jumps in the asset prices, we can regard the process as being a standard generalised geometric Brownian motion with constant drift and volatilities whilst it is in a particular state.

We shall now try and extend the above theorem to include the jump process as in model (2.2.7). To remind ourselves, when in state j the process can transit to any of the other n states or to itself, and for simplicity (although it can easily be generalised) for each such transition we will take there to be only one possible jump size in each of the asset prices. The probability of a transition from state j to state k at any time $t \leq T$ is $\lambda^{jk}dt$.

We will now develop a similar measure transformation as performed above to include the jump process as in Runggaldier [2003] (see also Brémaud [1981] and Björk et al. [1997]):

Theorem 2.3.2. *Consider a financial market described by equation (2.2.7). Let $(\psi^{j1}, \ldots, \psi^{jn})$ be an $\mathcal{F}_t$-predictable process where $\psi^{jk} \geq 0 \ \forall\, j, k \in \mathcal{S}$ so that $\forall\ t \leq T$ we have*

$$\int_{s=0}^{t} \sum_{j=1}^{n} \sum_{k=1}^{n} \psi^{jd} \lambda^{jk} I_s^j ds < \infty.$$

Define

$$L_t = L_t^{(1)} \cdot L_t^{(2)}$$

where $L_t^{(1)}$ satisfies equation (2.3.3) above and $L_t^{(2)}$ is given by

$$dL_t^{(2)} = \sum_{k=1}^{n} (\psi^{jk} - 1) L_{t-}^{(2)} (dN_t^{jk} - \lambda^{jk} dt). \tag{2.3.5}$$

Noting that we can have at most one jump in a time period of length dt, then from equation (1.3.2), where we ignore the negligible term we have

$$\begin{aligned} E\left[dN_t^{jk}\right] &= \Pr[Y_{t+dt} = k | Y_t = j] \\ &= \lambda^{jk} dt. \end{aligned}$$

We therefore also have that

$$E^P\left[L_t^{(2)}\right] = 1, \;\; L_0^{(2)} = 1,$$

which together with equation (2.3.4) gives us

$$E^P\left[L_t\right] = 1,$$

where $L_0 = 1$*, as the jump process and the Brownian motions are assumed to be independent. Using (2.3.3) and (2.3.5) and this independence property, we get for the Radon-Nikodým derivative* L_t

$$dL_t = d(L_t^{(1)} \cdot L_t^{(2)}) = L_{t-}^{(1)} dL_t^{(2)} + L_t^{(2)} dL_t^{(1)},$$

which becomes

$$dL_t = L_t \sum_{j=1}^{n} I_t^j \Theta_j' d\mathbf{W}(t) + L_{t-} \sum_{k=1}^{n} (\psi^{jk} - 1)(dN_t^{jk} - \lambda^{jk} dt),$$

and which can then be solved to give

$$L_t = exp\left\{ \int_{s=0}^{t} \sum_{j=1}^{n} I_s^j \left[\left(\sum_{k=1}^{n} (1 - \psi^{jk})\lambda^{jk} - \frac{1}{2}\|\Theta_j\|^2 \right) ds + \Theta_j' d\mathbf{W}(s) \right] \right\} \prod_{k=1}^{n} (\psi^{jk})^{N_t^{jk}}.$$

We now have an equivalent probability measure $\mathbf{Q}$ *given by*

$$dQ_t = L_t dP_t$$

under which not only do the drifts in each state change from $\mathbf{U}_j$ *to* $\hat{\mathbf{U}}_j$ *such that*

$$\hat{\mathbf{U}}_j = \mathbf{U}_j - \boldsymbol{\Sigma}_j \boldsymbol{\Theta}_j,$$

but the intensities of the jump process also undergo the transformation

$$\lambda^{jk} \rightarrow \lambda^{jk} \psi^{jk} \ \forall\, j, k.$$

□

Using this Girsanov-type transformation we thus have:

Corollary 2.3.3. *Under* $\mathbf{Q}$*, the discounted asset-price processes given in equation (2.2.8) become*

$$d\tilde{\mathbf{S}}_t = \sum_{j=1}^{n} I_t^j \mathbf{D}(\tilde{\mathbf{S}}_t) \left[(\mathbf{U}_j - \boldsymbol{\Sigma}_j \boldsymbol{\Theta}_j - \mathbf{1}^m r_j) dt + \boldsymbol{\Sigma}_j d\tilde{\mathbf{W}}_t + \boldsymbol{\Gamma}_j d\mathbf{N}_t^j \right], \quad (2.3.6)$$

where $\tilde{\mathbf{W}}_t$ *is a b-dimensional* $\mathbf{Q}$ *martingale, and where we also have that*

$$E\left[d\mathbf{N}_t^j \right] = \mathbf{D}(\boldsymbol{\Lambda}_j) \boldsymbol{\Psi}_j dt,$$

where

$$\boldsymbol{\Lambda}_j^{n \times 1} = \{\lambda^{jk}\}_{k=1 \ldots n} \quad (2.3.7)$$

$$\boldsymbol{\Psi}_j^{n \times 1} = \{\psi^{jk}\}_{k=1 \ldots n} \quad (2.3.8)$$

and once again $\mathbf{D}(\mathbf{a})$ *denotes the diagonal matrix with the vector* $\mathbf{a}$ *down the principal diagonal.*

□

2.3.2 Martingale Measure

We will now look at the conditions necessary for $\mathbf{Q}$ to be a *martingale measure*, that is that under $\mathbf{Q}$ we have

$$E^{Q}[d\tilde{\mathbf{S}}_t] = \mathbf{0} \ \forall \, t.$$

Taking expectations of equation (2.3.6) we can see that this condition becomes that for each state j we must have

$$\mathbf{U}_j - \mathbf{\Sigma}_j \mathbf{\Theta}_j - \mathbf{1}^m r_j + \mathbf{\Gamma}_j \mathbf{D}(\mathbf{\Lambda}_j)\mathbf{\Psi}_j = \mathbf{0}. \tag{2.3.9}$$

Let us define the $m \times (r+n)$ augmented matrix $\mathbf{B}_j$ as having entries b_j^{xy} for $x = 1, \ldots, m$, $y = 1, \ldots, (r+n)$ where

$$b_j^{xy} = \begin{cases} -\sigma_{x,y,j} & 1 < y \leq r \\ \gamma_x^{j(y-r)} \lambda^{j(y-r)} & r < y \leq r+n \end{cases}$$

so that

$$\underset{m \times (r+n)}{\mathbf{B}_j} = \left[\begin{array}{c} -\underset{m \times r}{\mathbf{\Sigma}_j} \end{array} \vdots \begin{array}{c} \underset{m \times n}{\mathbf{\Gamma}_j \mathbf{D}(\mathbf{\Lambda}_j)} \end{array} \right]. \tag{2.3.10}$$

Similarly define the $(r+n) \times 1$ augmented column vector $\mathbf{V}_j$ as having entries v_j^x for $x = 1, \ldots, (r+n)$ where

$$v_j^x = \begin{cases} \theta_{x,j} & 1 < x \leq r \\ \psi^{j(x-r)} & r < x \leq r+n \end{cases},$$

so that

$$\underset{(r+n)\times 1}{\mathbf{V}_j} = \begin{bmatrix} \underset{r\times 1}{\mathbf{\Theta}_j} \\ \cdots \\ \underset{n\times 1}{\mathbf{\Psi}_j} \end{bmatrix}.$$

We can now re-write equation (2.3.9) as

$$\mathbf{B}_j\mathbf{V}_j = [\mathbf{1}^m r_j - \mathbf{U}_j].$$

In order for there to be unique values for $\mathbf{\Theta}_j$ and $\mathbf{\Psi}_j$, and hence a unique $\mathbf{V}_j$, we therefore require that for all $j \in \mathcal{S}$ the matrix $\mathbf{B}_j$ be invertible. Given that $\mathbf{B}_j$ is an $m \times (r+n)$ matrix, this will clearly only be possible if we have exactly $r+n$ assets so that $m = r+n$, and that

$$\mathbf{Rank}\,(\mathbf{B}_j) = r+n. \tag{2.3.11}$$

This result is rather intuitive as it is requiring that we have one asset for each source of risk - the r Brownian motions and the n states that the model could jump to. Should the above conditions hold, this would then give us

$$\mathbf{V}_j = \mathbf{B}_j^{-1}[\mathbf{1}^m r_j - \mathbf{U}_j]. \tag{2.3.12}$$

We are then left with a unique martingale measure $\mathbf{Q}$, where by substituting (2.3.9) into (2.3.6) we see that under $\mathbf{Q}$, $\tilde{\mathbf{S}}_t$ follows the process

$$d\tilde{\mathbf{S}}_t = \sum_{j=1}^{n} I_t^j \mathbf{D}(\tilde{\mathbf{S}}_t) \left[\boldsymbol{\Sigma}_j d\tilde{\mathbf{W}}_t + \boldsymbol{\Gamma}_j \left(d\mathbf{N}_t^j - \mathbf{D}(\boldsymbol{\Lambda}_j)\boldsymbol{\Psi}_j dt\right)\right]. \tag{2.3.13}$$

Defining

$$d\tilde{\mathbf{N}}_t^j = d\mathbf{N}_t^j - \mathbf{D}(\boldsymbol{\Lambda}_j)\boldsymbol{\Psi}_j dt,$$

equation (2.3.13) then becomes

$$d\tilde{\mathbf{S}}_t = \sum_{j=1}^{n} I_t^j \mathbf{D}(\tilde{\mathbf{S}}_t) \left[\boldsymbol{\Sigma}_j d\tilde{\mathbf{W}}_t + \boldsymbol{\Gamma}_j d\tilde{\mathbf{N}}_t^j\right], \tag{2.3.14}$$

where $\tilde{\mathbf{W}}_t$ and $\tilde{\mathbf{N}}_t^j$ are both $\mathbf{Q}$ martingales for all $j \in \mathcal{S}$.

We shall now look at an example of this model.

Example 2.3.4. Consider a market where there are two possible states; a *bear market* represented by state 1 in which drifts tend to be lower, and a *bull market* represented by state 2 in which drifts tend to be higher. We shall set the interest rates $r_1 = r_2 = 0.03$. There are three assets in this market, whose price processes have dynamics given by (2.3.6) where there is only one Brownian motion. The drifts and volatilities under $\mathbf{Q}$ given by $\mathbf{U}_j$ and $\boldsymbol{\Sigma}_j$ for $j = 1, 2$, as well as the jump sizes given by $\boldsymbol{\Gamma}_j$ are shown in table 2.1 below. Suppose that we also have the jump intensities between each state (not transition intensities as explained in section 1.3.1) shown in table 2.2.

	Asset 1		Asset 2		Asset 3	
Parameters	**State 1**	**State 2**	**State 1**	**State 2**	**State 1**	**State 2**
Drift	-0.200	0.060	0.065	0.070	0.128	0.020
Volatility	0.070	0.090	0.052	0.100	0.070	0.120

Jump Sizes	Asset 2		Asset 3		Asset 4	
	State 1	**State 2**	**State 1**	**State 2**	**State 1**	**State 2**
State 1	-0.170	0.800	-0.010	0.200	-0.410	0.600
State 2	-0.450	0.600	-0.315	0.270	-0.350	0.390

Table 2.1: Asset parameters and jump sizes under 'real-world' measure.

Using these parameter values and employing equation (2.3.12), we then have a risk-neutral measure where the drifts and the jump intensities are transformed to the figures shown in table 2.3, where we have

$$\Theta_1 = 3.3620$$
$$\Theta_2 = -1.3846.$$

□

Jump Intensity	**State 1**	**State 2**
State 1	1.0000	1.0000
State 2	1.0000	1.0000

Table 2.2: Jump intensities between each state (λ^{jk}).

Drift	**Asset 1**	**Asset 2**	**Asset 3**
State 1	-0.435	-0.111	-0.107
State 2	0.185	0.208	0.186

Jump Intensity	**State 1**	**State 2**
State 1	0.7500	0.7412
State 2	0.9678	0.4682

Table 2.3: Risk-neutral drifts and jump intensities (ψ^{jk}) between each state.

Once the market is complete in the sense that there is a unique martingale-measure, we can invoke the *completeness theorem* which would imply that every contingent claim can be hedged by a self-financing portfolio, since there are only a finite number of jump sizes.

We shall now proceed to show the existence of such a portfolio in our market.

2.4 Replicating Portfolios

Let X_T be a contingent claim at time T, that is an $\mathcal{F}_t-$ measurable random variable with finite expected value. Denote the price process of this claim by

$$C_t = B_t E_Q[B_T^{-1} X_T \mid \mathcal{F}_t]$$

or alternatively the discounted price process

$$\tilde{C}_t = E_Q[B_T^{-1} X_T \mid \mathcal{F}_t] = E_Q[\tilde{X}_T \mid \mathcal{F}_t]. \tag{2.4.1}$$

Suppose that at time t the model is in state j, and we hold a portfolio consisting of $\eta_{t,j}$ units of the cash bond and $\phi^i_{t,j}$ units of asset i for all $i = 1, \ldots, m$. Let $\mathbf{\Phi}_{t,j}$ be the $m \times 1$ vector of asset holdings, so that

$$\mathbf{\Phi}_{t,j} = (\phi^1_{t,j} \cdots \phi^m_{t,j})'.$$

$\mathbf{\Phi}_{t,j}$ and $\eta_{t,j}$ must be $\mathcal{F}_t-$ predictable. The value of the portfolio under this strategy would then be given by

$$V_t^{\Phi} = \sum_{j=1}^{n} I_t^j \left[\eta_{t,j} B_t + \sum_{i=1}^{m} \phi^i_{t,j} S_{i,t} \right].$$

The value of the discounted portfolio is

$$\tilde{V}_t^{\Phi} = \sum_{j=1}^{n} I_t^j \left[\eta_{t,j} + \sum_{i=1}^{m} \phi^i_{t,j} \tilde{S}_{i,t} \right].$$

The strategy $\mathbf{\Phi}$ is said to be *self-financing* if

$$d\tilde{V}_t^{\Phi} = \sum_{j=1}^{n} I_t^j \left[\sum_{i=1}^{m} \phi^i_{t,j} d\tilde{S}_{i,t} \right],$$

or alternatively in matrix form

$$d\tilde{V}_t^{\Phi} = \sum_{j=1}^{n} I_t^j \mathbf{\Phi}'_{t,j} d\tilde{\mathbf{S}}_t. \tag{2.4.2}$$

A contingent claim X_T is said to be *attainable* if there exists some self-financing replicating portfolio $\mathbf{\Phi}$ for which

$$\tilde{V}_T^{\Phi} = \tilde{X}_T.$$

If any such contingent claim is being traded in the market, then in order to avoid arbitrage we must have that the price of this claim and its replicating portfolio are equal, i.e.

$$\tilde{V}_t^{\Phi} = \tilde{C}_t. \tag{2.4.3}$$

By inserting equation (2.4.3) into (2.4.2), we therefore have that in order for the claim to be replicable by a self-financing portfolio we must have that

$$d\tilde{C}_t = \sum_{j=1}^{n} I_t^j \mathbf{\Phi}'_{t,j} d\tilde{\mathbf{S}}_t. \tag{2.4.4}$$

Replacing (2.3.14) into (2.4.4), the strategy $\mathbf{\Phi}$ is a self-financing replicating portfolio of the discounted claim $\tilde{X}_t$ if and only if

$$d\tilde{C}_t = \sum_{j=1}^{n} I_t^j \mathbf{\Phi}'_{t,j} \mathbf{D}(\tilde{\mathbf{S}}_t)[\mathbf{\Sigma}_j d\tilde{\mathbf{W}}_t + \mathbf{\Gamma}_j d\tilde{\mathbf{N}}_t^j]. \tag{2.4.5}$$

Once the values of $\mathbf{\Phi}$ have been obtained, equation (2.4.5) can be used to determine the amount held of the cash bond by

$$\eta_t = \tilde{C}_t - \sum_{j=1}^{n} I_t^j \mathbf{\Phi}'_{t,j} d\tilde{\mathbf{S}}_t.$$

In section 2.5 we will calculate explicit formulas for the asset and bond holdings.

The market is said to be *complete* if every contingent claim is attainable. Noticing that the process (2.4.5) is also a **Q**-martingale, the *n-factor martingale representation theorem* tells us that we can represent the process $\tilde{C}_t$ as

$$\tilde{C}_t = E^Q[C_t] + \int_{s=0}^{t} \sum_{j=1}^{n} \left[\sum_{b=1}^{r} \varepsilon_{s,j}^b d\tilde{W}_s^b + \sum_{k=1}^{n} \zeta_s^{jk} d\tilde{N}_s^{jk} \right],$$

or similarly

$$d\tilde{C}_t = \sum_{j=1}^{n} I_t^j \left[\sum_{b=1}^{r} \varepsilon_{t,j}^b d\tilde{W}_t^b + \sum_{k=1}^{n} \zeta_t^{jk} d\tilde{N}_t^{jk} \right], \tag{2.4.6}$$

where $\varepsilon_{t,j}^b$ and ζ_t^{jk} are $\mathcal{F}_t-$ predictable for all b, j and k, and where $\tilde{W}_t$ and $\tilde{N}_t^{jk}$ are **Q** martingales. Let the $(r+n) \times 1$ column vector $\mathbf{\Pi}_{t,j}$ have entries π_t^x for $x = 1, \dots, (r+n)$, where

$$\pi_t^x = \begin{cases} \varepsilon_{t,j}^x & x \le r \\ \zeta_t^{j(x-r)} & r < x \le r+n. \end{cases}$$

To determine the values of $\varepsilon_{t,j}^b$ and ζ_t^{jk} we can compare coefficients in equation (2.4.5) to those in (2.4.6) to give us for all j

$$\mathbf{\Pi}_{t,j} = \mathbf{G}_j'\mathbf{D}(\tilde{\mathbf{S}}_t)'\mathbf{\Phi}_{t,j}, \tag{2.4.7}$$

where $'$ denotes the transpose, and $\mathbf{G}_j$ is the $m \times (r+n)$ augmented matrix

$$\mathbf{G}_j = \left[\mathbf{\Sigma}_j \ \vdots \ \mathbf{\Gamma}_j \right]. \tag{2.4.8}$$

For there to be a unique replicating portfolio for the contingent claim C_t, we therefore require the existence of a vector $\mathbf{\Phi}_{t,j}$ for every given $\mathbf{\Pi}_{t,j}$ such that the above equation holds. For that to be the case, we require that $\mathbf{\Pi}_{t,j}$ to be in the row space of $\mathbf{G}_j'\mathbf{D}(\tilde{\mathbf{S}}_t)$, or alternatively

$$\mathbf{Rank}\left(\mathbf{D}(\tilde{\mathbf{S}}_t)\mathbf{G}_j\right) = r + n. \tag{2.4.9}$$

As $D(\tilde{\mathbf{S}}_t)$ is an $m \times m$ diagonal matrix it clearly has rank m. When exploring the existence of a risk-neutral measure in section 2.3.2, we required that $m = r + n$, as well as the matrix $\mathbf{B}_j$ given in equation (2.3.10) to be invertible and hence to be of full rank. This therefore necessitates that

$$\begin{aligned} \mathbf{Rank}\left(\mathbf{\Sigma}_j\right) &= r & (2.4.10) \\ \mathbf{Rank}\left(\mathbf{\Gamma}_j\mathbf{D}(\mathbf{\Lambda}_j)\right) &= n. & (2.4.11) \end{aligned}$$

Noting that $\mathbf{D}(\mathbf{\Lambda}_j)$ is an $n \times n$ diagonal matrix and thus has rank n, equation (2.4.11) therefore implies that

$$\mathbf{Rank}\left(\mathbf{\Gamma}_j\right) = n. \tag{2.4.12}$$

Equations (2.4.10) and (2.4.12) therefore also imply that the rank of $\mathbf{G}_j$ is indeed equal to $r+n$. Condition (2.4.9) is therefore satisfied which tells us that a unique solution for $\mathbf{\Phi}_{t,j}$ exists, and hence a unique replicating strategy.

So we can see that completeness in the sense of the existence of a unique martingale measure stipulated in section 2.3.2, necessarily implies completeness in the form of the existence of a unique replicating portfolio of any $\mathcal{F}_T$ contingent claim.

2.5 Derivatives Pricing

Now that we have calculated the conditions necessary in order to have a complete market in which every T-claim can be uniquely replicated, we are left with the task of deriving equations to price them.

From equation (2.3.14) we can write the asset-price process for asset i, $i = 1, \dots m$, as

$$dS_{i,t} = \sum_{j=1}^{n} I_t^j S_{i,t} \left[r_j dt + \sum_{b=1}^{r} \sigma_{i,b,j} d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma_i^{jk} d\tilde{N}_t^{jk} \right]. \qquad (2.5.1)$$

If the market is arbitrage-free and complete, then the price of a contingent claim is given by (2.4.1). Let us consider T-claims represented by X_T that are a function of the state Y_T and the price of the single asset $S_{i,T}$, that is,

$$X_T = f(Y_T, S_{i,T}).$$

We can re-write (2.4.1) as

$$\tilde{C}_t = B_t^{-1} \sum_{j=1}^{n} I_t^j c^j(t, s), \qquad (2.5.2)$$

where

$$c^j(t,s) = B_t E_Q\left[B_T^{-1} X_T \mid Y_t = j\ ,\ S_{i,t} = s\right].$$

Assuming that the functions $c^j(t,s)$ are twice continuously differentiable and recalling that

$$d\tilde{N}_t^{jk} = dN_t^{jk} - \lambda^{jk}\psi^{jk}dt,$$

we can use Itô's lemma on (2.5.2), as well as an analogous lemma for the jump part, to obtain the following equation for whenever the process is in state j:

$$\begin{aligned} d\tilde{C}_t &= B_t^{-1}\left[-r_j c_t^j + r_j s\frac{\partial c_t^j}{\partial s} + \frac{1}{2}\sum_{b=1}^{r}\sigma_{i,b,j}^2 s^2 \frac{\partial^2 c_t^j}{\partial s^2} + \frac{\partial c_t^j}{\partial t}\right] dt \\ &+B_t^{-1}\sum_{b=1}^{r}\sigma_{i,b,j}s\frac{\partial c_t^j}{\partial s}d\tilde{W}_t^b \\ &+B_t^{-1}\sum_{k=1}^{n}[c^k(t,(1+\gamma_i^{jk})s) - c^j(t,s)]d\tilde{N}_t^{jk} \\ &+B_t^{-1}\sum_{k=1}^{n}\left[c^k(t,(1+\gamma_i^{jk})s) - c^j(t,s)\right]\lambda^{jk}\psi^{jk}dt. \end{aligned} \tag{2.5.3}$$

It was shown from equation (2.4.5) that this process is indeed a martingale, and so the drift term vanishes, leaving us with the partial differential equations

$$\begin{aligned} &-r_j c_t^j + r_j s\frac{\partial c_t^j}{\partial s} + \frac{1}{2}s^2\sum_{b=1}^{r}\sigma_{i,b,j}^2 \frac{\partial^2 c_t^j}{\partial s^2} + \frac{\partial c_t^j}{\partial t} \\ &+\sum_{k=1}^{n}\left[c^k(t,s(1+\gamma_i^{jk})) - c^j(t,s)\right]\lambda^{jk}\psi^{jk} = 0 \end{aligned} \tag{2.5.4}$$

for $j = 1 \dots n$. These equations are simply the standard generalised Black-Scholes formulas with an added term for the jumps. Solving these equations for the c^j's with the conditions

$$c^j(T, s) = f(j, s)$$

for $j = 1 \dots n$ will then leave us with the arbitrage price for the derivative. By replacing (2.5.4) into equation (2.5.3) we are left with the following stochastic differential equation when in state j:

$$d\tilde{C}_t = B_t^{-1} \left[\sum_{b=1}^{r} \sigma_{i,b,j} s \frac{\partial c_t^j}{\partial s} d\tilde{W}_t^b + \sum_{k=1}^{n} \left[c^k(t, (1 + \gamma_i^{jk})s) - c^j(t, s) \right] d\tilde{N}_t^{jk} \right]. \tag{2.5.5}$$

We can identify the replicating strategy by comparing the coefficients in (2.5.5) to those in (2.4.5), leaving us with the following equations when in state j:

$$\mathbf{G}_j' \mathbf{D}(\tilde{\mathbf{S}}_t)' \mathbf{\Phi}_{t,j} = \mathbf{Z}_j, \tag{2.5.6}$$

where $\mathbf{G}_j$ is as defined in equation (2.4.8) and $\mathbf{Z}_j$ is the $(r + n) \times 1$ column vector with entries z^x for $x = 1, \dots, (r + n)$, where

$$z^x = \begin{cases} S_{i,t} \sigma_{i,x,j} \frac{\partial c_t^j}{\partial s} & x \le r \\ c^{x-r}(t, (1 + \gamma_i^{j(x-r)}) S_{i,t})) - c^j(t, S_{i,t}) & r < x \le r + n. \end{cases}$$

It was shown in section 2.4 that since condition (2.4.9) holds, we are therefore able to find a unique solution for $\Phi_{t,j}$ from equation (2.5.6) and hence the replicating strategy.

In order to solve the set of stochastic differential equations (2.5.4) we will need to employ numerical methods as will be done in the next chapter. We will now try to the price the derivative using an alternative method.

Suppose that within the interval $[0,T]$ we start off in state j_1 and that there are x jumps. The jump sequence is represented by $\mathcal{J} = (j_1, \ldots, j_{x+1})$, and the jump sizes are given by $(\gamma_i^{j_1 j_2}, \ldots, \gamma_i^{j_x j_{x+1}})$. Conditioning on this jump sequence, and noting that the times at which these jumps occur do not affect the asset price at time T, we can therefore drop the jumps part from equation (2.5.1) to give us in exponential form

$$S_{i,t} = S_{i,0}\, exp\left\{ \int_{u=0}^{t} \sum_{j=1}^{n} I_u^j \left[(r_j - \frac{1}{2}\sum_{b=1}^{r} \sigma_{i,b,j}^2)du + \sum_{b=1}^{r} \sigma_{i,b,j} d\tilde{W}_u^b \right] \right\}, \quad (2.5.7)$$

where $S_{i,0} = s(\gamma_i^{j_1 j_2} \ldots \gamma_i^{j_x j_{x+1}})$. In appendix A corollary A.0.5 we show a methodology to derive the moment-generating function of the final stock price $S_{i,T}$ given the jump sequence $\mathcal{J}$, which will be denoted by $[M_{S_{i,T}}(r)|\mathcal{J}]$.

The probability of observing jump sequence $\mathcal{J}$ within a time T is given in corollary 1.3.3 as being

$$P[\mathcal{J}] = p^{j_1 \cdots j_{x+1}}(T).$$

	Asset 1	
Parameters	**State 1**	**State 2**
Drift	-0.435	0.185
Volatility	0.070	0.090

	Asset 1	
Jump Sizes	**State 1**	**State 2**
State 1	-0.170	0.800
State 2	-0.450	0.600

Table 2.4: Asset 1 parameters and jump sizes in each state.

Summing over all the possible jump sequences and number of jumps, we can calculate the unconditional moment generating function of S_T as

$$\begin{aligned} M_{S_{i,T}}(r) &= E[exp\{rS_{i,T}\}] \\ &= \sum_{x=0}^{\infty}\sum_{j_1=1}^{n}\cdots\sum_{j_{x+1}=1}^{n}[M_{S_{i,T}}(r)|\mathcal{J}]p^{j_1\cdots j_{x+1}}(T). \end{aligned} \quad (2.5.8)$$

Let us now look at an example of this methodology:

Example 2.5.1. Let us consider Asset 1 in example 2.3.4, which has risk-neutral parameters shown in table 2.4.

The interest rates for this model are $r_1 = r_2 = 0.03$, and the risk-neutral jump intensities are as shown in table 2.5.

The moments of the price of this asset after a time of 1 year calculated

	State 1	**State 2**
State 1	0.7500	0.7412
State 2	0.9678	0.4682

Table 2.5: Jump intensities between each state.

Moment	1	2	3	4	5
Value	1.030455	1.590614	4.5790747	29.32724	362.5219
Moment	6	7	8	9	10
Value	6382.791	132195.7	2962521	69352123	1669148415

Table 2.6: Moments of price of asset 1 after 1 year starting in state 1.

using the above methodology, given that at time 0 we are in state 1, are shown in table 2.6.

□

In order to price the derivative we need to calculate the expectation

$$E_Q[B_T^{-1} X_T \,|\mathcal{J}],$$

and for this in turn we will need to find the distribution of $S_{i,T}$ of which X_T is a deterministic function. To this end, we will be able to use the moments of $S_{i,T}$ to approximate its distribution[1]. In the following chapter we shall compare methods of doing this with varying numbers of moments.

However, since B_T^{-1} and $S_{i,T}$ are both dependent of the path of the Markov chain they will therefore be dependent, which means we will also need to

[1]Even though not every distribution may be uniquely determined by its moments (as first shown by Hausdorff [1921]), nevertheless with any set of moments we are still able to approximate its distribution.

derive the distribution of B_T^{-1} conditional on $S_{i,T}$. Rearranging (2.5.7) we get

$$\begin{aligned} B_T^{-1} &= exp\left\{-\int_{u=0}^{T}\sum_{j=1}^{n} I_u^j r_j du\right\} \\ &= \frac{S_{i,0}}{S_T} exp\left\{\int_{u=0}^{t}\sum_{j=1}^{n} I_u^j\left[\frac{1}{2}\sum_{b=1}^{r}\sigma_{i,b,j}^2 du - \sum_{b=1}^{r}\sigma_{i,b,j} d\tilde{W}_u^b\right]\right\}, \end{aligned}$$

and once again using corollary A.0.5 we can derive the moment-generating function of B_T^{-1} conditional on S_T, and then too approximate its distribution. We are now in a position to write the following:

Corollary 2.5.2. *The price of a derivative on asset number i with time 0 price s, where the contingent claim is $X(S_T)$ is given by*

$$c(0,s) = \int_{y=0}^{\infty}\int_{z=0}^{\infty} yX(z) f_{B_T^{-1}|S_{i,T}}(y) f_{S_{i,T}}(z) dy dz$$

where $f_{S_{i,T}}(z)$ and $f_{B_T^{-1}|S_{i,T}}$ are the distributions of the time T stock price and the time T discout rate conditional on the stock price. When the interest rate has constant value r in each state this simplifies to:

$$c(0,s) = e^{-rT}\int_{z=0}^{\infty} X(z) f_{S_{i,T}}(z) dz.$$

□

Even though the above methodology is not of closed form, it does allow for as much accuracy as required depending on how well we are able to approximate the distribution from its moments. How fast it will converge however is a potential problem, as an obvious drawback of the above methodology is the number of calculations that need to be performed in order to calculate the moments in equation (2.5.8), which as can be seen in corollary A.0.5 will be large. So whether or not such calculations can be performed in reasonable time will depend on the value of n, but more importantly the sizes of the transition intensities and the duration T. Also, for the case where the interest rate r is dependent on the state, we have the undesirable requirement that the density function $f_{B_T^{-1}|S_T}$ will need to be calculated for all values of $S_{i,T}$.

2.5.1 Interest-Rate Derivatives

We shall be considering more elaborate interest rate derivatives when we look at *Term-Structure Models* in Chapter 4. However, within our current framework we can calculate an explicit formula for a simple class of interest rate derivatives.

Consider an asset with price V_t such that

$$V_t = E\left[x_1\frac{B_t}{B_{t_1}} + x_2\frac{B_t}{B_{t_2}} + \ldots + x_h\frac{B_t}{B_{t_h}}\right]. \tag{2.5.9}$$

This payout represents that of a coupon bond which pays h coupons of value £x_1 at time t_1, £x_2 at time t_2 and so on, where we have that $t_1 < t_2 < \ldots < t_h$ and that $t < t_1$. The final coupon at time t_h will normally also include the nominal amount of the bond. The value of this bond is calculated by taking the expectation of the sum of the discounted coupon payments to the current time t. Let us denote the value of this bond given that we are currently in state j by V_t^j. We therefore have

$$V_t^j = \sum_{u=1}^{h} x_u E\left[exp\left\{-\int_{s=t}^{t_u} \sum_{k=1}^{n} I_s^k r_k ds\right\} \mid Y_t = j\right]. \qquad (2.5.10)$$

Applying Taylor's expansion to (2.5.10) we get

$$\begin{aligned} V_t^j &= \sum_{u=1}^{h} x_u \left\{E\left[\sum_{y=0}^{\infty}\left(-\int_{s=t}^{t_u}\sum_{k=1}^{n} I_s^k r_k ds\right)^y \mid Y_t = j\right]\right\} \\ &= \sum_{u=1}^{h} x_u \left\{\sum_{y=0}^{\infty}\left(-\sum_{k=1}^{n} r_k P_{t_u-t}^{jk}\right)^y\right\}, \end{aligned} \qquad (2.5.11)$$

where P_t^{jk} is defined in equation (1.3.6). For (2.5.11) to converge we require that

$$r_k P_{t_u-t}^{jk} < 1 \;\; \forall\, u, k.$$

Even though (2.5.11) is not of closed form, it does allow for as much accuracy as required by summing to a suitable value of y.

We will now try to price a call option on this bond at time T where $t_a < T < t_{a+1}$ such that $0 \le a < h$ and where $t_0 = t$. The value of the bond at time T will therefore be the expect discounted value of the remaining $h - a$ coupons. The strike price of this option is K, and we will denote the discounted value of the option by $\tilde{C}_t$, so that

$$\tilde{C}_t = B_t^{-1} \sum_{j=1}^{n} I_t^j c_t^j \tag{2.5.12}$$

where

$$c_t^j = B_t E[B_T^{-1} X_T | \, Y_t = j], \tag{2.5.13}$$

as well as

$$X_T = \sum_{j=1}^{n} I_T^j X_T^j$$

and

$$X_T^j = \max[V_T^j - K, 0].$$

Applying Itô's lemma to (2.5.12), we find

$$d\tilde{C}_t = B_t^{-1} \sum_{j=1}^{n} I_t^j \left[-r^j c_t^j + \frac{dc_t^j}{\partial t} + \sum_{k=1}^{n} (c_t^k - c_t^j) \lambda^{jk} \psi^{jk} \right] dt. \tag{2.5.14}$$

For $\tilde{C}_t$ to be a martingale we therefore require the right-hand side of (2.5.14) to be equal to zero, so that

$$\begin{aligned}\frac{\partial c_t^j}{\partial t} &= r_j c_t^j - \sum_{k=1}^{n}(c_t^k - c_t^j)\lambda^{jk}\psi^{jk} \\ &= (r_j + \lambda^{jj}\psi^{jj})c_t^j - \sum_{k=1}^{n} c_t^k \lambda^{jk}\psi^{jk}\end{aligned}$$

for $j = 1, \ldots n$, subject to $c_T^j = X_T^j$. Writing in matrix form, and noting that c_t^j is now only a function of t and hence it is not necessary to use partial derivatives, we therefore get that

$$\frac{d\mathbf{C}_t}{dt} = (\mathbf{D}(\mathbf{R}) - \mathbf{\Lambda}_j \mathbf{\Psi}_j')\mathbf{C}_t$$

with side condition

$$\mathbf{C}_T = \mathbf{X}_T,$$

where $\mathbf{\Lambda}_j$ and $\mathbf{\Psi}_j$ are defined in equations (2.3.7) and (2.3.8), and $\mathbf{1}^n$ is an $n \times 1$ column vector of 1's. $\mathbf{D}(\mathbf{R})$ once again represents the diagonal matrix with the elements of $\mathbf{R}$ along the principal diagonal, where $\mathbf{R}$ has entries

$$\mathbf{R} = \{r_j + \lambda^{jj}\psi^{jj}\}_{j=1\ldots n}.$$

We can solve this to get

$$\mathbf{C}_t = exp\{(\mathbf{\Lambda}_j \mathbf{\Psi}_j' - \mathbf{D}(\mathbf{R}))(T-t)\}\mathbf{X}_T. \qquad (2.5.15)$$

Chapter 3

Numerical Methods for the Equity Model

3.1 Introduction

In chapter 2 we looked at assets with price processes given by

$$dS_t = S_t \sum_{j=1}^{n} I_t^j \left[\mu_j dt + \sum_{b=1}^{r} \sigma_{b,j} dW_t^b + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right], \tag{3.1.1}$$

where we have state-dependent means and drifts μ_j and $\sigma_{b,j}$, and jump sizes γ^{jk} when the model jumps from state j to state k. We saw in section 2.5 that in order to price most derivatives of assets whose price processes follow this MCJD model, it is necessary to employ numerical methods. In this chapter we shall look at various methods of doing this. We shall begin by trying to numerically solve the partial differential equations given in (2.5.4) using finite-difference methods. We will then move on to consider parametric methods to approximate the distribution of S_T to price the derivatives, followed by tree-based methods and Monte Carlo simulation.

In the numerical work of this chapter, we will only be considering derivatives of one asset, and we therefore suppress the subscript i in equation (3.1.1) which was used in the previous chapter to denote the asset number. We can also reduce the generalised model in equation (3.1.1) to the following:

$$dS_t = S_t \sum_{j=1}^{n} I_t^j \left[\mu_j dt + \sigma_j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right], \tag{3.1.2}$$

where

$$\sigma_j = \sqrt{\sum_{b=1}^{r} \sigma_{b,j}^2},$$

so that we then have

$$\sigma_j dW_t \sim \sum_{b=1}^{r} \sigma_{b,j} dW_t^b,$$

where W_t is a $\mathbf{Q}$ Brownian motion. We can do this because the generalised model is needed only when we consider the relationship between different assets, but since we will now only be considering the behaviour of the price of one individual asset, the generalised model can be reduced to equation (3.1.2). For the rest of the chapter we shall assume that the parameter val-

	Asset 1	
Parameters	**State 1**	**State 2**
Drift	-0.435	0.185
Volatility	0.070	0.090

	Asset 1	
Jump Sizes	**State 1**	**State 2**
State 1	-0.170	0.800
State 2	-0.450	0.600

Table 3.1: Asset 1 parameters and jump sizes in each state.

ues of (3.1.2) are under a risk-neutral measure $\mathbf{Q}$, the existence of which was explored in section 2.3.2.

We shall be comparing the performance of each of the numerical methods on a particular example, the summary of which can be found at the end of the chapter in section 3.6. The example we will be calculating is the price of call options with strike price K, such that the payoff at time T is given by

$$X_T = \max[S_T - K, 0],$$

where we will take $T = 1$ year. We will assume that the model is in state 1 at time 0 and that $S_0 = 1$. For the asset-price process we will take Asset 1 in example 2.3.4, which had risk-neutral parameters in the two state market shown in table 3.1.

Jump Intensity	**State 1**	**State 2**
State 1	0.7500	0.7412
State 2	0.9678	0.4682

Table 3.2: Jump intensities between each state.

The interest rates for this model are $r_1 = r_2 = 0.03$, and the risk-neutral jump intensities are as shown in table 3.2. The moments of S_T where $T =$ 1 year were calculated in example 2.5.1.

3.2 Finite-Difference Methods

In section 2.5 we derived the partial differential equation for pricing a derivative whose discounted price is given by $\tilde{C}_t$, where

$$\tilde{C}_t = B_t^{-1} \sum_{j=1}^{n} I_t^j c^j(t,s)$$

and

$$c^j(t,s) = B_t E_Q \left[B_T^{-1} X_T \mid Y_t = j \ , \ S_{i,t} = s \right] .$$

To obtain the price of this derivative, it was shown in section 2.5 that we need to solve the following partial differential equation (where will abbreviate $c^j(t,s)$ to c_t^j):

$$\begin{aligned} -r_j c_t^j + r_j s \frac{\partial c_t^j}{\partial s} + \frac{1}{2} s^2 \sigma_j^2 \frac{\partial^2 c_t^j}{\partial s^2} + \frac{\partial c_t^j}{\partial t} \\ + \sum_{k=1}^{n} \left[c^k(t, s(1+\gamma_i^{jk})) - c^j(t,s) \right] \lambda^{jk} \psi^{jk} = 0 \qquad (3.2.1) \end{aligned}$$

for $j = 1, \ldots, n$, with the conditions

$$c^j(T,s) = \max\left[s - K, 0\right] .$$

We will now proceed to develop a numerical technique for solving this partial differential equation using *finite-difference methods.* As was mentioned in section 3.1, we will be valuing a time T call option with strike K. We shall follow the methodology set out in Hull and White [1990b].

We begin by dividing up the interval $[0, T]$ into h equal periods of length Δt so that $\Delta t = \frac{T}{h}$, and so we can now consider our interval as consisting of $h + 1$ discrete time points

$$0, \Delta t, 2\Delta t, \ldots, T.$$

Let us assume that we can subdivide the possible stock prices over this period into d values as follows. Firstly, for this model we need to specify the maximum value we will allow the stock price to have. Denote by S_{max} the maximum realistic value that the stock can take during this interval, so that the probability

$$P[S_T > S_{max}]$$

is so small that if we were to exclude the possibility that S_T is above this value, the effect on the value of the derivative would be negligible. Let us now set

$$\Delta S = \frac{S_{max}}{d},$$

so that we now are left with the $d + 1$ possible values for the stock

$$0, \Delta S, 2\Delta S, \ldots, S_{max}.$$

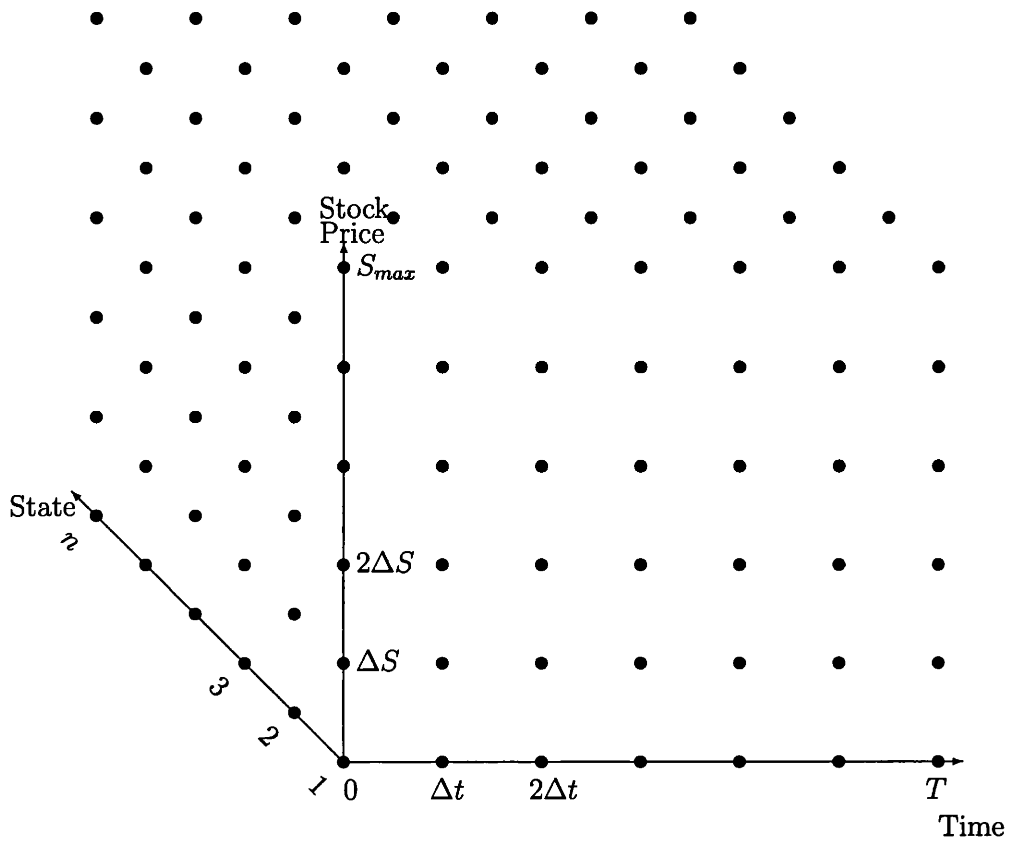

Figure 3.1: Cube for finite-difference approach

Finally, we recall that the derivative price will also depend on which of the n states the market is currently in. We can represent this by the three-dimensional diagram in figure 3.1. This cube consists of $(h+1) \times (d+1) \times n$ points, each of which we can label as point (x, y, z), where $x = 0, \ldots, h$, $y = 0, \ldots, d$ and $z = 1, \ldots, n$. Therefore x represents the time, y the state, and z represents the stock price. We will denote the value of the call option at each of these points by $c^z_{x,y}$.

There are two ways we can now proceed, the *implicit finite-difference method* and the *explicit finite-difference method.* We choose to follow the *implicit finite-difference method,* as fewer assumptions are made about the values of the partial derivatives.

Let us firstly begin by stating the fact that when the stock price hits 0 it cannot ever regain any value, and so the value of the option is therefore going to equal 0. For any other point (x, y, z) on the grid where $y > 0$, we can make the following approximation for the partial derivative $\frac{\partial c_{x,y}^z}{\partial s}$:

$$\frac{\partial c_{x,y}^z}{\partial s} = \frac{c_{x,y+1}^z - c_{x,y}^z}{\Delta S}, \tag{3.2.2}$$

where (3.2.2) is known as the *forward-difference approximation.* When $y = d$ the stock price has reached its maximum realistic value S_{max} and can therefore not increase, and so in this case we shall set

$$c_{x,y+1}^z = c_{x,y}^z.$$

We may use as an alternative to the *forward-difference approximation* the *backward-difference approximation* given by (3.2.3) below

$$\frac{\partial c_{x,y}^z}{\partial s} = \frac{c_{x,y}^z - c_{x,y-1}^z}{\Delta S}. \tag{3.2.3}$$

A third alternative would be to use an average of the two to give

$$\frac{\partial c^z_{x,y}}{\partial s} = \frac{c^z_{x,y+1} - c^z_{x,y-1}}{2\Delta S}. \tag{3.2.4}$$

We can similarly approximate $\frac{\partial c^z_{x,y}}{\partial t}$ using the forward differential equation

$$\frac{\partial c^z_{x,y}}{\partial t} = \frac{c^z_{x+1,y} - c^z_{x,y}}{\Delta t}. \tag{3.2.5}$$

The backward-difference equation for $\frac{\partial c^z_{x,y}}{\partial s}$ at the point (x, y, z) is given by (3.2.3). The backward difference value at the point $(x, y + 1, z)$ is given by

$$\frac{\partial c^z_{x,y+1}}{\partial s} = \frac{c^z_{x,y+1} - c^z_{x,y}}{\Delta S}.$$

We can therefore write the forward difference equation to approximate $\frac{\partial^2 c^z}{\partial s^2}$ at the point (x, y, z) as

$$\frac{\partial^2 c^z_{x,y}}{\partial s^2} = \left(\frac{c^z_{x,y+1} - c^z_{x,y}}{\Delta S} - \frac{c^z_{x,y} - c^z_{x,y-1})}{\Delta S} \right) / \Delta S,$$

which can be re-written as

$$\frac{\partial^2 c^z_{x,y}}{\partial s^2} = \frac{c^z_{x,y+1} + c^z_{x,y-1} - 2c^z_{x,y}}{\Delta S^2}. \tag{3.2.6}$$

Substituting these approximations into (3.2.1), and noting that $s = y\Delta S$ we then get

$$\begin{aligned}
-r_z c^z_{x,y} + r_z y\Delta S \frac{c^z_{x,y+1} - c^z_{x,y-1}}{2\Delta S} + \frac{1}{2}(y\Delta S)^2 \sigma_z^2 \frac{c^z_{x,y+1} + c^z_{x,y-1} - 2c^z_{x,y}}{\Delta S^2} & \\
+ \frac{c^z_{x+1,y} - c^z_{x,y}}{\Delta t} + \sum_{k=1}^{n} [c^k_x(t, y\Delta S(1+\gamma^{zk})) - c^z_{x,y}]\lambda^{zk}\psi^{zk} & = 0.
\end{aligned} \tag{3.2.7}$$

Using this methodology, we will need to represent any jump in the stock price as a vertical jump in the grid in figure 3.1. Therefore, we will assume that we have values q^{zk}_y for all y, z, k, such that we can make the following approximation:

$$y\Delta S(1+\gamma^{zk}) = q^{zk}_y \Delta S,$$

i.e. the effect of a jump from state z to state k is to cause the value of the stock to jump from $y\Delta S$ to $q^{zk}_y \Delta S$ where $q^{zk}_y \in \{0, 1, \dots, d\}$. We can thus re-write (3.2.7) as

$$\begin{aligned}
\frac{1}{2} y r_z (c^z_{x,y+1} - c^z_{x,y-1}) + \frac{1}{2} y^2 \sigma_z^2 \left(c^z_{x,y+1} + c^z_{x,y-1} - 2c^z_{x,y}\right) & \\
\frac{c^z_{x+1,y} - c^z_{x,y}}{\Delta t} - c^z_{x,y}(r_z + \sum_{k=1}^{n} \lambda^{zk}\psi^{zk}) + \sum_{k=1}^{n} c^k_{x,q^{zk}_y} \lambda^{zk}\psi^{zk} & = 0
\end{aligned} \tag{3.2.8}$$

for $x = 0, \dots, h$, $y = 1, \dots, d$ and $z = 1, \dots, n$. Equation (3.2.8) can be written as

$$c^z_{x+1,y} = a^z_y c^z_{x,y} + b^z_y c^z_{x,y+1} + e^z_y c^z_{x,y-1} + \Delta t \sum_{k=1}^{n} c^k_{x,q^{zk}_y} \lambda^{zk} \psi^{zk}, \tag{3.2.9}$$

where

$$\begin{aligned} a^z_y &= 1 + \Delta t(y^2\sigma^2_z + r_z + \sum_{k=1}^{n} \lambda^{zk}\psi^{zk}), \\ b^z_y &= -\frac{1}{2}\Delta t(y^2\sigma^2_z + yr_z), \\ e^z_y &= \frac{1}{2}\Delta t(yr_z - y^2\sigma^2_z). \end{aligned}$$

We have now set up the framework of the implicit finite-difference method and will proceed to price the derivative.

We shall begin by gathering all the values of the grid in figure 3.1 that we know from the outset. At time T (i.e. $x = h$) we have that

$$c^z_{h,y} = \max[y\Delta S - K, 0] \quad \forall y, z\,, \tag{3.2.10}$$

which is known at time 0. We also know that when the stock price has hit 0 it cannot change, hence we have that

$$c^z_{x,0} = 0 \quad \forall x, z. \tag{3.2.11}$$

The values known at time 0 are shown in figure 3.2. Equation (3.2.10) gives us the outer vertical plane of the cube, whilst (3.2.11) gives us the values

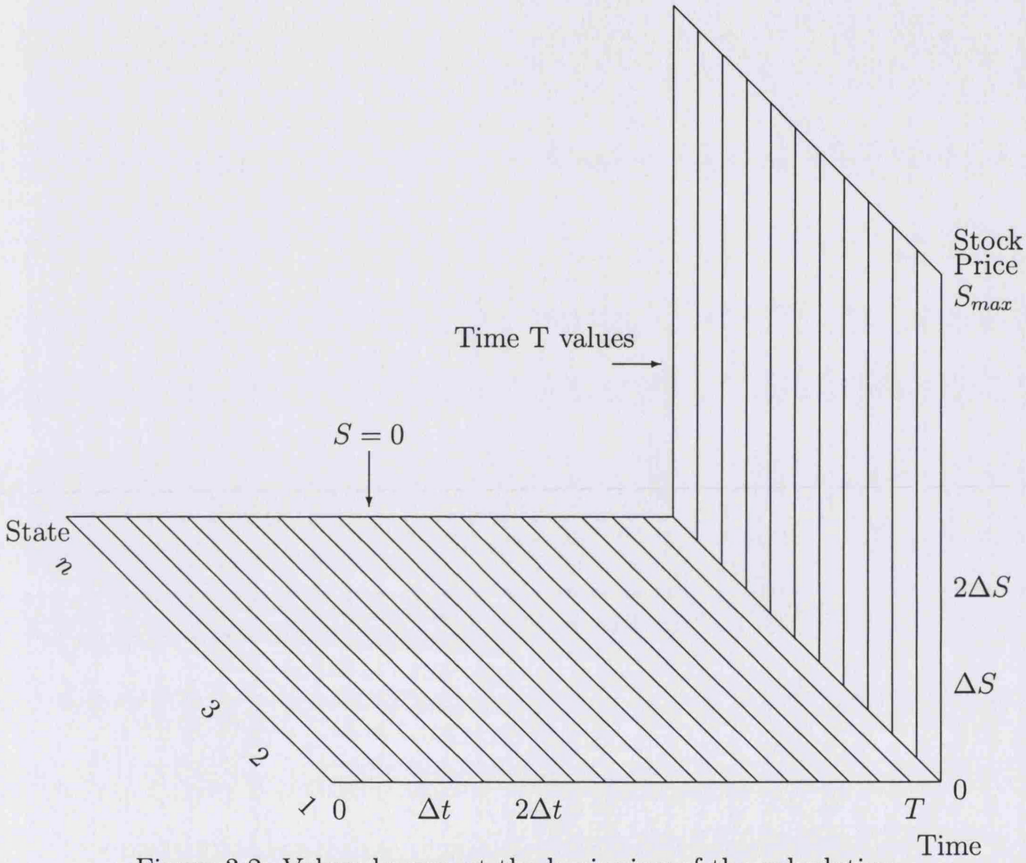

Figure 3.2: Values known at the beginning of the calculation

on the base of our cube. Using these starting values, we are now able to calculate all the values of the cube as follows. Firstly, we can calculate all values of the cube where $x = (h-1)\Delta t$ as shown in figure 3.3, by using equation (3.2.9) and setting $x = h - 1$. We thus have

$$c^z_{h,y} = a^z_y c^z_{h-1,y} + b^z_y c^z_{h-1,y+1} + e^z_y c^z_{h-1,y-1} + \Delta t \sum_{k=1}^{n} c^k_{h-1,q^{zk}_y} \lambda^{zk} \psi^{zk} \qquad (3.2.12)$$

for $y = 1, \ldots, d$ and $z = 1, \ldots, n$. We therefore have $d \times n$ unknowns and

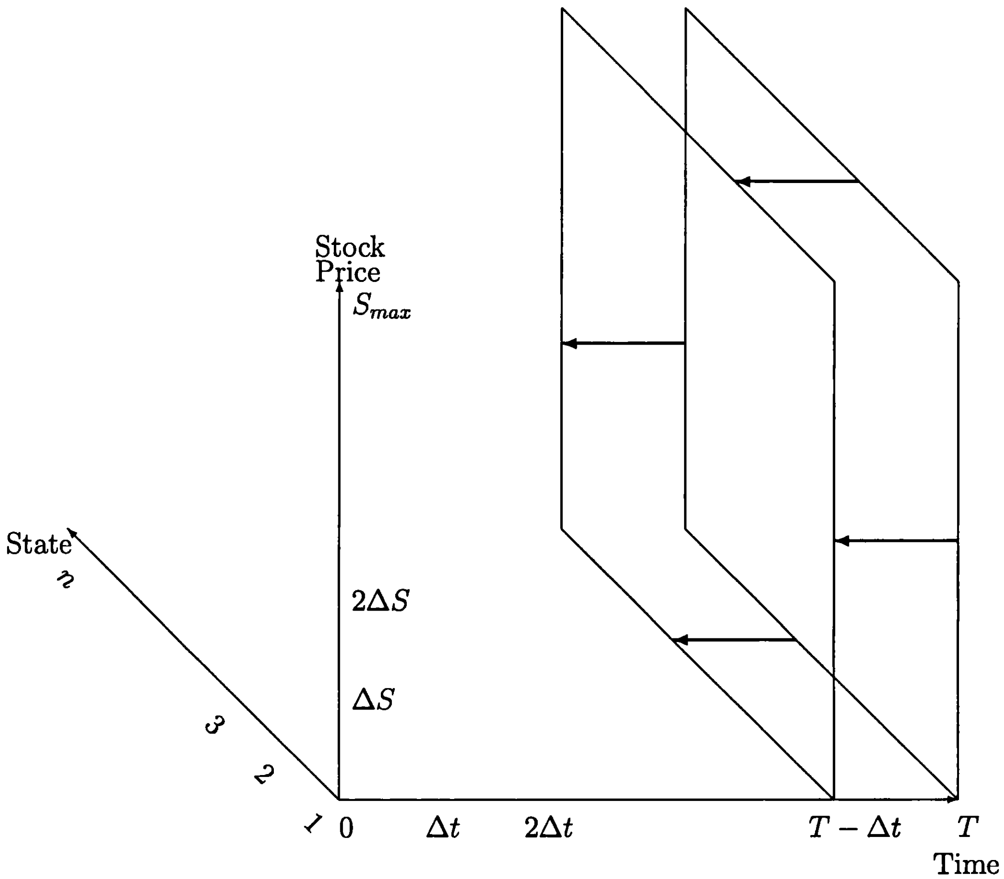

Figure 3.3: Estimating values at time $T - \Delta t$ from values at time T.

$d \times n$ equations with which to solve for them, which should provide a unique solution providing the equations are linearly independent. Once this has been done, we can then do the same to calculate the values for $x = h - 2$, and then for $x = h - 3$ and so on until all the values at time $x = 0$ are calculated. The value of the derivative would then depend on selecting the appropriate time 0 value for the starting state.

We can re-write (3.2.12) in matrix form as follows. Let us define the

$(d+1)\times n$ column vector $\mathbf{C}_h$ with entries

$$\mathbf{C}_h = (0, c^1_{h,1}, c^1_{h,2}, \dots, c^1_{h,d}, 0, c^2_{h,1}, \dots, c^2_{h,d}, \dots, c^n_{h,d}).$$

Let us also define the following $(d+1)n \times (d+1)n$ square matrices $\mathbf{A}$ and $\mathbf{B}$ with entries

$$\underset{(d+1)n\times(d+1)n}{\mathbf{A}} = \begin{pmatrix}
1 & 0 & & & & & & & & & \\
e^1_1 & a^1_1 & b^1_1 & & & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & & & \\
 & & e^1_{d-1} & a^1_{d-1} & b^1_{d-1} & & & & & & \\
 & & & e^1_d & a^1_d + b^1_d & 0 & & & & & \\
 & & & & 0 & 1 & 0 & & & & \\
 & & & & & e^2_1 & a^2_1 & b^2_1 & & & \\
 & & & & & & \ddots & \ddots & \ddots & & \\
 & & & & & & & e^n_{d-1} & a^n_{d-1} & b^n_{d-1} \\
 & & & & & & & & e^1_d & a^1_d + b^1_d
\end{pmatrix},$$

$$\underset{(d+1)n\times(d+1)n}{\mathbf{B}} = \begin{pmatrix}
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 \\
i^{11}_{11} & i^{11}_{12} & \dots & i^{11}_{1d} & i^{12}_{11} & i^{12}_{12} & \dots & i^{12}_{1d} & \dots & \dots & i^{1n}_{1d} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
i^{11}_{d1} & i^{11}_{d2} & \dots & i^{11}_{dd} & i^{12}_{d1} & i^{12}_{d2} & \dots & i^{12}_{dd} & \dots & \dots & i^{1n}_{dd} \\
0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & \dots & 0 \\
i^{21}_{11} & i^{21}_{12} & \dots & i^{21}_{1d} & i^{22}_{11} & i^{22}_{12} & \dots & i^{22}_{1d} & \dots & \dots & i^{2n}_{1d} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
i^{n1}_{d1} & i^{n1}_{d2} & \dots & i^{n1}_{dd} & i^{n2}_{d1} & i^{n2}_{d2} & \dots & i^{n2}_{dd} & \dots & \dots & i^{nn}_{dd}
\end{pmatrix},$$

where entries other than those shown are 0, and where

$$i_{yw}^{zk} = \begin{cases} \lambda^{zk}\psi^{zk}\Delta t & \text{if } q_y^{zk} = w \\ 0 & \text{otherwise.} \end{cases}$$

Finally, let us define the $dn \times dn$ matrix $\mathbf{K}$ as

$$\mathbf{K} = \mathbf{A} + \mathbf{B}.$$

We can now write equation (3.2.12) for $y = 1, \ldots, d$ and $z = 1, \ldots n$ in matrix form as

$$\mathbf{C}_h = \mathbf{K}\mathbf{C}_{h-1},$$

and so given the derivative prices for time period $h\Delta t$ we are able to calculate the time $(h-1)\Delta t$ values using the following recursion equation:

$$\mathbf{C}_{h-1} = \mathbf{K}^{-1}\mathbf{C}_h, \qquad (3.2.13)$$

provided that $\mathbf{K}$ is indeed invertible.

We shall now look at an example of this method.

Example 3.2.1. Suppose we wish to price call options on Asset 1 whose parameters are given in section 3.1. We shall take $d = 25$ and $\Delta s = 0.2$ so that $S_{max} = 5$, which is over 5 standard deviations above the mean. We will also set $h = 12$ so that $\Delta t = 0.0833$. The value of call options with various strike prices are shown in table 3.12 in section 3.6. The value of the call option with strike price £1 at each time period and state are shown in the tables 3.3 and 3.4 below. Reading off the tables, this method gives us the time 0 price of a call option with strike £1 as being £0.24 in state 1 and £0.21 in state 2. We can see the irregular behaviour of the option price when the price of the stock is above £4. This is due to the capping of the stock price at £5. This model's usefulness is therefore limited to stock prices of under £4. When the stock price is greater than this we will need to raise the maximum stock price allowed in the model.

□

Stock Price (£)	1	2	3	4	5	6	7	8	9	10	11	12	
5.0	6.26	6.54	6.84	7.14	7.43	7.70	7.92	8.03	7.97	7.64	6.93	5.74	4.00
4.8	0.24	0.26	0.28	0.32	0.40	0.52	0.71	1.00	1.42	1.97	2.62	3.29	3.80
4.6	3.42	3.56	3.69	3.81	3.91	3.97	4.00	3.97	3.89	3.78	3.67	3.60	3.60
4.4	1.53	1.61	1.70	1.82	1.96	2.13	2.32	2.54	2.76	2.97	3.16	3.31	3.40
4.2	2.46	2.53	2.61	2.67	2.74	2.79	2.85	2.90	2.95	3.01	3.08	3.14	3.20
4.0	1.83	1.92	2.01	2.12	2.23	2.35	2.47	2.59	2.70	2.80	2.89	2.96	3.00
3.8	2.04	2.09	2.15	2.22	2.28	2.35	2.42	2.50	2.57	2.65	2.71	2.77	2.80
3.6	1.82	1.89	1.96	2.03	2.11	2.18	2.26	2.34	2.41	2.48	2.54	2.58	2.60
3.4	1.82	1.87	1.92	1.97	2.03	2.08	2.14	2.20	2.26	2.32	2.36	2.39	2.40
3.2	1.71	1.76	1.81	1.87	1.92	1.97	2.02	2.07	2.11	2.15	2.19	2.20	2.20
3.0	1.66	1.70	1.74	1.78	1.82	1.86	1.90	1.94	1.97	2.00	2.01	2.02	2.00
2.8	1.61	1.65	1.69	1.73	1.76	1.79	1.82	1.84	1.85	1.86	1.86	1.84	1.80
2.6	1.53	1.57	1.61	1.65	1.68	1.71	1.72	1.73	1.73	1.72	1.69	1.65	1.60
2.4	1.35	1.38	1.41	1.44	1.46	1.48	1.50	1.51	1.51	1.50	1.48	1.45	1.40
2.2	1.21	1.23	1.25	1.26	1.28	1.29	1.29	1.29	1.29	1.28	1.26	1.24	1.20
2.0	1.02	1.03	1.05	1.06	1.06	1.07	1.07	1.08	1.07	1.06	1.05	1.03	1.00
1.8	0.86	0.87	0.87	0.87	0.87	0.87	0.87	0.87	0.86	0.85	0.84	0.82	0.80
1.6	0.74	0.74	0.74	0.73	0.73	0.72	0.71	0.70	0.68	0.66	0.64	0.62	0.60
1.4	0.54	0.54	0.53	0.52	0.52	0.51	0.50	0.49	0.47	0.46	0.44	0.42	0.40
1.2	0.38	0.37	0.35	0.34	0.33	0.31	0.29	0.28	0.26	0.25	0.23	0.21	0.20
1.0	0.24	0.23	0.22	0.21	0.19	0.18	0.16	0.14	0.11	0.09	0.06	0.03	0.00
0.8	0.13	0.13	0.12	0.12	0.11	0.10	0.09	0.08	0.07	0.06	0.04	0.02	0.00
0.6	0.04	0.04	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00
0.4	0.03	0.03	0.02	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00
0.2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Time (months)

Table 3.3: Call option prices with strike £1 for state 1.

Stock Price (£)	1	2	3	4	5	6	7	8	9	10	11	12	
5.0	0.78	0.85	0.93	1.03	1.15	1.29	1.47	1.69	1.97	2.32	2.75	3.30	4.00
4.8	1.07	1.16	1.27	1.40	1.56	1.75	1.96	2.21	2.50	2.82	3.17	3.51	3.80
4.6	1.32	1.43	1.57	1.72	1.88	2.07	2.28	2.51	2.75	2.99	3.22	3.43	3.60
4.4	1.49	1.61	1.74	1.89	2.05	2.22	2.39	2.58	2.76	2.94	3.11	3.26	3.40
4.2	1.61	1.72	1.84	1.97	2.11	2.25	2.40	2.54	2.69	2.83	2.96	3.09	3.20
4.0	1.64	1.74	1.84	1.95	2.07	2.19	2.31	2.43	2.55	2.67	2.78	2.90	3.00
3.8	1.59	1.67	1.77	1.86	1.96	2.06	2.16	2.27	2.38	2.49	2.60	2.70	2.80
3.6	1.53	1.61	1.69	1.77	1.86	1.95	2.04	2.14	2.24	2.34	2.43	2.52	2.60
3.4	1.43	1.50	1.57	1.64	1.72	1.80	1.89	1.98	2.07	2.16	2.25	2.33	2.40
3.2	1.34	1.40	1.47	1.54	1.61	1.69	1.77	1.85	1.93	2.01	2.08	2.15	2.20
3.0	1.28	1.33	1.39	1.46	1.53	1.60	1.67	1.74	1.80	1.86	1.92	1.96	2.00
2.8	1.28	1.33	1.38	1.44	1.49	1.55	1.59	1.64	1.68	1.72	1.75	1.78	1.80
2.6	1.21	1.25	1.29	1.33	1.37	1.41	1.45	1.48	1.51	1.54	1.56	1.58	1.60
2.4	1.13	1.16	1.19	1.22	1.25	1.27	1.30	1.32	1.34	1.36	1.37	1.39	1.40
2.2	0.99	1.02	1.04	1.06	1.08	1.10	1.11	1.13	1.15	1.16	1.18	1.19	1.20
2.0	0.87	0.89	0.90	0.92	0.93	0.94	0.96	0.97	0.98	0.99	0.99	1.00	1.00
1.8	0.73	0.74	0.74	0.75	0.76	0.76	0.77	0.77	0.78	0.78	0.79	0.79	0.80
1.6	0.59	0.60	0.60	0.61	0.61	0.61	0.62	0.62	0.61	0.61	0.61	0.60	0.60
1.4	0.47	0.47	0.47	0.47	0.46	0.46	0.46	0.45	0.44	0.43	0.42	0.41	0.40
1.2	0.34	0.34	0.34	0.33	0.32	0.32	0.31	0.30	0.28	0.27	0.25	0.23	0.20
1.0	0.21	0.20	0.19	0.18	0.17	0.15	0.14	0.12	0.10	0.08	0.06	0.03	0.00
0.8	0.10	0.10	0.09	0.08	0.07	0.06	0.06	0.05	0.04	0.03	0.02	0.01	0.00
0.6	0.05	0.05	0.04	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.00	0.00	0.00
0.4	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Time (months)

Table 3.4: Call option prices with strike £1 for state 2.

3.3 Parametric methods

We shall now develop methods for pricing derivatives by approximating the distribution of the 1-year stock price S_1. We will assume that we have as many moments of the distribution that we need, where they can be calculated using the methodology set out in section 2.5. We will be considering two different models: a translated gamma distribution where the first three moments will be the same as our time T stock price, and a more general methodology using polynomial splines that will match as many moments as required.

3.3.1 Translated Gamma

A translated gamma distribution can be fitted to approximate the distribution of a random variable, when we have the first three moments given by m_1, m_2 and m_3. A translated gamma distribution with translation x has density function

$$f_{S_T}(s; r, \lambda, x) = \frac{\lambda}{\Gamma(r)} \left[\lambda(s-x)\right]^{r-1} e^{-\lambda(s-x)} \quad s > x. \tag{3.3.1}$$

To solve for the parameter values x, r and λ we can equate the first 3 cumulants of the time T stock price with those of the translated gamma, so that

$$
\begin{aligned}
x + \frac{r}{\lambda} &= m_1 \\
\frac{r}{\lambda^2} &= m_2 - m_1^2 \\
\frac{2r}{\lambda^3} &= 2m_1^3 + m_3 - 3m_1 m_2,
\end{aligned}
$$

which gives us

$$
\begin{aligned}
\lambda &= \frac{2(m_2 - m_1^2)}{(2m_1^3 + m_3 - 3m_1 m_2)} \\
r &= (m_2 - m_1^2)\lambda^2 \\
x &= m_1 - \frac{r}{\lambda}.
\end{aligned}
$$

Pricing a call option with strike price K when the final stock price has the distribution given in (3.3.1) is equivalent to pricing a call option with strike price $K - x$ where the final stock price has distribution

$$
f_{S_T}(s; r, \lambda) = \frac{\lambda}{\Gamma(r)} [\lambda s]^{r-1} e^{-\lambda s} \quad s > 0. \tag{3.3.2}
$$

This is apparent by noting that the price of a call option will remain unchanged when there is a parallel shift of both the distribution and the strike price. The price of the call option will then be given by

$$
\begin{aligned}
C_t &= \int_{s=K-x}^{\infty} [s-K] f_{S_T}(s) ds \\
&= \int_{s=0}^{\infty} [s-K+x] ds - \int_{s=0}^{K-x} [s-K+x] f_{S_T}(s) ds.
\end{aligned}
$$

It can be easily shown that

$$
\int_{s=0}^{K-x} s f_{S_T}(s) ds = \frac{r}{\lambda} F_{S_T}(K-x;\ r+1, \lambda),
$$

which then gives us

$$
C_t = \frac{r}{\lambda}[1 - F_{S_T}(K-x;\ r+1,\lambda)] + (K-x)[F_{S_T}(K-x;\ r,\lambda) - 1]. \quad (3.3.3)
$$

Example 3.3.1. We shall now price call options on Asset 1 whose parameters are given in section 3.1. We find that the distribution of S_1, where we start in state 1 at time 0, can be approximated by a translated gamma where $x = 0.72822$, $r = 0.17275$ and $\lambda = 0.57157$. This distribution can be seen in figure 3.4. The value of call options with various strike values were calculated using equation (3.3.3) and are shown in table 3.12 in section 3.6.

□

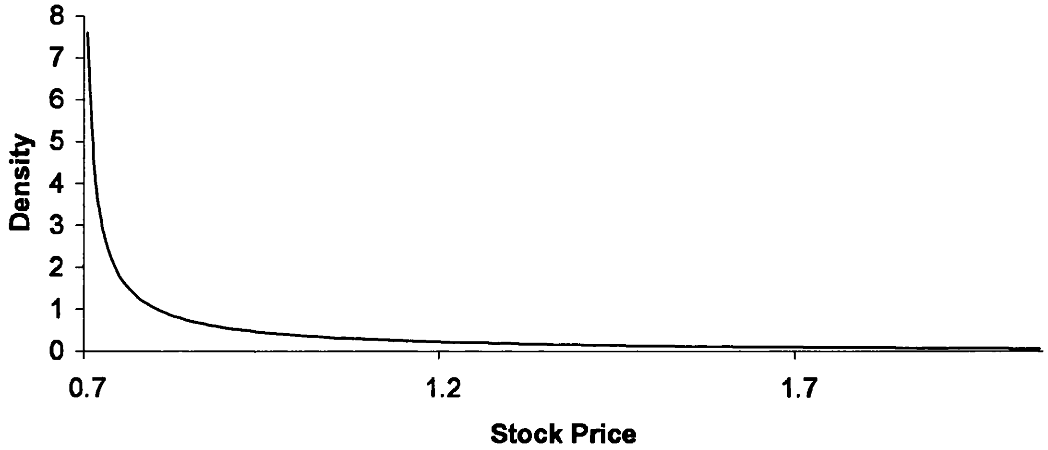

Figure 3.4: Chart showing approximated distribution of the price of Asset 1 after 1 year using a translated gamma distribution.

3.3.2 Polynomial Spline PDF Fitting

Using the above methodology we only took into consideration the first three moments of our time T stock price. We shall now develop an alternative parametric methodology, which will allow us to factor in many more moments to get a more accurate fit, and without having to make any assumptions on the underlying distribution of the stock price.

Suppose we have the density of $r+1$ points of a distribution $\{f(x) : x \in [-\infty, \infty]\}$, which we can represent by the set

$$\mathbf{P} = \{(x_1, f(x_1)), \ldots, (x_{r+1}, f(x_{r+1}))\},$$

such that $x_1 < \ldots < x_{r+1}$. Let us assume that x_1 and x_{r+1} are sufficiently extreme so that we can take $f(x) = 0$ for $x \notin [x_1, x_{r+1}]$. We will need to interpolate from these $r+1$ densities a smooth curve to create a complete density function. We can do this by fitting what is known as a *polynomial spline curve* (see Silverman [1985] or de Boor [1978]). To this end, we shall also assume that we have the first $r+1$ moments of the distribution that we wish to replicate given by $m_1, \ldots, m_{r+1}$.

A polynomial spline curve consists of piecewise polynomials which can be fitted to a series of data points, and has the property that it is the interpolating function which minimizes the integrated squared second derivative ($\int (f''(x))^2 dx$). The piecewise portions are defined so that at the *knots* (where the piecewise portions join, so that in our case these are at the points $x_2, \ldots, x_r$) the function and its first two derivatives are continuous, although the third derivatives may be discontinuous.

Within the range $x \in [x_i, x_{i+1}]$ for $i = 1, \ldots, r$ we can fit a polynomial curve (or spline) with equation

$$f(x) = a_4^i x^4 + a_3^i x^3 + a_2^i x^2 + a_1^i x + a_0^i \qquad x_i \leq x < x_{i+1}. \tag{3.3.4}$$

There are r such splines each consisting of 5 parameters, and so we are left with $5r$ parameters to solve for. To do this, we obtain $5r$ equations as follows.

Firstly we require that at each of the $r+1$ points the values of the spline curves are equal to the given densities. We therefore have the following $2r$ equations:

$$
\begin{aligned}
f(x_i) &= a_4^{i-1}x^4 + a_3^{i-1}x_i^3 + a_2^{i-1}x_i^2 + a_1^{i-1}x_i + a_0^{i-1} \qquad && i = 2, \ldots, r+1 \\
f(x_i) &= a_4^i x_i^4 + a_3^i x_i^3 + a_2^i x_i^2 + a_1^i x_i + a_0^i \qquad && i = 1, \ldots, r.
\end{aligned}
$$

Next, as mentioned above we ensure the smoothness of the curve by requiring that the first two derivatives are continuous at each of the knots. This leaves us with a further $2(r-1)$ equations

$$
\begin{aligned}
4a_4^{i-1}x_i^3 + 3a_3^{i-1}x_i^2 + 2a_2^{i-1}x_i + a_1^{i-1} &= 4a_4^i x_i^3 + 3a_3^i x_i^2 + 2a_2^i x_i + a_1^i \\
12a_4^{i-1}x_i^2 + 6a_3^{i-1}x_i + 2a_2^{i-1} &= 12a_4^i x_i^2 + 6a_3^i x_i + 2a_2^i
\end{aligned}
$$

for $i = 2, \ldots, r$. Next we require that the first $r+1$ moments of the curve are equal to $m_1, \ldots, m_{r+1}$, and that the density function integrates to 1. This can be done as follows.

We can calculate the y^{th} moment of the distribution as being

$$
E[x^y] = \sum_{i=1}^{r} \int_{x=x_i}^{x_{i+1}} x^y(a_4^i x^4 + a_3^i x^3 + a_2^i x^2 + a_1^i x + a_0^i)dx,
$$

which can be re-expressed as

$$
\begin{aligned}
E[x^y] &= \sum_{i=1}^{r}[\frac{a_4^i}{y+5}({x_{i+1}}^{y+5} - x_i^{y+5}) + \frac{a_3^i}{y+4}({x_{i+1}}^{y+4} - x_i^{y+4}) \\
&\quad + \frac{a_2^i}{y+3}({x_{i+1}}^{y+3} - x_i^{y+3}) + \frac{a_1^i}{y+2}({x_{i+1}}^{y+2} - x_i^{y+2}) \\
&\quad + \frac{a_0^i}{y+1}({x_{i+1}}^{y+1} - x_i^{y+1})]. \qquad (3.3.5)
\end{aligned}
$$

So by setting

$$E[x^y] = m_y \tag{3.3.6}$$

for $y = 0, \ldots, r+1$ where $m_0 = 1$, we now have a further $r+2$ equations leaving us with the $5r$ equations we require.

We are able to solve these $5r$ equations in matrix form as follows. Let us define the matrix $\mathbf{D_P}$, where the subscript denotes dependency on the set $\mathbf{P}$ defined at the beginning of the section, as having entries

$$\underset{5r\times 5r}{\mathbf{D_P}} = \begin{pmatrix} \mathbf{X}^1_{\mathbf{P}} & \hat{\mathbf{X}}^1_{\mathbf{P}} & & & & \\ & \mathbf{X}^2_{\mathbf{P}} & \hat{\mathbf{X}}^2_{\mathbf{P}} & & & \\ & & \mathbf{X}^3_{\mathbf{P}} & \hat{\mathbf{X}}^2_{\mathbf{P}} & & \\ & & & \ddots & \ddots & \\ & & & & \mathbf{X}^{r-1}_{\mathbf{P}} & \hat{\mathbf{X}}^{r-1}_{\mathbf{P}} \\ & & & & & \mathbf{X}^r_{\mathbf{P}} \\ \mathbf{Y}^1_{\mathbf{P}} & \cdots & \cdots & \cdots & \cdots & \mathbf{Y}^r_{\mathbf{P}} \end{pmatrix},$$

where entries other than those shown are 0, and where

$$\underset{4\times 5}{\mathbf{X}^i_{\mathbf{P}}} = \begin{pmatrix} x_i^4 & x_i^3 & x_i^2 & x_i & 1 \\ x_{i+1}^4 & x_{i+1}^3 & x_{i+1}^2 & x_{i+1} & 1 \\ 12x_{i+1}^2 & 6x_{i+1} & 2 & 0 & 0 \\ 4x_{i+1}^3 & 3x_{i+1}^2 & 2x_{i+1} & 1 & 0 \end{pmatrix} \qquad i = 1, \ldots, r-1,$$

$$\underset{4\times 5}{\hat{\mathbf{X}}^i_{\mathbf{P}}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ -12x_{i+1}^2 & -6x_{i+1} & -2 & 0 & 0 \\ -4x_{i+1}^3 & -3x_{i+1}^2 & -2x_{i+1} & -1 & 0 \end{pmatrix} \qquad i = 1, \ldots, r-1,$$

$$
\underset{2\times 5}{\mathbf{X}_{\mathbf{P}}^{r}} = \begin{pmatrix} 4x_r^4 & x_r^3 & x_r^2 & x_r & 1 \\ 4x_{r+1}^4 & x_{r+1}^3 & x_{r+1}^2 & x_{r+1} & 1 \end{pmatrix},
$$

$$
\underset{(r+2)\times 5}{\mathbf{Y}_{\mathbf{P}}^{i}} = \begin{pmatrix} \frac{x_{i+1}^5-x_i^5}{5} & \frac{x_{i+1}^4-x_i^4}{4} & \cdots & \cdots & \frac{x_{i+1}-x_i}{1} \\ \frac{x_{i+1}^6-x_i^6}{6} & \frac{x_{i+1}^5-x_i^5}{5} & \cdots & \cdots & \frac{x_{i+1}^2-x_i^2}{2} \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & & \vdots & \vdots & \vdots \\ \frac{x_{i+1}^{r+6}-x_i^{r+6}}{r+6} & \frac{x_{i+1}^{r+5}-x_i^{r+5}}{r+5} & \cdots & \cdots & \frac{x_{i+1}^{r+1}-x_i^{r+1}}{r+1} \end{pmatrix}.
$$

Now let us define the matrices

$$
\begin{aligned}
\underset{5r\times 1}{\mathbf{A}_{\mathbf{P}}} &= \begin{pmatrix} \mathbf{A}_{\mathbf{P}}^1 & \mathbf{A}_{\mathbf{P}}^2 & \ldots & \mathbf{A}_{\mathbf{P}}^r \end{pmatrix}^T \\
\underset{5r\times 1}{\mathbf{B}_{\mathbf{P}}} &= \begin{pmatrix} \mathbf{B}_{\mathbf{P}}^1 & \mathbf{B}_{\mathbf{P}}^2 & \ldots & \mathbf{B}_{\mathbf{P}}^{r+1} \end{pmatrix}^T,
\end{aligned}
$$

where

$$
\underset{1\times 5}{\mathbf{A}_{\mathbf{P}}^{i}} = \begin{pmatrix} a_4^i & a_3^i & a_2^i & a_1^i & a_0^i \end{pmatrix} \qquad i = 1,\ldots,r\,,
$$

$$
\underset{1\times 4}{\mathbf{B}_{\mathbf{P}}^{i}} = \begin{pmatrix} f(x_i) & f(x_{i+1}) & 0 & 0 \end{pmatrix} \qquad i = 1,\ldots,r-1\,,
$$

$$
\underset{1\times 2}{\mathbf{B}_{\mathbf{P}}^{r}} = \begin{pmatrix} f(x_i) & f(x_{i+1}) \end{pmatrix},
$$

$$
\underset{1\times(r+2)}{\mathbf{B}_{\mathbf{P}}^{r+1}} = \begin{pmatrix} m_0 & m_1 & \ldots & \ldots & m_r \end{pmatrix}.
$$

We can solve for the parameter values in the matrix $\mathbf{A_P}$ using the following equation:

$$\mathbf{A_P} = \mathbf{D_P^{-1}} \mathbf{B_P}, \tag{3.3.7}$$

provided that $\mathbf{D_P}$ is invertible.

We now have a polynomial spline curve approximation of our density function given by

$$y = \hat{f}_{\mathbf{P}}(x) \quad x_1 \leq x \leq x_{r+1}$$

and $y = 0$ elsewhere, where once again the subscript $\mathbf{P}$ denotes a dependency on the set $\mathbf{P}$. This curve has the property that the first $r+1$ moments are the same as that of our time T stock price. We shall now see an application of this methodology.

Example 3.3.2 (Standard Normal Distribution). We applied the above calculations using the first 6 moments of the standard normal distribution to obtain its polynomial spline approximation. For the set $\mathbf{P}$ we set $x_1 = -4$, $x_6 = 4$ and the rest of the x's equally spaced between these two values. For the densities we took $f(x_1) = f(x_6) = 0$ and $f(x_2) = \ldots = f(x_5) = 0.2$. The polynomial spline curve obtained is shown in figure 3.5.

□

Example 3.3.2 has demonstrated to us our next obstacle in this approximation, namely that even though our distribution will integrate to 1 and will

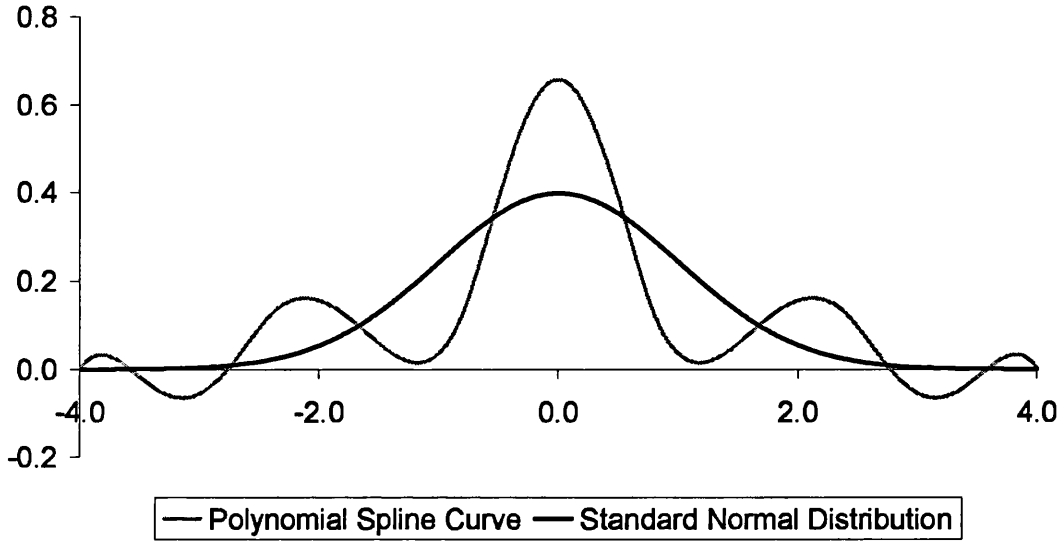

Figure 3.5: Chart showing standard normal distribution and 6-moment polynomial spline curve.

have moments as required, we have not ensured that the density values given by our polynomial spline curve are greater than 0 for all values of x. This condition could not be included in the above system of equations, and so we will have to try to find a solution that fulfills this condition numerically. We can do this as follows.

Firstly, let us define an error term for our set of points $\mathbf{P}$, denoted by $E_{\mathbf{P}}$, to represent the total area of our polynomial spline curve where the density is negative, so that

$$E_{\mathbf{P}} = \int_{x=x_1}^{x_{r+1}} \hat{f}_{\mathbf{P}}(x) I_{(\hat{f}(x)<0)} dx. \tag{3.3.8}$$

We therefore wish to find a set $\mathbf{P}$ which minimizes $E_{\mathbf{P}}$. As there may be more than one possible solution, we would like our resulting curve to be the 'smoothest', with preferably only one turning point. Let us define the variation $V_{\mathbf{P}}$ as the integrated squared second derivative of the curve, so that

$$V_{\mathbf{P}} = \int_{x=x_1}^{x_{r+1}} \left(\frac{\partial^2 \hat{f}_{\mathbf{P}}(x)}{\partial x^2} \right)^2 dx. \tag{3.3.9}$$

We would therefore like our set $\mathbf{P}$ to minimise $V_{\mathbf{P}}$ as well.

We are now left with the task of finding an algorithm which will find our solution. The algorithm we will use works by iteratively trying different values of $x_1, \ldots, x_{r+1}$ and $f(x_2), \ldots, f(x_r)$, each time accepting the new values only if the new curve generated reduces $E_{\mathbf{P}}$ and does not increase $V_{\mathbf{P}}$. This is repeated until $E_{\mathbf{P}} = 0$ or is at least minimised. This algorithm can be expressed by the following seven steps:

1. Initialise $\mathbf{P}$ by selecting values for x_1 and x_{r+1}, and spacing out all other x values equally between them. We set $f(x_1) = f(x_{r+1}) = 0$ and $f(x_i) = 1/(x_{r+1} - x_1)$ for $i = 2, \ldots, r$.

2. Calculate $\mathbf{A}_{\mathbf{P}}$, $E_{\mathbf{P}}$ and $V_{\mathbf{P}}$ numerically.

3. Try different values of $f(x_2)$ and calculate corresponding $\mathbf{A}_{\mathbf{P}}$, $E_{\mathbf{P}}$ and $V_{\mathbf{P}}$ to find value which minimises $E_{\mathbf{P}}$ whilst not increasing $V_{\mathbf{P}}$.

4. Repeat 3 for $f(x_3)$ to $f(x_r)$ and calculate overall reduction in $E_{\mathbf{P}}$ for this iteration.

5. Repeat 3-4 until overall reduction for the iteration is less than 1%.

6. Repeat 3-4 only changing values x_1 to x_{r+1} rather than $f(x_2)$ to $f(x_r)$ and calculate overall reduction in $E_{\mathbf{P}}$ for this iteration.

7. Repeat 3-6 until $E_{\mathbf{P}}$ is minimized.

We shall now show examples of the performance of the above algorithm on the standard normal distribution we attempted to approximate earlier on, as well as on a gamma distribution.

Example 3.3.3 (Standard normal distribution). We ran the above algorithm using between 4 and 10 moments of the standard normal distribution. For the starting value of the set $\mathbf{P}$, we took for x_1 and x_{r+1} values between -4 and -5 and 4 and 5 respectively. For each model, as well as computing the error value $E_{\mathbf{P}}$, we also computed the mean square error (MSE) between the polynomial spline curve and the standard normal distribution. The results can be seen in table 3.5.

We can therefore see that using 5 moments and above gives us a very good fit to the standard normal curve. Our algorithm yields the best results when 6 moments are used. The values x and $f(x)$ for each of the 6 points, as well as the parameter coefficients for the 5 splines are given in table 3.6.

Moments	$E_{\mathbf{P}}$ ($\times 10^{-5}$)	**MSE** ($\times 10^{-5}$)
4	0.00	3.30
5	0.00	0.56
6	0.00	0.32
7	0.00	1.02
8	0.00	0.54
9	0.00	0.90
10	0.00	0.46

Table 3.5: Results for the polynomial splines simulation of the standard normal distribution.

x	-4	-2.4	-0.8	0.8	2.4	4
f(x)	0.00000	0.02210	0.29000	0.28500	0.02010	0.00000

i	a_4^i	a_3^i	a_2^i	a_1^i	a_0^i
1	0.00438	0.06566	0.37156	0.94254	0.90642
2	-0.02001	-0.14615	-0.31055	-0.02019	0.40597
3	0.04798	0.00583	-0.20686	-0.00686	0.40024
4	-0.00764	0.06659	-0.13913	-0.11799	0.43747
5	0.01134	-0.15720	0.81654	-1.88707	1.64292

Table 3.6: Densities of 6 points used in the polynomial spline curve and spline parameter coefficients.

□

We can see how well the polynomial spline curve has approximated the standard normal distribution in figure 3.6. We shall now see how it performs for the gamma(2,3) distribution.

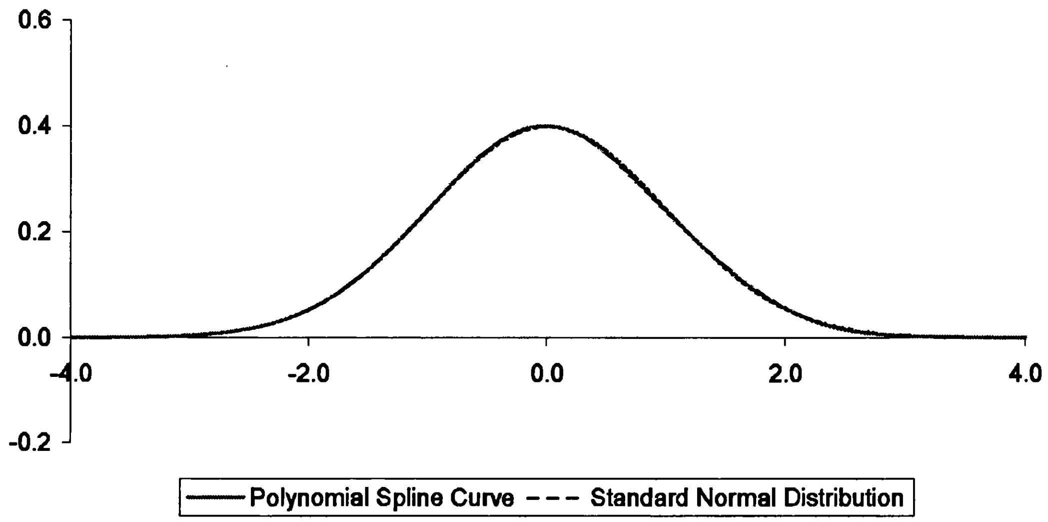

Figure 3.6: Chart showing standard normal distribution and 6-moment polynomial spline curve.

Example 3.3.4 (Gamma(2,3) distribution). As in the previous example, we ran the above algorithm using between 4 and 10 moments of the gamma(2,3) distribution. For the starting value of the set $\mathbf{P}$, we took $x_1 = 0$, whilst for x_{r+1} different starting values were tried between 4 and 6. We kept however $f(x_1) = 0$ constant throughout. For each model, as well as computing the error value $E_{\mathbf{P}}$ we also computed the MSE between the polynomial spline curve and the gamma(2,3) distribution. The results can be seen in table 3.7.

Moments	4	5	6	7	8	9	10
$E_{\mathbf{P}}(\times 10^{-5})$	0.00	0.00	0.00	0.00	0.11	0.00	5.43
MSE$(\times 10^{-5})$	92.07	12.02	8.14	2.59	6.54	0.31	178.89

Table 3.7: Results for the polynomial splines simulation of the gamma(2,3) distribution.

We can therefore see that using between 4 and 9 moments gives us a very good fit to the gamma(2,3) distribution, the best of which is when 9 moments are used. The parameter coefficients for the 8 splines in the 9 moment model are shown in table 3.8, and the values x and $f(x)$ for each of the 9 points are shown in table 3.9.

□

i	$\mathbf{a}_4^i$	$\mathbf{a}_3^i$	$\mathbf{a}_2^i$	$\mathbf{a}_1^i$	$\mathbf{a}_0^i$
1	-11.91299	27.80528	-24.75174	8.89989	0.00000
2	-0.67300	3.26439	-5.08131	2.09415	0.84617
3	0.82556	-5.38413	13.30164	-15.03080	6.76203
4	0.28979	-2.77118	9.90517	-15.72577	9.40369
5	-0.11770	1.32470	-5.53282	10.13494	-6.84096
6	-0.06264	0.85106	-4.31890	9.70240	-8.14089
7	-0.03452	0.57682	-3.60617	9.99526	-10.36059
8	0.00000	0.00000	0.00000	0.00000	0.00000

Table 3.8: Parameter coefficients for the 9-point polynomial spline curve.

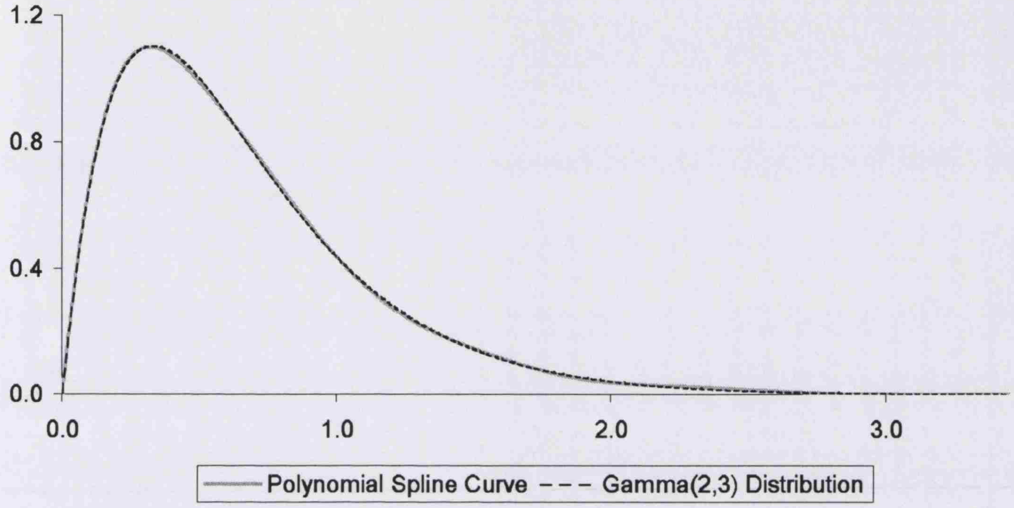

Figure 3.7: Chart showing gamma(2,3) distribution and 9-moment polynomial spline curve.

Figure 3.7 shows again how we have been able to approximate the gamma(2,3) distribution fairly accurately. We shall now apply this methodology to our MCJD model.

x	0	0.625	1.25	1.875	2.5
f(x)	0.00000	0.86440	0.25700	0.05540	0.01690

x	3.125	3.7501	4.375	27.67583
f(x)	0.00080	0.00140	0.00000	0.00000

Table 3.9: Densities of 9 points used in the polynomial spline curve.

3.3.3 Application to the MCJD model

We will now look at an application of this methodology by fitting a polynomial spline density curve to Asset 1, whose parameters are given in section 3.1. The moments of this asset are given in table 2.6 in section 2.5.

Using the algorithm above, we fitted a 6-point polynomial spline curve, where the first 6 moments are that of the 1-year asset price. Each of the points in the 6-moment model are given in table 3.10, as well as the parameter coefficients for the 5 splines in table 3.11.

x	0.00240	1.23041	3.05795	7.48815	8.24355	29.13993
f(x)	0.00000	0.58280	0.00520	0.00060	0.00010	0.00000

Table 3.10: Densities of the 6 points in the polynomial spline curve.

i	a_4^i	a_3^i	a_2^i	a_1^i	a_0^i
1	2.29012	-6.83373	5.13635	0.23411	-0.00059
2	0.01218	-0.23810	1.48183	-3.75568	3.37604
3	-0.00043	0.01030	-0.08901	0.32628	-0.41698
4	-0.00334	0.10657	-1.27273	6.74672	-13.39330
5	0.00000	0.00000	0.00000	-0.00006	0.00043

Table 3.11: Parameter coefficients for the 5 splines of the polynomial spline curve.

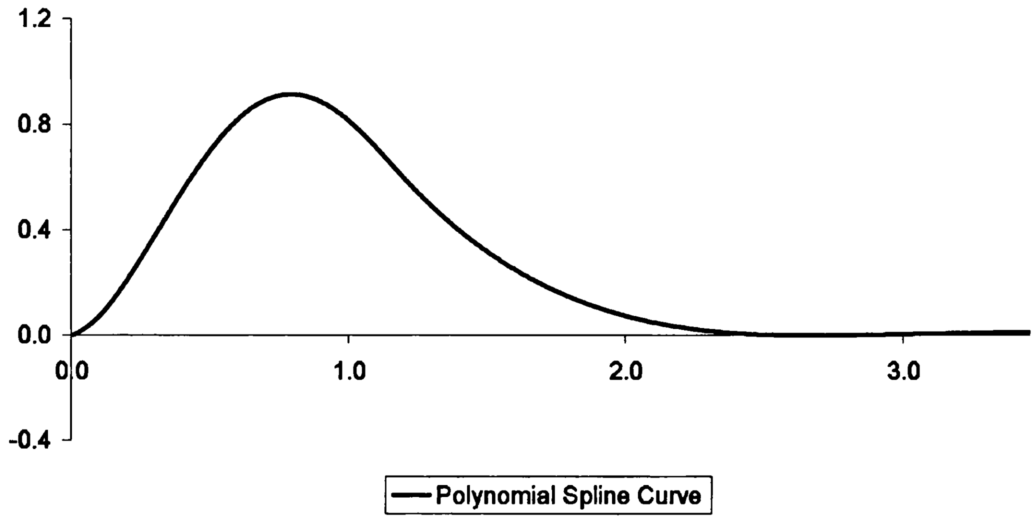

Figure 3.8: Polynomial spline curve fitted to 1-year ahead price of Asset 1.

The density function seen in figure 3.8 seems to resemble that of a gamma distribution, which is interesting as the translated gamma approximation in section 3.3.1 resulted in a density function looking very much like that of an exponential distribution. So we can see that factoring in more moments has radically transformed the shape of the density function.

The value of call options with various strike values were calculated using this density function and are shown in table 3.12 in section 3.6.

□

3.4 Tree-Based Methods

Tree-based methods are a type of simulation modelling, which involves trying to replicate the behaviour of the stock using more simplified models, which however retain the important characteristics of the initial model. Once this is done, derivatives can be valued using this simplified model. This is normally done by firstly developing a discrete-time analogue of the model, and then limiting the possible price movements within each time period. The simulation modelling techniques we shall use are *tree-based methods* in this section and *Monte Carlo methods* in section 3.5.

3.4.1 Trinomial Trees

When it comes to using tree-based methods to simulate the value of a stock or derivative, trinomial trees have probably become the benchmark tool employed by the financial world. We shall now explore its usefulness in our MCJD model described in section 3.1.

There are different ways in which a trinomial tree may be fitted. We shall employ a method which specifies at each node the ability for the stock price to rise to one value, fall to another value or remain constant as can be seen in figure 3.9.

Suppose we wish to estimate the time T value of the stock whose price process is governed by equation (3.1.1), where we are currently at time 0. To do this we shall begin by writing a discrete time analogue of equation (3.1.1).

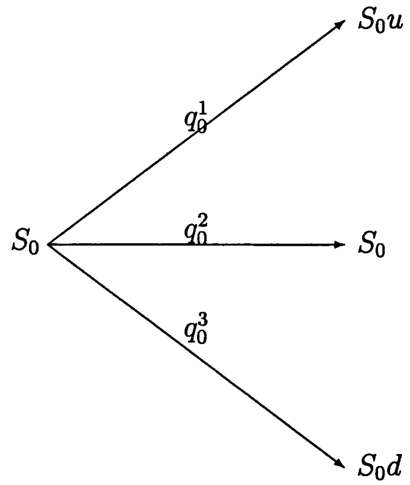

Figure 3.9: One-period trinomial tree for stock price.

We can divide the time period $[0, T]$ into h time periods of length Δt, where we therefore get $h = \frac{T}{\Delta t}$. During every such time period we will assume that the state of the Markov chain is constant, while there is a possibility of a transition of the state at the end of the period. The value of the stock at the end of the period is therefore a combination of a lognormally distributed random variable and a multinomial random variable as shown below, for when in state k

$$S_{t+\Delta t} = S_t e^{(r_k - \frac{1}{2}\sigma_k^2)\Delta t + \sigma_k \Delta W_t} \prod_{l=1}^{n} (1 + \gamma^{kl})^{\Delta N_t^{kl}}. \tag{3.4.1}$$

We shall now develop a useful lemma:

Lemma 3.4.1. *In the setting described above, we have that when in state k*

$$E[\prod_{l=1}^{n}(1+\gamma^{kl})^{\Delta N_t^{kl}}] = 1+\bar{\gamma}^k,$$

$$Var[\prod_{l=1}^{n}(1+\gamma^{kl})^{\Delta N_t^{kl}}] = \hat{\gamma}^k - (\bar{\gamma}^k)^2,$$

where

$$\bar{\gamma}^k = \sum_{l=1}^{n}\gamma^{kl}\lambda^{kl}\Delta t,$$

$$\hat{\gamma}^k = \sum_{l=1}^{n}(\gamma^{kl})^2\lambda^{kl}\Delta t.$$

Proof. This is easily shown by firstly noting that the probability of transiting from state k to state l in a time Δt is equal to $\lambda^{kl}\Delta t$, and by using standard results for two or more uncorrelated random variables (see for example Mood et al. [1974] chapter 5).

□

Using lemma 3.4.1 and equation (3.4.1), together with standard results of the cumulants of a lognormally distributed random variable, as well as the following results for two uncorrelated random variables:

$$E[XY] = E[X]E[Y],$$
$$Var[XY] = E[X]^2Var[Y] + E[Y]^2Var[X] + Var[X]Var[Y],$$

we have that the expected value of the stock at time $t + \Delta t$ conditional on the fact that we are in state k is given by

$$E[S_{t+\Delta t}] = S_t e^{r_k \Delta t}(1 + \bar{\gamma}^k).$$

We also have the variance being equal to

$$Var[S_{t+\Delta t}] \quad = \quad S_t^2 e^{2r_k \Delta t}[(1 + \hat{\gamma}^k + 2\bar{\gamma}^k)e^{\sigma_k^2 \Delta t} - (1 + \bar{\gamma}^k)^2].$$

Denote by I_t^k the indicator variable that the Markov chain is in state k at time t given that we start at time 0 in state j. We have that

$$\begin{aligned} E[I_t^k] &= p_t^{jk}, \\ Var[I_t^k] &= p_t^{jk}(1 - p_t^{jk}), \end{aligned}$$

where $0 \geq t < T$, and p_t^{jk} is the probability we are in any state k at time t given at time 0 we were in state j given in equation (1.3.5). Using the above we shall now develop another useful lemma:

Lemma 3.4.2. *Suppose we have a series of random variables $X_1, \ldots, X_n$ with means $\mu_1, \ldots, \mu_n$ and variances $\sigma_1^2, \ldots, \sigma_n^2$. Define another variable Z_t as follows:*

$$Z_t = I_t^1 X_1 + \ldots + I_t^n X_n.$$

The expected value of Z_t is given by

$$E[Z_t] = p_t^{j1}\mu_1 + \ldots + p_t^{jn}\mu_n.$$

The variance of Z_t will be equal to

$$\begin{aligned} Var[Z_t] &= p_t^{j1}\sigma_1^2 + \ldots + p_t^{jn}\sigma_n^2 \\ &\quad + p_t^{j1}(1-p_t^{j1})\mu_1^2 + \ldots + p_t^{jn}(1-p_t^{jn})\mu_n^2 \\ &\quad -2(p_t^{j1}p_t^{j2}\mu_1\mu_2 + \ldots + p_t^{j(n-1)}p_t^{jn}\mu_{n-1}\mu_n). \end{aligned}$$

Proof. This can be easily shown by taking expectations and squared expectations of each of the terms, and noting that as only one of the I_t^j's will have a value of one and all the rest will be zero, the covariances will therefore be zero.

□

Using lemma 3.4.2 we therefore have that the unconditional (with respect to the state we are in at time t) expected value of the stock as being

$$E[S_{t+\Delta t}] = \sum_{k=1}^{n} p_t^{jk} S_t e^{r_k \Delta t}(1+\bar{\gamma}^k), \tag{3.4.2}$$

and similarly the unconditional variance is equal to

$$\begin{aligned} Var[S_{t+\Delta t}] &= \sum_{k=1}^{n} p_t^{jk} S_t^2 e^{2r_k\Delta t}[(1+\hat{\gamma}^k + 2\bar{\gamma}^k)e^{\sigma_k^2\Delta t} - p_t^{jk}(1+\bar{\gamma}^k)^2] \\ &\quad -\sum_{k=1}^{n}\sum_{l=1}^{n} p_t^{jk} p_t^{l} S_t^2 e^{(r_k+r_l)\Delta t}(1+\bar{\gamma}^k)(1+\bar{\gamma}^l). \end{aligned} \tag{3.4.3}$$

Let us denote

$$\begin{aligned}
E_t &= \sum_{k=1}^{n} p_t^{jk} e^{r_k \Delta t}(1+\bar{\gamma}^k), \\
V_t &= \sum_{k=1}^{n} p_t^{jk} e^{2r_k \Delta t}[(1+\hat{\gamma}^k+2\bar{\gamma}^k)e^{\sigma_k^2 \Delta t} - p_t^{jk}(1+\bar{\gamma}^k)^2] \\
&\quad - \sum_{k=1}^{n}\sum_{l=1}^{n} p_t^{jk} p_t^{l} e^{(r_k+r_l)\Delta t}(1+\bar{\gamma}^k)(1+\bar{\gamma}^l),
\end{aligned}$$

so that

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

and

$$S_t^2 V_t = Var[S_{t+\Delta t}].$$

We will now construct a trinomial tree which will replicate the behaviour of the stock price given in (3.4.1). Suppose we have a one-period trinomial model as in figure 3.9, where at time t the time $t+\Delta t$ value of the stock price will be equal to $S_t u$ with probability q_t^1, S_t with probability q_t^2 and $S_t d$ with probability q_t^3. We can combine h such one-period models to approximate the stock price's behaviour in the interval $[0,T]$ (recalling that $h = \frac{T}{\Delta t}$). A three-period model can be seen in figure 3.10 where $u = \frac{1}{d}$.

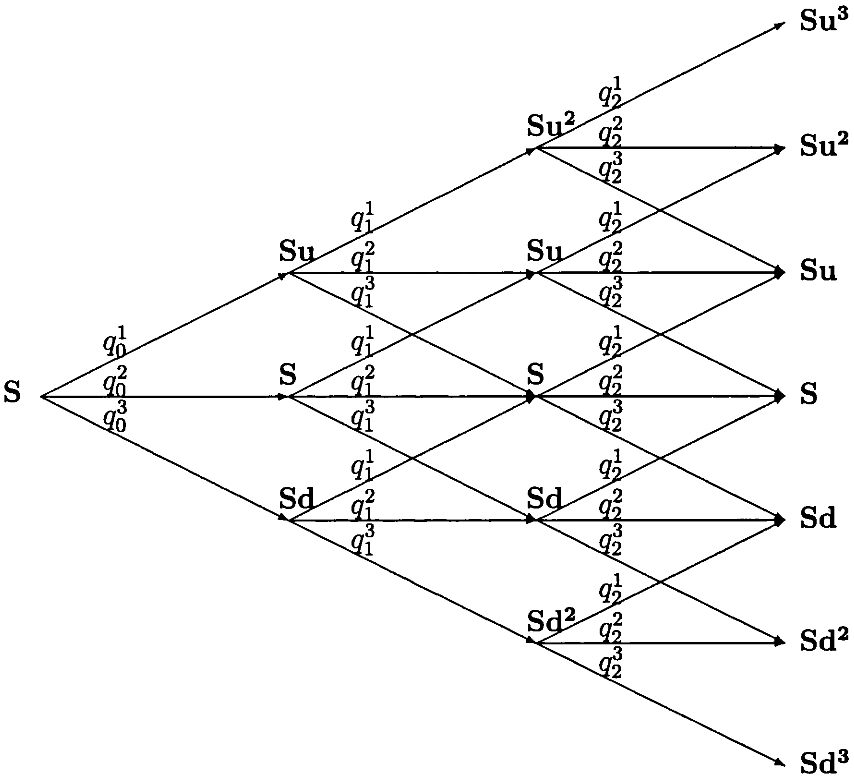

Figure 3.10: Trinomial stock price tree where $\Delta t = 1$.

We shall now calculate the moments of the trinomial tree model at each step. The expected value of the stock price at time $t + \Delta t$ is equal to

$$E[S_{t+\Delta t}] = S_t(q_t^1 u + q_t^2 + q_t^3 d). \tag{3.4.4}$$

The variance of $S_{t+\Delta t}$ can similarly be shown to be

$$\begin{aligned} Var[S_{t+\Delta t}] &= S_t^2[(q_t^1(1-q_t^1)u^2 + q_t^2(1-q_t^2) + q_t^3(1-q_t^3)d^2) \\ &\quad -2(q_t^1 q_t^2 u + q_t^1 q_t^3 ud + q_t^2 q_t^3 d)]. \end{aligned} \tag{3.4.5}$$

In order for the trinomial model to closely simulate the behavior of the

MCJD model, at each step we thus require the MCJD model and the trinomial model to have similar characteristics. We do this by equating the means and variances of both models. Starting with the means, equating equations (3.4.2) and (3.4.4) we get

$$E_t = q_t^1 u + q_t^2 + q_t^3 d.$$

A comparison of the variances in equations (3.4.3) and (3.4.5) gives us

$$V_t = q_t^1(1-q_t^1)u^2 + q_t^2(1-q_t^2) + q_t^3(1-q_t^3)d^2 - 2(q_t^1 q_t^2 u + q_t^1 q_t^3 ud + q_t^2 q_t^3 d).$$

Using these two equations, and given that we also know that

$$q_t^1 + q_t^2 + q_t^3 = 1,$$

we are left with three equations to solve for five unknowns $u, d, q_t^1, q_t^2, q_t^3$. We must therefore specify two of these values in order to get a unique solution. Let us set

$$\begin{aligned} u &= \frac{1}{d}, \\ q_t^2 &= \frac{2}{3}, \end{aligned}$$

which will allow us to solve for all the other unknowns. We are then left with the following two equations to solve for u and q_t^1:

$$q_t^1 = \frac{u[3E_t - 2] - 1}{3(u^2 - 1)},$$

and

$$q_t^1(1-q_t^1)u^4-\frac{4}{3}q_t^1u^3+((\frac{1}{3}+q_t^1)(\frac{2}{3}-q_t^1)-V_t)u^2+\frac{2}{3}(\frac{1}{3}-q_t)u+(\frac{1}{3}-q_t^1)(\frac{2}{3}+q_t^1) = 0.$$

These two equations must be solved numerically for the first time period $t = 0$. For subsequent periods $t = \Delta t, \ldots, T - \Delta t$ we can either solve again the above equations which will result in a different value for u and hence the tree will no longer be recombining, or alternatively we can keep the same value of u only relax the assumption that $q_t^2 = \frac{2}{3}$ and therefore solve for all the probabilities.

For an h-period tree we will then be left with $2h + 1$ final nodes at time T. The expected value of the stock can then be estimated as the weighted average of the stock value at each of these nodes, where the weights are the probabilities of the tree leading up to that node. This probability is calculated by taking for each path leading to that node the product of all the probabilities of the h one-period models, and summing over all the different paths. So for example, the probability of the stock price equaling Su^3 after three periods in the model shown in figure 3.10 is equal to $q_0^1q_1^1q_2^1$. The variance can be estimated using similar procedures.

Derivatives of the time T stock price can be calculated using the same tree. The only difference is that we replace the time T stock values at the end of each node and replace it with the value of the derivative. The weighted average of these values is then calculated, the result being the value of the

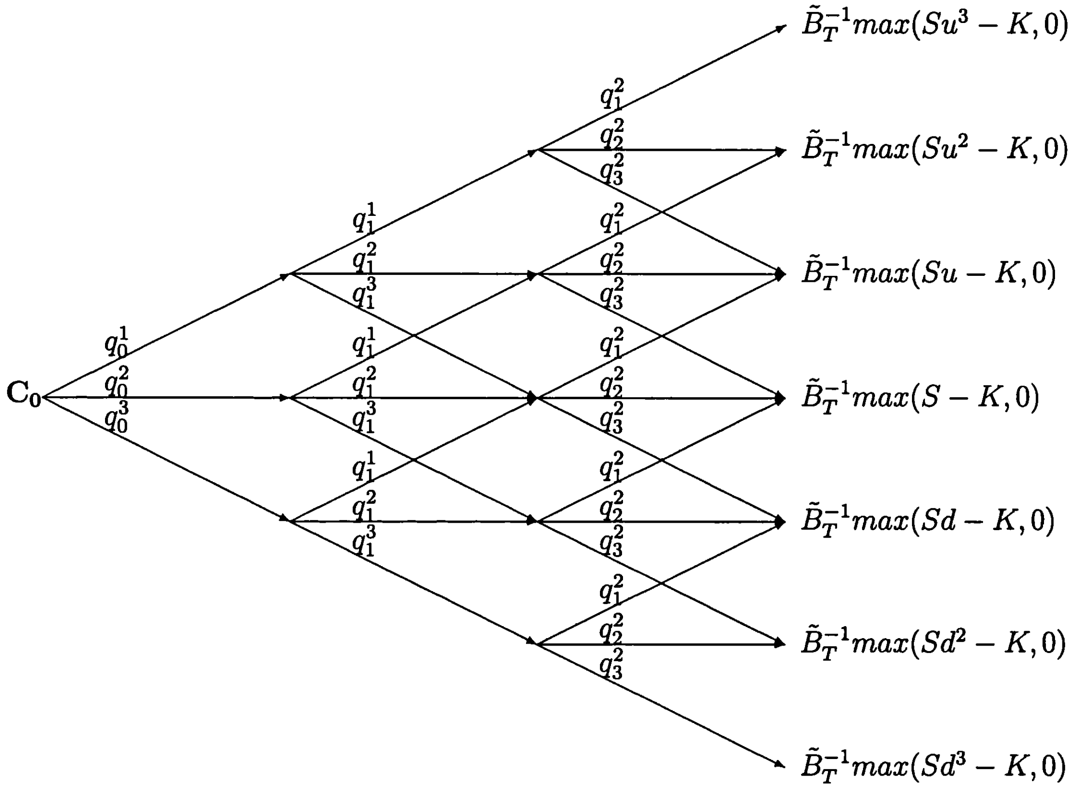

Figure 3.11: Pricing a three period call option using a trinomial tree.

derivative at time T. So for example, the estimate of a call option after three time periods with strike K whose payoff is given by

$$C_0 = B_T^{-1}max(S_T - K, 0)$$

can be calculated using figure 3.11.

In order to calculate the time 0 value of the derivative, we need to estimate the value of the discount factor

$$B_T^{-1} = e^{-\int_{t=0}^{T} r_t dt}.$$

This can be done by simply using the approximation $\tilde{B}_T^{-1}$ given by

$$\tilde{B}_T^{-1} = E[e^{-\int_{t=0}^{T} r_t dt}] = \prod_{i=0}^{h-1} \sum_{k=1}^{n} p_{i\Delta t}^{jk} e^{-r_k \Delta t},$$

where p_i^{jk} is once again the probability of being in state k at time $i\Delta t$ given that we started in state j.

The advantages of this model will be its fairly easy formulation and implementation. However, reducing an n dimensional process into just a three-branch tree will clearly sacrifice much of the original model's characteristics. We shall now look at an example of this methodology.

Example 3.4.3. Considering Asset 1 with parameters given in section 3.1, we have replicated the behaviour of this asset using a 12-node tree so that $\Delta t = 0.833$. The value of call options with various strike prices were calculated using this methodology and are shown in table 3.12 in section 3.6.

□

3.4.2 Multinomial Trees

We shall now try and represent our MCJD model in terms of a multinomial tree. Using multinomial trees we will be able to create a tree which will more closely resemble the behaviour of the stock, although this will lead to more computational complexity.

Let us assume that we are currently at time 0, and are interested in simulating the price of the stock at time T. We can divide this time into h sections of size Δt, so that $h = \frac{T}{\Delta t}$.

Firstly let us represent the stock price movement excluding the jumps by a binomial tree model as shown in figure 3.12. The probability of the stock price increasing by a factor u is given by q^j, and similarly the probability of the stock changing by a factor of d is given by $(1 - q^j)$.

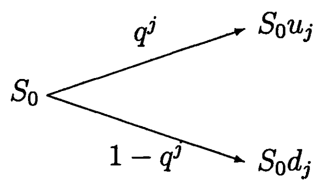

Figure 3.12: Stock price movement (excluding jumps) in time Δt when in state j.

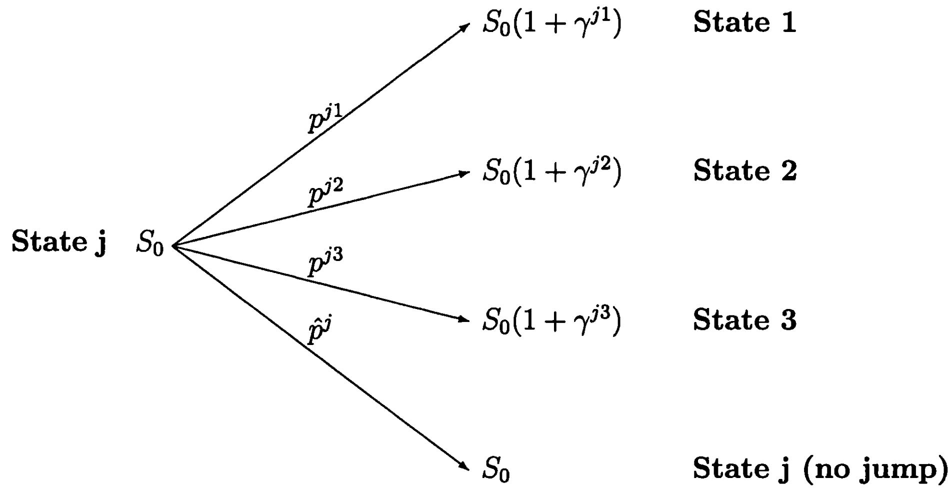

Figure 3.13: Stock price jumps in time Δt in a three-state model starting in state j.

Next we shall consider only the jumps part. For simplicity we will consider a model which has 3 states, although this can be easily generalised. To represent this we will need a 4-branch tree as shown in figure 3.13, one branch for each state the model can transit to and one branch representing no transition (this is due to the fact that when in state j we can either transit to state j in which case there is a jump, or alternatively we can remain in state j with no jump). We have the following probabilities:

$$
\begin{aligned}
p^{jk} &= \lambda^{jk}\Delta t \quad \forall j,k\,, \\
\hat{p}^j &= 1-\sum_{k=1}^{n}\lambda^{jk}\Delta t.
\end{aligned}
$$

We would like the variance implied by this binomial model to equal that implied from (3.4.1). We can see from (3.4.1) that when excluding the jumps S_t is lognormally distributed, meaning the expected value of the stock price after a change of time Δt when in state j is given by

$$E[S_{t+\Delta t}] = S_t e^{r_j \Delta t}.$$

The variance of the stock price over this period is given by

$$Var[S_{t+\Delta t}] = S_t^2 e^{2r_j \Delta t}(e^{\sigma_j^2 \Delta t} - 1).$$

Looking at the expected value and variance implied by the binomial model, we see that

$$\begin{aligned} E[S_{t+\Delta t}] &= S_t[q^j u_j + (1 - q^j)d_j], \\ Var[S_{t+\Delta t}] &= S_t^2[q^j u_j^2 + (1 - q^j)d_j^2]. \end{aligned}$$

In order for our binomial model to have the same mean and variance as the MCJD model (excluding the jumps) it is replicating, we therefore require

$$S_t e^{r_j \Delta t} = S_t[q^j u_j + (1 - q^j)d_j], \tag{3.4.6}$$

$$S_t^2 e^{2r_j \Delta t}(e^{\sigma_j^2 \Delta t} - 1) = S_t^2(q^j u_j^2 + (1 - q^j)d_j^2). \tag{3.4.7}$$

We now have two equations which we need to solve for three parameters, and so we will need to specify one of the parameters. We may set

$$u = \frac{1}{d}$$

as in the Cox-Ross and Rubenstein model (see Cox and Rubenstein [1985]). Equations (3.4.6) and (3.4.7) now become

$$\begin{aligned} S_t e^{r_j \Delta t} &= \frac{S_t}{u^j}[q^j u_j^2 + (1-q^j)] \\ S_t^2 e^{2r_j \Delta t}(e^{\sigma_j^2 \Delta t} - 1) &= (\frac{S_t}{u^j})^2 (q^j u_j^4 + (1-q^j)). \end{aligned}$$

These equations can be solved to give

$$\begin{aligned} q^j &= \frac{e^{r_j \Delta t} - d_j}{u_j - d_j}, \\ u_j &= \frac{1 + e^{a_j \Delta t} + [(1 + e^{a_j \Delta t})^2 - 4e^{2r_j \Delta t}]^{\frac{1}{2}}}{2e^{r_j \Delta t}}, \end{aligned}$$

where

$$a_j = 2r_j + \sigma_j^2.$$

For this to provide real solutions we need the constraint

$$\sigma_j^2 \geq \frac{ln(e^{-r_j \Delta t} - e^{-2r_j \Delta t})}{\Delta t}.$$

We can now replicate the behaviour of the MCJD model by combining the binomial model with the jumps model as in figure 3.14. We can use figure 3.14 to calculate the expected value of $S_{\Delta t}$ by taking the weighted average of the value at each of the end nodes, where the weights are the products of

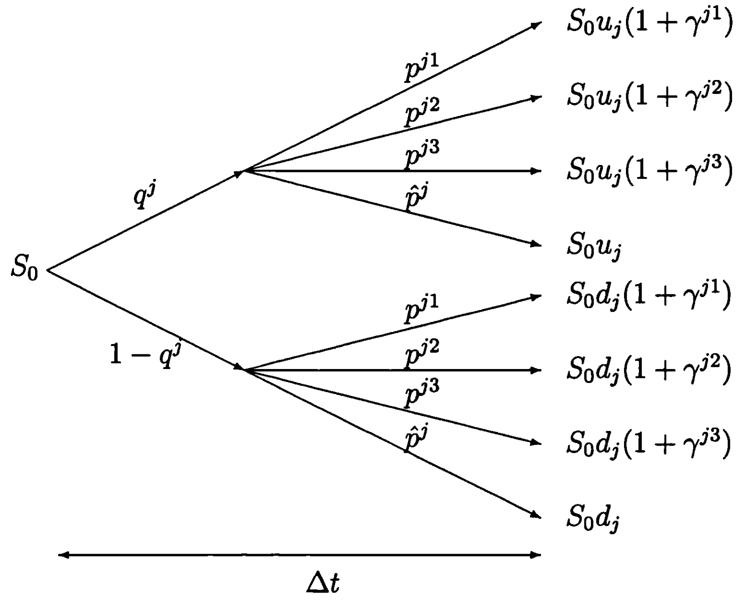

Figure 3.14: Stock price movement in time Δt when in state j.

the probabilities leading to that node.

We can similarly use this tree to price other derivatives of the stock price by replacing the values at the final nodes by the payoffs of the derivative. For example, figure 3.15 shows how we may use the tree to value a call option at time 0 on the stock at time Δt with strike value K.

So far we have calculated the expected value of the stock and valued derivatives of it over one time period Δt. If we wish to calculate the expected value of the stock after a time T, i.e. over h time periods where $h = \frac{T}{\Delta t}$, then we simply combine h of the above single-period trees. As will be fairly apparent, this will mean the number of final nodes will be fairly large even

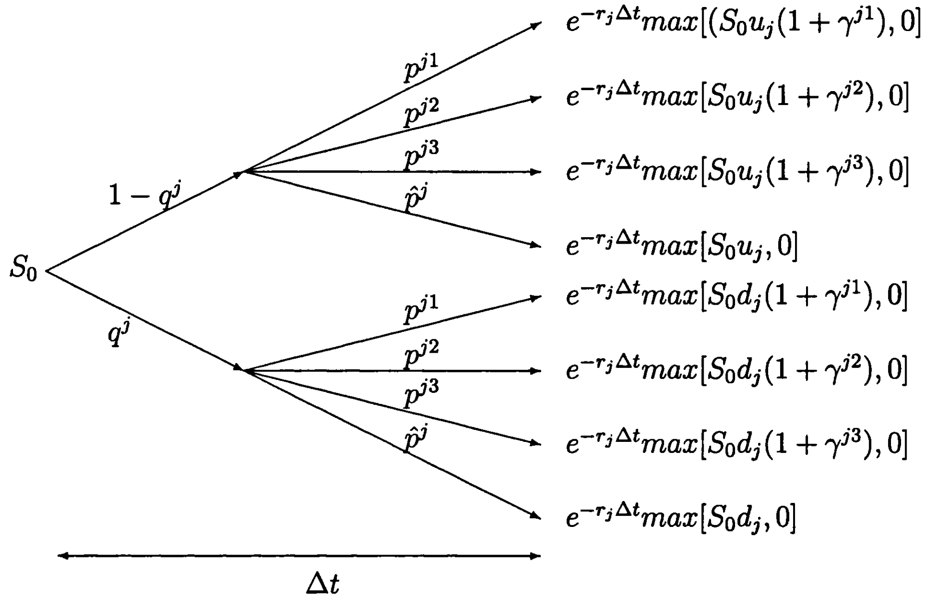

Figure 3.15: Valuing a call option at time 0 on $S_{\Delta t}$ when in state j.

for moderate values of h. The number of nodes after a time T represented by N_T is equal to

$$N_T = [2(n+1)]^h.$$

This methodology's usefulness will therefore be restricted to models where there are small values of n and h, and hence T.

Example 3.4.4. Once again considering Asset 1 with parameters given in section 3.1, we are able to replicate the behaviour of this asset using a 10 node tree so that $\Delta t = 0.1$. The value of call options with various strike prices were calculated using this methodology and are shown in table 3.12 in section 3.6.

□

3.5 Monte Carlo Simulation

We shall now attempt to estimate the value of the time T stock price and its derivatives using Monte Carlo simulation. The value at time 0 of a contingent claim at time T denoted by X_T is given by

$$C_0 = E[e^{-\int_{t=0}^{T} r_t dt} X_T], \tag{3.5.1}$$

where the expectation is under the risk-neutral measure $\mathbf{Q}$. There are many possible paths the stock price can take up until time T. We may generate randomly one of these paths and calculate the value of the derivative under this path. This is known as a *simulation trial.* The simulation involves repeating the simulation trial say N times and then aggregating the values. We shall now look at methods for simulating our MCJD model.

We shall once again divide the interval $[0, T]$ into h subsections of length Δt, so that

$$h = \frac{T}{\Delta t}.$$

We can write a discrete version of our MCJD model given in equation (3.1.1) whilst in state j as

$$\Delta S_t = S_t[r_j \Delta t + \sigma_j \epsilon_t \sqrt{\Delta t} + \sum_{k=1}^{n} \gamma^{jk} \Delta N_t^{jk}], \tag{3.5.2}$$

where ϵ_t is a sample from a standardised normal distribution, and ΔN_t^{jk} is a Bernoulli random variable with probability $\lambda^{jk} \Delta t$. We impose the restriction that

$$\Delta N_t^{j1} + \Delta N_t^{j2} + \ldots + \Delta N_t^{jn} \leq 1$$

for all j, i.e. at most one jump per time period. In order to simulate the value of S_t we must therefore generate each of the random variables $S_{\Delta t}, S_{2\Delta t}, \ldots, S_{h\Delta t}$ successively.

To generate these random samples we assume that we have an endless supply of uniform $[0,1]$ random variables $U_1, U_2, \ldots$, which can be easily generated by all computers. We also assume that we have samples from the standard normal distribution ϵ_t $\forall t$. To generate sample values of ΔN_t^{jk} for $k = 1, \ldots, n$, we can divide the interval $[0,1]$ into $n+1$ intervals $[0, I_1], [I_1, I_2], \ldots, [I_n, 1]$, where we have for $m = 1, \ldots, n$

$$I_m = \sum_{k=1}^{m} \lambda^{jk} \Delta t.$$

We then have that

$$\Delta N_t^{jk} = \begin{cases} 1 & U_i \in [I_{k-1}, I_k], \\ 0 & \text{otherwise}, \end{cases}$$

for all $k = 1, \ldots, n$, and where $I_0 = 0$.

We are therefore able to generate ΔS_t using one uniform $[0,1]$ random variable and one standard normal random variable, and hence each simulation trial using $2 \times h$ variables. In total we have N simulations where we need to generate $2 \times N \times h$ random variables. We can already see that the number of calculations that need to be performed is linear in h, whilst the number required using the multinomial tree method was exponential in h.

In order to value the derivative price given in (5.4.1), we will also need to value the bank account process B_T for each such simulation given by

$$B_T = e^{\sum_{t=0}^{T} r_t \Delta t}.$$

The accuracy of the estimate given by the Monte Carlo simulation is clearly going to be dependent on the number of trials N which are performed. The value of the derivative will be given by the mean of the values given by each simulation trial, and we will denote this by π. We are also able to calculate the standard deviation of these simulated values which we will denote by ω. Since these are i.i.d. trials, the variance of π will therefore be given by

$$Var(\pi) = \frac{\omega^2}{N}.$$

Using a normal approximation, we can thus write a 95% confidence interval for the price of the derivative C_T as follows:

$$\pi - \frac{1.96\omega}{\sqrt{N}} < C_T < \pi + \frac{1.96\omega}{\sqrt{N}}.$$

So the accuracy of our estimate will be proportional to the number of trials we perform.

Example 3.5.1. Once again considering Asset 1 with parameters given in section 3.1, we replicated the behaviour of this asset using 250,000 simulation trials. The value of call options with various strike prices were calculated using this methodology and are shown in table 3.12 in section 3.6.

□

3.6 Comparison of Performances

We shall now compare the performances of each of the methodologies described in this chapter. Table 3.12 shows the price of a call option with strike prices ranging from £0 to £5 using each of the above methodologies.

The tree-based methods seem to perform the worst, not being able to price even the zero-strike option correctly, as well as seemingly underpricing at higher strikes compared with the polynomial spline methodology and the Monte Carlo methods. The multinomal tree seems to be worse than the trinomial tree, although this is probably due to the fact that the the time T was divided into 12 subperiods in the trinomial model, but only 10 subperiods for the multinomial model. The finite-difference method also seems to be giving very low results at higher strikes, but this could be a result of the fact that we had capped the value of the stock price at £5.

Strike Price	Finite Difference	Translated Gamma	Polynomial Spline	Trinomial Tree	Multinomial Tree	Monte Carlo
0.0	1.0000	1.0000	1.0000	0.9929	0.9598	1.0009
0.2	0.8254	0.8059	0.8068	0.7993	0.7662	0.8068
0.4	0.6319	0.6118	0.6218	0.6180	0.5819	0.6145
0.6	0.4613	0.4177	0.4573	0.4680	0.4291	0.4415
0.8	0.3176	0.2605	0.3238	0.3511	0.3012	0.3149
1.0	0.2368	0.2035	0.2250	0.2576	0.2249	0.2322
1.2	0.1750	0.1648	0.1579	0.2006	0.1670	0.1712
1.4	0.1294	0.1358	0.1148	0.1458	0.1195	0.1306
1.6	0.0991	0.1131	0.0881	0.1175	0.0902	0.1042
1.8	0.0672	0.0949	0.0717	0.0893	0.0735	0.0829
2.0	0.0519	0.0801	0.0615	0.0673	0.0590	0.0662
2.2	0.0432	0.0679	0.0547	0.0558	0.0461	0.0533
2.4	0.0348	0.0578	0.0492	0.0444	0.0353	0.0440
2.6	0.0251	0.0493	0.0443	0.0329	0.0277	0.0374
2.8	0.0179	0.0422	0.0393	0.0264	0.0224	0.0321
3.0	0.0130	0.0363	0.0344	0.0227	0.0193	0.0274
3.2	0.0092	0.0312	0.0297	0.0190	0.0167	0.0234
3.4	0.0073	0.0269	0.0252	0.0152	0.0144	0.0200
3.6	0.0062	0.0232	0.0212	0.0115	0.0122	0.0172
3.8	0.0053	0.0200	0.0176	0.0090	0.0102	0.0148
4.0	0.0041	0.0173	0.0145	0.0081	0.0084	0.0129
4.2	0.0030	0.0150	0.0119	0.0071	0.0070	0.0114
4.4	0.0019	0.0130	0.0098	0.0061	0.0058	0.0102
4.6	0.0011	0.0113	0.0081	0.0051	0.0049	0.0092
4.8	0.0005	0.0098	0.0067	0.0042	0.0043	0.0084
5.0	0.0000	0.0086	0.0057	0.0032	0.0037	0.0076

Table 3.12: Prices of call options with various strike prices using each of the methodologies.

Chapter 4

Interest-Rate Theory

4.1 Introduction

So far we have looked at the pricing of derivatives on assets with price processes governed by our Markov chain jump-diffusion model. We will now look at financial instruments whose payoffs are determined solely by the value of interest rates over the period of the instrument. Such financial products are collectively known as the *Fixed-Income Market.* The main difference between these two financial markets is that when dealing with stocks the underlying is directly tradable, whilst with interest-rate derivatives it is not. This will lead to complications regarding *completeness*, as we shall see later on.

The main building block in pricing interest-rate derivatives is the *short rate* which we will denote by r_t. This is the interest rate for instantaneous borrowing at time t. From the short-rate process we can derive two other important process: the *forward-rate process* and the *zero-coupon price process.* The forward rate given by $f(t,T)$ is the interest rate set at time t for instantaneous borrowing at time T. The zero-coupon price given by $p(t,T)$ is the price paid at time t for receiving a fixed nominal amount, say £100, at time T. Unlike the short rate, forward-rate contracts and the zero-coupon bonds (which can be regarded as derivatives of the short rate) are freely traded on the market.

The *Term Structure of interest rates* is the name given to the relationship between the interest on a zero-coupon bond and its maturity. The *Yield Curve* is a graphical representation of this relationship where interest rates are plotted against maturity. The short rate r_t is the left-most point on the yield curve. The main objective of interest-rate modelling is to determine the yield curve, which can then be used to price interest-rate derivatives.

We will first begin by describing our Markov chain jump-diffusion model in the fixed-income context, and derive relationships between the short rate, forward rate and zero-coupon bond prices. We will then look at the two most common classes of models used to describe the short-rate dynamics: *Short Rate Models* and *HJM Models.*

For the rest of this chapter we shall mainly follow the methodologies set out in Bingham and Kiesel [2004] and Musiela and Rutkowski [1997].

4.2 The Model

We begin by assuming the following model for the short rate:

$$dr_t = \mu(t, Y_{t-})dt + \sigma(t, Y_{t-})dW_t + \gamma^r(t, Y_{t-}, Y_t)dN_t, \qquad (4.2.1)$$

where the drift and volatility functions $\mu()$ and $\sigma()$ are previsible, and we denote this by showing them as functions of Y_{t-}. The jump function $\gamma^r()$ is not previsible, but rather dependent on the state that we have transited from Y_{t-} and the state we are transiting to Y_t. Using the notation as in the previous chapters, we may also write the jump sizes as being γ_t^{jk} when $Y_{t-} = j$ and $Y_t = k$. We have also attached a superscript r to the jump function to show that it represents the jump in the short-rate process, as opposed to jumps in the other two process described below. W_t once again represent a Wiener process and N_t a counting process as described in section 1.3.1.

We also have the zero-coupon price process given by

$$dp(t,T) = p(t,T)[m(t,T,Y_{t-})dt + v(t,T,Y_{t-})dW_t + \gamma^p(t,T,Y_{t-},Y_t)dN_t] \qquad (4.2.2)$$

and the forward-rate dynamics given by

$$df(t,T) = a(t,T,Y_{t-})dt + b(t,T,Y_{t-})dW_t + \gamma^f(t,T,Y_{t-},Y_t)dN_t, \qquad (4.2.3)$$

where as above the drift and volatility functions $m()$, $a()$, $v()$ and $b()$ are previsible, whilst the jump functions $\gamma^p()$ and $\gamma^f()$ are not. We have the following well-known results:

$$\begin{aligned} df(t,T) &= -\frac{\partial log\, p(t,T)}{\partial T}, && (4.2.4) \\ r(t) &= f(t,t). && (4.2.5) \end{aligned}$$

We shall now find the relationships between the above processes. Let us begin with the zero-coupon price process and the forward-rate process.

Proposition 4.2.1. *If $p(t,T)$ satisfies (4.2.2) then the forward-rate dynamics are given by*

$$df(t,T) = a(t,T,Y_{t-})dt + b(t,T,Y_{t-})dW_t + \gamma^f(t,T,Y_{t-},Y_t)dN_t,$$

where

$$\begin{aligned} a(t,T,Y_{t-}) &= v_T(t,T,Y_{t-})v(t,T,Y_{t-}) - m(t,T,Y_{t-}), \\ b(t,T,Y_{t-}) &= -v_T(t,T,Y_{t-}), \\ \gamma^f(t,T,Y_{t-},Y_t) &= -\frac{\gamma^p_T(t,T,Y_{t-},Y_t)}{1+\gamma^p(t,T,Y_{t-},Y_t)}, \end{aligned}$$

and where by the subscript T we denote the partial derivative with respect to T.

Proof. Applying Itô's lemma to (4.2.2) as well as an analogous lemma for the jumps, we get

$$\begin{aligned} p(t,T) &= exp[(\int_{s=0}^{t}(m(s,T,Y_{s-}) - \frac{1}{2}v(s,T,Y_{s-})^2)ds + \int_{s=0}^{t} v(s,T,Y_{s-})dW_s \\ &\quad + \int_{s=0}^{t} log(1+\gamma^p(s,T,Y_{s-},Y_s)dN_s]. \end{aligned}$$

Taking the logarithm of the above equation and partially differentiating with respect to T we get

$$\begin{aligned}\frac{\partial log\, p(t,T)}{\partial T} &= \int_{s=0}^{t} (m_T(s,T,Y_{s-}) - v_T(s,T,Y_{s-})v(s,T,Y_{s-}))ds \\ &+ \int_{s=0}^{t} v_T(s,T,Y_{s-})dW_s + \int_{s=0}^{t} \frac{\gamma_T^p(s,T,Y_{s-},Y_s)}{1+\gamma^p(s,T,Y_{s-},Y_s)}dN_s.\end{aligned} \tag{4.2.6}$$

Inserting (4.2.6) into (4.2.3) completes the proof.

□

It is interesting to note that jumps in the zero-coupon price process only translate into jumps in the forward-rate process if γ^p is a function of T. This is a direct result of equation (4.2.4).

We will now look at the second relationship; between the forward-rate process and the short-rate process.

Proposition 4.2.2. *If $f(t,T)$ satisfies (4.2.3) then the short-rate dynamics are given by*

$$dr_t = \mu(t,Y_{t-})dt + \sigma(t,Y_{t-})dW_t + \gamma^r(t,Y_{t-},Y_t)dN_t, \tag{4.2.7}$$

where

$$\begin{aligned}\mu(t,Y_{t-}) &= f_T(t,t) + a(t,t,Y_{t-}),\\ \sigma(t,Y_{t-}) &= b(t,t,Y_{t-}),\\ \gamma^r(t,Y_{t-},Y_t) &= \gamma^f(t,t,Y_{t-},Y_t).\end{aligned}$$

Proof. We can integrate the forward-rate dynamics in (4.2.3) to get

$$f(t,t) = r(t) = f(0,t) + \int_{s=0}^{t} a(s,t,Y_{s-})ds + \int_{s=0}^{t} b(s,t,Y_{s-})dW_s$$

$$+ \int_{s=0}^{t} \gamma^f(s,t,Y_{s-},Y_s)dN_s. \tag{4.2.8}$$

Writing $a(\cdot)$, $b(\cdot)$ and $\gamma^f(\cdot)$ also in integrated form, we get

$$\begin{aligned}
a(s,t,Y_{s-}) &= a(s,s,Y_{s-}) + \int_{u=s}^{t} a_T(s,u,Y_{s-})du, \\
b(s,t,Y_{s-}) &= b(s,s,Y_{s-}) + \int_{u=s}^{t} b_T(s,u,Y_{s-})du, \\
\gamma^f(s,t,Y_{s-},Y_s) &= \gamma^f(s,s,Y_{s-},Y_s) + \int_{u=s}^{t} \gamma^f_T(s,u,Y_{s-},Y_s)du,
\end{aligned}$$

and on inserting into (4.2.8) we have

$$\begin{aligned}
r(t) &= f(0,t) + \int_{s=0}^{t} a(s,s,Y_{s-})ds + \int_{s=0}^{t}\int_{u=s}^{t} a_T(s,u,Y_{s-})duds \\
&+ \int_{s=0}^{t} b(s,s,Y_{s-})dW_s + \int_{s=0}^{t}\int_{u=s}^{t} b_T(s,u,Y_{s-})dudW_s \\
&+ \int_{s=0}^{t} \gamma^f(s,s,Y_{s-},Y_s)ds + \int_{s=0}^{t}\int_{u=s}^{t} \gamma^f_T(s,u,Y_{s-},Y_s)dudN_s.
\end{aligned} \tag{4.2.9}$$

Noting that

$$\begin{aligned} f_T(t,s) &= f_T(0,s) + \int_{s=0}^{t} a_T(s,t,Y_{s-})ds + \int_{s=0}^{t} b_T(s,t,Y_{s-})dW_s \\ &+ \int_{s=0}^{t} \gamma_T^f(s,t,Y_{s-},Y_s)dN_s, \end{aligned}$$

and after inserting into (4.2.10), we get

$$\begin{aligned} r(t) &= \int_{s=0}^{t} [a(s,s,Y_{s-}) + f_T(t,s)]ds \\ &+ \int_{s=0}^{t} b(s,s,Y_{s-})dW_s + \int_{s=0}^{t} \gamma^f(s,s,Y_{s-},Y_s)dN_s. \end{aligned} \quad (4.2.10)$$

On comparing coefficients between (4.2.7) and (4.2.10) the proof is complete.

□

Our final relationship is between the forward-rate process and the zero-coupon price process.

Proposition 4.2.3. *If $f(t,T)$ satisfies (4.2.3) then the dynamics for $p(t,T)$ are given by*

$$dp(t,T) = p(t,T)[m(t,T,Y_{t-})dt + v(t,T,Y_{t-})dW_t + \gamma^p(t,T,Y_{t-},Y_t)dN_t], \quad (4.2.11)$$

where

$$\begin{aligned} m(t,T,Y_{t-}) &= r(t) + A(t,T,Y_{t-}) + \frac{1}{2}|B(t,T)|^2, \\ v(t,T,Y_{t-}) &= B(t,T,Y_{t-}), \\ \gamma^p(t,T,Y_{t-},Y_t) &= exp\{\Gamma^f(t,T,Y_{t-},Y_t)\} - 1 \end{aligned}$$

and

$$A(t,T,Y_{t-}) = -\int_{s=t}^{T} a(t,s,Y_t-)ds, \tag{4.2.12}$$

$$B(t,T,Y_{t-}) = -\int_{s=t}^{T} b(t,s,Y_t-)ds, \tag{4.2.13}$$

$$\Gamma^f(t,T,Y_{t-},Y_t) = -\int_{s=t}^{T} \gamma^f(t,s,Y_t-,Y_t)ds. \tag{4.2.14}$$

Proof. Rearranging (4.2.4) we get

$$p(t,T) = exp[-\int_{s=t}^{T} f(t,s)ds],$$

and so we can therefore re-write (4.2.2) as

$$p(t,T) = exp\{Z(t,T)\},$$

where

$$Z(t,T) = -\int_{s=t}^{T} f(t,s)ds. \tag{4.2.15}$$

Writing (4.2.3) in integrated form, we have

$$f(t,s) = f(0,s)+\int_{u=0}^{t} a(u,s,Y_{u-})du+\int_{u=0}^{t} b(u,s,Y_{u-})dW_u+\int_{u=0}^{t} \gamma^f(u,s,Y_{u-},Y_u)dN_u,$$

and after inserting into (4.2.15) we get

$$\begin{aligned} Z(t,T) &= -\int_{s=t}^{T} f(0,s)ds - \int_{u=0}^{t}\int_{s=t}^{T} a(u,s,Y_{u-})dsdu \\ &\quad -\int_{u=0}^{t}\int_{s=t}^{T} b(u,s,Y_{u-})dsdW_u - \int_{u=0}^{t}\int_{s=t}^{T} \gamma^f(u,s,Y_{u-},Y_u)dsdN_u. \end{aligned}$$

Splitting the integrals and changing the order of integration gives us

$$\begin{aligned} Z(t,T) &= -\int_{s=0}^{T} f(0,s)ds - \int_{u=0}^{t}\int_{s=u}^{T} a(u,s,Y_{u-})dsdu \\ &\quad -\int_{u=0}^{t}\int_{s=u}^{T} b(u,s,Y_{u-})dsdW_u - \int_{u=0}^{t}\int_{s=u}^{T} \gamma^f(u,s,Y_{u-},Y_u)dsdN_u \\ &\quad +\int_{s=0}^{t} f(0,s)ds + \int_{u=0}^{t}\int_{s=u}^{t} a(u,s,Y_{u-})dsdu \\ &\quad +\int_{u=0}^{t}\int_{s=u}^{t} b(u,s,Y_{u-})dsdW_u + \int_{u=0}^{t}\int_{s=u}^{t} \gamma^f(u,s,Y_{u-},Y_u)dsdN_u, \end{aligned}$$

which can be written as

$$
\begin{aligned}
Z(t,T) &= Z(0,T) - \int_{u=0}^{t}\int_{s=u}^{T} a(u,s,Y_{u-})dsdu \\
&- \int_{u=0}^{t}\int_{s=u}^{T} b(u,s,Y_{u-})dsdW_u - \int_{u=0}^{t}\int_{s=u}^{T} \gamma^f(u,s,Y_{u-},Y_u)dsdN_u \\
&+ \int_{s=0}^{t} f(0,s)ds + \int_{s=0}^{t}\int_{u=0}^{s} a(u,s,Y_{u-})duds \\
&+ \int_{s=0}^{t}\int_{u=0}^{s} b(u,s,Y_{u-})dW_u ds + \int_{s=0}^{t}\int_{u=0}^{s} \gamma^f(u,s,Y_{u-},Y_u)dN_u ds.
\end{aligned}
$$

The last two lines of the above expression constitute the integrated form of the forward-rate dynamics, and since $r(s) = f(s,s)$ we can thus write

$$
\begin{aligned}
Z(t,T) &= Z(0,T) + \int_{s=0}^{t} r(s)ds - \int_{u=0}^{t}\int_{s=u}^{T} a(u,s,Y_{u-})dsdu \\
&- \int_{u=0}^{t}\int_{s=u}^{T} b(u,s,Y_{u-})dsdW_u - \int_{u=0}^{t}\int_{s=u}^{T} \gamma^f(u,s,Y_{u-},Y_u)dsdN_u.
\end{aligned}
$$

Replacing (4.4.6), (4.4.7) and (4.4.8) into the above equation we find

$$
dZ(t,T) = (r(t) + A(t,T,Y_{t-}))dt + B(t,T,Y_{t-})dW_t + \gamma^f(t,T,Y_{t-},Y_t))dN_t,
$$

and on application of Itô's lemma to $p(t,T) = exp\{Z(t,T)\}$ the proof is complete.

□

It is apparent from the above propositions, that knowledge of the zero-coupon process or of the forward-rate process is tantamount to knowing all three processes. The short-rate process however is not a one-to-one mapping of the zero-coupon prices or of the forward rates, which therefore means that specifying a process for the short rate in (4.2.1) will not determine the processes for the bond price and the forward rate in (4.2.2) and (4.2.3). This can be seen by noting that

$$f(t,t) = r_t$$

and so specifying r_s for $s \in [t,T]$ will only give us the points $f(s,s)$ and not the points $f(s,u)$ for $u > s$. So too from equation (4.2.4) it is clear that we will not be able to determine the bond prices either.

4.3 Short-Rate Models

Short rate models derive the evolution of the entire yield curve in terms of a single 1-dimensional state variable - the short rate. The first task that we are faced with is therefore the specification of the short rate.

Let consider the following model for the short rate:

$$dr_t = \mu(t, r_t, Y_{t-})dt + \sigma(t, r_t, Y_{t-})dW_t + \gamma^r(t, r_t, Y_{t-}, Y_t)dN_t \qquad (4.3.1)$$

with the functions $\mu()$,$\sigma()$ and $\gamma^r()$ being sufficiently regular, that is, satisfying (see Øksendal [2000])

$$|\mu(t, x, Y_{t-})| + |\sigma(t, x, Y_{t-})| \leq C(1 + |x|)$$

and

$$|\mu(t,x,Y_{t-}) - \mu(t,y,Y_{t-})| + |\sigma(t,x,Y_{t-}) - \sigma(t,y,Y_{t-})| \leq D|x-y|,$$

where $x, y \in \mathbf{R}$ and C and D are some constants. We also require the $\gamma^r(t, r_t, Y_{t-}, Y_t)$'s to be piecewise continuous. We shall begin by developing pricing methodologies for derivatives of this short rate, and we will subsequently discuss the *completeness* of such a model.

4.3.1 Derivatives pricing

We shall assume that the short-rate dynamics are described by (4.3.1) under the martingale measure $\mathbf{Q}$, which we assume exists. We shall discuss later the implications of modelling under this probability measure as opposed to the *real-world* probability measure $\mathbf{P}$. We can apply the *risk-neutral valuation* principle to obtain the value C_t of a sufficiently integrable T-contingent claim X_T as being

$$C_t = E_{\mathbf{Q}}[e^{-\int_t^T r(u)du} X_T | F_t]. \tag{4.3.2}$$

Let us consider T-claims of the form $X_T = f(Y_T, r(T))$. To develop the price process for this claim we first need to develop a modification to the *Feyman-Kac* formula as follows:

Theorem 4.3.1. *(Feyman-Kac formula modified). Suppose we have the following stochastic differential equation:*

$$dr_s = \mu(s, r_s, Y_{s-})ds + \sigma(s, r_s, Y_{s-})dW_s + \gamma^r(s, r_s, Y_{s-}, Y_s)dN_s, \quad t \leq s \leq T, \tag{4.3.3}$$

with initial conditions

$$\begin{aligned} r_t &= r, \\ Y_t &= j, \end{aligned}$$

where the functions $\mu(s, r_s, Y_{s-})$, $\sigma(s, r_s, Y_{s-})$ and $\gamma(s, r_s, Y_{s-}, Y_s)$ are sufficiently regular. The solution to the partial differential equation

$$\begin{aligned} &F_t(t,r,j) + \mu(t,r,j)F_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)F_{rr}(t,r,j) \\ &+ \sum_{k=1}^{n} \lambda^{jk} E\{F(t, r + \gamma(t,r,j,k), j) - F(t,r,j)\} = 0 \end{aligned} \tag{4.3.4}$$

with final condition

$$F(T, r_T, Y_T) = h(r_T, Y_T)$$

has the representation

$$F(t,r,j) = E(h(r_T, Y_T) |\, r_t = r,\ Y_t = j).$$

Proof. Consider a function of r_t given by $F(t, r_t, Y_t)$, and assume that we were just in state j, i.e. $Y_{t-} = j$. Using Itô's lemma and an analogous lemma for the jumps, we obtain

$$\begin{aligned} dF(t,r,j) &= (F_t(t,r,j) + \mu(t,r,j)F_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)F_{rr}(t,r,j))dt \\ &\quad + \sigma(t,r,j)F_r(t,r,j)dW_t + (F(t,r+\gamma(t,r,j,Y_t),j) - F(t,r,j))dN_t. \end{aligned}$$

Define for all j

$$d\tilde{N}_t = dN_t - \sum_{k=1}^{n} \lambda^{jk} dt,$$

where since

$$E[dN_t | Y_{t-} = j] = \sum_{k=1}^{n} \lambda^{jk} dt \qquad (4.3.5)$$

we have that $d\tilde{N}_t$ and hence $\tilde{N}_t$ are $\mathbf{Q}$ martingales. We can therefore write

$$\begin{aligned} dF(t,r,j) &= [F_t(t,r,j) + \mu(t,r,j)F_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)F_{rr}(t,r,j) \\ &\quad - \sum_{k=1}^{n} \lambda^{jk}(F(t,r+\gamma(t,r,j,k),j) - F(t,r,j))]dt \\ &\quad + \sigma(t,r,j)F_r(t,r,j)dW_t + [F(t,r+\gamma(t,r,j,Y_t),j) - F(t,r,j)]d\tilde{N}_t. \end{aligned}$$

If we have that

$$\begin{aligned} &F_t(t,r,j) + \mu(t,r,j)F_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)F_{rr}(t,r,j) \\ &+ \sum_{k=1}^{n} \lambda^{jk} E\{F(t,r+\gamma(t,r,j,k),j) - F(t,r,j)\} = 0 \end{aligned}$$

with boundary condition

$$F(T, r_T, Y_T) = h(r_T, Y_T)$$

we then get

$$E[dF(t,r,j)] = E[\sigma(t,r,j)F_r(t,r,j)dW_t + (F(t,r,j+\gamma(t,r,j,Y_t)) - F(t,r,j))d\tilde{N}_t].$$

Writing this in integral form, we have

$$\begin{aligned} E[F(s,r,Y_s)] &= E[F(t,r_t,j) + \int_{u=t}^{s} \sigma(u,r_u,j)F_r(u,r_u,j)dW_u \\ &\quad + \int_{u=t}^{s} (F(u,r_u,j+\gamma(u,r_u,Y_{u-},Y_u)) - F(u,r_u,Y_{u-}))d\tilde{N}_u]. \end{aligned}$$

With W_t and $\tilde{N}_u$ being martingales with initial values 0, we therefore have that

$$E(F(s,r,Y_s)|\, r_t = r,\ Y_t = j) = F(t,r_t,j), \qquad (4.3.6)$$

and so finally

$$F(t,r,j) = E(h(r_T,Y_T)|\, r_t = r,\ Y_t = j)$$

as required.

□

If we assume that X_T in (4.3.2) is of the form $X_T = \Phi(r_T, Y_T)$, we can use the above theorem to obtain the following proposition:

Proposition 4.3.2. *(Term Structure Equation). Suppose we have a T-contingent claim $X_T = f(r_T, Y_T)$, then the arbitrage-free value of price process C_t, where*

$$C_t = C(t, r_t, Y_t),$$

is given by (4.3.2), is a solution of the partial differential equation

$$-rC(t,r,j) + C_t(t,r,j) + \mu(t,r,j)C_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)C_{rr}(t,r,j)$$

$$+\sum_{k=1}^{n} \lambda^{jk} E\{C(t, r+\gamma(t,r,j,k), j) - C(t,r,j)\} = 0 \tag{4.3.7}$$

with terminal condition $C(T, r, Y_T) = f(r, Y_T)$ for all $r \in \Re$, $Y_T \in \mathcal{S}$.

□

We can now write the following result:

Corollary 4.3.3. *(T-bond prices) We can calculate the price of a T-bond $p(t,T)$ as being the solution to (4.3.7) with terminal condition $C(T, r, Y_T) = 1$.*

□

We will now price derivatives whose payoffs are functions of the T-bond prices, such as a European call option with payoff

$$X_T = \max\{p(S,T) - K, 0\},$$

where the maturity is S and the strike is K. To do this will require a two step-process. We must first calculate the value of the underlying T-bond by using corollary (4.3.3). We must then solve equation (4.3.7) for the value of the derivative $G(t, r_t, Y_t)$, giving us

$$-rG(t,r,j) + G_t(t,r,j) + \mu(t,r,j)G_r(t,r,j) + \frac{1}{2}\sigma^2(t,r,j)G_{rr}(t,r,j)$$
$$+\sum_{k=1}^{n} \lambda^{jk} E\{G(t, r + \gamma(t,r,j,k), j) - G(t,r,j)\} = 0 \tag{4.3.8}$$

with terminal condition

$$G(S,r,T) = \max\{p(S,T) - K, 0\}$$

for all $r \in \Re$, $Y_T \in \mathcal{S}$.

To solve these two partial differential equations may prove to be rather tricky. The standard way of ensuring that these equations remain solvable is by restricting the possible models for the short rate to those in which the bond prices possess an *affine term structure* (see Duffie [1992]). But as we shall see, even with models which exhibit such a structure pricing derivatives using the MCJD model is not totally straightforward.

4.3.2 Models of the Short Rate

There are many models which can be used to describe the behaviour of the short rate. Some of these models may allow an exact fit to the initial term structure or possess more preferable term structures than others. The advantages and disadvantages of various short rate models can be found in books such as Björk [1998] or Duffie [1992]. In our Markov chain jump-diffusion setting, we will consider extensions of two of the most predominant models. We will firstly show how to price derivatives of the short rate for each of the models without having to solve the partial differential equations in the previous section, and subsequently we will discuss how to estimate parameter values. In this section we shall assume that we are modelling under the risk-neutral measure **Q**. In the next section we shall explore the ramifications of such an assumption.

The Ho-Lee Model

We shall extend the Ho-Lee model first introduced in Ho and Lee [1986] so that the short rate has the following dynamics:

$$dr_t = \alpha_t dt + \sum_{j=1}^{n} I_t^j \left[\sigma_j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right], \tag{4.3.9}$$

where σ_j is previsible and strictly greater than 0. The values of α_t are chosen so that the short-rate model fits the initial term structure.

The Ho-Lee model is generally not the most preferred short rate model as it does not possess the mean-reverting property, as well as having the undesirable property that the short rate can be negative. However, the advantage it does possess over the other models is its ease of computation.

We shall consider a market where we have an empirical term structure $\{\hat{p}(0,T)\ \ 0 < t \leq T\}$ and so too an observed forward-rate curve $\{\hat{f}(0,t)\ \ 0 < t \leq T\}$. We shall now derive the values of α_t for $0 < t \leq T$ so that the model fits this initial term structure.

Lemma 4.3.4. *Suppose the short rate follows an extended version of the Ho-Lee model as given in equation (4.3.9). In order for this model to fit the initial term structure we require that*

$$\alpha_t = \hat{f}'(0,t) + M''(0,t), \qquad 0 \leq t \leq T, \tag{4.3.10}$$

where

$$M(t,T) = logE\left[exp\left\{-\int_{u=t}^{T}(\int_{s=0}^{u}\sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n}\sum_{k=1}^{n}\gamma^{jk}N_u^{jk})du\right\}\right]$$

and where $'$ represents the partial derivative with respect to t.

Proof. The value of the time t zero-coupon bond $p(0,t)$ is given by

$$p(0,t) = E\left[exp\left\{-\int_{u=0}^{t} r(u)du\right\}\right]. \tag{4.3.11}$$

Integrating (4.3.9), we get

$$r_t = r_0 + \int_{s=0}^{t} \alpha_s ds + \int_{s=0}^{t} \sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma^{jk} N_t^{jk}, \tag{4.3.12}$$

and inserting this into equation (4.3.11) we obtain

$$p(0,t) = E\left[exp\left\{-\int_{u=0}^{t}\left(r_0 + \int_{s=0}^{u} \alpha_s ds + \int_{s=0}^{u} \sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n}\sum_{k=1}^{n} \gamma^{jk} N_u^{jk}\right) du\right\}\right].$$

Writing

$$exp\{M(t,T)\} = E\left[exp\left\{-\int_{u=t}^{T}\left(\int_{s=0}^{u} \sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n}\sum_{k=1}^{n} \gamma^{jk} N_u^{jk}\right) du\right\}\right] \tag{4.3.13}$$

we therefore get

$$p(0,t) = exp\left\{-r_0 t - \int_{u=0}^{t}\int_{s=0}^{u} \alpha_s ds du + M(0,t)\right\}. \tag{4.3.14}$$

We then have

$$\begin{aligned} f(0,t) &= -\frac{\partial log\, p(0,t)}{\partial T} \\ &= r_0 + \int_{s=0}^{t} \alpha_s ds - M'(0,t). \end{aligned}$$

Differentiating once more, we find

$$f'(0,t) = \alpha_t - M''(0,t). \tag{4.3.15}$$

Rearranging (4.3.15) gives us (4.3.10). □

In appendix B we develop a methodology to calculate the value of $M(t, T)$, the result of which is given in corollary B.0.10.

Turning our attention to the zero-coupon bond prices, we have the following lemma:

Lemma 4.3.5. *The time* 0 *value of the zero-coupon bond* $p(t, T)$ *at time* t *is given by*

$$p(t, T) = p_0 exp\{M(t, T) - M(0, T) + M(0, t)\}, \tag{4.3.16}$$

where

$$p_0 = \frac{\hat{p}(0, T)}{\hat{p}(0, t)},$$

and where $\hat{p}(0, t)$ *once again refers to the empirical time* t *zero-coupon bond price.*

Proof. We have that

$$p(t, T) = E\left[exp\left\{-\int_{u=t}^{T} r(u)du\right\}\right],$$

and on inserting the short rate given by 4.3.12 we find

$$p(t, T) = E\left[exp\left\{-\int_{u=t}^{T}(r_0 + \int_{s=0}^{u} \alpha_s ds + \int_{s=0}^{u}\sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n}\sum_{k=1}^{n}\gamma^{jk} N_u^{jk})du\right\}\right].$$

Calculating this expected value we get

$$\begin{aligned} p(t,T) &= exp\left\{-(T-t)r_0 - \int_{u=t}^{T}\int_{s=0}^{u}\alpha_s ds + M(t,T)\right\} \\ &= exp\left\{-(T-t)r_0 - \int_{u=0}^{T}\int_{s=0}^{u}\alpha_s ds + \int_{u=0}^{t}\int_{s=0}^{u}\alpha_s ds + M(t,T)\right\}, \end{aligned}$$

and finally on inserting (4.3.14) we get (4.3.16).

□

The Hull-White (Extended Vasicek) Model

The second short-rate model we shall consider is an extension of the Hull-White (Extended Vasicek) model (see Hull and White [1990a]), where the short rate has the following dynamics:

$$dr_t = \sum_{j=1}^{n} I_t^j \left[(\alpha_t - \beta_t r_t)dt + \sigma_j dW_t + \sum_{k=1}^{n}\gamma^{jk} dN_t^{jk}\right] \qquad (4.3.17)$$

where σ_j is previsible and strictly greater than 0. The β_t term can be set as a function of the current state j, the time t or both. We shall consider models where β_t is a deterministic function of the time t only. The values of α_t are once again chosen so that the short-rate model fits the initial term structure.

The Hull-White model is generally preferred to the Ho-Lee as it possesses the mean reverting property characterised above by the term β_t. We shall again consider a market where we have an empirical term structure $\{\hat{p}(0,T) \ \ 0 < t \leq T\}$ and so too an observed forward rate curve $\{\hat{f}(0,t) \ \ 0 < t \leq T\}$. Let us now derive the values of α_t for $0 < t \leq T$ so that the model fits this initial term structure.

Lemma 4.3.6. *Suppose the short rate follows an extended version of the Hull-White model as given in equation (4.3.17). In order for this model to fit the initial term structure we require that*

$$\alpha_t = f'(0,t) + \beta_t f(0,t) + \beta_t M'(0,t) + M''(0,t), \qquad 0 \leq t \leq T, \quad (4.3.18)$$

where

$$\begin{aligned} M(t,T) &= logE\left[exp\left\{-\int_{u=0}^{t} e^{-K(u)}\left(\int_{s=0}^{u}\sum_{j=1}^{n} I_s^j e^{K(s)}\sigma_j dW_s\right.\right.\right. \\ &\quad \left.\left.\left. + \int_{s=0}^{u}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{K(s)}\gamma^{jk} dN_s^{jk}\right) du\right\}\right] \end{aligned}$$

and where $'$ once again represents the partial derivative with respect to t.

Proof. The value of the time t zero-coupon bond $p(0,t)$ is given by

$$p(0,t) = E\left[exp\left\{-\int_{u=0}^{t} r(u)du\right\}\right]. \qquad (4.3.19)$$

We can re-write (4.3.17) as

$$dr_t = \sum_{j=1}^{n} I_t^j \left[(\alpha_t - \beta_t r_t)dt + \sigma_j dW_t + \sum_{k=1}^{n}\gamma^{jk} dN_t^{jk}\right]. \qquad (4.3.20)$$

Integrating (4.3.20) we get (see Rogers [1995])

$$r_t = e^{-K(t)} \left\{ r_0 + \int_{s=0}^{t} e^{K(s)} \left(\alpha_s ds + \sum_{j=1}^{n} I_s^j \sigma_j dW_s + \sum_{j=1}^{n} \sum_{k=1}^{n} \gamma^{jk} dN_s^{jk} \right) \right\}, \tag{4.3.21}$$

where

$$k(t) = \int_{u=0}^{t} \beta_u du,$$

which is known at time 0 as β_t is a deterministic function of t as stated above.

Inserting (4.3.21) into equation (4.3.19) we find

$$\begin{aligned} p(0,t) &= E\left[exp\left\{ -\int_{u=0}^{t} e^{-K(u)} (r_0 + \int_{s=0}^{u} e^{K(s)} \alpha_s ds + \int_{s=0}^{u} \sum_{j=1}^{n} I_s^j e^{K(s)} \sigma_j dW_s \right.\right. \\ &\left.\left. + \int_{s=0}^{u} \sum_{j=1}^{n} \sum_{k=1}^{n} e^{K(s)} \gamma^{jk} dN_s^{jk}) du \right\} \right]. \end{aligned}$$

Writing

$$\begin{aligned} exp\{M(t,T)\} &= E\left[exp\left\{ -\int_{u=0}^{t} e^{-K(u)} (\int_{s=0}^{u} \sum_{j=1}^{n} I_s^j e^{K(s)} \sigma_j dW_s \right.\right. \\ &\left.\left. + \int_{s=0}^{u} \sum_{j=1}^{n} \sum_{k=1}^{n} e^{K(s)} \gamma^{jk} dN_s^{jk}) du \right\} \right] \end{aligned} \tag{4.3.22}$$

we therefore get

$$p(0,t) = exp\left\{-\int_{u=0}^{t} e^{-k(u)} r_0 dt - \int_{u=0}^{t}\int_{s=0}^{u} e^{k(s)}\alpha_s ds du + M(0,t)\right\}. \tag{4.3.23}$$

We then have

$$\begin{aligned} f(0,t) &= -\frac{\partial log\, p(0,t)}{\partial T} \\ &= e^{-k(t)}\{r_0 + \int_{s=0}^{t} e^{k(s)}\alpha_s ds\} - M'(0,t). \end{aligned}$$

Differentiating once more we find

$$f'(0,t) = -\beta_t f(0,t) - \beta_t M'(0,t) + \alpha_t - M''(0,t). \tag{4.3.24}$$

Rearranging (4.3.24) gives us (4.3.18).

□

The calculation of $M(t,T)$ in equation (4.3.22) in the Hull-White model, can be seen to be somewhat more complicated than that in equation (4.3.13) for the Ho-Lee model, and will therefore need to be calculated numerically.

Turning our attention to the zero-coupon bond prices we have the following lemma:

Lemma 4.3.7. *The time* 0 *value of the zero-coupon bond* $p(t,T)$ *at time* t *is given by*

$$p(t,T) = p_0 exp\{M(t,T) - M(0,T) + M(0,t)\}, \tag{4.3.25}$$

where

$$p_0 = \frac{\hat{p}(0,T)}{\hat{p}(0,t)}.$$

Proof. We have that

$$p(t,T) = E\left[exp\left\{-\int_{u=t}^{T} r(u)du\right\}\right],$$

and on inserting the short rate given by 4.3.20 we find

$$\begin{aligned} p(t,T) &= E\left[exp\left\{-\int_{u=t}^{T} e^{-K(u)}(r_0 + \int_{s=0}^{u} e^{K(s)}\alpha_s ds + \int_{s=0}^{u}\sum_{j=1}^{n} I_s^j e^{K(s)}\sigma_j dW_s\right.\right. \\ &\left.\left.+\int_{s=0}^{u}\sum_{j=1}^{n}\sum_{k=1}^{n} e^{K(s)}\gamma^{jk} dN_s^{jk})du\right\}\right]. \end{aligned}$$

Calculating this expectation gives us

$$\begin{aligned} p(t,T) &= exp\left\{-\int_{u=t}^{T} e^{-K(u)}r_0 du - \int_{u=t}^{T}\int_{s=0}^{u} e^{K(s)}\alpha_s dsdu + M(t,T)\right\} \\ &= exp\left\{-\int_{u=0}^{T} e^{-K(u)}r_0 du - \int_{u=0}^{T}\int_{s=0}^{u} e^{K(s)}\alpha_s dsdu\right. \\ &\left.+\int_{u=0}^{t} e^{-K(u)}r_0 du + \int_{u=0}^{t}\int_{s=0}^{u} e^{K(s)}\alpha_s dsdu + M(t,T)\right\}, \end{aligned}$$

and finally on inserting (4.3.23) we get (4.3.25).

□

4.3.3 Market Completeness

In the previous section we assumed that the short rate was modelled under the risk-neutral measure $\mathbf{Q}$. The drawback of such an assumption is that in practice we only observe data under the *real-world* probability measure $\mathbf{P}$, and so we are clearly going to encounter difficulties in parameter estimation. An alternative would have been to assume that short-rate dynamics are given by (4.3.1) under the measure $\mathbf{P}$. However, this would lead us into other difficulties, as we shall now see.

Modelling under P

Let us assume now that the short rate is modelled by (4.3.1) under the measure $\mathbf{P}$. Define as in theorem (2.3.2)

$$\begin{aligned} L_t &= exp\left\{ -\sum_{k=1}^{n} \int_0^t (1-\psi^{jk})\lambda^{jk} I_s^j ds - \frac{1}{2}\int_0^t \|\theta(s,r_s,Y_s)\|^2 ds \right. \\ &\quad \left. + \int_0^t \theta(s,r_s,Y_s)dW_s \right\} \prod_{k=1}^{n} (\psi^{jk})^{N_t^{jk}}. \end{aligned}$$

As in theorem (2.3.2), this Girsanov density will yield an equivalent measure $\tilde{\mathbf{P}}$ under which equation (4.3.1) becomes

$$\begin{aligned} dr_t &= (\mu(t,r_t,Y_{t-}) - \sigma(t,r_t,Y_{t-})\theta(t,r_t,Y_{t-}))dt \\ &\quad + \sigma(t,r_t,Y_{t-})d\bar{W}_t + \gamma^r(t,r_t,Y_{t-},Y_t)d\bar{N}_t, \qquad (4.3.26) \end{aligned}$$

where $\bar{W}_t$ is a $\bar{\mathbf{P}}$ Brownian motion, and the transition intensities of the Markov chain undergo the transformation $\lambda^{jk} \longrightarrow \lambda^{jk}\psi^{jk}\ \forall\ k$ under which the counting process $\bar{N}_t$ is defined. Using theorem 4.3.1, we can see that any T-contingent claim $X_T = f(Y_T, r_T)$ which has present value $F_t(t, r, j)$ given by (4.3.2), must solve the partial differential equation

$$
\begin{aligned}
&F_t(t,r,j) + (\mu(t,r,j) - \sigma(t,r_t,j)\theta(t,r_t,j))F_r(t,r,j) \\
&+\frac{1}{2}\sigma^2(t,r,j)F_{rr}(t,r,j) + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}E\{F(t,r+\gamma(t,r,j,k),j) - F(t,r,j)\} = 0
\end{aligned}
$$

with final condition

$$F(T, r_T, Y_T) = f(r(T), Y_T).$$

The difficulty here arises since $\theta(t, r_t, Y_{t-})$ and the ψ^{jk}'s are not specified within the model. This therefore leads us to an *incomplete market* situation. This is due to the fact that the short rate is not actually tradable, which leaves us only with the risk-neutral bank account to try to set up replicating portfolios. Therefore, all we will be able to achieve when pricing bonds whose prices are determined by the short rate, or alternatively which are derivatives of the short rate, is to ensure that given we have a portfolio of bonds under which $\theta(t, r_t, Y_{t-})$ and the ψ^{jk}'s can be determined uniquely, then we can ensure that all other bonds are priced consistently. Alternatively, we can use statistical methods similar to those in the next subsection to calibrate the 'best' parameter values given some objective function.

Modelling under Q

As mentioned earlier, modelling under **Q** leaves us the problem of having to calibrate the parameter values which are observed under **P**. We therefore need to develop methods for determining the parameter values. For this we shall use a similar methodology to that described in Björk [1997] and brought in Bingham and Kiesel [2004].

The process of calibration of a term-structure model involves solving the following system of equations:

$$\hat{p}(0,T) = p(0,T;\zeta) \tag{4.3.27}$$

for all $T \geq 0$, where $\hat{p}(0,T)$ are the observed bond prices and $p(0,T;\zeta)$ are the estimated values using the parameter vector ζ. In order to solve exactly for ζ we require the number of parameters to be equal to the number of bonds in the market. Theoretically, (4.3.27) represents an infinite system of equations, and so only short-rate models with an infinite-dimensional parameter vector will be able to give us an exact fit. We would therefore need to find the 'best' fitting parameter vector using a given objective function.

In practice however there are not an infinite number of bonds. Suppose there are m bonds in the market, and we choose to model the short rate using an extension of the Hull-White extended Vasicek model, where the short-rate dynamics are given by

$$dr_t = \sum_{j=1}^{n} I_t^j \left[(\alpha^j - \beta^j r)dt + \sigma^j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \right] \tag{4.3.28}$$

for $j = 1, \ldots, n$. This means the parameter vector ζ contains 3 parameters for each state, α^j, β^j and σ^j, as well as n parameters for each transition possibility, γ^{jk} for $k = 1, \ldots, n$. This leaves us with a total of $3n + n^2$ parameters to solve for. So if we have that $m = 3n + n^2$ then there will be an equal number of parameters and assets (and so too equations), and there can therefore be an exact fit. If we have more or less bonds on the market, then the model will not even price current bonds on the market correctly, let alone other derivatives.

4.4 HJM Models

There are many well-documented drawbacks in using short-rate models, such as the modelling the yield curve which is an infinite-dimensional space using only one factor which will generally prove to be inadequate, as well as the assumption that the yields of all maturities are perfectly correlated. These deficiencies as well as others have motivated the development of models which contain more than one explanatory variable.

At the other extreme, the Heath-Jarrow-Morton (HJM) methodology tries to improve on the short rate models in the previous section by directly modelling the evolution of the entire yield curve simultaneously. We will now apply this methodology to our Markov chain jump-diffusion model.

We will model the forward rates as having the following dynamics:

$$df(t,T) = a(t,T,Y_{t-})dt + \sum_{b=1}^{r} s_b(t,T,Y_{t-})dW_t^b + \gamma^f(t,T,Y_{t-},Y_t)dN_t \quad (4.4.1)$$

for all $T < T^*$, where the W_T^b's are Brownian motions under the objective probability measure $\mathbf{P}$. The bond-price dynamics are then given by

$$dp(t,T) = p(t,T)\left[m(t,T,Y_{t-})dt + \sum_{b=1}^{r} v_b(t,T,Y_{t-})dW_t^b + \gamma^p(t,T,Y_{t-},Y_t)dN_t\right], \quad (4.4.2)$$

where using proposition 4.2.3 we have

$$m(t,T,Y_{t-}) = r(t) + A(t,T,Y_{t-}) + \frac{1}{2}||S(t,T)||^2, \quad (4.4.3)$$

$$v_b(t,T,Y_{t-}) = S_b(t,T,Y_{t-}), \quad (4.4.4)$$

$$\gamma^p(t,T,Y_{t-},Y_t) = exp\{\Gamma^f(t,T,Y_{t-},Y_t)\} - 1, \quad (4.4.5)$$

and

$$A(t,T,Y_{t-}) = -\int_{u=t}^{T} a(t,u,Y_{t-})du, \quad (4.4.6)$$

$$S_b(t,T,Y_{t-}) = -\int_{u=t}^{T} s_b(t,u,Y_{t-})du, \quad (4.4.7)$$

$$\Gamma^f(t,T,Y_{t-},Y_t) = -\int_{u=t}^{T} \gamma^f(t,u,Y_{t-},Y_t)du. \quad (4.4.8)$$

4.4.1 Martingale Measure

For a numéraire we shall take the bank account process given by

$$B_T = exp\{\int_{u=0}^{t} f(u,u)du\} = exp\{\int_{u=0}^{t} r(u)du\}.$$

We must now find a measure so that

$$Z(t,T) = \frac{p(t,T)}{B_t} \tag{4.4.9}$$

is a martingale for every $0 \leq T \leq T^*$, where $p(t,T)$ is given by (4.4.2) under the probability measure $\mathbf{P}$. The theorem below explores the existence of such a measure.

Theorem 4.4.1. *Assume that the family of forward rates is determined by (4.4.1). Whilst in state j, let $(N_t^{j1}, \ldots, N_t^{jn})$ be an n-point process with intensities $(\lambda^{j1}, \ldots, \lambda^{jn})$, where*

$$N_t^{jk} = |\{u; 0 < u < t, Y_u = k, Y_{u-} = j\}|.$$

Let $(\psi^{j1}, \ldots, \psi^{jn})$ be an $\mathcal{F}_t$ predictable process where $\psi^{jk} \geq 0 \ \forall \, j,k$ so that $\forall \ t \leq T$ we have

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\int_{u=0}^{t} \psi^{jk}\lambda^{jk}I_u^j du < \infty.$$

Suppose there exist predictable processes $\Theta(t,T,Y_t)^{r\times 1}$ which satisfy the usual regularity condition (see theorem 2.3.2). Define

$$L_t = exp\left\{\int_{u=0}^{t}\sum_{j=1}^{n} I_u^j\left[-\left(\sum_{k=1}^{n}(1-\psi^{jk})\lambda^{jk}+\frac{1}{2}\|\Theta_u\|^2\right)du+\Theta_u' d\mathbf{W}(u)\right]\right\}\prod_{k=1}^{n}(\psi^{jk})^{N_t^{jk}},$$

where

$$E^P[L_t]=1,\ \ L_0=1.$$

If we also have for each state j

$$\begin{aligned} a(t,T,j) &= \sum_{i=1}^{d} s_b(t,T,j)\{S_b(t,T,j)+\theta_b(t,T,j)\} \\ &\quad -\sum_{k=1}^{n}\lambda^{jk}\psi^{jk}\gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\} \quad (4.4.10) \end{aligned}$$

for $j=1,\dots,n$, then there exists a risk-neutral martingale measure.

Proof. From theorem (2.3.2) we can see that under such a measure $\mathbf{Q}$ we induce a change both in the drift and the transition intensities of the Markov chain. Under this probability measure we can re-write (4.4.2) whilst in state j as

$$\begin{aligned} dp(t,T) &= p(t,T)\left[(m(t,T,j)-\sum_{b=1}^{r}v_b(t,T,j)\theta_b(t,T,j)dt)\right. \\ &\quad \left.+\sum_{b=1}^{r}v_b(t,T,j)d\tilde{W}_t+\sum_{k=1}^{n}\gamma^p(t,T,j,k)dN_t^{jk}\right], \quad (4.4.11) \end{aligned}$$

where

$$E[dN_t^{jk}] = \lambda^{jk}\psi^{jk}dt.$$

Defining

$$d\tilde{N}_t^{jk} = dN_t^{jk} - \lambda^{jk}\psi^{jk}dt$$

and similarly

$$d\tilde{N}_t = dN_t - \sum_{j=1}^{n}\sum_{k=1}^{n} I_t^j \lambda^{jk}\psi^{jk}dt,$$

we see that $\tilde{N}_t$ is a martingale under the measure $\mathbf{Q}$. We can now express (4.4.11) as

$$\begin{aligned} dp(t,T) &= p(t,T)[(m(t,T,j) - \sum_{b=1}^{r} v_b(t,T,j)\theta_b(t,T,j) + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}\gamma^p(t,T,j,k))dt \\ &\quad + \sum_{b=1}^{r} v_b(t,T,j)d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma^p(t,T,j,k)d\tilde{N}_t^{jk}]. \end{aligned}$$

From (4.4.9) we can see that

$$dZ_t = dB_t^{-1}p(t,T) + B_t^{-1}dp(t,T),$$

which on applying Itô's lemma can be written in terms of the forward coefficients as follows:

$$
\begin{aligned}
dZ_t &= Z_t \Bigg[\Bigg\{ A(t,T,j) + \frac{1}{2}||S(t,T,j)||^2 - \sum_{b=1}^{r} S_b(t,T,j)\theta_b(t,T,j) \\
&\quad + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}(exp\{\Gamma^f(t,T,j,k)\} - 1) \Bigg\} dt \\
&\quad + \sum_{b=1}^{r} S_b(t,T,j)d\tilde{W}_t^b + \sum_{k=1}^{n}(exp\{\Gamma^f(t,T,j,k)\} - 1)d\tilde{N}_t^{jk} \Bigg].
\end{aligned}
$$

For Z_t to be a $\mathbf{Q}$-martingale we therefore require that

$$
\begin{aligned}
A(t,T,j) + \frac{1}{2}||S(t,T,Y_j)||^2 - \sum_{b=1}^{r} S_b(t,T,j)\theta_b(t,T,j)& \\
+ \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}(exp\{\Gamma^f(t,T,j,k)\} - 1) &= 0, \quad (4.4.12)
\end{aligned}
$$

and on differentiating with respect to T we are left with the condition

$$
\begin{aligned}
-a(t,T,j) - \sum_{b=1}^{r} s_b(t,T,j) \int_{u=t}^{T} s_b(t,u,j)du + \sum_{b=1}^{r} s_b(t,T,j)\theta_b(t,T,j)& \\
- \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}\gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\} &= 0.
\end{aligned}
$$

Rearranging gives us (4.4.10).

□

In the HJM model all objects are specified directly under the risk-neutral measure $\mathbf{Q}$. In order for us to do this we must set $\theta_i = 0 \quad \forall i$, as well as $\psi^{jk} = 1 \ \forall j, k$. We are then left with the following model.

Corollary 4.4.2. *(Extended HJM) Assume we are modelling under the risk-neutral measure* $\mathbf{Q}$*, and that the forward-rate dynamics are given by*

$$df(t,T) = a(t,T,Y_{t-})dt + \sum_{b=1}^{r} s_b(t,T,Y_{t-})d\tilde{W}_t^b + \gamma^f(t,T,Y_{t-},Y_t)d\tilde{N}_t. \quad (4.4.13)$$

We then have the necessary drift condition whilst in state j *given by*

$$a(t,T,j) = \sum_{b=1}^{r} s_b(t,T,j) \int_{u=t}^{T} s_b(t,u,j)du - \sum_{k=1}^{n} \lambda^{jk}\gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\}. \quad (4.4.14)$$

Integrating the forward-rate dynamics, the short-rate process can then be expressed as

$$\begin{aligned} r(T) &= f(0,T) + \int_{u=0}^{T} a(u,T,Y_{u-})du + \int_{u=0}^{T} \sum_{b=1}^{r} s_b(u,T,Y_{u-})d\tilde{W}_u^b \\ &\quad + \int_{u=0}^{T} \gamma^f(u,T,Y_{u-},Y_u)d\tilde{N}_u, \end{aligned} \quad (4.4.15)$$

and the bond-price dynamics are given by

$$dp(t,T) = p(t,T)[r_t dt + \sum_{b=1}^{r} S_b(t,T,Y_{t-})d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma^p(t,T,j,k)d\tilde{N}_t^{jk}]. \quad (4.4.16)$$

□

4.4.2 Forward Rate Measure

For interest-rate models it is often more appropriate to use a bond maturing at date T^* with price $p(t, T^*)$ as the numéraire. The associated martingale measure using this bond as a numéraire is then called the *T^*-forward martingale measure* $\mathbf{Q}^*$, such that under $\mathbf{Q}^*$ the process

$$Z^*(t,T) = \frac{p(t,T)}{p(t,T^*)} \tag{4.4.17}$$

is a martingale for all $T < T^*$, where $p(t,T)$ is given by (4.4.2) under the probability measure $\mathbf{P}$.

To find such a measure we can amend theorem 4.4.1 as follows :

Theorem 4.4.3. *As in theorem 4.4.1, the Girsanov density L_t for the change of measure to $\mathbf{Q}^*$ induces a change in the drift of $\theta_b(t, T, Y_{t-})$ for each Brownian motion b, as well as change in the transition intensities from λ^{jk} to $\psi^{jk}\lambda^{jk}$. The necessary drift condition for the process $Z^*(t,T)$ given by (4.4.17) to be a $\mathbf{Q}^*$-martingale now becomes*

$$\begin{aligned} a(t,T,j) &= \sum_{b=1}^{r} s_b(t,T,j)\{S_b(T,T^*,j) + \theta_b(t,T,j)\} + \\ & \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}(\gamma^f(t,T,j,k) exp\{\Gamma^f(t,T,j,k)\} - \\ & \gamma^f(t,T^*,j,k) exp\{\Gamma^f(t,T^*,j,k)\}). \end{aligned} \tag{4.4.18}$$

Proof. As in theorem 4.4.1 we have

$$\begin{aligned} dp(t,T) &= p(t,T)\Bigg[\Bigg\{m(t,T,j) - \sum_{b=1}^{r} v_b(t,T,j)\theta_b(t,T,j) \\ &\quad + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}\gamma^p(t,T,j,k)\Bigg\} dt \\ &\quad + \sum_{b=1}^{r} v_b(t,T,j)d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma^p(t,T,j,k)d\tilde{N}_t^{jk}\Bigg]. \end{aligned}$$

From equation (4.4.17) we can see that

$$dZ_t^* = dp^{-1}(t,T^*)p(t,T) + p^{-1}(t,T^*)dp(t,T),$$

where on applying Itô's lemma and writing in terms of the forward rate coefficients given by (4.4.3), (4.4.4) and (4.4.5) gives us whilst in state j

$$
\begin{aligned}
dZ_t^* \;=\; Z_t^* \Bigg[\Bigg\{ & A(t,T,j) - A(t,T^*,j) + \frac{1}{2}(||S(t,T,j)||^2 - ||S(t,T^*,j)||^2) \\
& - \sum_{b=1}^{r} (S_b(t,T,j) - S_b(t,T^*,j))\theta_b(t,T,j) \\
& + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}(exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(t,T^*,j,k)\}) \Bigg\} dt \\
& + \sum_{b=1}^{r} \{S_b(t,T,j) - S_b(t,T^*,j)\} d\tilde{W}_t^b \\
& + \sum_{k=1}^{n} (exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(t,T^*,j,k)\}) d\tilde{N}_t^{jk} \Bigg].
\end{aligned}
$$

For Z_t^* to be a $\mathbf{Q}$-martingale we therefore require that

$$
\begin{aligned}
& A(t,T,j) - A(t,T^*,j) + \frac{1}{2}(||S(t,T,j)||^2 - ||S(t,T^*,j)||^2) \\
& - \sum_{b=1}^{r} (S_b(t,T,j) - S_b(t,T^*,j))\theta_b(t,T,j) \\
& + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}[exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(t,T^*,j,k)\}] \quad = 0,
\end{aligned}
$$

and on inserting (4.4.6), (4.4.7) and (4.4.8) this condition can be expressed as

$$\int_{u=T}^{T^*} a(t,u,j)du + \frac{1}{2}\left\|\sum_{b=1}^{r}\int_{u=T}^{T^*} s_b(t,u,j)du\right\|^2 - \sum_{b=1}^{r}\theta_b(t,T,j)\int_{u=T}^{T^*} s_b(t,u,j)du$$

$$+\sum_{k=1}^{n}\lambda^{jk}\psi^{jk}[exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(t,T^*,j,k)\}] = 0.$$

Differentiating with respect to T we are left with the condition

$$a(t,T,j) + \sum_{b=1}^{r} s_b(t,T,j)\int_{u=T}^{T^*} s_b(t,u,j)du - \sum_{b=1}^{r} s_b(t,T,j)\theta_b(t,T,j) +$$

$$\sum_{k=1}^{n}\lambda^{jk}\psi^{jk}[\gamma^f(t,T^*,j,Y_t)exp\{\Gamma^f(t,T^*,j,k)\} - \gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\}] = 0.$$

Rearranging gives us (4.4.18).

□

Once again we assume for the HJM model that all objects are specified directly under the risk-neutral measure $\mathbf{Q}^*$ so that $\theta_i = 0 \;\; \forall i$, as well as $\psi^{jk} = 1 \;\; \forall j,k$. We then have the following drift condition:

Corollary 4.4.4. *The HJM drift condition whilst in state j under $\mathbf{Q}^*$ is given by*

$$\begin{aligned} a(t,T,j) &= \sum_{b=1}^{r} s_b(t,T,j)S_b(T,T^*,j) \\ &+\sum_{k=1}^{n} \lambda^{jk}(\gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\} \\ &-\gamma^f(t,T^*,j,k)exp\{\Gamma^f(t,T^*,j,k)\}), \end{aligned} \tag{4.4.19}$$

and we are then left with the $\mathbf{Q}^*$-martingale

$$\begin{aligned} dZ^* &= Z^*[\sum_{b=1}^{r}\{S_b(t,T,j)-S_b(t,T^*,j)\}d\tilde{W}_t^b \\ &+\sum_{k=1}^{n}(exp\{\Gamma^f(t,T,j,k)\}-exp\{\Gamma^f(t,T^*,j,k)\})d\tilde{N}_t^{jk}]. \end{aligned}$$

□

We have developed the main mathematical machinery that we will use to price these interest-rate derivatives. Let us now turn our attention to developing pricing formulas.

4.4.3 Derivatives Pricing

The first and most important derivative that we need to price is the zero-coupon bond. As was mentioned earlier, we can price these bonds by the following formula:

$$p(t,T) = E\left[exp\left\{\int_{u=t}^{T} r(u)du\right\}\right], \tag{4.4.20}$$

where the process $r(u)$ is given by (4.4.15). Finding the solution to this will often be very difficult, which will mean that we may have to resort to using numerical methods. Alternatively, there may be situations where the HJM model will correspond to a particular short-rate model, and so we may then use the pricing techniques developed in the previous section. We shall now give a couple of examples which can be translated into short-rate models.

Example 4.4.5. Suppose we have parameter values

$$\begin{aligned} s_b(t,T,j) &= s_{b,j} \times exp\{-\kappa^s(T-t)\}, & (4.4.21)\\ \gamma^p(t,T,j,k) &= \gamma^{jk} \times exp\{-\kappa^\gamma(T-t)\} - 1, & (4.4.22)\end{aligned}$$

where $\kappa^s > 0$ and $\kappa^\gamma > 0$ so that the volatility and jump sizes are exponentially damped. On differentiating (4.4.5), we see that

$$\frac{\partial}{\partial T}\gamma^p(t,T,Y_{t-},Y_t) = -\gamma^f(t,T,Y_{t-},Y_t)exp\{\Gamma^f(t,T,Y_{t-},Y_t)\}, \tag{4.4.23}$$

and together with (4.4.21) and (4.4.22) we can see that that the drift in state j given by (4.4.14) becomes

$$
\begin{aligned}
a(j,t,T) &= \sum_{b=1}^{r} s_{b,j} exp\{-\kappa^s(T-t)\} \int_{u=t}^{T} s_{b,j} exp\{-\kappa^s(u-t)\} du \\
&\quad + \kappa^\gamma \sum_{k=1}^{n} \lambda^{jk} \gamma^{jk} exp\{-\kappa^\gamma(T-t)\} \\
&= \frac{1}{\kappa^s} exp\{-\kappa^s(T-t)\}[1 - exp\{-\kappa^s(T-t)\}](\sum_{b=1}^{r} s_{b,j}^2) \\
&\quad + \kappa^\gamma exp\{-\kappa^\gamma(T-t)\}(\sum_{k=1}^{n} \lambda^{jk} \gamma^{jk}). \qquad (4.4.24)
\end{aligned}
$$

Using proposition 4.2.1 we see that

$$
\begin{aligned}
\gamma^f(t,T,Y_{t-},Y_t) &= -\frac{\gamma_T^p(t,T,Y_{t-},Y_t)}{1+\gamma^p(t,T,Y_{t-},Y_t)} \\
&= \kappa^\gamma.
\end{aligned}
$$

The forward dynamics in (4.4.13) can therefore be written

$$
\begin{aligned}
df(t,T) &= \left[\frac{1}{\kappa^s} exp\{-\kappa^s(T-t)\}[1 - exp\{-\kappa^s(T-t)\}](\sum_{b=1}^{r} s_{b,j}^2) \right. \\
&+ \left. \kappa^\gamma exp\{-\kappa^\gamma(T-t)\}(\sum_{k=1}^{n} \lambda^{jk}\gamma^{jk}) \right] dt \\
&+ exp\{-\kappa^s(T-t)\} \sum_{b=1}^{r} s_{b,j} d\tilde{W}_t^b \\
&+ \kappa^\gamma d\tilde{N}_t^{jk}.
\end{aligned}
$$

The short rate is therefore given by

$$
\begin{aligned}
r_t &= f(0,t) + \int_{u=0}^{t} \left[\frac{1}{\kappa^s} exp\{-\kappa^s(t-u)\}[1 - exp\{-\kappa^s(t-u)\}](\sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j}^2) \right. \\
&+ \left. \kappa^\gamma exp\{-\kappa^\gamma(t-u)\}(\sum_{j=1}^{n}\sum_{k=1}^{n} I_u^j \lambda^{jk}\gamma^{jk}) \right] du \\
&+ \int_{u=0}^{t} exp\{-\kappa^s(t-u)\} \sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j} d\tilde{W}_u^b \\
&+ \kappa^\gamma N_t.
\end{aligned}
$$

Denoting

$$m(t) = f(0,t) + \int_{u=0}^{t} [\frac{1}{\kappa^s} exp\{-\kappa^s(t-u)\}[1 - exp\{-\kappa^s(t-u)\}](\sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j}^2)$$

$$+ \quad \kappa^\gamma exp\{-\kappa^\gamma(t-u)\}(\sum_{j=1}^{n}\sum_{k=1}^{n} I_u^j \lambda^{jk}\gamma^{jk})]du,$$

we get

$$r_t = m(t) + \int_{u=0}^{t} exp\{-\kappa^s(t-u)\} \sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j} d\tilde{W}_u^b$$

$$+ \quad \kappa^\gamma N_t,$$

which finally gives us

$$dr_t = [\bar{m}(t) - \kappa^s r_t]dt + \sum_{b=1}^{r} s_{b,j} d\tilde{W}_t^b + \kappa^\gamma d\tilde{N}_t, \qquad (4.4.25)$$

where

$$\bar{m}_t = \frac{\partial}{\partial t} m_t + \kappa^s m_t + \kappa^s \kappa^\gamma N_t.$$

We can see that (4.4.25) is an extended version of the Hull-White (extended Vasicek) model given in equation (4.3.17). We can therefore use the methods described in section 4.3.2 to price zero-coupon bonds where the short rate is described by this model.

□

Example 4.4.6. In this example we will set the volatility to be dependent only on the state on the Markov chain, but will allow the jump sizes to be exponentially damped. We have the following parameter values:

$$s_b(t,T,j) = s_{b,j}, \tag{4.4.26}$$

$$\gamma^p(t,T,j,k) = exp[-\gamma^{jk}(T-t)] - 1. \tag{4.4.27}$$

The drift in state j given by (4.4.14) becomes

$$a(j,t,T) = (T-t)\sum_{b=1}^{r} s_{b,j}^2 + \sum_{k=1}^{n} \lambda^{jk}\gamma^{jk} exp[-\gamma^{jk}(T-t)]. \tag{4.4.28}$$

Using proposition 4.2.1, we see that

$$\begin{aligned} \gamma^f(t,T,Y_{t-},Y_t) &= -\frac{\gamma_T^p(t,T,Y_{t-},Y_t)}{1+\gamma^p(t,T,Y_{t-},Y_t)} \\ &= \gamma^{jk}. \end{aligned}$$

The forward dynamics in (4.4.13) can therefore be written as

$$\begin{aligned} df(t,T) &= \{(T-t)\sum_{b=1}^{r} s_{b,j}^2 + \sum_{k=1}^{n} \lambda^{jk}\gamma^{jk} exp[-\gamma^{jk}(T-t)]\}dt \\ &\quad + \sum_{b=1}^{r} s_{b,j} d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma^{jk} d\tilde{N}_t^{jk}. \end{aligned}$$

The short rate is therefore given by

$$r_t = f(0,t) + \int_{u=0}^{t} \{(t-u)\sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j}^2 + \sum_{k=1}^{n} \lambda^{jk}\gamma^{jk} exp[-\gamma^{jk}(t-u)]\} du$$

$$+ \int_{u=0}^{t} \sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j} d\tilde{W}_u^b + \int_{u=0}^{t} \sum_{k=1}^{n} \gamma^{jk} d\tilde{N}_t^{jk}.$$

Setting

$$m(t) = f(0,t) + \int_{u=0}^{t} \{(t-u)\sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j}^2 + \sum_{k=1}^{n} \lambda^{jk}\gamma^{jk} exp[-\gamma^{jk}(t-u)]\} du$$

we get that

$$r_t = m(t) + \int_{u=0}^{t} \sum_{j=1}^{n}\sum_{b=1}^{r} I_u^j s_{b,j} d\tilde{W}_u^b + \int_{u=0}^{t} \sum_{k=1}^{n} \gamma^{jk} d\tilde{N}_t^{jk},$$

which finally gives us

$$dr_t = m'(t)dt + \sum_{b=1}^{r} s_{b,j} d\tilde{W}_t^b + \sum_{k=1}^{n} \gamma^{jk} d\tilde{N}_t^{jk}. \tag{4.4.29}$$

We can see that (4.4.29) is an extended version of the Ho-Lee model given in equation (4.3.9), and bond prices can be priced using the methods described in section 4.3.2.

□

Once we have obtained the price of a zero-coupon bond, we may then wish to price call options on this bond such that at time T the payout is given by

$$X_T = \max[p(T,T^*) - K, 0].$$

The time 0 price of this derivative will be given by

$$C_0 = E[p(0,T)X_T|\mathcal{F}_0],$$

where the expectation is under the T-forward martingale measure $\mathbf{Q}^T$. We can then write

$$C_0 = p(0,T)E[p(T,T^*)1_A] - Kp(0,T)\mathbf{Q}^T(A), \tag{4.4.30}$$

where 1_A is the indicator variable for event A where $A = \{\omega \ : p(T,T^*) > K\}$. We have the following proposition (see Bingham and Kiesel [2004], Geman et al. [1995]):

Proposition 4.4.7. *Let the value of a contingent claim X_T discounted by the numéraire $Z(t)$ be a $\mathbf{Q}^Z$-martingale. Suppose we have another numéraire $Y(t)$ such that $Y(t)/Z(t)$ is also a $\mathbf{Q}^Z$-martingale. There then exists a probability measure $\mathbf{Q}^Y$ defined by the Radon-Nikodým derivative*

$$\eta_t = \frac{d\mathbf{Q}^Y}{d\mathbf{Q}^Z} = \frac{Y(t)}{Y(0)X(0)}$$

such that

$$Z(t)E_{\mathbf{Q}^Z}\left[\frac{X}{Z(T)}|\mathcal{F}_t\right] = Y(t)E_{\mathbf{Q}^Y}\left[\frac{X}{Y(T)}|\mathcal{F}_t\right]. \tag{4.4.31}$$

□

Using proposition 4.4.7 we can now write the following very useful result:

Corollary 4.4.8. *The time 0 price of a call option on the zero-coupon bond $p(T,T^*)$ is given by*

$$C_0 = p(0,T^*)\mathbf{Q}^{T^*}(A) - Kp(0,T)\mathbf{Q}^T(A), \tag{4.4.32}$$

where $\mathbf{Q}^T$ and $\mathbf{Q}^{T^*}$ are the T and T^* forward martingale measures respectively, and $A = \{\omega\ : p(T,T^*) > K\}$.

Proof. Using proposition 4.4.7, where we use the T and T^* bonds as numéraires, we have that

$$p(0,T)E_{\mathbf{Q}^T}\left[\frac{p(T,T^*)}{p(T,T)}1_A\right] = p(0,T^*)E_{\mathbf{Q}^{T^*}}\left[\frac{p(T,T^*)}{p(T,T^*)}1_A\right] = p(0,T^*)\mathbf{Q}^{T^*}(A),$$

where $p(T,T) = p(T^*,T^*) = 1$, and which on replacing into (4.4.30) gives us (4.4.32).

□

We are now left with the task of calculating $\mathbf{Q}^T(A)$ and $\mathbf{Q}^{T^*}(A)$. As in (4.4.17) we have

$$Z^*(t,T) = \frac{p(t,T)}{p(t,T^*)}, \tag{4.4.33}$$

which was shown in corollary 4.4.4 to have $\mathbf{Q}^{T^*}$-dynamics

$$\begin{aligned} dZ^*(t,T) &= Z^*(t,T)\left[\sum_{b=1}^{r}\{S_b(t,T,Y_{t-}) - S_b(t,T^*,Y_{t-})\}d\tilde{W}_t^b\right. \\ &\left. +\sum_{k=1}^{n}(exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(t,T^*,j,k)\})d\tilde{N}_t^{jk}\right]. \end{aligned} \tag{4.4.34}$$

We then have that

$$\begin{aligned} Z^*(T,T) &= Z^*(0,T)exp[-\frac{1}{2}\int_{u=0}^{T}\sum_{b=1}^{r}\{S_b(u,T,Y_{u-}) - S_b(u,T^*,Y_{u-})\}^2 dt \\ &+\int_{u=0}^{T}\sum_{b=1}^{r}\{S_b(u,T,Y_{u-}) - S_b(u,T^*,Y_{u-})\}d\tilde{W}_u^b \\ &+\int_{u=0}^{T}\sum_{j=1}^{n}\sum_{k=1}^{n} log(exp\{\Gamma^f(t,T,j,k)\} - exp\{\Gamma^f(u,T^*,j,k)\} + 1)d\tilde{N}_u^{jk}]. \end{aligned} \tag{4.4.35}$$

Using the following manipulation:

$$\mathbf{Q}^{T^*}(p(T,T^*) \geq K) = \mathbf{Q}^{T^*}(\frac{p(T,T)}{p(T,T^*)} \leq \frac{1}{K}) = \mathbf{Q}^{T^*}(Z^*(T,T^*) \leq \frac{1}{K}), \tag{4.4.36}$$

together with (4.4.35), we are now able to calculate $\mathbf{Q}^{T^*}(A)$ using the process $Z^*(T,T^*)$. We shall see an example later where this result will be useful.

Similarly we can write

$$\hat{Z}(t,T) = \frac{p(t,T^*)}{p(t,T)}, \tag{4.4.37}$$

which has $\mathbf{Q}^T$-dynamics

$$\begin{aligned} d\hat{Z}(t,T) &= \hat{Z}(t,T^*)\left[\sum_{b=1}^{r}\{S_b(t,T^*,Y_{t-}) - S_b(t,T,Y_{t-})\}d\tilde{W}_t^b \right. \\ &\left. + \sum_{k=1}^{n}(exp\{\Gamma^f(t,T^*,j,k)\} - exp\{\Gamma^f(t,T,j,k)\})d\tilde{N}_t^{jk}\right], \end{aligned} \tag{4.4.38}$$

so that

$$\begin{aligned} \hat{Z}(T,T) &= \hat{Z}(0,T^*)exp\left[-\frac{1}{2}\int_{u=0}^{T}\sum_{b=1}^{r}\{S_b(u,T^*,Y_{u-}) - S_b(u,T,Y_{u-})\}^2 dt \right. \\ &+ \int_{u=0}^{T}\sum_{b=1}^{r}\{S_b(u,T^*,Y_{u-}) - S_b(u,T,Y_{u-})\}d\tilde{W}_u^b \\ &\left. + \int_{u=0}^{T}\sum_{j=1}^{n}\sum_{k=1}^{n} log(exp\{\Gamma^f(t,T^*,j,k)\} - exp\{\Gamma^f(u,T,j,k)\} + 1)d\tilde{N}_u^{jk}\right]. \end{aligned} \tag{4.4.39}$$

We can also write

$$\mathbf{Q}^T(p(T,T^*) \geq K) = \mathbf{Q}^T(\frac{p(T,T^*)}{p(T,T)} \geq K) = \mathbf{Q}^T(\hat{Z}(T,T) \geq K), \quad (4.4.40)$$

and so we can once again use (4.4.39) together with (4.4.40) to calculate $\mathbf{Q}^T(A)$, only this time using the process $\hat{Z}(T,T)$.

Whether or not such calculations are easy to compute will depend on our choice of parameters. We shall now consider a third example where we will use the formula in corollary 4.4.8 to price call options on the bonds.

Example 4.4.9. Suppose this time we have parameter values

$$\begin{aligned} s_b(t,T,j) &= s_{b,j}, && (4.4.41)\\ \gamma^p(t,T,j,k) &= \gamma^{jk}(T-t). && (4.4.42) \end{aligned}$$

We can see from (4.4.7) that

$$\begin{aligned} S_b(t,T,j) &= -\int_{u=t}^{T} s_b(t,u,j)du \\ &= -s_{b,j}(T-t), && (4.4.43) \end{aligned}$$

and similarly form (4.4.5) that

$$\begin{aligned} exp\{\Gamma^f(t,T^*,j,k)\} &= \gamma^p(t,T,j,k)+1 \\ &= \gamma^{jk}(T-t)+1. && (4.4.44) \end{aligned}$$

Once again we have

$$\hat{Z}(t,T^*) = \frac{p(t,T^*)}{p(t,T)},$$

and replacing (4.4.43) and (4.4.44) into (4.4.38), we get that under the measure $\mathbf{Q}^T$, the dynamics of $\hat{Z}(T,T^*)$ becomes

$$d\hat{Z}(t,T^*) = \hat{Z}(t,T)[\sum_{j=1}^{n}\sum_{b=1}^{r} I_t^j s_{b,j}(T-T^*)d\tilde{W}_t^b + \sum_{j=1}^{n}\sum_{k=1}^{n}\gamma^{jk}(T^*-T)d\tilde{N}^{jk}].$$

Similarly we have

$$Z^*(t,T^*) = \frac{p(t,T)}{p(t,T^*)},$$

where from equation (4.4.34) we find that $Z^*(t,T^*)$ has dynamics

$$d\hat{Z}(t,T^*) = \hat{Z}(t,T)[\sum_{j=1}^{n}\sum_{b=1}^{r} I_t^j s_{b,j}(T^*-T)d\tilde{W}_t^b + \sum_{j=1}^{n}\sum_{k=1}^{n}\gamma^{jk}(T-T^*)d\tilde{N}^{jk}].$$

In appendix C we develop a methodology to calculate the probabilities $\mathbf{Q}^T(p(T,T^*) \geq K)$ and $\mathbf{Q}^{T^*}(p(T,T^*) \geq K)$, where the result is given in corollary C.0.12. We can then price a call option on the zero-coupon bond $p(T,T^*)$ using equation (4.4.32).

□

4.4.4 Market Completeness

We will now once again explore the question of market completeness. As discussed in section 4.3.3, the bond market theoretically contains an infinite number of bonds in which we can trade. However, unlike in the case of the short-rate models, HJM models have an infinite number of parameters, and so potentially there can be a complete market situation. In practice however we have a finite number of bonds, and so we will try to discover the conditions necessary for such a market to be complete, in the sense of there being a risk-neutral measure and lack of arbitrage when trading in only a given set of assets. To do this, we shall expand on the methodology set out in Bingham and Kiesel [2004].

For the market to be complete in this sense, we require there to be unique solutions for θ_b for $b = 1, \ldots, r$ and for ψ^{jk} for $k = 1, \ldots, n$ in equation (4.4.12). Let us assume there are m bonds in the market with maturities $t < T^1 < T^2 < \ldots < T^m$. We therefore have the following series of equations, where we use the subscript i where $i = 1, \ldots, m$ to denote the bond number:

$$A_i + \frac{1}{2}||S_i||^2 + \sum_{b=1}^{r} S_{i,b}\theta_b + \sum_{k=1}^{n} \lambda^{jk}\psi^{jk}(exp\{\Gamma^f(t,T,j,k)\} - 1) = 0. \quad (4.4.45)$$

Define

$$
\begin{aligned}
\mathbf{\Theta}_j^{r\times 1} &= \{\theta_b\}_{b=1\ldots r},\\
\mathbf{\Lambda}_j^{n\times 1} &= \{\lambda^{jk}\}_{k=1\ldots n},\\
\mathbf{\Psi}_j^{n\times 1} &= \{\psi^{jk}\}_{k=1\ldots n},\\
\mathbf{S}_j^{m\times r} &= \{S_{i,b}\}_{i=1\ldots m\ b=1\ldots r},\\
\mathbf{G}_j^{m\times n} &= \{exp\{\Gamma^f(t,T_i,j,k)\}-1\}_{i=1\ldots m\ k=1\ldots n},\\
\mathbf{M}_j^{m\times 1} &= \{A_i+\frac{1}{2}||B_i||^2\}_{i=1\ldots m}.
\end{aligned}
$$

Equation (4.4.45) can now be written in matrix form as

$$\mathbf{S}_j\mathbf{\Theta}_j+\mathbf{G}_j\mathbf{D}(\mathbf{\Lambda}_j)\mathbf{\Psi}_j=-\mathbf{M}_j, \tag{4.4.46}$$

where once again $\mathbf{D}(\mathbf{a})$ denotes the diagonal matrix with the vector $\mathbf{a}$ down the principle diagonal. Let us define the $m\times(r+n)$ augmented matrix as

$$\underset{m\times(r+n)}{\mathbf{U}_j} = \left[\ \underset{m\times r}{\mathbf{S}_j}\ \ \vdots\ \underset{m\times n}{\mathbf{G}_j\mathbf{D}(\mathbf{\Lambda}_j)}\ \right]. \tag{4.4.47}$$

Similarly define the $(r+n)\times 1$ augmented column vector $\mathbf{V}_j$ as

$$\underset{(r+n)\times 1}{\mathbf{V}_j} = \underset{(r+n)\times 1}{\begin{bmatrix}\mathbf{\Theta}_\mathrm{j}\\ \cdots\\ \mathbf{\Psi}_j\end{bmatrix}}.$$

We can now re-write equation (4.4.46) as

$$\mathbf{U}_j \mathbf{V}_j = -\mathbf{M}^j.$$

In order for there to be unique processes $\mathbf{\Theta}_j$ and $\mathbf{\Psi}_j$ and hence a unique $\mathbf{V}_j$, we therefore require that for all j the matrix $\mathbf{U}_j$ be invertible. This therefore necessitates that

$$m = r + n, \tag{4.4.48}$$

$$m = \text{Rank}(\mathbf{U}_j). \tag{4.4.49}$$

We are then left with

$$\mathbf{V}_j = -\mathbf{U}_j^{-1} \mathbf{M}_j. \tag{4.4.50}$$

4.4.5 Replicating Portfolios

We have so far discovered the conditions necessary for the existence of a unique risk-neutral measure when there are a given number of assets in the market. We shall now consider the extent to which contingent claims can be replicated using these assets.

Let us try to replicate a contingent claim X_T where $X_T \in \mathcal{F}_T$ is bounded. We assume that we are able to invest in m bonds with maturities $t < T^1 \leq T^2 < \ldots < T^m$ as well as the risk-free bank-account process. We make the restriction that $T \leq T^1$, so that throughout the time until the claim is effected we can trade in all m assets. At time t we hold a portfolio $\Phi_t \in \mathbf{R}^{m+1}$, where

$$\Phi_t = (\eta_t, \phi_t^1, \dots, \phi_t^m),$$

where η_t represents the holdings of the risk-free bank account whilst ϕ_t^i is the holding of the bond with maturity T^i. η_t and ϕ_t^i are predictable and bounded for $t \leq T$ and for all i. The value of this portfolio at time t is given by

$$V_t^\Phi = \eta_t B_t + \sum_{i=1}^{m} \phi_t^i p(t, T^i),$$

or alternatively the discounted value process

$$\tilde{V}_t^\Phi = \eta_t + \sum_{i=1}^{m} \phi_t^i \tilde{p}(t, T^i) \tag{4.4.51}$$

where

$$\tilde{p}(t, T^i) = B_t^{-1} p(t_T^i).$$

The portfolio Φ_t is said to be self-financing if

$$d\tilde{V}_t^\Phi = \sum_{i=1}^{m} \phi_t^i d\tilde{p}(t, T^i). \tag{4.4.52}$$

We can re-write (4.4.16) as

$$d\tilde{p}(t-, T^i) = \tilde{p}(t, T^i)[\sum_{i=1}^{d} S_i(t, T, Y_{t-}) d\tilde{W}_t^i + \sum_{k=1}^{n} exp\{\Gamma^f(t, T, j, k)\} - 1) d\tilde{N}_t^{jk}]. \tag{4.4.53}$$

So in order to be able to replicate X_T we need to find solutions for all the ϕ_t^i's such that

$$\begin{aligned}
\tilde{V}_t &= \tilde{V}_0 + \int_{s=0}^{t} \sum_{i=1}^{m} \phi_t^i \tilde{p}(t, T^i) \sum_{i=1}^{d} S_i(t, T, Y_{t-}) d\tilde{W}_t^i \\
&+ \int_{s=0}^{t} \sum_{i=1}^{m} \phi_t^i \tilde{p}(t, T^i) \sum_{k=1}^{n} exp\{\Gamma^f(t, T, j, k)\} - 1) d\tilde{N}_t^{jk}], \\
\tilde{V}_T &= \tilde{X}_T,
\end{aligned}$$

where

$$\tilde{X}_T = B_T^{-1} X_T.$$

Under the risk-neutral measure $\mathbf{Q}$ we have the following martingale:

$$M_t = E[\tilde{X}_t | \mathcal{F}_t], \tag{4.4.54}$$

which can be re-expressed using the martingale representation theorem as

$$M_t = M_0 + \int_{s=0}^{t} \sum_{b=1}^{r} v_s^b d\tilde{W}_t + \int_{s=0}^{t} \sum_{k=1}^{n} g_s^{jk} d\tilde{N}_t^{jk}, \tag{4.4.55}$$

where the v's and g's are predictable. Setting

$$V_0 = M_0$$

and

$$\begin{aligned}
v_s^b &= \phi_s^i \tilde{p}(s, T^i) S_b(s, T, Y_{s-}) \ b = 1, \ldots, r, \\
g_s^{jk} &= \phi_s^i \tilde{p}(s, T^i) exp\{\Gamma^f(s, T, j, k)\} - 1) \ k = 1, \ldots, n,
\end{aligned}$$

we therefore have a system of (r+n) equations. If we have that $m = r + n$ then a solution exists, and moreover if equation (4.4.49) holds, then it is unique.

4.5 Credit Derivatives

So far we have only considered bonds where the payoffs are guaranteed. We will now venture into the world of credit derivatives where the payoffs are subject to *default risk*, and whose price will therefore be determined by the credit-worthiness of the institution providing the payoff. We will look at two such instruments: *corporate bonds* and *credit default swaps.*

4.5.1 Corporate Bonds

Corporate bonds are bonds issued by corporations much like the government bonds discussed in earlier sections. The main difference is that buying a corporate bond, one runs the risk of the corporation defaulting on the bond or any of the coupon payments. We will therefore need to extend the model derived in the above sections to encapsulate this risk.

Working within the HJM framework of section 4.4, we can model the behaviour of the corporate bond prices as follows:

$$\frac{dp^c(t,T)}{p^c(t,T)} = \begin{cases} m^c(t,T,Y_{t-})dt + \sum_{b=1}^{r} v_i^c(t,T,Y_{t-})dW_t^i + \gamma^c(t,T,Y_{t-},Y_t)dN_t, & t < \tau, \\ 1 - R(\tau,T,Y_{\tau-}), & t = \tau, \\ 0, & t > \tau, \end{cases}$$

for all $T < T^*$, where the W_T^i's are Brownian motions under the objective probability measure $\mathbf{P}$. We have introduced a stopping time τ which is the time when the bond defaults ($\tau = \infty$ if no default occurs), at which time the bond loses a proportion of its value given by $R(\tau, T, Y_{\tau-})$, and henceforth remains at this value. Later on we will express this model as being similar to the government bonds in the previous sections, only with an additional state in the market corresponding to default. For now though we will continue with the above formulation, which will enable us to develop necessary conditions for the existence of a risk-neutral measure, when these corporate assets are introduced into our already complete bond market.

Assuming default has not occurred, we can express the dynamics of the corporate bond price by

$$\begin{aligned} dp^c(t,T) &= p^c(t,T)\{m^c(t,T,Y_{t-})dt + \sum_{b=1}^{r} v_i^c(t,T,Y_{t-})dW_t^i \\ &\quad + \gamma^c(t,T,Y_{t-},Y_t)dN_t - [1 - R(t,T,Y_{t-})]dU_t\}, \qquad (4.5.1) \end{aligned}$$

where U_t is the default indicator variable

$$U_t = \begin{cases} 0 & \text{no default,} \\ 1 & \text{default.} \end{cases}$$

Let the probability of default in the interval $[t, t+dt]$ whilst in state j be given by

$$P[dU_t = 1 | Y_t = j] = d^j dt.$$

We can therefore write

$$d\tilde{U}_t = dU_t - d^j dt,$$

so that $\tilde{U}_t$ is a martingale.

Let us assume we have a probability measure $\mathbf{Q}$ under which the bond market in section 4.4 is a martingale, as described in theorem 4.4.1. Define

$$Z^c(t,T) = \frac{p^c(t,T)}{B_t},$$

which has dynamics whilst in state j (assuming the bond has not defaulted)

$$\begin{aligned} dZ^c &= dB_t^{-1} p^c(t,T) + B_t^{-1} dp^c(t,T) \\ &= Z^c \left[\left\{ -r_t + m^c(t,T,j) + \sum_{k=1}^{n} \lambda^{jk} \gamma^c(t,T,j,k) - [1 - R(t,T,Y_{t-})] d^j \right\} dt \right. \\ &\quad \left. + \sum_{b=1}^{r} v_i^c(t,T,j) d\tilde{W}_t^i + \sum_{k=1}^{n} \gamma^c(t,T,j,k) d\tilde{N}_t^{jk} - [1 - R(t,T,Y_{t-})] d\tilde{U}_t \right], \end{aligned}$$

(4.5.2)

where $\tilde{N}_t$, $\tilde{W}_t$ and $\tilde{U}_t$ are **Q** martingales. Let us define the excess return of the corporate bonds over the government bonds by

$$\bar{m}(t,T,Y_{t-}) = m^c(t,T,Y_{t-}) - m(t,T,Y_{t-}),$$

and similarly for the jumps

$$\bar{\gamma}^p(t,T,j,k) = \gamma^c(t,T,j,k) - \gamma^p(t,T,j,k).$$

Equation (4.5.2) can now be written as

$$\begin{aligned} dZ^c &= Z^c[\{-r_t + m(t,T,j) + \bar{m}(t,T,Y_{t-}) \\ &\quad + \sum_{k=1}^{n} \lambda^{jk}[\gamma^p(t,T,j,k) + \bar{\gamma}^p(t,T,j,k)] - [1 - R(t,T,Y_{t-})]d^j\}dt \\ &\quad + \sum_{b=1}^{r} v_i^c(t,T,j)d\tilde{W}_t^i + \sum_{k=1}^{n} \gamma^c(t,T,j,k)d\tilde{N}_t^{jk} - (1 - R(t,T,Y_{t-})d\tilde{U}_t]. \end{aligned} \tag{4.5.3}$$

We can re-write (4.4.12) with the HJM conditions $\theta_b = 0$ $\forall b$ and $\psi^{jk} = 1$ $\forall\, j,k$ to give us

$$m(t,T,j) - r_t + \sum_{k=1}^{n} \gamma^p(t,T,j,k)d\tilde{N}_t^{jk} = 0,$$

so that for (4.5.3) to be a martingale we require

$$\bar{m}(t,T,Y_{t-}) + \sum_{k=1}^{n} \lambda^{jk}\bar{\gamma}^{p}(t,T,j,k) - (1 - R(t,T,Y_{t-}))d^{j} = 0. \qquad (4.5.4)$$

So given the excess returns and jumps of the corporate bonds over the government bonds, we can use (4.5.4) to calculate the implied default probabilities of the bond. Alternatively, the default probabilities may be exogenous in the model, and (4.5.4) can be used to derive the excess return required by the corporate bond over the government bond.

To price the corporate bonds and their derivatives we shall express the bond dynamics in a slightly different way. Our model is currently represented by an n-state Markov chain which has transition intensities λ^{jk} when transiting from state j to state k. Let us add another state which we will enumerate as state $n+1$ and which will correspond to default. This state will clearly be absorbing as once defaulted the bond can no longer be paid, so we have that

$$\lambda^{(n+1)j} = 0 \quad \forall\, j.$$

We also have that

$$\gamma^{p}(t,T,j,n+1) = -(1 - R(t,T,j)).$$

Given this formulation, we can now readily use the methodologies developed in section 4.4 to price derivatives of these bonds.

4.5.2 Credit Default Swaps

Credit default swaps (CDS) are used by companies or investors to hedge their credit exposure to other companies. The investor will pay a certain fee say annually, and in return will receive a payout should the company he has insured against default on their payments. Let us illustrate this with an example with a CDS which specifies *physical settlement.* For other types of CDS's see Hull [2003].

Suppose Barclays is exposed to £100 million worth of exposure to the Ford Motor Company, and wished to completely hedge this risk. They can therefore buy a CDS on Ford from a CDS seller, say Lloyd's, for which they will pay x basis points (hundredths of a percent) of the 100 million *notional principal* of the bond per year. Should Ford default on their payments, Barclays will be reimbursed for £100 million by Lloyd's. This annual premium is paid at the end of the year, and should Ford default during the year then Barclays would have to pay the proportion of the premium accrued up until that time, e.g. should default occur after 1 month they would have to pay $\frac{x}{12}$ basis points. They would then also immediately get paid their £100 million, and Lloyd's would receive any recovery payments Ford paid out.

To value such swaps, i.e. the annual premium charged by the seller, we need to calculate the expected present value of the premiums paid by the buyer, as well as the expected present value of the payout by seller, and subsequently employ a risk-neutral argument to equate the two. Let us define:

N Notional amount of the CDS.

T Length of the CDS in years.

x_i Premium charged for period year $i-1$ to year i.

R Recovery amount.

p_t Probability of default not having occurred up to time t.

We will assume that the probabilities of default are those under the risk-neutral measure developed in the previous section. Also, as with the corporate bonds above, we will take the recovery rates to be a function of the time the default took place, the length of the CDS and the state the market was in at the time of default, so that $R = R(t, T, Y_{t-})$ where we use Y_{t-} to show that it is predictable. We again denote U_t as the default indicator variable given by

$$U_t = \begin{cases} 0 & \text{no default,} \\ 1 & \text{default.} \end{cases}$$

The probability of default in the interval $[t, t+dt]$ whilst in state j is given by

$$P[dU_t = 1 | Y_t = j] = d^j dt.$$

We can separate the expected present value of the premiums paid by the buyer, into the sum of end of year premiums paid where default hasn't occurred and the premium payment due on default. So for year i we have

$$E[\text{End of year premium}] = \hat{B}_i^{-1} p_i x_i$$

for $i = 1, \ldots, T$, where $\hat{B}_t$ is the expected value of the bank-account process at time t. We also have

$$E[\text{Default premium}] = \int_{u=0}^{1} \hat{B}_{i+u}^{-1} p_{(i+u)-}(1 - p_{(i+u)}) u x_i du,$$

for $i = 1, \ldots, T$. The expected present value of the payout by the seller of the CDS is given by

$$E[\text{Payout}] = \int_{u=0}^{1} \hat{B}_{i+u}^{-1} p_{(i+u)-}(1 - p_{(i+u)})[1 - \hat{R}(i+u, T, Y_{(i+u)-})] N du,$$

for $i = 1, \ldots, T$, where

$$\hat{R}(i+u, T, Y_{t-}) = E[R(i+u, T, Y_{t-})].$$

Assuming that probabilities of default are under a risk-neutral probability measure, we get for each year i the premium is given by

$$x_i = \frac{\int_{u=0}^{1} \hat{B}_{i+u}^{-1} p_{(i+u)-}(1 - p_{(i+u)})[1 - \hat{R}(i+u, T, Y_{(i+u)-})] N du}{\hat{B}_i^{-1} p_i + \int_{u=0}^{1} \hat{B}_{i+u}^{-1} p_{(i+u)-}(1 - p_{(i+u)}) u du}. \tag{4.5.5}$$

To evaluate (4.5.5) we firstly begin by conditioning on the jump sequence between times 0 and T. Suppose the model starts in state j_1, and up until time T there are x jumps, after each of which we are in state j_l for $l = 1, \ldots, x+1$. We can express this sequence by the set $\mathcal{J} = (j_1, j_1, \ldots, j_{x+1})$.

We therefore have the probability of being in state k at time t given by $p_t^{j_1k}$ whose value is given in equation (1.3.5). We can thus write

$$
\begin{aligned}
p_t &= p_{t-}(1 - \sum_{k=1}^{n} p_t^{j_1k} d^k dt) = \int_{u=0}^{t} \sum_{k=1}^{n} p_t^{j_1k} d^k dt \\
&= \sum_{k=1}^{n} P_t^{j_1k} d^k, \qquad (4.5.6)
\end{aligned}
$$

where P_t^{jk} is the total time spent in state k up to time t given that we are currently in state j, and is given by equation (1.3.6). For the bank-account process we have

$$B_t = E\left[exp\left\{\int_{u=0}^{t} \sum_{j=1}^{n} I_t^j r_j dt\right\}\right].$$

Specifying the recovery rates as

$$R(t, T, j) = R^j$$

for $j = 1, \ldots, n$ so that it is dependent on the state alone, we get

$$\hat{R}(t, T, j) = \sum_{j=1}^{n} p_t^j R^j. \qquad (4.5.7)$$

Inserting (4.5.6) and (4.5.7) into (4.5.5) we obtain the value of $x_i^{\mathcal{J}}$, where the superscript $\mathcal{J}$ is to denote that this is the premium conditional on the jump sequence $\mathcal{J}$. To obtain the the value of the premium we must then sum over all jump sequences to get

$$x_i = \sum_{x=0}^{\infty} \sum_{j_1=1}^{n} \cdots \sum_{j_x=1}^{n} x_i^{\mathcal{J}} p^{j_1 \cdots j_x}(T),$$

where $p^{j_1 \cdots j_x}(T)$ is the probability that we obtain jump sequence $(j_1 \dots j_x)$ in time T as is defined lemma 1.3.3.

Chapter 5

Numerical Methods for Interest-Rate Models

5.1 Introduction

We shall now develop methods for pricing interest-rate derivatives for the short-rate models and the HJM models discussed in chapter 4.

In section 4.3 we developed various short-rate models and derived formulas for pricing zero-coupon bonds, as well as partial differential equations that need to be solved to price other derivatives. In this chapter we shall develop numerical methods for pricing such derivatives. For the short rate we will consider both models developed in section 4.3.2. Firstly we have the Hull-White extended Vasicek model given by equation (4.3.17), where when in state j the short rate given by r_t has dynamics

$$dr_t = (\alpha_j - \beta_j r_t)dt + \sigma_j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \tag{5.1.1}$$

for $j = 1, \ldots, n$, where α_j, β_j and σ_j are all functions of the current state j. Secondly, the Ho-Lee model is given by equation (4.3.9), where when in state j the short rate has dynamics

$$dr_t = \alpha_j dt + \sigma_j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk}. \tag{5.1.2}$$

Derivatives of other short rate models can be priced using similar techniques as will be developed for these models.

In section 4.4 we developed an HJM model using our MCJD framework. It was shown how certain HJM models can be expressed as short-rate models, and so in these instances derivatives can be priced using the same methodologies as for the short-rate models. It was also shown in section 4.4 how we may price derivatives using techniques involving the forward-rate measure. We shall now develop numerical methods for pricing these derivatives. For much of the numerical work on HJM models we will follow Clewlow and Strickland [1998].

The model we shall be considering is that described in corollary 4.4.2, where under the risk-neutral measure $\mathbf{Q}$ the forward-rate dynamics are given by

$$df(t,T) = a(t,T,Y_{t-})dt + \sum_{b=1}^{r} s_b(t,T,Y_{t-})d\tilde{W}_t^b + \gamma^f(t,T,Y_{t-},Y_t)d\tilde{N}_t, \quad (5.1.3)$$

and where the necessary drift condition whilst in state j is

$$a(t,T,j) = \sum_{b=1}^{r} s_b(t,T,j)\int_{u=t}^{T} s_b(t,u,j)du - \sum_{k=1}^{n} \lambda^{jk}\gamma^f(t,T,j,k)exp\{\Gamma^f(t,T,j,k)\}. \quad (5.1.4)$$

The short-rate process is then given by

$$\begin{aligned} r(T) &= f(0,T) + \int_{u=0}^{T} a(u,T,Y_{u-})du + \int_{u=0}^{T}\sum_{b=1}^{r} s_b(u,T,Y_{u-})d\tilde{W}_u^b \\ &+ \int_{u=0}^{T} \gamma^f(u,T,Y_{u-},Y_u)d\tilde{N}_u, \end{aligned} \quad (5.1.5)$$

and the bond-price dynamics was shown to be

$$dp(t,T) = p(t,T)[r_t dt + \sum_{b=1}^{r} S_b(t,T,Y_{t-})d\tilde{W}_t^b + \sum_{k=1}^{n}\gamma^p(t,T,j,k)d\tilde{N}_t^{jk}], \quad (5.1.6)$$

where all the terms are defined in section 4.4.

As outlined in Clewlow and Strickland [1998], we may value at time t a call option with expiry T on the bond maturing at time T^*, where $T < T^*$, by taking discrete points on the forward-rate curve obtained from the market, and simulating from these their time T values. The bond price will then be given by

$$P(T,T^*) = \exp\left\{-\int_{u=T}^{T^*} f(T,u)du\right\}. \quad (5.1.7)$$

The drawback with this method is the large number of points that would be required. We shall therefore be working directly with pure discount bond dynamics given in (5.1.6). In this context, the price of a call option at time t with payout at time T on a bond maturing at time T^*, where the strike is K, is given by

$$C_t = E\left[exp\left\{-\int_{u=t}^{T} r(u)du\right\}\max[0, p(T,T^*) - K]\right], \tag{5.1.8}$$

where the expectation is under the risk-neutral measure $\mathbf{Q}$.

5.2 Trinomial Trees for Short-Rate Models

As was done in section 3.4 for the equity model, we may simulate the behaviour of short-rate models using trinomial trees. Suppose we are interested in derivatives of this short rate at a time T given we are currently at time 0 and in state j. We can divide the time interval $[0,T]$ into h subsections of length Δt where $\Delta t = \frac{T}{h}$. We will assume that the market can change state only at the time periods $\Delta t, 2\Delta t, \ldots, T - \Delta t$.

We will use the trinomial tree methodology to price derivatives where the short rate follows the Hull-White extended Vasicek model described in section 5.1. We shall begin by writing a discrete-time analogue to (5.1.1) so that when in state k we have

$$\Delta r_t = (\alpha_k - \beta_k r_t)\Delta t + \sigma_k \Delta W_t + \sum_{l=1}^{n} \gamma^{kl} \Delta N_t^{kl} \tag{5.2.1}$$

for $k = 1, \ldots, n$ and $l = 1, \ldots, n$, where I_t^k is the indicator variable with value 1 when in state k and 0 otherwise. Using the fact that the expected value of ΔN_t^{kl} is given by

$$E[\Delta N_t^{kl}] = \gamma^{kl}\lambda^{kl}\Delta t,$$

where λ^{kl} is the transition intensity for transiting from state k to state l, we therefore have that the value of the short rate at time $t + \Delta t$ has expected value

$$E[r_{t+\Delta t}] = r_t + \left(\alpha_k - \beta_k r_t + \sum_{l=1}^{n} \gamma^{kl}\lambda^{kl}\right)\Delta t.$$

The variance of $r_{t+\Delta t}$ can easily be calculated as the sum of the variance of the Brownian motion and the variance of the point process (as the two are independent), and is thus given by

$$Var[r_{t+\Delta t}] = \sigma_k^2\Delta t + \sum_{l=1}^{n}(\gamma^{kl})^2\lambda^{kl}\Delta t(1 - \lambda^{kl}\Delta t).$$

The unconditional mean is given by

$$E[r_{t+\Delta t}] = r_t + \sum_{k=1}^{n} p_t^{jk}\left(\alpha_k - \beta_k r_t + \sum_{l=1}^{n} \gamma^{kl}\lambda^{kl}\right)\Delta t, \tag{5.2.2}$$

where p_t^{jk} is the probability of being in state k at time t given in equation (1.3.5), and using lemma 3.4.2 the unconditional variance of $r_{t+\Delta t}$ is given by

$$
\begin{aligned}
Var[r_{t+\Delta t}] &= \sum_{k=1}^{n}\left[p_t^{jk}\left\{\sigma_k^2\Delta t+\sum_{l=1}^{n}(\gamma^{kl})^2\lambda^{kl}\Delta t(1-\lambda^{kl}\Delta t)\right\}\right.\\
&\left.+p_t^{jk}(1-p_t^{jk})\left\{r_t(1-\beta_k\Delta t)+\left(\alpha_k+\sum_{l=1}^{n}\gamma^{kl}\lambda^{kl}\right)\Delta t\right\}^2\right].
\end{aligned}
\tag{5.2.3}
$$

We shall now use the trinomial tree shown in figure 5.1 to simulate the short-rate dynamics given in (5.2.1). In order to do this we must calculate the moments of the trinomial tree model at each step and equate them to the discrete-time MCJD model. The expected value of the stock price at time $t+\Delta t$ under the trinomial tree model is equal to

$$
E[r_{t+\Delta t}] = r_t(q_t^1 u + q_t^2 + q_t^3 d), \tag{5.2.4}
$$

where q_t^1, q_t^2 and q_t^3 are the probabilities of an upward jump by a factor of u, no jump or a downward jump by a factor of d respectively. The variance of $S_{t+\Delta t}$ can similarly be shown to be

$$
\begin{aligned}
Var[r_{t+\Delta t}] &= r_t^2[(q_t^1(1-q_t^1)u^2+q_t^2(1-q_t^2)+q_t^3(1-q_t^3)d^2)\\
&-2(q_t^1q_t^2u+q_t^1q_t^3ud+q_t^2q_t^3d)].
\end{aligned}
\tag{5.2.5}
$$

By equating the means of the trinomial tree model with that of the discrete time MCJD model in equations (5.2.2) and (5.2.4) we get

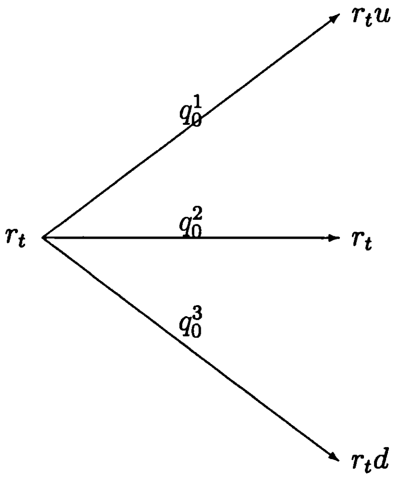

Figure 5.1: One period trinomial tree for the short rate.

$$r_t + \sum_{k=1}^{n} p_t^{jk} \left(\alpha_k - \beta_k r_t + \sum_{l=1}^{n} \gamma^{kl} \lambda^{kl} \right) \Delta t = r_t (q_t^1 u + q_t^2 + q_t^3 d).$$

A comparison of the variances in equations (5.2.3) and (5.2.5) gives us

$$\begin{aligned}
&\sum_{k=1}^{n} [p_t^{jk} \{\sigma_k^2 \Delta t + \sum_{l=1}^{n} (\gamma^{kl})^2 \lambda^{kl} \Delta t (1 - \lambda^{kl} \Delta t)\} + p_t^{jk} (1 - p_t^{jk}) \{r_t (1 - \beta_k \Delta t) \\
&+(\alpha_k + \textstyle\sum_{l=1}^{n} \gamma^{kl} \lambda^{kl}) \Delta t\}^2] = \; r_t^2 [(q_t^1 (1 - q_t^1) u^2 + q_t^2 (1 - q_t^2) + q_t^3 (1 - q_t^3) d^2) \\
&\qquad -2(q_t^1 q_t^2 u + q_t^1 q_t^3 u d + q_t^2 q_t^3 d)].
\end{aligned}$$

Using these two equations, and given that we also know that

$$q_t^1 + q_t^2 + q_t^3 = 1,$$

we are left with three equations to solve for five unknowns $u, d, q_t^1, q_t^2, q_t^3$. We must therefore specify two of these values in order to get a unique solution. Let us set

$$d = \frac{1}{u},$$

$$q_t^2 = \frac{2}{3},$$

which will allow us to solve for all the other unknowns. We are then left with the following two equations to solve for u and q_t^1:

$$q_t^1 = \frac{uE - r_t}{3r_t(u^2 - 1)},$$

and

$$q_t^1(1 - q_t^1)u^4 - \frac{4}{3}q_t^1 u^3 + (\frac{2}{9} - V)u^2 + \frac{1}{9}(12q_t^1 - 4)u + \frac{2}{9} - q_t^1(q_t^1 + \frac{1}{3}) = 0,$$

where

$$
\begin{aligned}
E &= 1 - 3\sum_{k=1}^{n} p_t^{jk}\left(\alpha_k - \beta_k r_t + \sum_{l=1}^{n}\gamma^{kl}\lambda^{kl}\right)\Delta t \\
V &= \sum_{k=1}^{n}\left[p_t^{jk}\left\{\sigma_k^2\Delta t + \sum_{l=1}^{n}(\gamma^{kl})^2\lambda^{kl}\Delta t(1-\lambda^{kl}\Delta t)\right\}\right. \\
&\quad \left. + p_t^{jk}(1-p_t^{jk})\left\{r_t(1-\beta_k\Delta t) + \left(\alpha_k + \sum_{l=1}^{n}\gamma^{kl}\lambda^{kl}\right)\Delta t\right\}^2\right].
\end{aligned}
$$

These two equations must be solved numerically for each time period $t = 0, \Delta t, \ldots, T - \Delta t$. For an h-period tree we will then be left with $2h+1$ final nodes at time t. More importantly, we now have 3^h paths the short rate can take, and for each such path we can calculate the value of a derivative of the short rate had the short rate followed that path.

After these derivative values have been calculated for each path, the overall value of the derivative is then given by the weighted average of these values, the weights being the probabilities of observing those paths. These probabilities are calculated by taking for each path the product of all the probabilities of the h one-period models.

The most common derivative of the short rate is the zero-coupon bond with value

$$p(0,T) = E[e^{-\int_{t=0}^{T} r_t dt}].$$

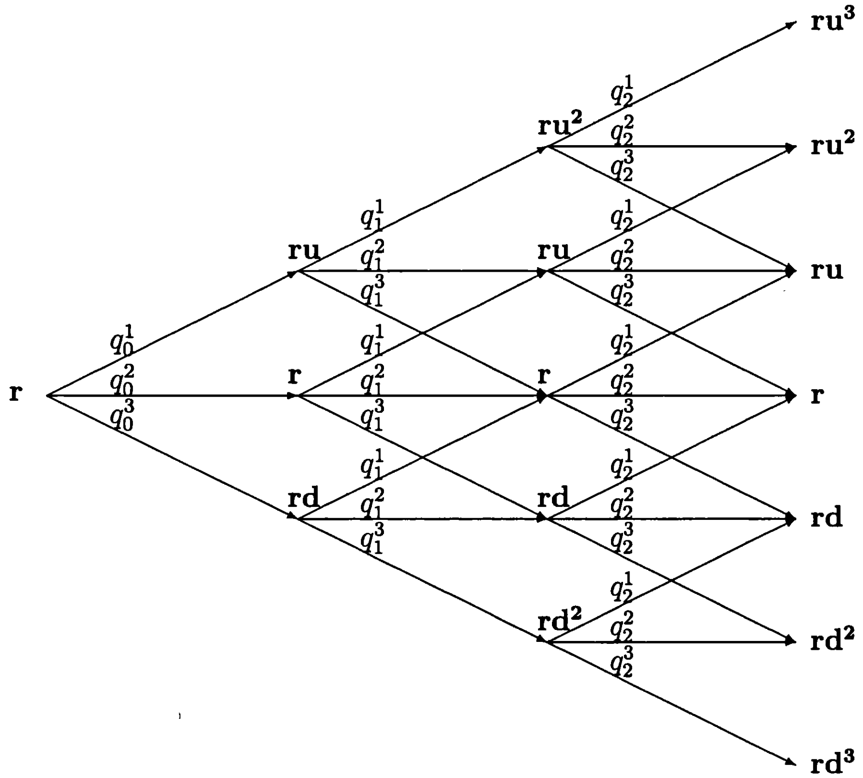

Figure 5.2: Three-period trinomial tree to simulate the behaviour of the short rate.

This can be estimated using our discrete time model by

$$p(0,T) = E[e^{-\sum_{i=0}^{h} r_{i\Delta t}\Delta t}],$$

which can be calculated using the tree in figure 5.2. For each path we calculate the value $e^{-\sum_{i=0}^{h} r_{i\Delta t}\Delta t}$, as well as the probability of observing that path given by the product of the probabilities for each node, and then aggregate to find the unconditional price of the derivative. So for example, the value of the zero-coupon bond for the model shown in figure 5.2 is given by

$$p(0,T) = q_0^1 q_1^1 q_2^1 e^{-(r+ru+ru^2)\Delta t} + \ldots + q_0^3 q_1^3 q_2^3 e^{-(r+rd+rd^2)\Delta t}.$$

Other derivatives can be priced using a similar method.

The advantages of this model are its fairly easy formulation and implementation. However, reducing an n dimensional process into just a three-branch tree will clearly sacrifice much of the original model's characteristics. The number of final nodes is exponential in the number of time periods h, and so it may not be practical to implement models with large values of h.

5.3 Multinomial Trees for Short-Rate Models

We will now setup a multinomial tree to replicate the behaviour of r_t as was done in section 3.4.2 for the equity model. For this we will consider the Ho-Lee model given in section 5.1. Assume we are currently at time 0, and that we are firstly interested in pricing the zero-coupon bond expiring at time T denoted by $p(0,T)$, and subsequently pricing derivatives of this bond. We can section the time interval $[0,T]$ into h subsections of length Δt where $\Delta t = \frac{T}{h}$. We will assume that market can change state only at the time periods $\Delta t, 2\Delta t, \ldots, T - \Delta t$.

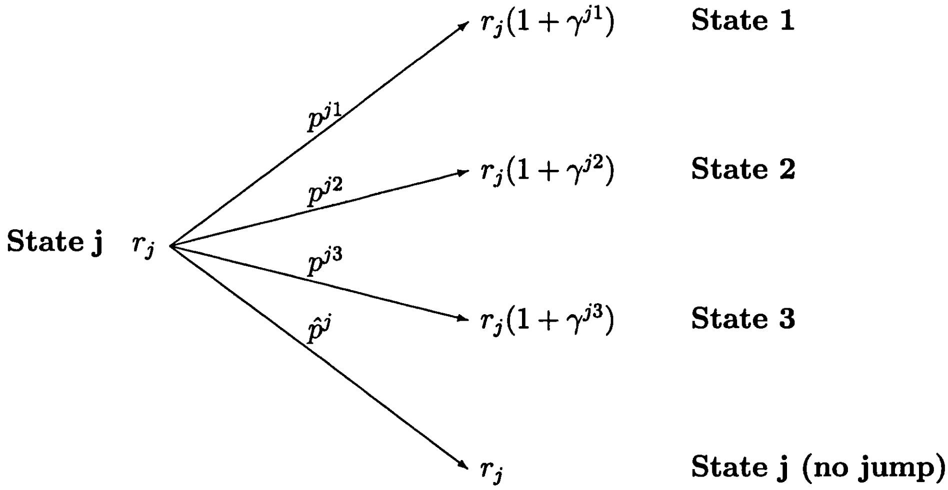

Figure 5.3: Jumps in the short-rate in time Δt in a three state model starting in state j.

Let us once again consider a market where there are three states so that $n = 3$. We can represent the jumps part of the model for a period Δt by the four-branched tree in figure 5.3, where one branch is for each state the model can transit to and 1 branch represents no transition. We also have the following probabilities $\forall j, k$:

$$
\begin{aligned}
p^{jk} &= \lambda^{jk}\Delta t, \\
\hat{p}^j &= 1 - \sum_{k=1}^{n} \lambda^{jk}\Delta t.
\end{aligned}
$$

We will now represent the short-rate movement excluding the jumps by a binomial-tree model as shown in figure 5.4. The probability of the stock

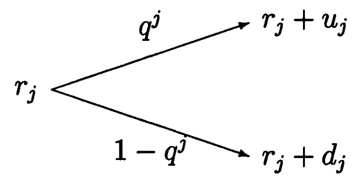

Figure 5.4: Stock-price movement (excluding jumps) in time Δt when in state j.

price increasing by a factor u is given by q^j, and similarly the probability of the stock changing by a factor of d is given by $(1 - q^j)$.

We want the variance implied by this binomial model to equal that implied from (5.1.2). We can see from (5.1.2) that $(r_{t+\Delta t} - r_t)$, when in state j and excluding the jumps, is normally distributed with expected value

$$E[r_{t+\Delta t}] = r_t + \alpha_j \Delta t.$$

The variance of the short rate over this period is given by

$$Var[r_{t+\Delta t}] = \sigma_j^2 \Delta t.$$

Looking at the expected value implied by the binomial model we have that

$$E[r_{t+\Delta t}] = r_t + q^j u_j + (1 - q^j) d_j,$$

and the variance is given by

$$Var[r_{t+\Delta t}] = q_j(1-q_j)(u_j-d_j)^2.$$

In order for our binomial model to have the same mean and variance as the MCJD model (excluding the jumps) it is replicating, we therefore require

$$r_t + \alpha_j \Delta t = r_t + q^j u_j + (1-q^j) d_j \tag{5.3.1}$$

$$\sigma_j^2 \Delta t = q_j(1-q_j)(u_j-d_j)^2. \tag{5.3.2}$$

We thus have two equations with which we need to solve for three parameters, and so we will need to specify one of the parameters. Firstly, we can set the probability

$$q^j = 0.5.$$

Equations (5.3.1) and (5.3.2) become

$$\alpha_j \Delta t = \frac{1}{2}(u_j + d_j),$$

$$\sigma_j^2 \Delta t = \frac{1}{4}(u_j - d_j)^2,$$

and this can be solved to give

$$u_j = \alpha_j \Delta t + \sigma_j \sqrt{\Delta t},$$

$$d_j = \alpha_j \Delta t - \sigma_j \sqrt{\Delta t}.$$

Unlike in section 3.4.2 there are no restrictions on the value of α_j or σ_j. Alternatively we can set

$$d_j = -u_j.$$

Equations (5.3.1) and (5.3.2) become

$$\begin{aligned} \alpha_j \Delta t &= u_j(2q_j - 1), \\ \sigma_j^2 \Delta t &= 4q_j(1 - q_j)u_j^2, \end{aligned}$$

and this can be solved to give

$$\begin{aligned} q_j &= \frac{1}{2}K, \\ u_j &= \frac{\alpha_j \Delta t}{K - 1}, \\ d_j &= \frac{\alpha_j \Delta t}{1 - K}, \end{aligned}$$

where

$$K = 1 - \sqrt{1 - \frac{\sigma_j^2 \Delta t}{\sigma_j^2 \Delta t + \alpha_j^2 \Delta t^2}}.$$

Again there are no restrictions on the value of α_j or σ_j. We can now replicate the behaviour of the MCJD model by combining the binomial model with the jumps model as in figure 5.5.

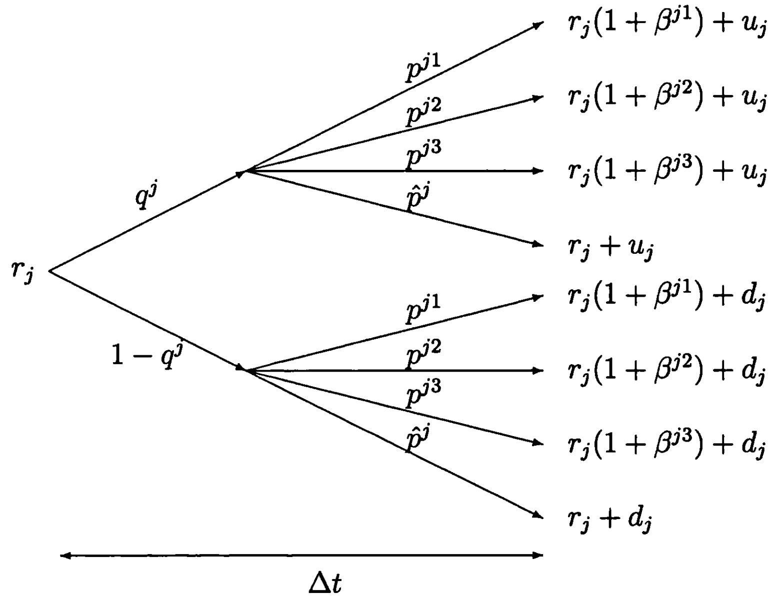

Figure 5.5: Short-rate movement in time Δt when in state j.

The value of a zero-coupon bond is given by

$$p(t,T) = E\left[exp\left\{-\int_{s=t}^{T} r(s)ds\right\}\right], \tag{5.3.3}$$

where the expectation is under the risk-neutral measure $\mathbf{Q}$. In order to approximate this we shall use a discrete-time version of equation (5.3.3), so that

$$p(t,T) = E\left[exp\left\{-\sum_{s=0}^{h-1} r_{t+s\Delta t}\Delta t\right\}\right]. \tag{5.3.4}$$

We can use figure 5.6 to estimate the value of the two-period-ahead zero-

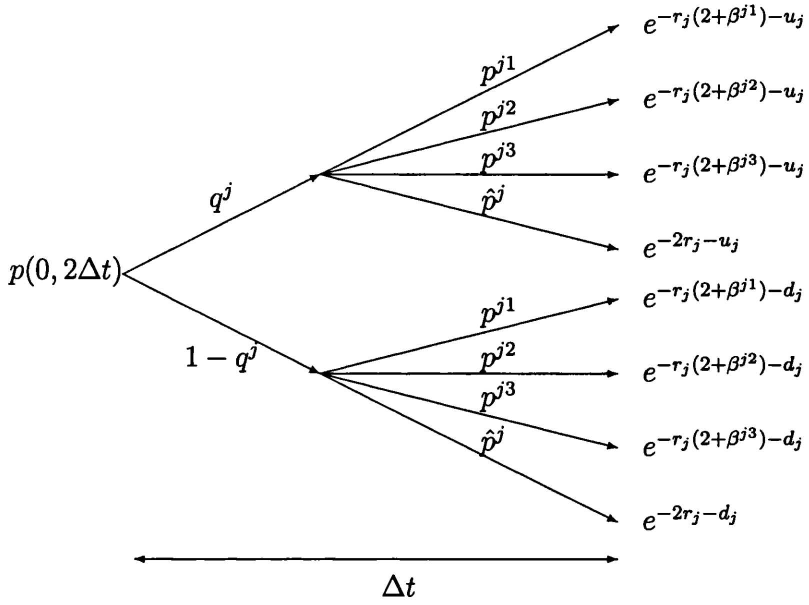

Figure 5.6: Valuing a two period zero-coupon bond.

coupon bond as follows. Firstly we note that the short rate for any time period is previsible, so that the time 0 value of the short rate r_j is effective until the time Δt, and the subsequent value determined at time Δt applies until the time $2\Delta t$. To value the bond we start by calculating for each node the value of the zero-coupon bond price should the short rate follow the path leading to that node, and we then take a weighted average of these values where the weights are given by the product of the probabilities leading to that node. We thus have

$$p(0,2\Delta t) = q_j p^{j1} e^{-(r_j(2+\beta^{j1})+u_j)\Delta t} + \ldots + (1-q_j)\hat{p}^{j1} e^{-(2r_j+d_j)\Delta t},$$

where we assume that the values q_j are under the risk-neutral measure $\mathbf{Q}$.

We can similarly use this tree to price other derivatives of the short rate or the zero-coupon bond, by replacing the values at the final nodes with the payoffs of the derivative. The value of the derivative would then once again be the weighted average of these payoffs.

So far we have calculated the expected value of the short rate and valued derivatives of it over two time periods. If we wish to extend this to a time T, i.e. over h time periods where $h = \frac{T}{\Delta t}$, we simply combine h of the above single-period trees, and then calculate the derivative values in the same way as in the single-period model. As with the equity model, the number of final nodes will be large even for moderate values of h. The number of nodes after a time T represented by N_T is equal to

$$N_T = [2(n+1)]^h.$$

This methodology's usefulness will therefore be restricted to models where there are small values of n and h (hence T).

5.4 Monte Carlo Simulation for Short-Rate Models

We shall now attempt to estimate the value of derivatives of the short rate using the Monte Carlo simulation technique. Unlike in the case of stocks, the value of the derivative is going to be determined by all the values of the short rate over the life of the derivative, and not just the final value. We will consider valuing zero-coupon bonds, although this method can easily be extended to cope with other derivatives. The value of a zero-coupon bond for time T at time 0 is given by

$$p(0,T) = E\left[exp\left\{-\int_{t=0}^{T} r_t dt\right\}\right], \tag{5.4.1}$$

where the expectation is under the risk-neutral measure $\mathbf{Q}$. There are many possible paths the short can take up until time T. We may generate randomly one of these paths and calculate the value of the zero-coupon bond under this path. This is known as a *simulation trial.* The simulation involves repeating this say N times and then aggregating the values. We shall now look at methods for simulating our MCJD model.

We shall once again divide the interval $[0, T]$ into h subsections of length Δt, so that

$$h = \frac{T}{\Delta t}.$$

Let us model the short rate using the Hull-White extended Vasicek model described in section 5.1, where when in state j the short rate given by r_t has dynamics

$$dr_t = (\alpha_j - \beta_j r_t)dt + \sigma_j dW_t + \sum_{k=1}^{n} \gamma^{jk} dN_t^{jk} \tag{5.4.2}$$

for $j = 1, \ldots, n$. We can write a discrete-time analogue to (5.4.2), so that when in state k we have

$$\Delta r_t = (\alpha_k - \beta_k r_t)\Delta t + \sigma_k \epsilon_t \sqrt{\Delta t} + \sum_{l=1}^{n} \gamma^{kl} \Delta N_t^{kl} \tag{5.4.3}$$

for $k = 1, \ldots, n$ and $l = 1, \ldots, n$. Also, ϵ_t is a sample from a standardised normal distribution, and ΔN_t^{jk} is a Bernoulli random variable with probability $\lambda^{jk}\Delta t$, although we impose the restriction that

$$\Delta N_t^{j1} + \Delta N_t^{j2} + \ldots + \Delta N_t^{jk} \leq 1$$

for all j, i.e. at most one jump per time period. In order to simulate the value of the zero-coupon bond, we must therefore generate each of the random variables $r_{\Delta t}, r_{2\Delta t}, \ldots, r_{h\Delta t}$ successively.

To generate these random samples we assume that we have an endless supply of uniform $[0, 1]$ random variables $U_1, U_2, \ldots$ which can be easily generated by all computers, as well as those from a standardised normal distribution. To generate sample values of ΔN_t^{jk} for $k = 1, \ldots, n$, we divide the interval $[0, 1]$ into $n + 1$ intervals $[0, I_1], [I_1, I_2], \ldots, [I_n, I_{n+1}]$, where we have that

$$I_m = \begin{cases} \sum_{k=1}^{m} \lambda^{jk}\Delta t & m \leq n, \\ 1 - I_n & m = n + 1. \end{cases}$$

We then have that

$$\Delta N_t^{jk} = \begin{cases} 1 & U_i \in [I_{k-1}, I_k], \\ 0 & \text{otherwise}, \end{cases}$$

for all $k = 1, \ldots, n$, and where $I_0 = 0$. Should the value of $\Delta N_t^{jk} = 1$ for any k, i.e. we have transitioned to state k, then the next simulation trial will be as above only with parameters for state k not j.

We are therefore able to generate Δr_t using one uniform $[0,1]$ random variable and one standardised normal random variable, and hence each simulation trial using $2 \times h$ random variables. In total we have N simulations and so need to generate $2 \times N \times h$ random variables. We can already see that the number of calculations that need to be performed is linear in h, whilst the number required using the multinomial tree method was exponential in h.

The accuracy of the estimate given by the Monte Carlo simulation is clearly going to be dependent on the number of trials N which are performed. The value of the derivative will be given by the mean of the simulation values, which we will denote π. We are also able to calculate the standard deviation of these simulated values which we will denote by ω. Since these are i.i.d. trials, the variance of π will therefore be given by

$$Var(\pi) = \frac{\omega^2}{N}.$$

Using a normal approximation, we can thus write a 95% confidence interval for the price of the derivative C_T as

$$\pi - \frac{1.96\omega}{\sqrt{N}} < C_T < \pi + \frac{1.96\omega}{\sqrt{N}}.$$

So the accuracy of our estimate will be proportional to the number of trials we perform.

5.5 Monte Carlo Simulation for HJM Models

We shall now use Monte Carlo methods to price derivatives under the HJM model described in section 5.1. Using a similar methodology to Carverhill and Pang [1995] we can write the following lemma:

Lemma 5.5.1. *The price of a call option on a pure discount bond given in (5.1.8) can be written as*

$$C_t = E[\max(0, p(t,T^*)H(t,T,T^*) - Kp(t,T)H(t,T,T))], \tag{5.5.1}$$

where

$$\begin{aligned} H(t,T,\tau) &= \exp\{\sum_{b=1}^{r}\int_{u=t}^{T}[S_b(u,\tau,Y_{u-})d\tilde{W}_u^b - \frac{1}{2}S_b^2(u,\tau,Y_{u-})du] \\ &+ \sum_{j=1}^{n}\sum_{k=1}^{n}\int_{u=t}^{T}\log(1+\gamma^p(u,\tau,j,k))d\tilde{N}_u^{jk}\}. \end{aligned}$$

Proof. We can immediately see that (5.1.8) can be written as

$$\begin{aligned} C_t &= E[exp\{-\int_{u=t}^{T} r(u)du\}\max(0, p(T,T^*) - K)] \\ &= E[\max(0, p(T,T^*)exp\{-\int_{u=t}^{T} r(u)du\} - Kexp\{-\int_{u=t}^{T} r(u)du\})]. \end{aligned} \tag{5.5.2}$$

Applying Itô's lemma to (5.1.6) as well as analogous lemma for the jumps we get

$$p(T,T^*) = \exp\left\{\int_{u=T}^{T^*}\left(r_u - \frac{1}{2}\sum_{b=1}^{r} S_b^2(u,T^*,Y_{u-}\right) du + \sum_{b=1}^{r}\int_{u=T}^{T^*} S_b(u,T^*,Y_{u-})d\tilde{W}_u^b$$

$$+ \sum_{j=1}^{n}\sum_{k=1}^{n}\int_{u=T}^{T^*}\log(1+\gamma^p(u,T^*,j,k))d\tilde{N}_u^{jk}\right\}.$$

It is therefore straightforward to show that

$$p(T,T^*)exp\{-\int_{u=t}^{T} r(u)du\} = p(t,T^*)H(t,T,T^*) \qquad (5.5.3)$$

$$exp\{-\int_{u=t}^{T} r(u)du\}) = p(t,T)H(t,T,T). \qquad (5.5.4)$$

Inserting (5.5.3) and (5.5.4) int (5.5.2) completes the proof.

□

Using lemma 5.5.1 we can value the price of the call option using Monte Carlo methods with $l = 1, \ldots, M$ simulations by

$$C_t = \frac{1}{M}\sum_{l=1}^{M}\max(0, p(t,T^*)H_l(t,T,T^*) - Kp(t,T)H_l(t,T,T)), \qquad (5.5.5)$$

where when in state j we have

$$\begin{aligned} H_l(t,T,\tau) &= \exp\{\sum_{i=1}^{N}\sum_{b=1}^{r}[S_b(t+i\Delta t,\tau,j)\epsilon_{i,b}\sqrt{\Delta t} - \frac{1}{2}S_b^2(t+i\Delta t,\tau,j)\Delta t] \\ &+ \sum_{i=1}^{N}\sum_{k=1}^{n}\log(1+\gamma^p(t+i\Delta t,\tau,j,k))\Delta\tilde{N}^{jk}\}, \end{aligned}$$

where $N = \frac{T}{\Delta t}$, $\epsilon_{i,b} \overset{\text{iid}}{\sim} N(0,1)$ and $\Delta\tilde{N}^{jk}$ are Bernoulli random variables with probability $\lambda^{jk}\Delta t$. We once again impose the restriction that

$$\Delta N_t^{j1} + \Delta N_t^{j2} + \ldots + \Delta N_t^{jk} \leq 1$$

for all j, i.e. at most one jump per time period. We may generate these values as we did in the short rate simulations in section 5.4.

5.6 Binomial Trees for HJM Models

We shall now discuss the pricing of Derivatives under the MCJD HJM framework using binomial trees. We have the bond-price dynamics given by (5.1.6), only now we shall be taking the savings account as numéraire to avoid having to maintain the short rate at every node. We can write a discrete time analogue for this process as follows:

$$\Delta\tilde{p}(t,T) = \sum_{j=1}^{n} I_t^j \tilde{p}(t,T)[\sum_{b=1}^{r} S_b(t,T,j)\Delta\tilde{W}_t^b + \sum_{k=1}^{n}\gamma(t,T,j,k)\Delta\tilde{N}_t^{jk}]. \quad (5.6.1)$$

We can see from 5.6.1 that $\Delta p(t,T)$ clearly has expected value 0 as it is a martingale. Its variance is given by

$$Var[\Delta\tilde{p}(t,T)] = \sum_{j=1}^{n} p_t^j \tilde{p}(t,T)^2 \Delta t \left[\sum_{b=1}^{r} S_b^2(t,T,j) + \sum_{k=1}^{n} \gamma^2(t,T,j,k)\lambda^{jk}(1-\lambda^{jk}\Delta t)\right]. \tag{5.6.2}$$

In our binomial model, the bond price can either go up to value $\tilde{p}_u(t+\Delta t,T)$ or down to $\tilde{p}_d(t+\Delta t,T)$ with equal probability $p=0.5$. For this to accurately represent our model we must equate mean and variance to get

$$\tilde{p}_u(t+\Delta t,T) + \tilde{p}_d(t+\Delta t,T) = \tilde{p}(t,T)$$

and

$$\frac{1}{4}[\tilde{p}_u(t+\Delta,T) - \tilde{p}_d(t+\Delta,T)]^2 = Var[\Delta\tilde{p}(t,T)].$$

Solving this we get

$$\begin{aligned}\tilde{p}_u(t+\Delta,T) &= \frac{\tilde{p}(t,T)}{2} + \sqrt{Var[\Delta\tilde{p}(t,T)]}\\ \tilde{p}_d(t+\Delta,T) &= \frac{\tilde{p}(t,T)}{2} - \sqrt{Var[\Delta\tilde{p}(t,T)]}.\end{aligned}$$

The bond price is then evolved over the N periods, where $N = \frac{T}{\Delta T}$, until the all the 2^N final node values are calculated and the derivative may be priced as in section 5.2.

Appendix A

A Distributional Result for the Equity Model

In this appendix we will derive the moment generating function of the time T stock price S_T defined in section 2.5.

We shall begin with the following lemma (see Whittaker and Watson [1946]).

Lemma A.0.1 (Dirichlet's integral). *Suppose we have the following integral:*

$$R_{x_1 \ldots x_z} = \int_{x_2=x_1}^{T} \cdots \int_{x_z=x_{z-1}}^{T} (x_2 - x_1)^{y_2} \ldots (x_z - x_{z-1})^{y_z} dx_2 \ldots dx_z$$

where all the superscripts are integers. We find that

$$R_{x_1 \ldots x_z} = \frac{y_2! \ldots y_z!}{(y_2 + \ldots + y_z + z - 1)!} (T - x_1)^{(y_2 + \ldots + y_z + z - 1)}. \tag{A.0.1}$$

Proof. We shall use a proof by induction. Suppose (A.0.1) is true for a given value of x_1 and z. We can calculate $R_{x_0 \ldots x_z}$ as follows:

$$\begin{aligned} R_{x_0 \ldots x_z} &= \int_{x_1=x_0}^{T} R_z (x_1 - x_0)^{y_1} dx_1 \\ &= \int_{x_1=x_0}^{T} \frac{y_2! \ldots y_z!}{(y_2 + \ldots + y_z + z - 1)!} (T - x_1)^{(y_2 + \ldots + y_z + z - 1)} (x_1 - x_0)^{y_1} dx_1. \end{aligned} \tag{A.0.2}$$

It can easily be shown by taking recursive integrals that

$$\int_{x_1=x_0}^{T} (T - x_1)^a (x_1 - x_0)^{y_1} dx_1 = \frac{y_1! a!}{(y_1 + a + 1)!} (T - x_0)^{y_1 + a + 1}, \tag{A.0.3}$$

and so on inserting (A.0.3) into (A.0.2) we find

$$R_{x_0 \ldots x_z} = \frac{y_1! \ldots y_z!}{(y_1 + \ldots + y_z + z)!} (T - x_0)^{(y_1 + \ldots + y_z + z)}.$$

Taking trivial values for x_1 for and z completes the proof. Note that we have shown that for a given value of x_1 and z the result holds true for x_0 and z as well. However, as the symmetry of the integral will suggest, this is equivalent to taking values of x_1 and $z + 1$.

□

Corollary A.0.2. *Using lemma A.0.1 we can now easily calculate the following integral*

$$R = \int_{x_1=t}^{T} \int_{x_2=x_1}^{T} \cdots \int_{x_z=x_{z-1}}^{T} x_1^{y_1} (x_2 - x_1)^{y_2} \ldots (x_z - x_{z-1})^{y_z} (T - x_z)^{y_z + 1} dx_1 dx_2 \ldots dx_z$$

by taking successive integrals to be equal to

$$R = \sum_{i=0}^{y_1} \frac{i! y_2! \ldots y_{z+1}!}{(i + y_2 + \ldots + y_{z+1} + z)!} t^{y_1 - i} (T - t)^{(i + y_2 + \ldots + y_{z+1} + z)}.$$

This result will be usefull in the following lemma.

□

Lemma A.0.3. *Suppose we have a Markov chain process Y_t where we condition on the path $\mathcal{J}$ as described in section 2.5. Let us define*

$$I^{k_1 k_2 \ldots k_z} = E[\int_{s_1=0}^{T} \int_{s_2=0}^{T} \cdots \int_{s_z=0}^{T} I_{s_1}^{k_1} I_{s_2}^{k_2} \ldots I_{s_z}^{k_z} ds_1 \ldots ds_z] \qquad \text{(A.0.4)}$$

where I_t^j is the indicator variable that $Y_t = j$. In terms of the Markov chain D_t given in the section 1.3.2 with generator $\mathbf{Q}_{\mathcal{J}}$, and defining the set $\mathcal{G}$ such that

$$\mathcal{G} = \{w_1, \ldots, w_z : \{j_{w_1}, \ldots, j_{w_z}\} = \{k_1, \ldots, k_z\},\ w_i \le w_{i+1}\ i = 1, \ldots, z\},$$

we then have

$$I^{k_1 k_2 \ldots k_z} = z! \sum_{\mathcal{G}} \frac{\sum_{y_1=0}^{\infty} \cdots \sum_{y_{z+1}=0}^{\infty} A(\mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y_1} \mathbf{1}^{w_1}) \ldots (\mathbf{1}^{w_{z-1}\prime} \mathbf{Q}_{\mathcal{J}}^{y_z} \mathbf{1}^{w_z})(\mathbf{1}^{w_z\prime} \mathbf{Q}_{\mathcal{J}}^{y_{z+1}} \mathbf{1}^{x+1})}{(\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1})}$$

(A.0.5)

where

$$A = \frac{T^{(y_1 + y_2 + \ldots + y_{z+1} + z)}}{(y_1 + y_2 + \ldots + y_{z+1} + z)!}.$$

Proof. Let the indicator variable U_t^w be such that

$$U_t^w = \begin{cases} 1 & \text{if } D_t = w, \\ 0 & \text{otherwise.} \end{cases} \tag{A.0.6}$$

We can therefore rewrite (A.0.4) in terms of the Markov chain D_t rather than Y_t as follows:

$$I^{k_1 k_2 \ldots k_z} = E[\sum_{\mathcal{G}} \int_{s_1=0}^{T} \int_{s_2=0}^{T} \cdots \int_{s_z=0}^{T} U_{s_1}^{w_1} U_{s_2}^{w_2} \ldots U_{s_z}^{w_z} ds_1 \ldots ds_z]. \tag{A.0.7}$$

Take successive conditional expectations of (A.0.7) and note that since $w_i \leq w_{i+1}$ therefore $s_i \leq s_{i+1}$ $\forall i$ so that

$$I^{k_1 k_2 \ldots k_z} = \frac{1}{p_T^{w_1 w_{z+1}}} \sum_{\mathcal{G}} \int_{s_1=0}^{T} \int_{s_2=s_1}^{T} \cdots \int_{s_z=s_{z-1}}^{T} p_{s_1}^{w_0 w_1} p_{s_2-s_1}^{w_1 w_2} \cdots p_{s_z-s_{z-1}}^{w_{z-1} w_z} p_{T-s_z}^{w_z w_{z+1}} ds_1 \ldots ds_z$$

where once again

$$p_t^{w_i w_{i+1}} = P[D_t = w_{i+1} | \, D_0 = w_i],$$

and set $w_0 = 1$ and $w_{z+1} = x + 1$. Using corollary (1.3.2) we get

$$I^{k_1 k_2 \ldots k_z} = \sum_{\mathcal{G}} \int_{s_1=0}^{T} \cdots \int_{s_z=s_{z-1}}^{T} \Psi \; ds_1 \ldots ds_z, \tag{A.0.8}$$

where

$$\Psi = \frac{(\sum_{y=0}^{\infty} \frac{s_1{}^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^y \mathbf{1}^{w_1}) \ldots (\sum_{y=0}^{\infty} \frac{(s_z - s_{z-1})^y}{y!} \mathbf{1}^{w_{z-1}\prime} \mathbf{Q}_{\mathcal{J}}^y \mathbf{1}^{w_z})(\sum_{y=0}^{\infty} \frac{(T-s_z)^y}{y!} \mathbf{1}^{w_z\prime} \mathbf{Q}_{\mathcal{J}}^y \mathbf{1}^{x+1})}{(\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^y \mathbf{1}^{x+1})}.$$

This can be re-written as

$$\Psi = \frac{\sum_{y_1=0}^{\infty} \cdots \sum_{y_{z+1}=0}^{\infty} A(\mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y_1} \mathbf{1}^{w_1}) \ldots (\mathbf{1}^{w_{z-1}\prime} \mathbf{Q}_{\mathcal{J}}^{y_z} \mathbf{1}^{w_z})(\mathbf{1}^{w_z\prime} \mathbf{Q}_{\mathcal{J}}^{y_z} \mathbf{1}^{x+1})}{(\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^y \mathbf{1}^{x+1})}, \tag{A.0.9}$$

where

$$A = \frac{s_1{}^{y_1} \ldots (s_z - s_{z-1})^{y_z} (T - s_z)^{y_{z+1}}}{y_1! \ldots y_{z+1}!}.$$

Using corollary A.0.2 we have

$$\int_{s_1=0}^{T} \cdots \int_{s_z=s_{z-1}}^{T} A \, ds_1 \ldots ds_z = \frac{T^{(y_1+y_2+\ldots+y_{z+1}+z)}}{(y_1 + y_2 + \ldots + y_{z+1} + z)!}. \tag{A.0.10}$$

On inserting (A.0.9) together with (A.0.10) into (A.0.8) we obtain (A.0.5), noting that since we have assumed $s_i \leq s_{i+1}$ for $i = 1, \ldots z$ we therefore need to multiply by $z!$ to obtain all combinations.

□

We are now in a position to write the following lemma:

Lemma A.0.4. *Suppose we have a stochastic process whose time t value is given by X_t for $0 \leq t < T$, and has dynamics*

$$dX_t = \sum_{j=1}^{n} I_t^j \left[\mu_j dt + \sigma_j dW_t\right], \tag{A.0.11}$$

where W_t is a standard Wiener processes, μ_j and σ_j^2 are the drift and variance respectively both of which are dependent on the state of Markov chain Y_t. We also have that I_t^j is the indicator variable that $Y_t = j$. Even though the models looked at in section 2.5 had multiple Brownian motions, the model for each asset individually can be expressed using a single Brownian motion as shown in section 3.1. We condition on the path of the Markov chain following the jump sequence $\mathcal{J}$ in the time $[0,T]$ as described in section 2.5. Given $X_0 = 0$ we have that the moment generating function of X_T is given by:

$$[M_{X_T}(r)|\mathcal{J}] = 1 + m_1 r + \frac{1}{2!} m_2 r^2 + \dots \tag{A.0.12}$$

where

$$m_d = \sum_{z=||d/2||}^{d} \sum_{k_1=1}^{n} \cdots \sum_{k_z=1}^{n} \frac{d! I^{k_1 \dots k_z}}{z! 2^{d-z}} \left\{ \sum_{b_1=1}^{z} \cdots \sum_{b_{d-z}=1}^{z} \left(\prod_{\substack{c=1 \\ c \neq b_1, \dots, b_{d-z}}}^{z} \mu_{k_c} \sigma_{k_{b_1}}^2 \cdots \sigma_{k_{b_{d-z}}}^2 \right) \right\}$$

for $d = 1, \dots, \infty$, where all the subscripts above are integers and we denote by $||z/2||$ the smallest integer above z.

Proof. The moment generating for X_T is defined as

$$M_{X_T}(r) = E\left[exp\left\{\int_{s=0}^{T}\sum_{j=1}^{n} rI_s^j\left[\mu_j dt + \sigma_j dW_t\right]\right\}\right].$$

Conditioning on the path of the Markov chain, i.e. assuming that I_s^j are known for all s, we find that

$$[M_{X_T}(r) \mid \mathcal{J},\, I_s^j\, \forall s] = exp\left\{\int_{s=0}^{T}\sum_{j=1}^{n}(I_s^j dt)[\mu_j r + \frac{1}{2}r^2\sigma_j^2]\right\},$$

where once we uncondition we have

$$[M_{X_T}(r) \mid \mathcal{J}] = E\left[exp\left\{\int_{s=0}^{T}\sum_{j=1}^{n}(I_s^j dt)[\mu_j r + \frac{1}{2}r^2\sigma_j^2]\right\}\right].$$

Applying Taylor's expansion and using lemma A.0.3 we get

$$
\begin{aligned}
M_{X_T}(r) &= 1+\sum_{k_1=1}^{n} I^{k_1}(\mu_{k_1}r+\frac{1}{2}r^2\sigma_{k_1}^2) \\
&+ \frac{1}{2}\sum_{k_1=1}^{n}\sum_{k_2=1}^{n} I^{k_1k_2}(\mu_{k_1}r+\frac{1}{2}r^2\sigma_{k_1}^2)(\mu_{k_2}r+\frac{1}{2}r^2\sigma_{k_2}^2) \\
&+ \ldots\ldots\ldots\ldots \\
&+ \frac{1}{z!}\sum_{k_1=1}^{n}\cdots\sum_{k_z=1}^{n} I^{k_1\ldots k_z}(\mu_{k_1}r+\frac{1}{2}r^2\sigma_{k_1}^2)\ldots(\mu_{k_z}r+\frac{1}{2}r^2\sigma_{k_z}^2) \\
&+ \ldots\ldots\ldots\ldots
\end{aligned}
$$

On rearranging we obtain (A.0.12).

□

Corollary A.0.5. *Suppose we have a random variable X_T with $X_0 = 0$ and which has dynamics given by (A.0.11). Let us define the variable S_T*

$$S_T = S_0 exp\{X_T\}. \tag{A.0.13}$$

We can derive from this the moment generating function of S_T using Taylor's *expansion as follows:*

$$\begin{aligned} M_{S_T}(r) &= E\left[exp\left\{S_0 r e^{X_T}\right\}\right] \\ &= E\left[1 + S_0 r e^{X_T} + \frac{S_0^2 r^2 e^{2X_T}}{2!} + \ldots\right] \\ &= 1 + S_0 r M_{X_T}(1) + \frac{S_0^2 r^2 M_{X_T}(2)}{2!} + \ldots \end{aligned}$$

where the values $M_{X_T}(u)$ are given by equation (A.0.12).

□

Appendix B

A Distributional Result for Short-Rate Models

In this appendix we shall develop a methodology to calculate the value of $M(t,T)$ and $\hat{M}(t,T)$ used for pricing derivatives of the short rate in section 4.3.2.

Let us begin by writing an extended version of lemma A.0.1:

Lemma B.0.6. *Suppose we have the following integral:*

$$R_{x_1 \dots x_z} = \int_{x_2=x_1}^{u_2} \dots \int_{x_z=x_{z-1}}^{u_z} (x_2 - x_1)^{y_2} \dots (x_z - x_{z-1})^{y_z} dx_2 \dots dx_z$$

where $u_i < u_{i+1}$ for $i = 2, \dots, z-1$ and all the superscripts are integers. We find that

$$R_{x_1 \ldots x_z} = \sum_{h_z=0}^{y_z} \cdots \sum_{h_2=0}^{y_2} [\prod_{d=3}^{z} \frac{y_d!(u_d - u_{d-1})^{y_d+h_d-h_{d+1}+1}}{(y_d + h_d - h_{d+1} + 1)!}] \frac{y_2!(u_2 - x_1)^{y_2-h_2}}{(y_2 - h_2)!} \tag{B.0.1}$$

where $h_{z+1} = -1$.

Proof. We shall use a proof by induction. Suppose (A.0.1) is true for a given value of x_1 and z. We can calculate $R_{x_0 \ldots x_z}$ as follows:

$$R_{x_0 \ldots x_z} = \int_{x_1=x_0}^{u_1} R_z (x_1 - x_0)^{y_1} dx_1. \tag{B.0.2}$$

It can easily be shown by taking recursive integrals that

$$\int_{x_1=x_0}^{u_1} (u_2 - x_1)^a (x_1 - x_0)^{y_1} dx_1 = \sum_{h_1=0}^{y_1} \frac{y_1! a! (u_2 - u_1)^{a+h_1+1}}{(y_1 - h_1)!(a + h_1 + 1)!} (u_1 - x_0)^{y_1-h_1}, \tag{B.0.3}$$

and so on inserting (B.0.3) into (B.0.2) we find

$$R_{x_0 \ldots x_z} = \sum_{h_z=0}^{y_z} \cdots \sum_{h_1=0}^{y_1} [\prod_{d=2}^{z} \frac{y_d!(u_d - u_{d-1})^{y_d+h_d-h_{d+1}+1}}{(y_d + h_d - h_{d+1} + 1)!}] \frac{y_1!(u_1 - x_0)^{y_1-h_1}}{(y_1 - h_1)!}.$$

Taking trivial values for x_1 for and z completes the proof. Note that we have shown that for a given value of x_1 and z the result holds true for x_0 and z as well. However, as the symmetry of the integral will suggest, this is equivalent to taking values of x_1 and $z + 1$.

□

Corollary B.0.7. *Using lemma B.0.6 we can now easily calculate the following integral:*

$$R = \int_{x_1=0}^{u_1} \int_{x_2=x_1}^{u_2} \cdots \int_{x_z=x_{z-1}}^{u_z} x_1(x_2-x_1)^{y_2} \ldots (x_z - x_{z-1})^{y_z} dx_2 \ldots dx_z$$

where $u_i < u_{i+1}$ *for* $i = 1, \ldots, z-1$ *to be equal to*

$$R = \sum_{h_z=0}^{y_z} \cdots \sum_{h_1=0}^{y_1} \left[\prod_{d=2}^{z} \frac{y_d!(u_d - u_{d-1})^{y_d+h_d-h_{d+1}+1}}{(y_d + h_d - h_{d+1} + 1)!} \right] \frac{y_1! u_1^{y_1-h_1}}{(y_1 - h_1)!}.$$

Moreover, using corollary A.0.2 we have

$$\int_{u_1=t}^{T} \int_{u_2=u_1}^{T} \cdots \int_{u_z=u_{z-1}}^{T} R du_1 \ldots du_z =$$

$$\sum_{h_1=0}^{y_1} \cdots \sum_{h_z=0}^{y_z} \sum_{h_0=0}^{y_1-h_1} \frac{(h_0! y_1! \ldots y_z!) t^{y_1-h_1-h_0} (T-t)^{h_0+h_2+y_2+\ldots+y_z+2z-2}}{(y_1-h_1)!(h_0+h_2+y_2+\ldots+y_z+2z-2)!}.$$

□

We can now derive the following lemma:

Lemma B.0.8. *Suppose we have a Markov chain process* Y_t*, where we condition on the path* $\mathcal{J}$ *as described in section 1.3.2 when there are* x *jumps. Let us define*

$$\begin{aligned} I^{k_1 k_2 \ldots k_z}(i_1, \ldots, i_e) \; = \; & E[\int_{u_1=t}^{T} \cdots \int_{u_z=0}^{T} \int_{s_1=0}^{u_1} \cdots \int_{s_z=0}^{u_z} \\ & I_{s_1}^{k_1} \ldots I_{s_z}^{k_z} dN_{s_{i_1}} \ldots dN_{s_{i_e}} ds_{r_1} \ldots ds_{r_{z-e}} du_1 \ldots du_z], \end{aligned} \tag{B.0.4}$$

where

$$\{i_1, \ldots, i_e\} \in \{1, \ldots, z\}$$

and

$$\{i_1, \ldots, i_e\} \cup \{r_1, \ldots, r_{z-e}\} = \{1, \ldots, z\},$$

and where I_t^j *is the indicator variable that* $Y_t = j$. *In terms of the Markov chain* D_t *as also described in section 1.3.2 with generator* $\mathbf{Q}_{\mathcal{J}}$, *and defining the set* $\mathcal{G}$ *such that*

$$\mathcal{G} = \{w_1, \ldots, w_{z+1} : \{j_{w_1}, \ldots, j_{w_z}\} = \{k_1, \ldots, k_z\}, w_i \leq w_{i+1} i = 1, \ldots, z{+}1\},$$

where w_{z+1} *can take any value in* $[1, x+1]$. *We then have*

$$I^{k_1 k_2 \ldots k_z}(i_1, \ldots, i_e) = \sum_{\mathcal{G}} \frac{\Psi}{\Delta}, \tag{B.0.5}$$

where

$$\Psi = \sum_{y_1=0}^{\infty} \cdots \sum_{y_z=0}^{\infty} A(\mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y_1} \mathbf{1}^{w_1}) \prod_{d=1}^{e} \lambda^{w_{i_d} w_{i_d+1}} (\mathbf{1}^{w_{i_d}\prime} \mathbf{Q}_{\mathcal{J}}^{y_{i_d}+1} \mathbf{1}^{w_{i_d+1}}) \prod_{\substack{h=1 \\ j_h \neq z}}^{z-e} (\mathbf{1}^{w_{j_h}\prime} \mathbf{Q}_{\mathcal{J}}^{y_{j_h}+1} \mathbf{1}^{w_{j_h+1}}),$$

and

$$\Delta = \Big(\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1}\Big)^{z+e},$$

as well as

$$A = \sum_{h_1=0}^{y_1} \cdots \sum_{h_z=0}^{y_z} \sum_{h_0=0}^{y_1-h_1} \frac{(h_0! y_1! \ldots y_z!) t^{y_1-h_1-h_0} (T-t)^{h_0+h_2+y_2+\ldots+y_z+2z-2}}{(y_1-h_1)!(h_0+h_2+y_2+\ldots+y_z+2z-2)!}.$$

Proof. We again have the indicator variable U_t^w where:

$$U_t^w = \begin{cases} 1 & \text{if } D_t = w, \\ 0 & \text{otherwise.} \end{cases} \tag{B.0.6}$$

We can therefore rewrite (B.0.4) in terms of the Markov chain D_t rather than Y_t as follows:

$$\begin{aligned} I^{k_1 k_2 \ldots k_z}(i_1, \ldots, i_e) &= E[\sum_{\mathcal{G}} \int_{u_1=t}^{T} \cdots \int_{u_z=t}^{T} \int_{s_1=0}^{u_1} \cdots \int_{s_z=0}^{u_z} \\ & U_{s_1}^{w_1} \ldots U_{s_z}^{w_z} dN_{s_{i_1}} \ldots dN_{s_{i_e}} ds_{r_1} \ldots ds_{r_{z-e}} du_1 \ldots du_z]. \end{aligned} \tag{B.0.7}$$

Taking successive conditional expectations of (B.0.7) and noting that since $w_i \leq w_{i+1}$ $\forall i$ therefore $s_i \leq s_{i+1}$ $\forall i$, we have that

$$\begin{aligned} I^{k_1 k_2 \ldots k_z}(i_1, \ldots, i_e) &= \sum_{\mathcal{G}} \int_{u_1=t}^{T} \int_{u_2=u_1}^{T} \cdots \int_{u_z=u_{z-1}}^{T} \int_{s_1=0}^{u_1} \int_{s_2=s_1}^{u_2} \cdots \int_{s_z=s_{z-1}}^{u_z} \\ & p_{s_1}^{w_0 w_1} \prod_{d=1}^{e} \bar{\lambda}^{w_{i_d} w_{i_d+1}} p_{s_{i_d+1}-s_{i_d}}^{w_{i_h+1} w_{i_h+1}} \prod_{\substack{h=1 \\ j_h \neq z}}^{z-e} p_{s_{j_h+1}-s_{j_h}}^{w_{j_h} w_{j_h+1}} ds_1 \ldots ds_z du_1 \ldots du_z, \end{aligned}$$

where once again

$$p_t^{w_i w_{i+1}} = P[D_t = w_{i+1} |\, D_0 = w_i],$$

and setting $w_0 = 1$ and $p^{w_z w_{z+1}}_{s_z+1-s_z} = 1$, and where finally

$$\bar{\lambda}^{w_j w_k} = \frac{\lambda^{w_j w_k}}{p_T^{w_0 w_{z+1}}}.$$

Using lemma (A.0.3) we get

$$I^{k_1 k_2 \dots k_z}(i_1, \dots, i_e) = \sum_{\mathcal{G}} \int_{u_1=t}^{T} \int_{u_2=u_1}^{T} \cdots \int_{u_z=u_{z-1}}^{T} \int_{s_1=0}^{u_1} \int_{s_2=s_1}^{u_2} \cdots \int_{s_z=s_{z-1}}^{u_z} \frac{\Psi}{\Delta}\, ds_1 \dots ds_z du_1 \dots du_z, \tag{B.0.8}$$

where

$$\Psi = \left(\sum_{y=0}^{\infty} \frac{{s_1}^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{w_1} \right) \prod_{d=1}^{e} \lambda^{w_{i_d} w_{i_d+1}} \left(\sum_{y=0}^{\infty} \frac{(s_{i_d+1} - s_{i_d})^y}{y!} \mathbf{1}^{w_{i_d}\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{w_{i_d+1}} \right).$$

$$\prod_{\substack{h=1\\ j_h \neq z}}^{z-e} \left(\sum_{y=0}^{\infty} \frac{(s_{j_h+1} - s_{j_h})^y}{y!} \mathbf{1}^{w_{j_h}\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{w_{j_h+1}} \right)$$

$$\Delta = \left(\sum_{y=0}^{\infty} \frac{T^y}{y!} \mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y} \mathbf{1}^{x+1} \right)^{z+e}.$$

We can re-write Ψ as

$$\Psi = \sum_{y_1=0}^{\infty} \cdots \sum_{y_z=0}^{\infty} K(\mathbf{1}^{1\prime} \mathbf{Q}_{\mathcal{J}}^{y_1} \mathbf{1}^{w_1}) \prod_{d=1}^{e} \lambda^{w_{i_d} w_{i_d+1}} (\mathbf{1}^{w_{i_d}\prime} \mathbf{Q}_{\mathcal{J}}^{y_{i_d}+1} \mathbf{1}^{w_{i_d+1}}) \prod_{\substack{h=1\\ j_h \neq z}}^{z-e} (\mathbf{1}^{w_{j_h}\prime} \mathbf{Q}_{\mathcal{J}}^{y_{j_h}+1} \mathbf{1}^{w_{j_h+1}}), \tag{B.0.9}$$

where

$$K = \frac{s_1{}^{y_1} \dots (s_z - s_{z-1})^{y_z}}{y_1! \dots y_z!}.$$

Using corollary B.0.7 we have

$$\int_{u_1=t}^{T} \int_{u_2=u_1}^{T} \cdots \int_{u_z=u_{z-1}}^{T} \int_{s_1=0}^{u_1} \int_{s_2=s_1}^{u_2} \cdots \int_{s_z=s_{z-1}}^{u_z} K \, ds_1 \dots ds_z du_1 \dots du_z =$$

$$\sum_{h_1=0}^{y_1} \cdots \sum_{h_z=0}^{y_z} \sum_{h_0=0}^{y_1-h_1} \frac{(h_0! y_1! \dots y_z!) t^{y_1-h_1-h_0} (T-t)^{h_0+h_2+y_2+\dots+y_z+2z-2}}{(y_1-h_1)!(h_0+h_2+y_2+\dots+y_z+2z-2)!},$$

(B.0.10)

where on inserting (B.0.9) together with (B.0.10) into (B.0.8) the proof is complete.

□

Lemma B.0.9. *Suppose we have a stochastic process whose time t value is given by $H(t,T)$ for $0 \le t < T$, where*

$$H(t,T) = E\left[exp\left\{-\int_{u=t}^{T} \int_{s=0}^{u} \sum_{j=1}^{n} I_s^j (\sigma_j dW_s + \sum_{k=1}^{n} \gamma^{jk} dN_s^{jk}) du\right\}\right]$$

and where we condition on the path $\mathcal{J}$ as described in section 1.3.2. We therefore get

$$H(t,T) = \sum_{d=0}^{\infty} \frac{(-1)^d}{d!} X^d(t,T), \qquad \text{(B.0.11)}$$

where

$$X^d(t,T) = \sum_{k_1=1}^{n} \cdots \sum_{k_d=1}^{n} \sum_{e=0}^{d} \sum_{\mathcal{C}^e} \sum_{\mathcal{G}} \frac{\Psi}{2^e \Delta} \gamma^{w_{i_1} w_{i_1+1}} \cdots \gamma^{w_{i_e} w_{i_e+1}} \sigma_{j_1}^2 \cdots \sigma_{j_{d-e}}^2 \tag{B.0.12}$$

where Ψ, Δ *and the set* $\mathcal{G}$ *are as defined as in lemma B.0.8, and where*

$$\mathcal{C}^e = \{i_1, \ldots, i_e : \{i_1, \ldots, i_e\} \in \{1, \ldots, d\}, i_r < r+1, \;\; r = 1, \ldots, d-1\}$$

and also

$$\{j_1, \ldots, j_{d-e} : \{j_1, \ldots, j_{d-e}\} \cup \{i_1, \ldots, i_e\} = \{1, \ldots, d\}, j_r < j_{r+1}, \; r = 1, \ldots, d-e-1\}.$$

Proof. Conditioning on the state of the Markov chain during the interval $[t,T]$, i.e. that we know I_s^j for $t \leq s \leq T$, we have

$$[H(t,T)|I_s^j] = exp\left\{-\int_{u=t}^{T}\int_{s=0}^{u}\sum_{j=1}^{n} I_s^j(\frac{1}{2}\sigma_j^2 ds + \sum_{k=1}^{n}\gamma^{jk} dN_s^{jk})du\right\},$$

and once we uncondition we get

$$H(t,T) = E\left[exp\left\{-\int_{u=t}^{T}\int_{s=0}^{u}\sum_{j=1}^{n} I_s^j(\frac{1}{2}\sigma_j^2 ds + \sum_{k=1}^{n}\gamma^{jk} dN_s^{jk})du\right\}\right].$$

Let us write

$$X(t,T) \;\; = \;\; \int_{u=t}^{T}\int_{s=0}^{u}\sum_{j=1}^{n} I_s^j(\frac{1}{2}\sigma_j^2 ds + \sum_{k=1}^{n}\gamma^{jk} dN_s^{jk})du,$$

so that on applying Taylor's expansion we get

$$H(t,T) = E[1 - X(t,T) + \frac{1}{2}X(t,T)^2 - \frac{1}{6}X(t,T)^3 \ldots].$$

Denote by $X^d(t,T)$ the following expectation:

$$X^d(t,T) = E[X(t,T)^d],$$

so that $H(t,T)$ is now given by (B.0.11). We can now calculate the value of $X^d(t,T)$ as follows:

$$\begin{aligned} X^d(t,T) &= E\left[\int_{u_1=t}^{T} \cdots \int_{u_d=t}^{T} \int_{s_1=0}^{u_1} \cdots \int_{s_d=0}^{u_d} \left\{ \sum_{k_1=1}^{n} \cdots \sum_{k_d=1}^{n} I_{s_1}^{k_1} \ldots I_{s_d}^{k_d} \right.\right. \\ &\left.\left. (\frac{1}{2}\sigma_{k_1}^2 ds_1 + \sum_{l=1}^{n} \gamma^{k_1 l} dN_{s_1}^{k_1 l}) \ldots (\frac{1}{2}\sigma_{k_d}^2 ds_d + \sum_{l=1}^{n} \gamma^{k_d l} dN_{s_d}^{k_d l}) \right\} du_1 \ldots du_d \right]. \end{aligned}$$

Using the definition of C^e given above as well as employing lemma B.0.8 we can see that $X^d(t,T)$ is given by (B.0.12).

□

Corollary B.0.10. *Using lemma B.0.9 The value of $M(t,T)$ used in section 4.3.2 is then given by the sum of $logH(t,T)$ over all jump sequences $\mathcal{J}$, so that*

$$M(t,T) = \sum_{x=0}^{\infty} \sum_{j_1=1}^{n} \cdots \sum_{j_{x+1}=1}^{n} logH(t,T) p^{j_1 \cdots j_{x+1}}(T),$$

where $p^{j_1 \cdots j_{x+1}}(T)$ is given in corollary 1.3.3.

□

Appendix C

A Distributional Result for HJM Models

In this appendix we shall develop a methodology for calculating $\hat{Z}(T,T)$ and $Z^*(T,T)$ in Example 4.4.9.

We have the following lemma:

Lemma C.0.11. *Let us use the same parameter values as in Example 4.4.9, i.e.*

$$
\begin{aligned}
s_b(t,T,j) &= s_{b,j}, \\
\gamma^p(t,T,j,k) &= \gamma^{jk}(T-t).
\end{aligned}
$$

Conditioning on the path $\mathcal{J} = \{j_1, \ldots, j_{x+1}\}$ of Y_t as described in section 2.5, *we have that the moment generating function of $\hat{Z}(T,T)$ is given by*

$$M_{[\hat{Z}(T,T)\,|\mathcal{J}]}(r)=\sum_{d=0}^{\infty}\sum_{x=0}^{\infty}\sum_{k_1=1}^{n}\cdots\sum_{k_d=1}^{n}\frac{d^x(d-1)^x}{2^x x!}I^{k_1\ldots k_d}\bar{T}^{2x}s_{k_1}^2\ldots s_{k_d}^2 r^d,$$

where $I^{k_1\ldots k_d}$ is given by lemma A.0.3, $\bar{T}=T-T^*$, and where we have that

$$\gamma^{j_1\ldots j_x}=\prod_{i=1}^{x}(1-\gamma^{j_i j_{i+1}}\bar{T}). \tag{C.0.1}$$

Proof. We can see from (C.0.1) and (4.4.7) that

$$\begin{aligned} S_b(t,T,j) &= -\int_{u=t}^{T}s_b(t,u,j)du \\ &= -s_{b,j}(T-t), \end{aligned}$$

and

$$\begin{aligned} exp\{\Gamma^f(t,T^*,j,k)\} &= \gamma^p(t,T,j,k)+1 \\ &= \gamma^{jk}(T-t)+1. \end{aligned}$$

Inserting this into (4.4.39) we get

$$\begin{aligned} \hat{Z}(T,T) &= \hat{Z}(0,T)exp\{\bar{T}\int_{u=0}^{T}\sum_{j=1}^{n}I_u^j[-\frac{1}{2}\sum_{b=1}^{r}\bar{T}s_{b,j}^2du+\sum_{b=1}^{r}s_{b,j}d\tilde{W}_u^b] \\ &\quad +\sum_{j=1}^{n}\sum_{k=1}^{n}log[1-\gamma^{jk}\bar{T}]d\tilde{N}^{jk}\}, \end{aligned} \tag{C.0.2}$$

where $\bar{T} = T - T^*$. Let us define s_j for all j such that

$$s_j = \sqrt{\sum_{b=1}^{r} s_{b,j}^2},$$

so that

$$s_j d\tilde{W}_t \sim \sum_{b=1}^{r} s_{b,j} d\tilde{W}_t^b,$$

where $\tilde{W}_t$ is a $\mathbf{Q}$ brownian motion. Equation (C.0.2) then becomes

$$\begin{aligned}\hat{Z}(T,T) &= \hat{Z}(0,T)exp\{\bar{T}\int_{u=0}^{T}\sum_{j=1}^{n} I_u^j[-\frac{1}{2}\bar{T}s_j^2 du + s_j d\tilde{W}_u] \\ &+\sum_{j=1}^{n}\sum_{k=1}^{n} log[1-\gamma^{jk}\bar{T}]d\tilde{N}^{jk}\}.\end{aligned}$$

Conditioning on the path $\mathcal{J}$, we have that

$$[\hat{Z}(T,T)\,|\mathcal{J}] = \hat{Z}(0,T)^r\gamma^{j_1\cdots j_x}exp\{\bar{T}\int_{u=0}^{T}\sum_{j=1}^{n} I_u^j[-\frac{1}{2}\bar{T}s_j^2 du + s_j d\tilde{W}_u]\},$$

where $\gamma^{j_1\cdots j_x}$ is given by (C.0.1). Conditioning now also on the path I_s^j $0 \leq s \leq T$, we can calculate the following expectation:

$$\begin{aligned}E[\hat{Z}(T,T)^d\,|\mathcal{J}, I_T^j] &= \hat{Z}(0,T)^d\gamma^{j_1\cdots j_x d}E[exp\{d\bar{T}\int_{u=0}^{T}\sum_{j=1}^{n} I_u^j[\frac{1}{2}s_j^2 du - s_j d\tilde{W}_u]\}] \\ &= \hat{Z}(0,T)^d\gamma^{j_1\cdots j_x d}exp\{\frac{d(d-1)}{2}\bar{T}^2\int_{u=0}^{T}\sum_{j=1}^{n} I_u^j s_j^2 du\}.\end{aligned}$$

Unconditioning, we find

$$E[\hat{Z}(T,T)^d \,|\mathcal{J}] = \hat{Z}(0,T^*)^d \gamma^{j_1 \dots j_x d} E[exp\{\frac{d(d-1)}{2}\bar{T}^2 \int_{u=0}^{T} \sum_{j=1}^{n} I_u^j s_j^2 du\}]. \tag{C.0.3}$$

Using Taylor's expansion as well as lemma A.0.3, we have

$$E[exp\{\frac{d(d-1)}{2}\bar{T}^2 \int_{u=0}^{T} \sum_{j=1}^{n} I_u^j s_j^2 du\}] = \sum_{x=0}^{\infty} \frac{d^x (d-1)^x}{2^x x!} \bar{T}^{2x} \sum_{k_1=1}^{n} \dots \sum_{k_d=1}^{n} I^{k_1 \dots k_d} s_{k_1}^2 \dots s_{k_d}^2, \tag{C.0.4}$$

where $I^{k_1 \dots k_r}$ is given by (A.0.5). Inserting (C.0.4) into (C.0.3) completes the proof.

□

Corollary C.0.12. *We can derive the density function of* $[\hat{Z}(T,T)\,|\mathcal{J}]$ *given by* $f_{[\hat{Z}(T,T)\,|\mathcal{J}]}(z)$ *using lemma C.0.11, and hence calculate the probability* $\mathbf{Q}^T(p(T,T^*) \geq K)$ *using equation (4.4.40) so that*

$$\mathbf{Q}^T(p(T,T^*) \geq K) = \int_{z=k}^{\infty} \sum_{\mathcal{J}}^{\infty} f_{[\hat{Z}(T,T)\,|\mathcal{J}]}(z) dz. \tag{C.0.5}$$

It can be shown in a similar manner using (4.4.36) that $\mathbf{Q}^{T^*}(p(T,T^*) \geq K)$ *is given by*

$$\mathbf{Q}^{T^*}(p(T,T^*) \geq K) = \int_{z=-\infty}^{\frac{1}{k}} \sum_{\mathcal{J}}^{\infty} f_{[\hat{Z}^*(T,T)\,|\mathcal{J}]}(z) dz, \tag{C.0.6}$$

where the moment generating function of $f_{[\hat{Z}^*(T,T)\,|\mathcal{J}]}$ *is given by*

$$M_{[\hat{Z}^*(T,T)\,|\mathcal{J}]}(r) = \sum_{d=0}^{\infty}\sum_{x=0}^{\infty}\sum_{k_1=1}^{n}\cdots\sum_{k_d=1}^{n}\frac{d^x(d-1)^x}{2^x x!}I^{k_1\ldots k_d}\tilde{T}^{2x}s_{k_1}^2\ldots s_{k_d}^2 r^d$$

where all the terms are as in the previous lemma as well as $\tilde{T} = T^* - T$.

□

Bibliography

L. Andersen and J. Andersen. Jump diffusion models: volatility and smile fitting and numerical methods for pricing. *Review of Derivatives Research*, 4:231–262, 2000.

D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, 2004.

S. Asmussen. Matrix-analytic models and their analysis. *Scandinavian Journal of Statistics*, 27:193–226, 2000.

G. Bakshi, C. Cao, and Z. Chen. Empirical performance of alternative option pricing models. *Journal of Finance*, 52(5):2003–2049, 1997.

C.A. Ball and W.N. Torous. On jumps in common stock prices and their impact on call option pricing. *Journal of Finance*, 40(1):155–173, 1985.

N.H. Bingham and R. Kiesel. *Risk-neutral valuation - pricing and hedging of financial derivatives.* Springer Finance, second edition, 2004.

T. Björk. *Financial mathematics.* Springer, 1997.

T. Björk. *Arbitrage theory in continous time.* Oxford University Press, 1998.

T. Björk, Yu. Kabanov, and W.J. Runggaldier. Bond market structure in the presence of marked point processes. *Mathematical Finance*, 40(1):211–239, 1997.

F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Fixed Political Economy*, 81:637–654, 1973.

P. Brémaud. *Point processes and queues: martingale dynamics.* Springer Verlag, New York, 1981.

A.P. Carverhill and K. Pang. Efficient and flexible bond option valuation in the heath, jarrow and morton framework. *Journal of Fixed Income*, 5: 70–77, 1995.

L Clewlow and C Strickland. *Implementing derivative models.* John Wiley & Sons, 1998.

J.C. Cox and M. Rubenstein. *Options markets.* Prentice Hall, 1985.

C. de Boor. *A practical guide to splines.* Springer-Verlag New York, 1978.

D. Duffie. *Dynamic asset pricing theory.* Princeton University Press, 1992.

P.H. Dybvig and C. Huang. Non-negative wealth, absence of arbitrage, and feasible consumption plans. *Review of Financial Studies 1*, pages 377–401, 1988.

J.P. Fouque, G. Papanicolaou, and K.R. Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, 2000.

H. Geman, N. El Karoui, and J-C. Rochet. Changes of numéraire, changes of probability measure and option pricing. *Journal of Applied Probability*, 32:443–458, 1995.

G. Grimmett and D. Stirzaker. *Probability and random processes.* Oxford University Press, third edition, 2001.

J.M. Harrison and D.M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

J.M. Harrison and S.R. Pliska. Martingales and and stochastic integrals in the theory of continuous trading. *Journal of Stochastic Processes and Applications*, 11:215–260, 1981.

F. Hausdorff. Summationsmethoden und momentfolgen. *Math. Z.*, 9:74–109,280–299, 1921.

D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: a new methodology of contingent claim valuation. *Econometrica*, 60:77–105, 1992.

S. Heston. A closed-form solution for options with stochastic volatility with applications to bond currency options. *Review of Financial Studies*, 6: 327–343, 1993.

T.S.Y Ho and S.B. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41:1011–29, 1986.

J. Hull. *Options, futures, and other derivatives.* Prentice Hall, fifth edition, 2003.

J. Hull and A. White. The pricing of options with stochastic volatilities. *Journal of Finance*, 42:281–300, 1987.

J. Hull and A. White. Pricing interest rate derivative securities. *Review of Financial Studies*, 3:573–592, 1990a.

J. Hull and A. White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantative Analysis*, 25:87–100, 1990b.

J. Hull and A. White. Numerical procedures for implementing term structure models i: single factor models. *Journal of Derivatives*, 2:7–16, 1994.

P. Jorion. On jump processes in the foreign exchange and stock markets. *Review of Financial Studies*, 1(4):427–445, 1988.

D. Madan. Financial modelling with discontinuous price processes. In O. Barndorff-Nielsen, editor, *Levy processes - theory and applications.* Boston:Birkhäuser, 2001.

R.C. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Sciences*, 4:141–183, 1973.

R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

A. Mood, F. Graybill, and D. Boes. *Intoduction to the theory of Staistics.* McGraw-Hill Inernational Editions, third edition, 1974.

M. Musiela and M. Rutkowski. *Martingale Methods in Financial Modelling.* Springer, 1997.

R. Norberg. The markov chain market. *ASTIN Bulletin*, pages 265–287, 2003.

C.A. O'Cinneade. Characterization of phase-type distributions. *Comm. Statist. Stochastic Models*, 6:1–57, 1990.

B. Øksendal. *Stochastic differential equations - an introduction with applications.* Springer, fifth edition, 2000.

L.C.G. Rogers. Which model of the term structure of interest rates should one use? In M.H.A. Davis et al., editor, *Mathematical Finance*, volume 65 of IMA. Springer, Berlin Heidelberg New York, 1995.

S.A. Ross. The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13:341–360, 1976.

S.A. Ross. A simple approach to the valuation of risky streams. *Journal of Business*, 51:453–475, 1978.

W.J. Runggaldier. Jump-diffusion models. In S.T. Rachev, editor, *Handbook of heavy-tailed distributions in finance, Handbooks in finance Book 1.* W. Ziemba Series, 2003.

B.W. Silverman. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. *Journal of the Royal Statistical Society B*, 47:1–52, 1985.

E. Stein and J. Stein. Stock price distributions with stochastic volatility. *Review of Financial Studies*, 4:727–752, 1991.

E.T. Whittaker and G.N. Watson. *Modern Analysis.* Cambridge University Press, fourth edition, 1946.