THE UNIVERSITY OF CALGARY

Numerical Methods for Modeling Energy Spot Prices

by

Bradley Dale Tifenbach

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA

September , 2000

Abstract

We develop a general numerical method for solving stochastic differential equations by constructing trinomial trees. The traditional trinomial tree is introduced and then a significant improvement is adapted to the method to create faster trees. The methods are illustrated by applying them to a class of mean reverting models suitable for modeling energy spot prices. We illustrate the models with crude oil, natural gas, and electricity data. The geometry of the trees is such that the futures prices match the expected future spot prices. This renders the probabilities risk neutral, in that investment opportunities in the energey market have no net present value. Hence, the trees can and are used to evaluate derivative securities. Numerical methods are developed to estimate the parameters in the mean reverting equations.

Acknowledgements

This thesis benefited greatly from the mathematical advice of Doctors Ali Sadeghi and Tony Ware. I would like to express my gratitude to them for sharing their expert knowledge of stochastic differential equations.

I would also like to thank my professor Doctor Ali Lari-Lavassani for his boundless enthusiasm and support for this project.

Finally, a special thank you goes to Garth Renee and Shawn Hatch (formerly) of TransAlta for providing the natural gas, crude oil, and electricity data used to illustrate the models.

This thesis evolved out of a project sponsored and partially funded by TransAlta Energy Marketing Corp and by the Networks of Centers of Excellence for the Mathematics of Information Technology and Complex Systems (MITACS).

Dedication

In fond memory of Barbara Moore, principal dancer of the Alberta Ballet Company, and a personal hero whose mental, physical, and spiritual perfection and beauty has been a constant delight and inspiration to me.

Table of Contents

Approval Page ii
Abstract iii
Acknowledgements iv
Dedication v
Table of Contents vi
List of Tables ix
List of Figures xi

CHAPTER 1 INTRODUCTION 1

CHAPTER 2 STOCHASTIC CALCULUS 5
2.1 Brownian Motion 5
2.1.1 Geometric Brownian Motion 7
2.1.2 Ito Process 9
2.2 Energy Prices 11
2.2.1 Log Normality 14
2.3 Mean and Variance 15

CHAPTER 3 ENERGY PRICE DYNAMICS 20
3.1 Mean Reversion 20
3.1.1 1-Factor Pilipovic Equation 20
3.1.2 Generalized 1-Factor Pilipovic Equation 21
3.2 Systems of Equations 23
3.2.1 2-Factor Pilipovic Equation 24
3.2.2 Generalized 2-Factor Pilipovic Equation 26

CHAPTER 4 PARAMETER ESTIMATION 29
4.1 1-Factor Model 30
4.1.1 Long Term Mean Estimator 31
4.1.2 Discrete Time Series 35
4.1.3 Gaussian Likelihood 36
4.1.4 Minimum Variance Estimators 36
4.2 2-Factor (a) Model (Two Correlated Assets) 40
4.2.1 Minimum Variance Correlation 41
4.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean) 42
4.3.1 Average Long Term Mean Estimator 44
4.3.2 Discrete Time Series 46
4.3.3 Minimum Variance Estimators 47
4.3.4 Deconvolution 49
4.3.5 Standard Deviation 53

CHAPTER 5 FUTURES PRICES 60
5.1 Futures Contracts in the Energy Market 60
5.2 Financial Pricing Theory 61
5.2.1 Risk and Return 62
5.2.2 The Risk in a Futures Position 62
5.3 Statistical Analysis 63
5.3.1 Hypothesis Testing 64
5.3.2 Futures Prices and Expected Future Spot Prices 68

CHAPTER 6 NUMERICAL SOLUTIONS 70
6.1 Trinomial Trees 71
6.1.1 Tree Structure 74
6.1.2 Branching Probabilities 77
6.1.3 Discrete Probability Distribution 80
6.1.4 Matching Futures Prices 82
6.2 Log Normal Distribution 83
6.3 Fast Trinomial Trees 88
6.4 3 Dimensional Trees 92
6.4.1 Induce the Correlation 93

CHAPTER 7 MEAN REVERTING MODELS 99
7.1 1-Factor Model 99
7.1.1 Log Transformation 101
7.1.2 Separation of Deterministic and Stochastic Parts 101
7.1.3 Preliminary Tree 105
7.1.4 Branching Probabilities 107
7.1.5 Final Tree 111
7.1.6 Matching Futures Prices 112
7.2 2-Factor (a) Model (Two Correlated Assets) 114
7.2.1 Construct Separate Trees 115
7.2.2 Induce the Correlation 117
7.2.3 Fast Trinomial Tree 118
7.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean) 118
7.3.1 Constant Standard Deviations 119
7.3.2 Decouple the Equations 119
7.3.3 Construct Separate Trees 120
7.3.4 Induce the Correlation 121
7.3.5 Matching Futures Prices 122
7.3.6 Fast Trinomial Tree 124

CHAPTER 8 ENERGY OPTIONS 127
8.1 European Option Evaluation 129
8.1.1 1-Factor Model 129
8.1.2 2-Factor (a) Model (2 Correlated Assets) 135
8.1.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean) . . . 139
8.2 American Option Evaluation 144
8.2.1 1-Factor Model 145
8.2.2 2-Factor (a) Model (2 Correlated Assets) 147
8.2.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean) . . . 151

CHAPTER 9 CONCLUSION 158

APPENDIX A. PARAMETER ESTIMATION 166

APPENDIX B. FUTURES PRICES 170

APPENDIX C. OPTION VALUES 174

List of Tables

4.1 1-Factor Model Minimum Variance Estimators (1997) 37
4.2 1-Factor Model Minimum Variance Estimators (1998) 37
4.3 1-Factor Model Minimum Variance Estimators (1999) 38
4.4 1-Factor Model Parameter Estimation Experiment 39
4.5 2-Factor (a) Model Minimum Variance Correlations (1997) 41
4.6 2-Factor (a) Model Minimum Variance Correlations (1998) 42
4.7 2-Factor (a) Model Minimum Variance Correlations (1999) 42
4.8 2-Factor (b) Model Minimum Variance Estimators (1997) 54
4.9 2-Factor (b) Model Minimum Variance Estimators (1998) 57
4.10 2-Factor (b) Model Minimum Variance Estimators (1999) 57
4.11 2-Factor (b) Model Parameter Estimation Experiment ($\alpha = 0.15$) 58
4.12 2-Factor (b) Model Parameter Estimation Experiment ($\alpha = 0.2$) 59
5.1 Null Hypothesis t-Test 66
8.1 1-F European Call Natural Gas (January 1, 1997) 131
8.2 1-F European Call Crude Oil (May 1, 1998) 132
8.3 1-F European Call Electricity (September 1, 1999) 134
8.4 2-F (a) European Spread Oil-Electricity (January 1, 1999) 136
8.5 2-F (a) European Spread Electricity-Gas (May 1, 1997) 137
8.6 2-F (a) European Spread Gas-Oil (September 1, 1998) 139
8.7 2-F (b) European Call Natural Gas (January 1, 1997) 141
8.8 2-F (b) European Call Crude Oil (May 1, 1999) 143
8.9 2-F (b) European Call Electricity (September 1, 1997) 144
8.10 1-F American Call Natural Gas (January 1, 1997) 146
8.11 1-F American Call Crude Oil (May 1, 1998) 147
8.12 1-F American Call Electricity (September 1, 1999) 149
8.13 2-F (a) American Call Oil-Electricity (January 1, 1999) 151
8.14 2-F (a) American Call Electricity-Gas (May 1, 1997) 153
8.15 2-F (a) American Call Gas-Oil (September 1, 1998) 154
8.16 2-F (b) American Call Natural Gas (January 1, 1997) 156
8.17 2-F (b) American Call Crude Oil (May 1, 1999) 156
8.18 2-F (b) American Call Electricity (September 1, 1997) 157
9.1 Ordinary and Fast Tree Option Price Accuracy 160
9.2 Ordinary and Fast Tree CPU Time Ratios 161

List of Figures

2.1 Geometric Brownian Motion 8
2.2 Log Normal Distribution 12
3.1 Expected Spot Price 23
3.2 1-Factor Mean Reversion 24
3.3 2-Factor (b) Mean Reversion 27
4.1 Expected Spot Price Experiment 32
4.2 Long Term Mean Experiment 34
4.3 Expected Long Term Mean Experiment 45
4.4 Natural Gas 1997 Parameter Estimation 55
4.5 Natural Gas 1998 Parameter Estimation 55
4.6 Natural Gas 1999 Parameter Estimation 56
5.1 January 1, 1997 Natural Gas Futures 67
5.2 May 1, 1999 Crude Oil Futures 67
5.3 May 1, 1999 Electricity Futures 68
6.1 Ordinary Trinomial Tree Structure 75
6.2 Trinomial Tree Node Branching Process 77
6.3 Log Normal Fast Tree Construction 87
6.4 Preliminary Trinomial Tree Structure 91
7.1 Mean Reverting Fast Tree Construction 106
7.2 Alternative Branching Processes 108
8.1 Option Evaluation Procedure 128
8.2 1-Factor Model European Call Option Errors 133
8.3 1-Factor Model European Call Option Run Times 133
8.4 Crude Oil 1998 1-Factor Model European Call Options 135
8.5 2-Factor (a) Model European Call Errors 138
8.6 2-Factor (a) Model European Call Option Run Times 138
8.7 Electricity-Gas 1997 2-Factor (a) European Call Options 140
8.8 2-Factor (b) Model European Call Errors 142
8.9 2-Factor (b) Model European Call Optoin Run Times 142
8.10 Crude Oil 1998 2-Factor (b) European Call Options 145
8.11 1-Factor Model American Call Option Errors 148
8.12 1-Factor Model American Call Option Run Times 148
8.13 Crude Oil 1998 1-Factor Model American Call Options 150
8.14 2-Factor (a) Model American Call Errors 150
8.15 2-Factor (a) American Call Option Run Times 152
8.16 Electricity-Gas 1997 2-Factor (a) American Call Options 152

8.17 2-Factor (b) Model American Call Errors 155
8.18 2-Factor (b) Model American Call Option Run Times 155
8.19 Crude Oil 1998 2-Factor (b) American Call Options 157
A..1 Crude Oil 1997 Parameter Estimation 167
A..2 Crude Oil 1998 Parameter Estimation 167
A..3 Crude Oil 1999 Parameter Estimation 168
A..4 Electricity 1997 Parameter Estimation 168
A..5 Electricity 1998 Parameter Estimation 169
A..6 Electricity 1999 Parameter Estimation 169
B..1 September 1, 1998 Natural Gas Futures 171
B..2 May 1, 1999 Natural Gas Futures 171
B..3 March 1, 1998 Crude Oil Futures 172
B..4 July 1, 1998 Crude Oil Futures 172
B..5 November 1, 1997 Electricity Futures 173
B..6 July 1, 1998 Electricity Futures 173
C..1 Natural Gas 1997 1-Factor Model European Call Options 175
C..2 Natural Gas 1-Factor Model American Call Options (1997) 175
C..3 Electricity 1999 1-Factor Model European Call Options 176
C..4 Electricity 1999 1-Factor Model American Call Options 176
C..5 Oil-Electricity 1999 2-Factor (a) Model European Call Options 177
C..6 Oil-Electricity 1999 2-Factor (a) American Call Options 177
C..7 Gas-Oil 1998 2-Factor (a) Model European Call Options 178
C..8 Gas-Oil 1998 2-Factor (a) American Call Options 178
C..9 Natural Gas 1997 2-Factor (b) European Call Options 179
C..10Natural Gas 1997 2-Factor (b) American Call Options 180
C..11Electricity 1999 2-Factor (b) European Call Options 181
C..12Electricity 1999 2-Factor (b) American Call Options 181

CHAPTER 1
INTRODUCTION

This thesis offers a complete industrial application in financial modeling. The model is designed to capture the seasonal nature of the spot prices of commodities. For the purpose of illustration, it is specifically applied to the energies natural gas, crude oil, and electricity. The seasonal shape of the spot price fluctuations is obtained by matching the expected future spot prices with the futures prices quoted in the energy market. Thus the model is risk neutral and we can and do use it to evaluate derivative securities such as American style options.

The idea of obtaining a risk neutral model by matching the seasonal shape was inspired by Hull & White in the context of modeling the term structure of interest rates. The objective of this thesis is to generalize their methods to the exciting arena of energy trading. We take the futures curve as the energy equivalent of the term structure of interest rates in order to create risk neutral probabilities. Furthermore, we use a variation of the mean reverting model developed by Pilipovic for the express purpose of modeling energy. The result is a coherent and efficient means of modeling spot prices and evaluating options in the energy industry.

We begin this thesis with a review of some basic elements of stochastic calculus in Chapter (2). We give the mathematical definition of a Brownian motion and use this natural phenomena to create stochastic differential equations. We then see how random processes like energy spot prices can be described by stochastic differential equations.

Next we introduce the concept of time varying mean reversion in Chapter (3) and explain why it is a suitable model for capturing the seasonal shape of energy prices. In essence, we hypothesize that the spot prices are influenced by a hidden economic factor called the long term mean. We introduce the mean reverting equations proposed by Pilipovic and extend them by choosing one of the parameters to be an unknown function of time. This allows us to dictate the shape of the expectation of future values to come. In this thesis we consider the following versions of the Pilipovic equation:

(1) 1-Factor Model

(2) 2-Factor (a) Model (Two Correlated Assets)

(3) 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

In Chapter (4) we consider parameter estimation. Only one of the parameters in the Pilipovic equation is selected to be the unknown function of time that is calculated so that the futures prices are matched. The other parameters must be estimated using historical data. This involves the maximum likelihood estimation of unobservable deterministic variables. Also, we develop a robust and efficient deconvolution technique for recovering hidden random variables using limited information.

We formally define the mechanics of futures contracts in Chapter (5). We investigate a simple portfolio to establish an important relationship between futures prices and expected future spot prices.

In Chapter (6) we extend the work of Hull & White, who developed a trinomial tree implementation together with a linear mean reverting model appropriate for inter-

est rates. We extend and modify their procedures so that they may be applied to the area of energy modeling. In fact it is fair to say that this thesis attempts to do for the energy industry what Hull & White accomplished for interest rate derivative securities. The following list briefly describes the main innovations offered here:

(1) Following a suggestion made in [Hull & White 1990], we develop a drift adapted method of arranging the geometry of trinomial trees that speeds convergence by minimizing errors near the median nodes of all the branches in a tree.

(2) We introduce the notion of creating a forward looking model by matching the futures price curve available in the energy market. This is the energy equivalent of the method of matching the initial interest rate term structure discovered in [Hull & White 1993].

(3) We extend the method of fast trees introduced in [Hull & White 1994 (a)] by generalizing the procedure to admit random processes with nonlinear drifts. This innovation decomposes a trinomial tree into a preliminary tree that is transformed into a final tree by arranging the geometry of each branch so that the futures prices are precisely matched. This has the pleasing effect of separating the deterministic and stochastic parts of the random process. Thus each component can be dealt with one at a time instead of tackling the whole problem all at once. The fast trees are infinitely easier to understand and implement. Furthermore, they run faster (hence our name for them) and tend to converge more rapidly than ordinary trinomial trees.

(4) The nonlinear Pilipovic 2-factor models are decoupled by making sub-

stitutions similar to the ones found in [Hull & White 1994 (b)]. This completes our nonlinear analysis of all the linear methods developed by Hull & White.

We illustrate the numerical methods by applying them to our mean reverting models in Chapter (7). We develop step by step algorithms of the various models under consideration for both ordinary and fast trees.

In Chapter (8) we complete our industrial application by using our models to evaluate derivative securities and present an extensive numerical comparison. After giving a brief definition of a financial option, we compare the values calculated by ordinary and fast trees. Furthermore, we analyze their run times and convergence rates. To accomplish all this, we use a broad range of New York Mercantile Exchange (NYMEX) data including natural gas, crude oil, and electricity over the period 1997-1999. The results are exciting and informative, not to mention the fact that they support everything we have said regarding the elegance and efficiency of fast trinomial trees. All computer programs were written in Matlab and executed on Pentium III personal computers and Sun work stations.

CHAPTER 2
STOCHASTIC CALCULUS

In this chapter we will review some of the elementary concepts from stochastic calculus. Our ultimate objective is to model the unpredictability of energy prices as stochastic differential equations.

A Brownian motion is a particularly simple model that was originally applied to the motion of a particle undergoing random molecular collisions. For example, a mote of dust floating in water is thought to capture the essence of a Brownian motion. What is this essence? We imagine a random process that continuously exhibits unpredictable fluctuations in all possible magnitudes and directions. The idea was proposed by Brown and later analyzed by Einstein.

2.1 Brownian Motion

Let (Ω, A, P) be a probability space and consider the time interval $[0, T]$.

Definition 2.1 *A **random (stochastic) process** is a function $x : \Omega \times [0, T]$ so that $x(\cdot, t)$ is measurable for all $t \in [0, T]$ (where the set of real numbers R is equipped with its Borel σ-algebra).*

We will usually write $x(t)$ or x_t instead of $x(\cdot, t)$ for the sake of brevity.

Definition 2.2 *A **Brownian motion** is a real valued random process z satisfying:*

(1) $\Delta z = \epsilon\sqrt{\Delta t}$

where $\Delta z = z(u) - z(t)$ *and* $\Delta t = u - t$ *for some* $0 \leq t < u \leq T$ *and* ϵ *is a random drawing from a standard normal random variable* $\phi(0, 1)$.

(2) The values of Δz *for any different small time intervals* Δt *are independent from one another (this implies that* z *follows a Markov process*[1]*).*

(3) $z(t)$ *is a continuous function of time* t *almost surely (a.s.).*

(4) $z(0) = 0$ *a.s.*

Recall that any linear combination of normal random variables is another normal random variable. Hence $z(t) - z(0)$ is normally distributed with mean zero and variance t for all $t \in [0, T]$.

The limit as $\Delta t \to 0$ describes the continuous motion and we write

$$dz = \epsilon\sqrt{dt} \tag{2.1}$$

We are now in a position to examine the most basic instance of a stochastic differential equation.

[1] A Markov process is one that has no memory. That is, the outcome realized in the next moment of time depends only on the present and not on the past.

2.1.1 Geometric Brownian Motion

Definition 2.3 *A random process x that follows a* ***geometric Brownian motion*** *satisfies the equation*

$$dx = a\ dt + b\ dz \tag{2.2}$$

where z is a Brownian motion and a and b are constant parameters called the ***drift*** *and* ***standard deviation*** *of the random process x.*

Then $x(t) - x(0)$ is normally distributed with mean at and variance b^2t.

Observe that a stochastic differential equation is made up of 2 parts:

(1) A deterministic part $a\ dt$ that is representable by an ordinary differentiable equation.

(2) A stochastic part $b\ dz$ that involves an elementary Brownian motion.

When $b = 0$ the variable x is no longer random and its evolution is governed by the ordinary differential equation $x'(t) = a$. Alternatively, when $a = 0$ the random process is commonly referred to as a (simple) *diffusion*. Figure (2.1) shows a typical realization of a geometric Brownian motion.

An obvious generalization of a geometric Brownian motion is one where the parameters a and b are not constrained to be constants. The following excerpt is taken from Definition (1) of Section (3.2) in [Kloeden & Platen 1999].

Suppose we have a probability space (Ω, A, P) and a Brownian motion $z(t)$. We imagine there is an increasing family $\{A_t : t \in [0,T]\}$ of sub σ-algebras of A so that

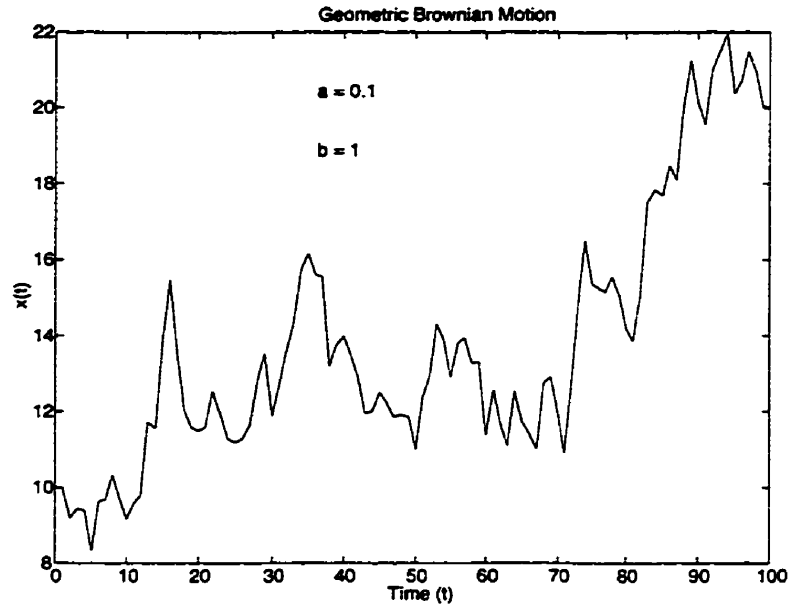

Figure 2.1: Geometric Brownian Motion

$z(t)$ is A_t measurable where A_t may be thought of as a collection of events that are detectable prior to or at time t. Let $x(t)$ be a random variable that is A_t-measurable. In other words, the random process $x(t)$ is completely determined by the realizations of $z(t)$ for all $t \in [0,T]$.

Definition 2.4 *Define L_T^2 to be the set of jointly $(A \times L)$-measurable functions b : $\Omega \times [0,T] \rightarrow R$ satisfying*

$$\int_0^T E[b^2(\cdot,t)]dt < \infty$$

$$E[b^2(\cdot,t)] < \infty \text{ for all } t \in [0,T]$$

and $b(\cdot,t)$ is A_t-measurable for all $t \in [0,T]$ where L is understood to be the set of Lebesgue subsets of the interval $[0,T]$.

In fact, under the Euclidean norm

$$\|b\| = \sqrt{\int_0^T \mathrm{E}[b^2(\cdot,t)]dt}$$

L_T^2 is a Hilbert space (so long as we identify functions that are equal everywhere except possibly on a set of measure zero).

2.1.2 Ito Process

Definition 2.5 *A variable x that follows an **Ito process** satisfies an equation of the form*

$$dx = a(x,t)\,dt + b(x,t)\,dz \tag{2.3}$$

*where z is a Brownian motion and $\sqrt{|a(\cdot,t)|}$ and $b(\cdot,t)$ are functions in L_T^2 that satisfy the usual regularity conditions. For example, see [Kloeden & Platen 1999]. Naturally a and b are called the **drift** and **standard deviation functions** of the random process x.*

Then $x(t) - x(0)$ need not be normally distributed. However, we will prove later that $x(t) - x(0)$ is well approximated by a normal distribution with mean $a(x_0,0)\,t$ and variance $b^2(x_0,0)\,t$ for small t.

It is also common to write Equation (2.3) in integral form

$$x(T) = x_0 + \int_0^T a(x,t)\,dt + \int_0^T b(x,t)\,dz \tag{2.4}$$

where the integral on the right involving the Brownian motion $z(t)$ is referred to as an Ito integral and we write

$$I(b) = \int_0^T b(x,t)\,dz \tag{2.5}$$

Before proceeding further, we write down a theorem that establishes a couple of useful properties of Ito integrals.

Theorem 2.6 *For any $b \in L_T^2$, the Ito integral (2.5) satisfies*

(1) $E[I(b)] = 0$

(2) $E[I^2(b)] = \|b\|^2$

Proof. This is Theorem 3.2.3 of [Kloeden & Platen 1999]. ■

The realization of a Brownian motion is everywhere continuous but nowhere differentiable. However, stochastic differential equations can still have solutions. As it happens, the solutions are themselves composed of Brownian motions. Of course - the same as with ordinary differential equations - an analytic solution does not always necessarily exist. Indeed this is the case with the equations we will study in this thesis.

When an analytic solution is impossible to obtain we turn to numerical methods. One such technique involves the approximation of the probability distribution (of the solution of a stochastic differential equation) by a trinomial tree. We will construct trinomial trees later when we have developed more tools.

Let us now look at some of the ways that Ito processes can be used to model the uncertain evolution of energy prices.

2.2 Energy Prices

The most basic model for energy prices is the so-called log normal distribution $l(t)$. It assumes that the natural logarithm of the price process $l(t)$ is normally distributed. We shall denote a log normal random variable by $l(t)$ for reasons that will become clear later in this thesis.

Definition 2.7 *The stochastic differential equation describing a log normal random variable l takes the form*

$$dl = \beta\ l\ dt + \tau\ l\ dy \tag{2.6}$$

or equivalently

$$\frac{dl}{l} = \beta\ dt + \tau\ dy$$

The number τ is known as the *volatility* of the price process. Figure (2.2) shows a typical log normal distribution.

Observe that this is an Ito process as the drift and standard deviation are both proportional to the energy price $l(t)$. An obvious question now poses itself to us: How can we prove that the log of $l(t)$ is normal? We shall require the use of a fundamental result known as the Ito formula.

Theorem 2.8 (Ito Formula) *Let $g(x,t)$ be a function of a variable x (that may also depend explicitly on time t) that satisfies the Ito process*

$$dx = a(x,t)\,dt + b(x,t)\,dz$$

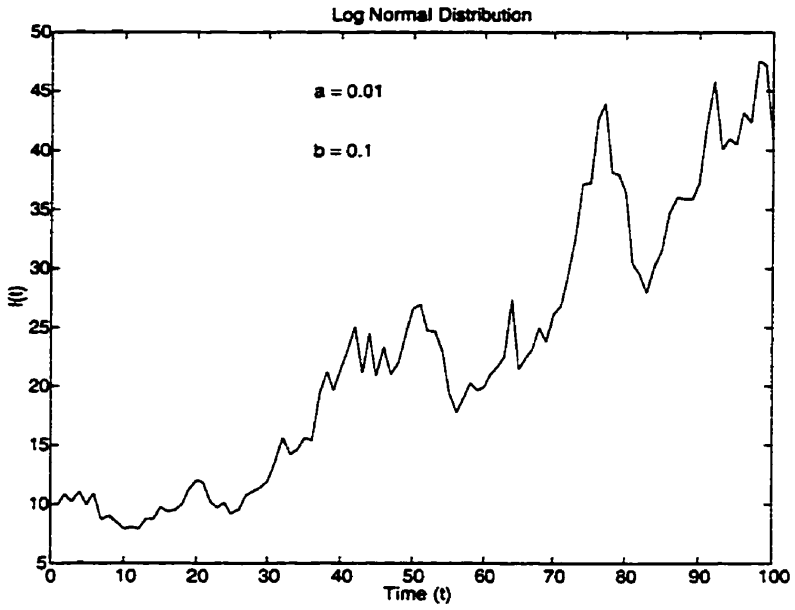

Figure 2.2: Log Normal Distribution

Then g is another random variable that satisfies the Ito process

$$dg = [\dot{g}(x,t) + a(x,t)\, g'(x,t) + \frac{b^2(x,t)}{2} g''(x,t)]dt + b(x,t)\, g'(x,t)\, dz$$

where the prime and dot symbols (respectively) indicate differentiation with respect to x and t.

Proof. This is the differential form of Theorem 3.3.2 (Ito Formula) of [Kloeden & Platen 1999].

■

This result allows us to transform relatively complex stochastic differential equations into ones that are simpler and easier to use. Eventually we will also require the use of a higher dimensional Ito formula. Refer to [Kloeden & Platen 1999] for a fully

generalized version of the higher dimensional Ito formula. The following relatively simple specific case due to Wilmott - see Equation (4.9) in [Wilmott 1998] - is sufficient for our needs in this work.

Theorem 2.9 (Higher Dimensional Ito Formula) *Let $g(x,v,t)$ be a function of a pair of correlated Ito processes x and v (that may also depend explicitly on time t) that satisfy the system of stochastic differential equations*

$$\begin{aligned} dx &= a(x,v,t)\,dt + b(x,v,t)\,dz \\ dv &= c(x,v,t)\,dt + d(x,v,t)\,dy \end{aligned}$$

Denote the correlation between the Brownian motions z and y by ρ. Then g is another random variable that satisfies the Ito process

$$dg = [\dot{g} + \frac{1}{2}b^2 g_{xx} + \rho b d g_{xv} + \frac{1}{2}d^2 g_{vv}]dt + g_x dx + g_v dv$$

where the subscripts and dot symbol (respectively) indicate differentiation with respect to x, v, and t and we have neglected to write the dependence of the standard deviation functions b and d, and all derivatives of g on x, v, and t for the sake of clarity.

Observe that in the higher dimensional case, the random function $g(x,v,t)$ evolves according to the variations of both Brownian motions z and y.

2.2.1 Log Normality

Suppose we make the substitution $L(l) = \ln(l)$. The required derivatives are $L'(l) = \frac{1}{l}$ and $L''(l) = -\frac{1}{l^2}$. In Equation (2.6) the drift and standard deviation are $a = \beta l$ and $b = \tau l$. Therefore by the Ito formula we have

$$dL = (\beta - \frac{\tau^2}{2})dt + \tau\, dy$$

This reveals that the log of the price process $l(t)$ is in fact a geometric Brownian motion. Hence $L(T) - L(0)$ is normally distributed with mean $(\beta - \frac{\tau^2}{2})T$ and variance $\tau^2 T$.

It is now desirable to use our exact knowledge of the simplified process $L(t)$ to predict the behavior of the energy price $l(t)$. In other words

$$\ln[l(t)] \sim \phi[\ln(l_0) + (\beta - \frac{\tau^2}{2})t, \tau\sqrt{t}]$$

where $l_0 = l(0)$.

We can and will show that the mean and variance of $l(t)$ are $\mathrm{E}[l(t)] = l_0 \mathrm{e}^{\beta t}$ and $\mathrm{var}[l(t)] = l_0^2 \mathrm{e}^{2\beta t}(\mathrm{e}^{\tau^2 t} - 1)$.

This means that the expected value of a log normal random variable $l(t)$ exhibits exponential growth (provided β is positive). This is qualitatively correct behavior for a stock price. However, in reality energy prices do not follow a log normal distribution. They are strongly affected by the temperature and in particular the changing of the seasons. A far more realistic equation for energy prices is the so-called mean reverting model.

2.3 Mean and Variance

The expected value of a random process can be obtained by its associated stochastic differential equation in a particularly elegant way. As it happens the expected value of an Ito process is the solution of the ordinary differential equation obtained by ignoring the stochastic part.

Definition 2.10 *Consider a random process $s(t)$ evolving over the time interval $[0, T]$. We shall adopt the notation*

$$\overline{s}(t) = E[s(t) \mid s(0)]$$

Similarly, when we are working with a short time interval of the form $[t, t + \Delta t]$, we will write

$$\overline{s}(u) = E[s(u) \mid s(t)]$$

For the remainder of this thesis, when we mention the expectation of a random process, we really are referring to the expectation conditional on the value of the random variable at the left endpoint of the time interval under study. This should be obvious from the individual contexts in which expectations arise.

Theorem 2.11 *Suppose a variable s satisfies an Ito process*

$$ds = a(s,t)\,dt + b(s,t)\,dz$$

Put $\varphi(t) = \overline{s}(t)$ *for the expected value of* $s(t)$. *Then* $\varphi(t)$ *is the solution of the ordinary integral equation*

$$\varphi(t) = s_0 + \int_0^t E\{a[s(u), u]\}du \tag{2.7}$$

where $s_0 = s(0)$.

Furthermore, if $a(s,t)$ *has the functional form* $a(s,t) = f(t)s + g(t)$ *where* $f(t)$ *and* $g(t)$ *are deterministic functions of time* t, *then* $\varphi(t)$ *satisfies the ordinary differential equation*

$$\dot{\varphi}(t) = a[\varphi(t), t] \tag{2.8}$$

Proof. The integral equation for $s(t)$ is

$$s(t) = s_0 + \int_0^t a[s(u), u]\, du + \int_0^t b[s(u), u]\, dz$$

By Theorem (2.6) the expectation of the Ito integral vanishes

$$\mathrm{E}\{\int_0^t b[s(u), u]\, dz\} = 0$$

We can use this result and take expectations across the stochastic integral equation (2.4) to get

$$\begin{aligned} \varphi(t) &= s_0 + \mathrm{E}\{\int_0^t a[s(u), u]\, du\} \\ &= s_0 + \int_0^t \mathrm{E}\{a[s(u), u]\}du \end{aligned}$$

Now if $a(s,t)$ is of the form $a(s,t) = f(t)s + g(t)$, then

$$\begin{aligned} \mathrm{E}\{a[s(u),u]\} &= \mathrm{E}[f(u)s(u) + g(u)] \\ &= f(u)\mathrm{E}[s(u)] + g(u) \\ &= a[\varphi(u),u] \end{aligned}$$

Hence

$$\varphi(t) = s_0 + \int_0^t a[\varphi(u),u]du$$

By the Fundamental Theorem of Calculus we obtain an ordinary differential equation for the expected value

$$\dot{\varphi}(t) = a[\varphi(t),t]$$

■

Let us verify this result by calculating the mean and variance of a log normal random variable directly. Suppose that in the stochastic differential equation for $l(t)$ the drift and standard deviation are given by $a(l,t) = \beta l$ and $b(l,t) = \tau l$.

Put $\psi(t) = \bar{l}(t)$. The solution of equation (2.8) reveals the expected value

$$\psi(t) = l_0 e^{\beta t}$$

Now consider the parabola l^2. By the Ito formula we obtain

$$d(l^2) = (2\beta + \tau^2) l^2 dt + 2\tau l^2 dy$$

By Theorem (2.11) we similarly obtain an ordinary differential equation for the expected value $\mathrm{E}(l^2)$

$$\frac{d\mathrm{E}[l^2(t)]}{dt} = \left(2\beta + \tau^2\right) \mathrm{E}[l^2(t)]$$

The solution of this equation reveals the expectation of the square

$$\mathrm{E}[l^2(t)] = l_0^2 \mathrm{e}^{\left(2\beta+\tau^2\right)t}$$

Finally the variance of $s\,(t)$ is given by the formula

$$\begin{aligned} \mathrm{var}[l\,(t)] &= \mathrm{E}[l^2(t)] - \psi^2\,(t) \\ &= l_0^2 \mathrm{e}^{2\beta t}(\mathrm{e}^{\tau^2 t} - 1) \end{aligned}$$

This completes our analysis of the log normal distribution.

Theorem 2.12 *Let $X(t)$ be an Ito process (2.3)*

$$dX = a\,(X,t)\,dt + b\,(X,t)\,dz$$

Then $X(t+\Delta t) - X(t)$ is approximated by a normal random variable with mean M and variance V given by

$$\begin{aligned} M &= a[X(t),t]\Delta t + O(\Delta t^2) \\ V &= b^2[X(t),t]\ \Delta t + O(\Delta t^2) \end{aligned}$$

Proof. Put $x\,(u) = \overline{X}\,(u)$ for all $u \in [t, t+\Delta t]$. Then $x(t) = X(t)$ and by Theorem (2.11) the expected value of $X(t+\Delta t)$ is

$$x(t+\Delta t) = x(t) + \int_t^{t+\Delta t} \mathrm{E}\{a[X(u),u]\}du \tag{2.9}$$

The expectation of a continuous random variable is a continuous function. Furthermore, it is obvious that $\mathrm{E}\{a[X(t),t]\} = a[X(t),t]$. By the Fundamental Theorem of Calculus we obtain

$$\begin{aligned} x(t+\Delta t) &= x(t) + E\{a[X(t),t]\}\Delta t + \mathrm{O}(\Delta t^2) \\ &= x(t) + a[X(t),t]\Delta t + \mathrm{O}(\Delta t^2) \end{aligned}$$

Thus the mean is

$$\begin{aligned} M &= x(t+\Delta t) - x(t) \\ &= a[X(t),t]\Delta t + \mathrm{O}(\Delta t^2) \end{aligned}$$

Now by Theorem (2.6) we find the variance to be

$$\begin{aligned} V &= \mathrm{E}\{[X(t+\Delta t) - x(t+\Delta t)]^2\} \\ &= \mathrm{E}(\{\int_t^{t+\Delta t} b[X(u),u]dz\}^2) + \mathrm{O}(\Delta t^2) \\ &= \int_t^{t+\Delta t} \mathrm{E}\{b^2[X(u),u]\}du + \mathrm{O}(\Delta t^2) \end{aligned}$$

Similarly $\mathrm{E}\{b^2[X(t),t]\} = b^2[X(t),t]$ and so by the Fundamental Theorem of Calculus and our previous remarks it follows that the variance is

$$V = b^2[X(t),t]\Delta t + \mathrm{O}(\Delta t^2)$$

Finally we observe that $X(t+\Delta t) - X(t)$ is well approximated by a normal random variable $\phi(M,V)$ since the basic building block of the stochastic differential equation (2.3) is a Brownian motion z. Therefore in the limit as $\Delta t \to 0$ the distribution of $X(t+\Delta t) - X(t)$ becomes perfectly normal. ■

CHAPTER 3
ENERGY PRICE DYNAMICS

We now investigate a series of models appropriate for energy prices. The themes explored are mean reversion and seasonality. These are two of the main features rendering energy prices fundamentally different from relatively simple stock prices.

3.1 Mean Reversion

A widely used class of models for energy prices consists of processes that are mean reverting. The Ito process that describes this behavior can take the form

3.1.1 1-Factor Pilipovic Equation

$$ds = \alpha(l - s)dt + \sigma s\, dz \tag{3.1}$$

where the constant parameter α is called the *strength of mean reversion.* Observe that the noise term is proportional to s. This differentiates the Pilipovic equation from the Hull & White equation, which has what we call additive noise[1] and is appropriate for modeling interest rates. In order to work with the Pilipovic equation, we will find out later that it is necessary to transform to a new random variable that does have additive noise. This new variable has an exponential term in the drift. It is in this sense that

[1]The 1-Factor Hull & White equation for interest rates is $dr = [\theta(t) - ar]dt + \sigma\, dz$. Observe that the noise term is independent of the interest rate r.

the Pilipovic equation is nonlinear and consequently it is much harder to analyze than the Hull & White equation. Refer to [Pilipovic 1997] for a more complete discussion of the Pilipovic equation and why it is appropriate for modeling energy prices.

The number l is known as the long term mean and it represents the quantity that the price process $s(t)$ is trying to revert to. It is easy to show that l is a fixed point of the dynamical system $\dot{\varphi} = \alpha(l - \varphi)$. In all practical applications, the strength of mean reversion α is positive. If $s(t)$ is above l then the drift is negative, while if $s(t)$ is below l then the drift is positive. In both cases, the drift always acts to inexorably draw the energy price back toward the long term mean.

Although the long term mean can be a constant parameter this is too simple to be of any use. In practice the long $l(t)$ term mean is allowed to be an unknown function of time. Why unknown? The long term mean $l(t)$ should be allowed to capture the essence of the economic trend driving the energy market. An important factor affecting $l(t)$ is probably the temperature which is itself a function of the changing of the seasons. This highly realistic model for energy prices therefore exhibits one of the singular qualities that make energy prices unique: namely the concept of seasonality.

3.1.2 Generalized 1-Factor Pilipovic Equation

$$ds = \alpha[l(t) - s]dt + \sigma s\, dz \tag{3.2}$$

For the remainder of this thesis, Equation (3.2) will be referred to as the 1-

factor model. Instead of growing exponentially as stock prices do, energy prices move up and down with the changing of the seasons. They tend to display peaks when the temperature is at an extreme value. For example energy prices tend to peak in Summer when it is hot and in Winter when it is cold. The long term mean is specifically designed to capture this economic trend. Then the mean reverting model will successfully mimic the seasonality clearly inherent in the prices of energy commodities like oil, natural gas, and electricity.

Put $\varphi(t) = \overline{s}(t)$. By Theorem (2.11) the expected value of a mean reverting random variable $s(t)$ satisfies the ordinary differential equation

$$\dot{\varphi}(t) = \alpha[l(t) - \varphi(t)]$$

This is a linear equation and is easily solved to yield

$$\varphi(t) = \alpha e^{-\alpha t} \int_0^t e^{\alpha u} l(u)\, du$$

Let us consider a suitable example for the sake of illustration. Put $l(t) = a + b\sin(\frac{2\pi t}{P})$ where the period $P = 1$ year. This is perhaps the most basic seasonal function. Then the above solution becomes

$$\varphi(t) = a + B[\alpha \sin(\frac{2\pi t}{P}) - \frac{2\pi}{P}\cos(\frac{2\pi t}{P})] + (s_0 - a + \frac{2\pi B}{P})e^{-\alpha t} \tag{3.3}$$

where $B = \frac{b\alpha}{\alpha^2 + \frac{4\pi^2}{P^2}}$ and $s_0 = s(0)$.

Figure (3.1) depicts the situation for a typical value of the strength of mean reversion α.

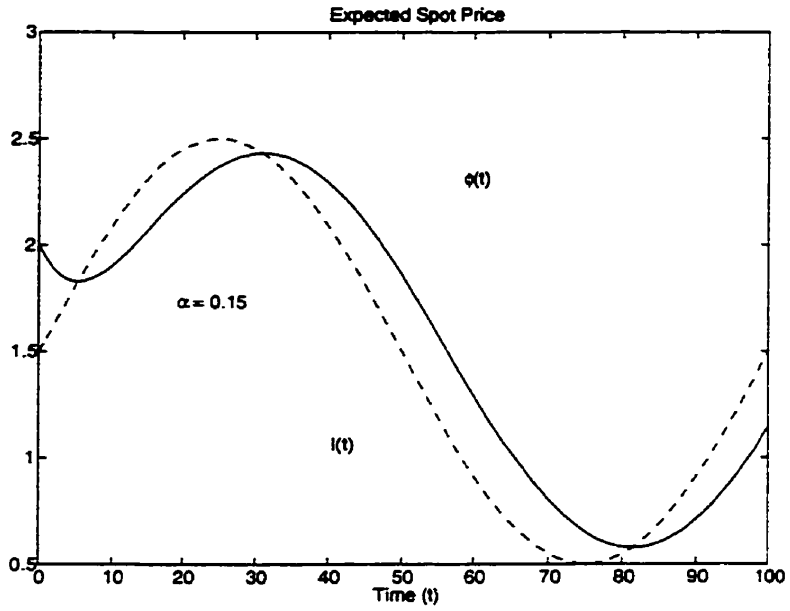

Figure 3.1: Expected Spot Price

This gives a nice image of how the mean reverting model can also capture the natural behavior of energy prices. Observe that the expected spot price $\varphi(t)$ follows along behind the long term mean $l(t)$: this is the nature of mean reversion. Now look at Figure (3.2) to see a sample path satisfying Equation (3.2) when the long term mean is $l(t) = a + b\sin(\frac{2\pi t}{P})$.

3.2 Systems of Equations

We now move on to study systems of stochastic differential equations. In this thesis we consider systems involving only 2 stochastic variables.

A system of stochastic differential equations is obtained by allowing the long

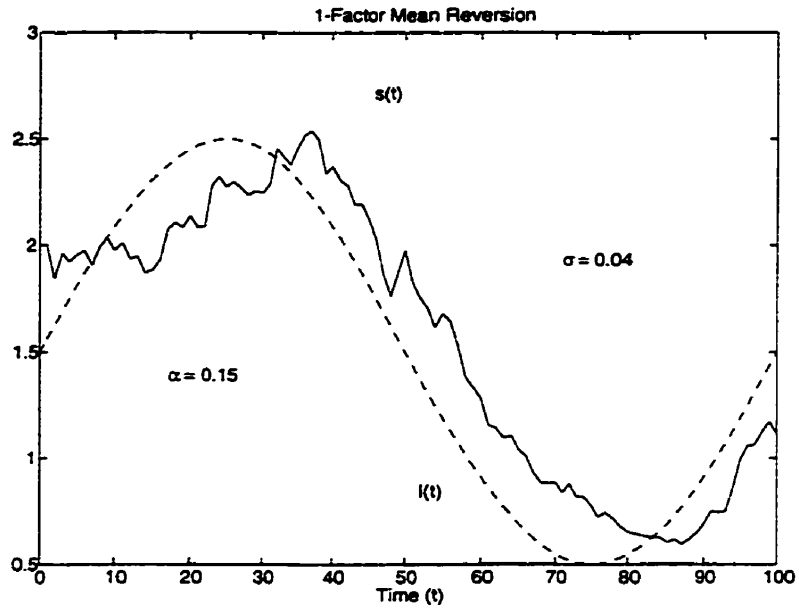

Figure 3.2: 1-Factor Mean Reversion

term mean $l(t)$ to be a log normal random variable instead of an unknown function of time.

3.2.1 2-Factor Pilipovic Equation

$$\begin{aligned} ds &= \alpha(l-s)dt + \sigma s\, dz \\ dl &= \beta l\, dt + \tau l\, dy \end{aligned} \tag{3.4}$$

For the sake of simplicity we will assume that the random components of $s(t)$ and $l(t)$ are uncorrelated[2]. The number β can be a constant parameter. In this

[2]To be more precise the Brownian motions $z(t)$ and $y(t)$ are uncorrelated (remember that the Brownian motions are the sources of unpredictability in a stochastic differential equation).

case $l(t)$ is a standard log normal random variable and the expected value $\psi(t) = \overline{l}(t)$ exhibits exponential growth. If we also put $\varphi(t) = \overline{s}(t)$ then it is derived in [Lari-Lavassani, Simchi & Ware 2000] that the trajectory of these expected values admit a decomposition in terms of the stable and unstable manifolds of the system

$$\begin{pmatrix} \varphi(t) \\ \psi(t) \end{pmatrix} = (s_0 - \frac{\alpha}{\alpha+\beta} l_0) e^{-\alpha t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + l_0 e^{\beta t} \begin{pmatrix} \frac{\alpha}{\alpha+\beta} \\ 1 \end{pmatrix}$$

where $s_0 = s(0)$ and $l_0 = l(0)$. Note that as $t \to \infty$, $\varphi(t) \to \frac{\alpha}{\alpha+\beta} l_0 e^{\beta t}$, or equivalently, $\varphi(t) \to \frac{\alpha}{\alpha+\beta}\psi(t)$. In other words, as $t \to \infty$, the expectation $\varphi(t)$ tends to a multiple of the expected value $\psi(t)$ of the long term mean, whence the term 'mean reversion'.

Although this is a fine model, it is too simple to capture the seasonality of an energy price. Indeed we would like the long term mean to move up and down as the seasons change - just as energy prices do. This can be done by allowing the parameter $\beta(t)$ to be an unknown function of time.

Put $\varphi(t) = \overline{s}(t)$ and $\psi(t) = \overline{l}(t)$. By Theorem (2.11) the expected value of the long term mean $l(t)$ is the solution of the ordinary differential equation

$$\dot{\psi}(t) = \beta(t)\psi(t) \tag{3.5}$$

Thus the function $\beta(t)$ can be chosen so that the long term mean behaves precisely the way we want - and that is exactly what we do later in this thesis!

3.2.2 Generalized 2-Factor Pilipovic Equation

$$\begin{aligned} ds &= \alpha(l-s)dt + \sigma s\, dz \\ dl &= \beta(t)\, l\, dt + \tau l\, dy \end{aligned} \tag{3.6}$$

For the remainder of this thesis, Equation (3.6) will be referred to as the 2-factor (b) model. If a mean reverting random variable $s(t)$ and its (stochastic) long term mean $l(t)$ satisfy Equation (3.6), then by Theorem (2.11) their expected values satisfy the ordinary system

$$\begin{aligned} \dot{\varphi}(t) &= \alpha[\psi(t) - \varphi(t)] \\ \dot{\psi}(t) &= \beta(t)\,\psi(t) \end{aligned} \tag{3.7}$$

The solution is

$$\begin{aligned} \varphi(t) &= s_0 + \alpha e^{-\alpha t}\int_0^t e^{\alpha r}\psi(r)\, dr \\ \psi(t) &= l_0 e^{\int_0^t \beta(r)dr} \end{aligned}$$

One can exogenously introduce seasonality by imposing an appropriate functional form for $\beta(t)$. For example, this could be

$$\beta(t) = \frac{2\pi}{P}\frac{b\cos(\frac{2\pi t}{P})}{a + b\sin(\frac{2\pi t}{P})} \tag{3.8}$$

In this case, the expectation of the long term mean is $\psi(t) = a + b\sin(\frac{2\pi t}{P})$. Remember that this is the simple shape we used in Subsection (3.1.2) to illustrate the

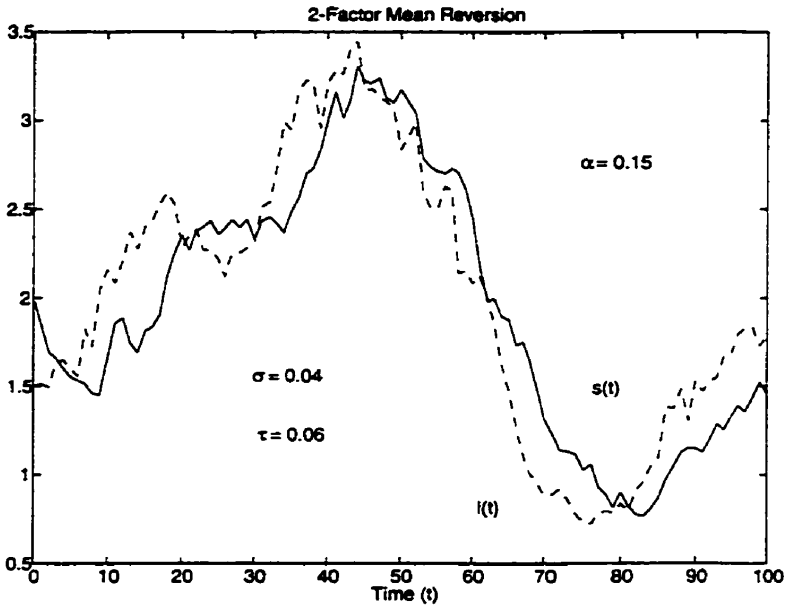

Figure 3.3: 2-Factor (b) Mean Reversion

concept of seasonality. Then the expected value of the energy price $s(t)$ is given by Equation (3.3) and the relationship between $\varphi(t)$ and $\psi(t)$ looks the same as Figure (3.1) with $l(t)$ replaced by $\psi(t)$. What is the difference? Remember we are considering expected values here: the long term mean $l(t)$ is actually a stochastic variable in Equation (3.6). This does indicate that the 2-factor (b) model has a richer structure than the 1-factor model. Figure (3.3) shows a sample path satisfying Equation (3.6) when the drift parameter is given by Equation (3.8)

We shall also consider systems of 2 correlated mean reverting energy prices later - this is the previously mentioned 2-factor (a) model - but the details will not be given here.

This concludes our review of the elements of stochastic calculus. We shall now proceed to examine the mean reverting model for energy prices in greater detail, beginning with an expose on the statistical estimation of the parameters in the Pilipovic equation.

CHAPTER 4
PARAMETER ESTIMATION

An important and difficult problem that arises in the financial modeling of energy prices is the estimation of the parameters in the stochastic differential equations. If we have available historical data - and in practice we do - the parameters can be estimated by maximizing the likelihood (or minimizing the variance) of the sample paths. The proportional noise in the Pilipovic equations precludes the possibility of using a traditional Kalman filter to estimate the parameters. Instead we propose some approximate solutions whose numerical implementations yield reasonably good and stable results. The data used as input arguments for the computer programs are natural gas, crude oil, and electricity spot prices quoted on NYMEX over the period 1997-1999.

Recall that in our generalized models, we allow one of the parameters to be an unknown function of time in order to capture the seasonal shape of energy spot prices. Although this unknown function of time is estimated along with the other parameters in the Pilipovic equations it is never used in the numerical solution of a stochastic differential equation. Why? The shape of the spot price distribution will later be designed to coincide with the futures prices quoted in the energy market. In other words, the models we are developing are forward looking in that they incorporate the future beliefs of traders into the calculation of the unknown function of time. Therefore we focus on estimating the constant parameters with historical data. The techniques developed for calculating the unknown function of time will be used later, when we

have built the theoretical framework that enables futures curve matching in Chapter (5). Although it is interesting to observe what has gone before, we are really interested in what the future holds in store for us. This combination of historical calibration and forward looking curve alignment is one of the innovations of this thesis.

4.1 1-Factor Model

As a prelude to the relatively complex 2-factor (a) and (b) models we first look at the generalized 1-factor Pilipovic equation (3.2)

$$ds = \alpha\,[l(t) - s]\,dt + \sigma s\,dz$$

where

- s Spot price
- $l(t)$ Long term mean (unknown function of time)
- α Strength of mean reversion
- σ Volatility
- z Brownian motion

We assume that the discrete time series for the spot prices $s\,(t)$ is available by whatever means. This sample path $\{s_i\}$ represents the sum total of our knowledge of the energy price process. The long term mean $l(t)$ is an unobservable deterministic variable. It is assumed to capture the economic trend in the energy market. Evidently, it is necessary to estimate this hidden quantity, by some statistical means.

4.1.1 Long Term Mean Estimator

As it happens the underlying economic trend $l(t)$ can be estimated by solving an ordinary differential equation. Put $\varphi(t) = \overline{s}(t)$. The behavior of the expected value of the spot price is given by Equation (2.8), which now becomes

$$\dot{\varphi}(t) = \alpha\left[l(t) - \varphi(t)\right] \tag{4.1}$$

However, the expected value $\varphi(t)$ is not observable: we know only the spot price $s(t)$. We now make an ergodicity type of assumption: the expected value $\varphi(t)$ may be approximated by taking a moving average of the sample path $\{s_i\}$. As it happens, a moving average is a special case of a convolution.

Definition 4.13 *A **convolution function** $c(t)$ is an even, continuous, and real valued function that satisfies $c(t) \geq 0$ and*

$$\int_{-\infty}^{\infty} c(t)dt = 1$$

Definition 4.14 *A **convolution** of a curve $s(t)$ is another function $\varphi(\mathbf{t})$ defined by*

$$\varphi(t) = \int_{-\infty}^{\infty} c(r)s(t-r)dr$$

where $c(t)$ is a convolution function.

Assumption1 *The expectation $\varphi(t)$ can be (approximately) calculated by taking a convolution of the sample path $\{s_i\}$*

$$\varphi_i = \sum_{j=-m}^{m} c_j s_{i-j} \tag{4.2}$$

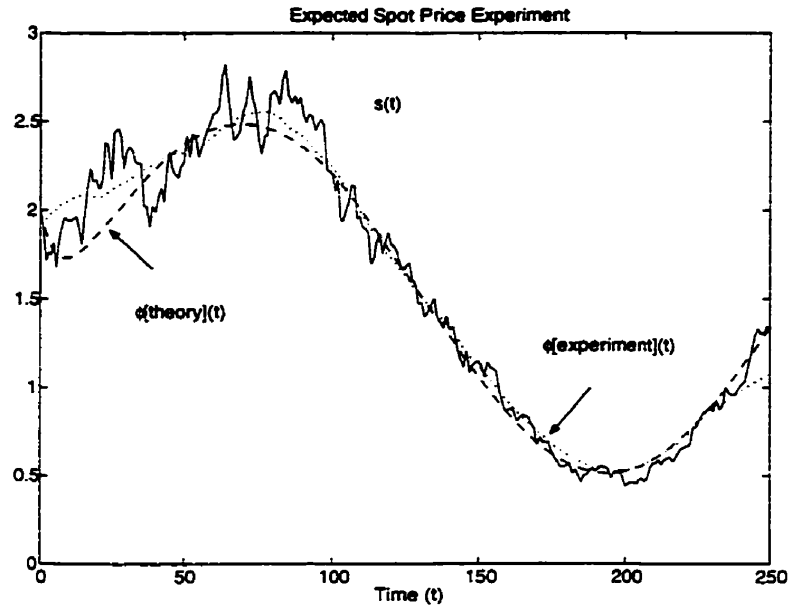

Figure 4.1: Expected Spot Price Experiment

When $c_j = \frac{1}{2m+1}$ for all m the convolution is a moving average of the data points. We find that a monthly moving average ($m = 20$ trading days) works well in practice. Figure (4.1) displays a sample path satisfying Equation (3.2) that was simulated by a random number generator[1], where the dashed line depicts the theoretical expectation (3.3) and the dotted line represents the moving average (4.2) of the data points. This picture appears to justify the use of Assumption (1), at least for a mean reverting process.

Before proceeding further, we mention in passing why a convolution of a random variable is a reliable approximator of its expected value. Recall that a random variable

[1]The long term mean used in the simulation was $l(t) = a + b\sin(\frac{2\pi t}{P})$.

is composed of a deterministic part and a stochastic part that in some sense represent the expectation with an addition of noise. A convolution has the effect of eliminating the noise while leaving the expected value behind in its wake. Obviously the technique is not perfect because information is destroyed by the smoothing process. However, the differences between the approximate expectation and the real one tend to cancel out - especially when the process moves up and down in a seasonal way.

Ideally we would like to take a large number of random samples of the probability distribution of the spot price $s(t)$. We realize that this probability distribution is a (seasonal) function of time. However, if the distribution varies slowly over time, then the probability distribution of $s(t+\Delta t)$ is approximately equal to that of $s(t)$ for small values of Δt. This is the philosophy behind the convolution technique of estimating the expected value $\varphi(t)$.

By Taylor's theorem the derivative of $\varphi(t)$ at time index i is easily calculated to a high degree of accuracy by the formula

$$\dot{\varphi}_i = \frac{8(\varphi_{i+1} - \varphi_{i-1}) - (\varphi_{i+2} - \varphi_{i-2})}{12\ \Delta t} \tag{4.3}$$

for $i = 3, ..., (n-2)$. The remaining derivatives are given by

$$\begin{aligned} \dot{\varphi}_1 &= \frac{\varphi_2 - \varphi_1}{\Delta t} \qquad \dot{\varphi}_2 = \frac{\varphi_3 - \varphi_1}{2\ \Delta t} \\ \dot{\varphi}_{n-1} &= \frac{\varphi_n - \varphi_{n-2}}{2\ \Delta t} \qquad \dot{\varphi}_n = \frac{\varphi_n - \varphi_{n-1}}{\Delta t} \end{aligned} \tag{4.4}$$

Note that we have adopted the (trading) day as our unit of time. Therefore the length of a small time interval is $\Delta t = 1$.

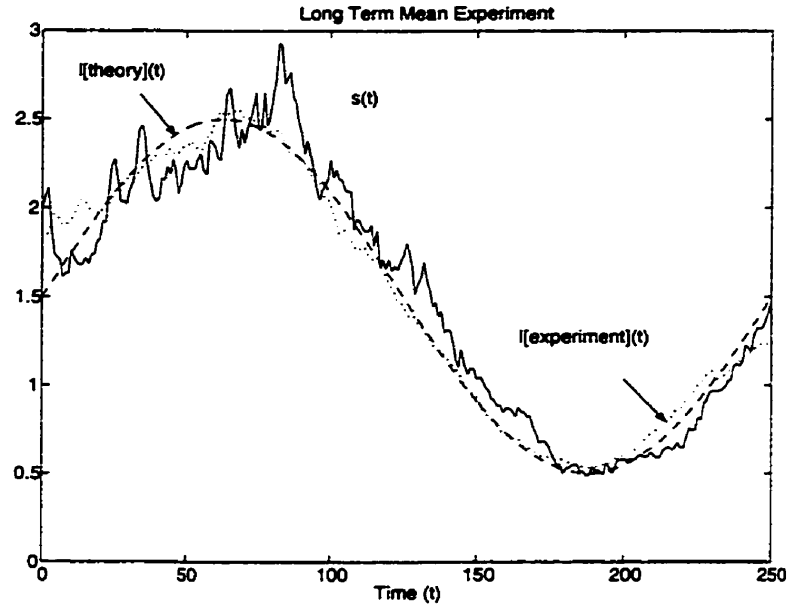

Figure 4.2: Long Term Mean Experiment

Now it is possible to solve Equation (4.1) for the hidden quantity

$$l(t) = \varphi(t) + \frac{\dot{\varphi}(t)}{\alpha} \tag{4.5}$$

Then Equation (4.5) is a reliable estimator for the unobservable deterministic variable $l(t)$. Figure (4.2) shows the long term mean (dotted line) resulting from this technique (4.5) compared to the sine curve (dashed line) that was used to generate the sample path. Again we perceive the close agreement between experiment and theory.

4.1.2 Discrete Time Series

Before conducting the time series analysis proper it is necessary to discretize the 1-factor model to get

$$s_{i+1} = s_i + \alpha(l_i - s_i)\Delta t + s_i z_i \tag{4.6}$$

for $i = 1, ..., (n-1)$.

The sequence of random variables $\{z_i\}$ is assumed to be a normal stationary time series[2] with an expected value of zero and standard deviation $\sigma\sqrt{\Delta t}$. The stochastic component of the random process can be isolated by solving (4.6) for z_i

$$z_i = \frac{s_{i+1} - s_i - \alpha(l_i - s_i)\Delta t}{s_i} \tag{4.7}$$

Now by our earlier result (4.5) this becomes

$$z_i = \frac{s_{i+1} - s_i - [\alpha(\varphi_i - s_i) + \dot{\varphi}_i]\Delta t}{s_i} \tag{4.8}$$

Recall that z_i is a normal random variable with mean zero and standard deviation $\sigma\sqrt{\Delta t}$. It is now an easy matter to compute the variance V of z_i.

$$V = \frac{1}{n-1}\sum_{i=1}^{n-1} z_i^2 \tag{4.9}$$

Observe that the z_i's and consequently V are functions only of the strength of mean reversion α.

[2]Do not confuse z, which is a Brownian motion with mean zero and standard deviation $\sqrt{\Delta t}$, with $\{z_i\}$, which is a simple diffusion with mean zero and standard deviation $\sigma\sqrt{\Delta t}$. The number σ has been absorbed into the z_i's because it is more convenient to estimate the parameters in this way.

4.1.3 Gaussian Likelihood

The numbers $\{z_i\}$ are assumed to be independent random drawings from a normal distribution $\phi\left(0, \sigma\sqrt{\Delta t}\right)$. Therefore the Gaussian likelihood of the realization $\{s_i\}$ is

$$L(\alpha, \sigma) = \frac{\exp(-\frac{1}{2\sigma}\sum_{i=1}^{n-1} z_i^2)}{(2\pi\sigma)^{\frac{n-1}{2}}} \tag{4.10}$$

Proposition 4.15 *The likelihood L is maximized when the variance V is minimized.*

Proof. Put $\sigma = W(\alpha) = \sqrt{\frac{V(\alpha)}{\Delta t}}$. Then the likelihood L can be expressed purely as a function of α

$$L(\alpha) = \frac{\exp\{-[\frac{(n-1)\Delta t}{2}]W(\alpha)\}}{[2\pi W(\alpha)]^{\frac{n-1}{2}}}$$

Observe that $W > 0$. Otherwise we would have s_i equal to some (constant) number for all $i = 1, ..., n$: we tacitly assume that this a statistical impossibility.

We maximize the likelihood by differentiating L with respect to α

$$L'(\alpha) = -(\frac{n-1}{2})W'(\alpha)[\Delta t + \frac{1}{W(\alpha)}]L(\alpha)$$

Evidently the only way this derivative can equal zero is if $W'(\alpha) = 0$ in which case we must also have $V'(\alpha) = 0$. This is precisely the condition that minimizes V.

■

4.1.4 Minimum Variance Estimators

The maximum likelihood estimators of α and σ are obtained by minimizing the variance V. This can be done by differentiating (4.9) with respect to α and setting the derivative

equal to zero

$$\widehat{\alpha} = \sum_{i=1}^{n-1} \frac{(s_{i+1} - s_i - \dot{\varphi}_i \Delta t)(\varphi_i - s_i)}{s_i^2} [\sum_{i=1}^{n-1} \frac{(\varphi_i - s_i)^2 \Delta t}{s_i^2}]^{-1} \tag{4.11}$$

Hence $V'(\alpha) = 0$ has a unique solution $\widehat{\alpha}$. The variance $V(\widehat{\alpha})$ must then be a global minimum as we can make V arbitrarily large by picking negative values of α. Once the best estimator for α has been found, the best estimator for σ is calculated by

$$\widehat{\sigma} = W(\widehat{\alpha}) \tag{4.12}$$

Tables (4.1), (4.2), and (4.3) show the minimum variance estimators $\widehat{\alpha}$ and $\widehat{\sigma}$ for natural gas, crude oil, and electricity over the period 1997-1999.

Table 4.1: 1-Factor Model Minimum Variance Estimators (1997)

	$\widehat{\alpha}$	$\widehat{\sigma}$
Natural Gas	0.232	0.0331
Crude Oil	0.141	0.0174
Electricity	0.188	0.0418

Table 4.2: 1-Factor Model Minimum Variance Estimators (1998)

	$\widehat{\alpha}$	$\widehat{\sigma}$
Natural Gas	0.151	0.0349
Crude Oil	0.170	0.0283
Electricity	0.172	0.0479

Table 4.3: 1-Factor Model Minimum Variance Estimators (1999)

	$\widehat{\alpha}$	$\widehat{\sigma}$
Natural Gas	0.158	0.0301
Crude Oil	0.146	0.0212
Electricity	0.154	0.0454

At this point there is an obvious question: how do we know that our minimum variance estimators are accurate? Furthermore, what is the precision of our measurements? We now propose a numerical test as evidence of the validity of our work. We simulate sample paths satisfying Equation (3.2) (using a random number generator) with a specific set of parameters and then attempt to recover those parameters using our method. By repeating this experiment a large number of times and taking the mean and standard deviation of the recovered parameters, we can see whether our technique reveals unbiased estimators of the true parameters. Also, twice the standard deviation is evidently a good estimator of the absolute error.

Table (4.4) shows the results of this experiment wherein $\overline{\alpha}$ and $\overline{\sigma}$ are the sample means of $\widehat{\alpha}$ and $\widehat{\sigma}$ while s_α and s_σ are their sample standard deviations. For all possible pairs of the parameter sets $\alpha = 0.15$ and $\alpha = 0.2$ and $\sigma = 0.02$, $\sigma = 0.03$, and $\sigma = 0.04$, 100 sample paths consisting of 250 data points were simulated.

We perceive that the measurements of $\widehat{\sigma}$ are excellent while those for $\widehat{\alpha}$ are (on average) consistently too high by an amount approximately equal to 0.025. Also,

Table 4.4: 1-Factor Model Parameter Estimation Experiment

		σ		
		0.02	0.03	0.04
$\alpha = 0.15$	$\overline{\alpha}$	0.175	0.175	0.178
	s_α	0.028	0.032	0.029
	$\overline{\sigma}$	0.0198	0.0298	0.0397
	s_σ	0.0007	0.0009	0.0013
$\alpha = 0.2$	$\overline{\alpha}$	0.226	0.226	0.226
	s_α	0.027	0.029	0.030
	$\overline{\sigma}$	0.0198	0.0298	0.0394
	s_σ	0.0007	0.0009	0.0013

we remark in passing that upper bounds on the absolute errors associated with the measurements of $\widehat{\alpha}$ and $\widehat{\sigma}$ may be approximated by $2s_\alpha$ and $2s_\sigma$. We will not do an analysis of the errors here although we display (approximate) error bars for interests sake.

$$\Delta\alpha = 0.06 \qquad \Delta\sigma = 0.002 \tag{4.13}$$

4.2 2-Factor (a) Model (Two Correlated Assets)

Consider a pair of correlated 1-factor energy prices with each one individually obeying the generalized 1-factor Pilipovic equation

$$\begin{aligned} ds &= \alpha\,[l(t) - s]\,dt + \sigma s\,dz \\ dv &= \beta\,[m(t) - v]\,dt + \tau v\,dw \end{aligned} \tag{4.14}$$

where

(s, v) Spot prices of two correlated assets

$l(t), m(t)$ Long term means (unknown functions of time)

(α, β) Strengths of mean reversion

(σ, τ) Volatilities

(z, w) Correlated Brownian motions

Let the correlation between the two related assets be denoted by $\rho = corr(z, w)$. We shall henceforth refer to Equation (4.14) as the 2-factor (a) model.

The parameters (α, σ) and (β, τ) for s and v can obviously be estimated one at a time by minimizing the individual variances of the simple diffusions $\{z_i\}$ and $\{w_i\}$ using the procedure we just developed in Subsection (4.1.4).

Put $\psi(t) = \overline{v}(t)$. Evidently the discrete equation for the simple diffusion $\{w_i\}$ is

$$w_i = \frac{v_{i+1} - v_i - [\beta(\psi_i - v_i) + \dot{\psi}_i]\Delta t}{v_i} \tag{4.15}$$

It merely remains to calculate the best estimator of the correlation ρ.

4.2.1 Minimum Variance Correlation

Suppose $(\widehat{\alpha}, \widehat{\sigma})$ and $\left(\widehat{\beta}, \widehat{\tau}\right)$ are the minimum variance estimators given by Equations (4.11) and (4.12). Then the minimum variance estimator of the correlation is

$$\widehat{\rho} = \frac{\sum_{i=1}^{n-1} z_i(\widehat{\alpha})\, w_i(\widehat{\beta})}{(n-1)\,\widehat{\sigma}\,\widehat{\tau}} \tag{4.16}$$

where $z_i(\widehat{\alpha})$ and $w_i(\widehat{\beta})$ are calculated by Equations (4.8) and (4.15).

Tables (4.5), (4.6), and (4.7) show the correlations between various energies over the period 1997-1999.

Table 4.5: 2-Factor (a) Model Minimum Variance Correlations (1997)

	Natural Gas	Crude Oil	Electricity
Natural Gas	1	0.203	0.283
Crude Oil	0.203	1	-0.035
Electricity	0.283	-0.035	1

This evidence suggests that natural gas is positively correlated with both crude oil and electricity while crude oil and electricity are probably uncorrelated.

Table 4.6: 2-Factor (a) Model Minimum Variance Correlations (1998)

	Natural Gas	Crude Oil	Electricity
Natural Gas	1	0.132	0.138
Crude Oil	0.132	1	-0.051
Electricity	0.138	-0.051	1

Table 4.7: 2-Factor (a) Model Minimum Variance Correlations (1999)

	Natural Gas	Crude Oil	Electricity
Natural Gas	1	0.167	0.140
Crude Oil	0.167	1	0.056
Electricity	0.140	0.056	1

4.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

Consider next the more interesting generalized 2-factor Pilipovic equation (3.6)

$$ds = \alpha(l - s)\,dt + \sigma s\,dz$$

$$dl = \beta(t)\,l\,dt + \tau l\,dy$$

where

s Spot price

l Long term mean

α Strength of mean reversion

$\beta(t)$ Drift of long term mean (unknown function of time)

σ Volatility of spot price

τ Volatility of long term mean

(z, y) Uncorrelated Brownian motions

The spot price obeys the same stochastic differential equation as before. However the long term mean is now also considered to be a stochastic variable and is ruled by an equation of its own. The spot price is still assumed to revert to the long term mean while the natural logarithm of the long term mean is a geometric Brownian motion (2.2). For the sake of simplicity we assume that the spot price and long term mean are uncorrelated[3].

Our objective is to estimate the parameters α, σ, and τ using historical data. The drift parameter $\beta(t)$ is not necessary to model energy prices as it is selected in such a way that the expected value of the spot price evolves according to predictions made by insightful forecasting such as futures price matching. Now it will be necessary to calculate the best estimator of $\beta(t)$ in order to determine the volatility τ of the long term mean $l(t)$.

The important thing to keep in mind is that the long term mean is unobservable. It represents an underlying economic factor driving the market and we have no way to

[3]In other words the Brownian motions z and y that contribute to the unpredictability of s and l are uncorrelated.

measure it directly. In practice we only have a historical time series for the spot prices.

Put $\psi(t) = \bar{l}(t)$. We can use the same method we developed in Subsection (4.1.1) to estimate the expected value of the long term mean. We then replace $l(t)$ by $\psi(t)$ in Equation (4.7) (hopefully) without altering the distribution of $z(t)$ too much. Hence we can find satisfactory estimators for $\widehat{\alpha}$ and $\widehat{\sigma}$ by treating the 2-factor (b) system as though it were in fact 1-dimensional[4]. Finally we attempt to recover the unobservable random variable $l(t)$ using a deconvolution of $\psi(t)$. This enables us to find $\widehat{\tau}$ by calculating the standard deviation of the Brownian motion $y(t)$.

4.3.1 Average Long Term Mean Estimator

By Theorem (2.11) the expected values $\varphi(t)$ and $\psi(t)$ of the spot price $s(t)$ and long term mean $l(t)$ satisfy the ordinary system (3.7)

$$\begin{aligned} \dot{\varphi}(t) &= \alpha[\psi(t) - \varphi(t)] \\ \dot{\psi}(t) &= \beta(t)\psi(t) \end{aligned}$$

Recall that the expected value $\varphi(t)$ can be safely approximated by an appropriate moving average of $s(t)$. Furthermore, the derivative $\dot{\varphi}(t)$ can be calculated by Equations (4.3) and (4.4). Hence we can solve Equation (3.7) for to reveal the expected

[4]In other words, we obtain estimators for α and σ under the supposition that $\tau = 0$ even though we know that $l(t)$ is a stochastic variable. This seems to work well in practice.

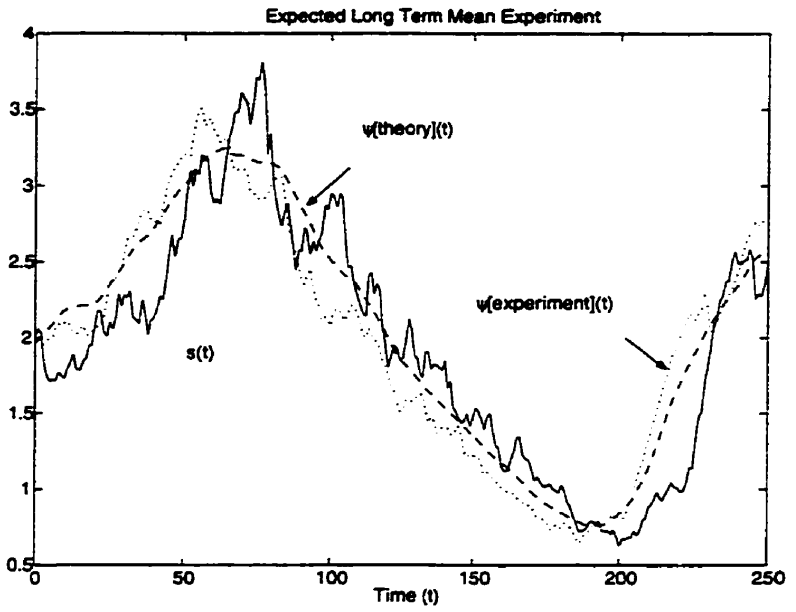

Figure 4.3: Expected Long Term Mean Experiment

value of the long term mean $l(t)$

$$\psi(t) = \varphi(t) + \frac{\dot{\varphi}(t)}{\alpha} \tag{4.17}$$

We test the validity of Equation (4.17) by simulating a pair of sample paths $\{s_i\}$ and $\{l_i\}$ obeying the 2-factor (b) model. Figure (4.3) compares an experimental expected long term mean $\{\psi_i\}$ (dotted line) calculated by Equation (4.17), to the theoretical one (dashed line) calculated by taking a moving average of the realized long term mean $\{l_i\}$. It would appear that Equation (4.17) does an excellent job of recovering the expected value $\psi(t)$ of the unobservable stochastic variable $l(t)$.

4.3.2 Discrete Time Series

The next step is to convert the continuous system (3.6) into an equivalent discrete model

$$\begin{aligned} s_{i+1} &= s_i + \alpha(l_i - s_i)\Delta t + s_i z_i \\ l_{i+1} &= (1 + \beta_i \Delta t) l_i + l_i y_i \end{aligned} \tag{4.18}$$

for $i = 1, ..., (n-1)$.

The sequences of random variables $\{z_i\}$ and $\{y_i\}$ are assumed to be (uncorrelated) normal stationary time series with expected values of zero and standard deviations of $\sigma\sqrt{\Delta t}$ and $\tau\sqrt{\Delta t}$.

We now replace $l(t)$ by its expectation $\psi(t)$ in Equation (4.8) to obtain an estimator of the simple diffusion $\{z_i\}$

$$\widehat{z}_i = \frac{s_{i+1} - [s_i + \alpha(\psi_i - s_i)\Delta t]}{s_i} \tag{4.19}$$

Assumption 2 *The estimator $\widehat{z}$ defined by Equation (4.19) has (approximately) the same probability distribution as z, namely $\phi(0, \sigma\sqrt{\Delta t})$.*

Observe that

$$\widehat{z}_i = z_i - \frac{\alpha(\psi_i - l_i)\Delta t}{s_i}$$

Since $s(t)$ and $l(t)$ are uncorrelated by our assumption, we see that the mean of

$\widehat{z}$ vanishes

$$\begin{aligned} \mathrm{E}(\widehat{z}_i) &= \mathrm{E}(z_i) - \frac{\alpha[\psi_i - \mathrm{E}(l_i)]\Delta t}{\mathrm{E}(s_i)} \\ &= 0 \end{aligned}$$

Next consider the variance of

$$\mathrm{var}(\widehat{z}_i) = \mathrm{var}(z_i) - \alpha\Delta t\mathrm{E}[\frac{(\psi_i - l_i)z_i}{s_i}] + \alpha^2\Delta t^2\mathrm{E}[\frac{(\psi_i - l_i)^2}{s_i^2}]$$

Now by a purely heuristic argument it can be shown that the fraction $\frac{(\psi_i - l_i)}{s_i}$ is approximately equal to the simple diffusion of the long term mean y_i. This is not precise; indeed, it is at this point that our attempted 'proof' of Assumption (2) breaks down. However, proceeding under the reasonable premise $y_i \sim \frac{(\psi_i - l_i)}{s_i}$, we get

$$\begin{aligned} \mathrm{var}(\widehat{z}_i) &\sim \mathrm{var}(z_i) - \alpha\Delta t\mathrm{E}[y_i z_i] + \alpha^2\Delta t^2\mathrm{E}[y_i^2] \\ &= \mathrm{var}(z_i) + \alpha^2\Delta t^2\mathrm{var}(y_i) \\ &\sim \mathrm{var}(z_i) \end{aligned}$$

when Δt is small. This concludes our informal scrutiny of Assumption (2).

4.3.3 Minimum Variance Estimators

Working under Assumption.(2), the estimators $\widehat{\alpha}$ and $\widehat{\sigma}$ may be found using precisely the same procedure that we developed in Subsection (4.1.4) for a 1-factor model. For the purpose of estimating the parameters α and σ in the 2-factor (b) model, the deter-

ministic process $\psi(t)$ would seem to be a valid substitution for the stochastic process $l(t)$.

We still have to find the volatility τ of the long term mean $l(t)$. We begin by solving Equation (4.18) for the simple diffusion $\{y_i\}$

$$y_i = \frac{l_{i+1} - l_i}{l_i} - \beta_i \Delta t \tag{4.20}$$

The parameter $\beta(t)$ can be calculated by Equation (3.7)

$$\dot{\psi}(t) = \beta(t)\psi(t)$$

where the expected value $\psi(t)$ of the long term mean $l(t)$ is given by Equation (4.17) with the strength of mean reversion α replaced by its minimum variance estimator $\widehat{\alpha}$

$$\psi(t) = \varphi(t) + \frac{\dot{\varphi}(t)}{\widehat{\alpha}} \tag{4.21}$$

That is, the drift parameter is

$$\beta(t) = \frac{\dot{\varphi}(t) + \frac{\ddot{\varphi}(t)}{\widehat{\alpha}}}{\varphi(t) + \frac{\dot{\varphi}(t)}{\widehat{\alpha}}} \tag{4.22}$$

where the derivatives at each discrete time $i\ \Delta t$ are calculated by the finite difference equations written down in Subsection (4.1.1). It now remains to develop a method of recovering the hidden variable $l(t)$.

4.3.4 Deconvolution

Assumption 3 *The expectation $\psi(t)$ is a convolution of the long term mean $l(t)$*

$$\psi_i = \sum_{j=-m}^{m} c_j l_{i-j} \tag{4.23}$$

This is essentially the same as Assumption (1) but for $l(t)$ instead of $s(t)$. However, a moving average is no longer appropriate. Alternatively, we now require that the convolution function satisfy the condition $c(u) < c(t)$ whenever $|u| > |t|$. This will ensure that the linear transformation representing the convolution is diagonally dominant. It is then possible to do a deconvolution by inverting the transformation.

Numerical investigations reveal that a suitable choice for $c(t)$ is the Euler distribution

$$c(t) = \frac{1}{1 + \gamma t^2}$$

where γ is a convolution parameter that may be selected to provide the most satisfactory results. We find that choosing $\gamma = 0.6$ works well in practice for the energy data sets we have.

The convolution (4.23) can be written conveniently in matrix form

$$\begin{pmatrix} \psi_0 \\ \\ . \\ \\ \psi_n \end{pmatrix} = \begin{pmatrix} c_{-m} & \cdot & c_m & 0 & \cdot & & 0 \\ 0 & c_{-m} & \cdot & c_m & 0 & \cdot & 0 \\ . & & & . & & & . \\ 0 & \cdot & 0 & c_{-m} & \cdot & c_m & 0 \\ 0 & & \cdot & 0 & c_{-m} & \cdot & c_m \end{pmatrix} \begin{pmatrix} l_{-m} \\ . \\ l_0 \\ \\ . \\ \\ l_n \\ . \\ l_{n+m} \end{pmatrix} \tag{4.24}$$

where n is the number of data points in the sample path $\{s_i\}$.

Observe that we have $n+1$ equations for $\psi_0, ..., \psi_n$ in $n+2m+1$ unknowns $l_{-m}, ..., l_{n+m}$. Evidently we are only really interested in solving for $l_0, ..., l_n$. To this end, put

$$\mathbf{l} = \begin{pmatrix} l_0 \\ \\ . \\ \\ l_n \end{pmatrix}$$

and consider the numbers $l_{-m}, ..., l_{-1}$ and $l_{n+1}, ..., l_{n+m}$ to be control parameters that will be selected so that our solution is as well behaved as possible.

To be more precise, the control parameters will be chosen so that the distance $\|\mathbf{l} - \boldsymbol{\psi}\|$ between the solution and the expectation is minimized. Why would we calculate the control parameters in this fashion? A deconvolution is notorious for exhibiting highly unstable behavior. Recall that a convolution has the effect of eliminating noise by a smoothing process. A deconvolution attempts to inject the lost noise back into $l(t)$ by amplifying any abnormalities or irregularities present in $\psi(t)$. The deconvolution has a tendency to magnify these deviations way out of proportion. Hence by constraining our solution $l(t)$ to be as close as possible to $\psi(t)$, we ensure that we obtain the most desirable solution.

Define a square matrix by

$$A = \begin{pmatrix} c_0 & \cdot & c_m & 0 & & & 0 \\ c_{-1} & c_0 & \cdot & c_m & 0 & \cdot & 0 \\ & & \cdot & & & & \\ \cdot & & & \cdot & & & \cdot \\ & & & & \cdot & & \\ 0 & \cdot & 0 & c_{-m} & \cdot & c_0 & c_1 \\ 0 & & & 0 & c_{-m} & \cdot & c_0 \end{pmatrix}$$

along with vectors

$$\mathbf{C}_j = \begin{pmatrix} c_{-j} \\ . \\ c_{-m} \\ . \\ 0 \end{pmatrix}$$

for $j = 1, ..., m$ and

$$\mathbf{C}_j = \begin{pmatrix} 0 \\ . \\ c_m \\ . \\ c_{j-m} \end{pmatrix}$$

for $j = (m+1), ..., 2m$ and finally relabel the control parameters by identifying them with $L_j = l_{-j}$ for $j = 1, ..., m$ and $L_j = l_{n-m+j}$ for $j = (m+1), ..., 2m$. Then the system (4.24) can be written more succinctly as

$$\psi = A\mathbf{l} + \sum_{j=1}^{2m} L_j \mathbf{C}_j$$

The solution of this matrix equation is

$$\mathbf{l} = A^{-1}\psi + \sum_{j=1}^{2m} L_j \mathbf{B}_j$$

where $\mathbf{B}_j = -A^{-1}\mathbf{C}_j$ for $j = 1, ..., 2m$.

Put $D = \|\mathbf{1} - \boldsymbol{\psi}\|^2$ as it is more convenient to minimize the square of the distance rather than the norm itself. The derivative of D with respect to L_k is

$$D^{(k)} = 2\sum_{i=1}^{n} B_{ik}(\sum_{j=1}^{2m} L_j B_{ij} + \mathbf{D}_i - \boldsymbol{\psi}_i)$$

where $\mathbf{D} = A^{-1}\boldsymbol{\psi}$ and B_{ij} is understood to be the i^{th} component of $\mathbf{B}_j$. Taking all of the derivatives $D^{(k)}$ for $k = 1, ..., 2m$ and setting them equal to zero we obtain a linear system of equations for the control parameters $M\mathbf{L} = \mathbf{N}$ where

$$M_{jk} = \sum_{i=1}^{n} B_{ij} B_{ik}$$

and

$$\mathbf{N}_j = \sum_{i=1}^{n} B_{ij}(\boldsymbol{\psi}_i - \mathbf{D}_i)$$

Therefore by choosing $\mathbf{L} = M^{-1}\mathbf{N}$ we ensure that $\|\mathbf{1} - \boldsymbol{\psi}\|$ is minimized.

4.3.5 Standard Deviation

Now we can calculate the volatility τ of the long term mean $l(t)$. It is given by dividing the standard deviation of the simple diffusion $\{y_i\}$ by Δt. If we denote the variance of the y_i's by U then

$$U = \frac{\sum_{i=1}^{n-1} y_i^2}{n-1} \tag{4.25}$$

where y_i is calculated by Equation (4.20). The best estimator of τ is

$$\widehat{\tau} = \sqrt{\frac{U(\widehat{\alpha})}{\Delta t}} \tag{4.26}$$

This procedure provides the desired estimators $\widehat{\alpha}$, $\widehat{\sigma}$, and $\widehat{\tau}$. Figures (4.4), (4.5), and (4.6) show the minimum variance estimators $\widehat{\alpha}$, $\widehat{\sigma}$, and $\widehat{\tau}$ for natural gas over the period 1997-1999. Also included are plots of the expected value $\psi(t)$ of the long term mean, along with an average value for the unknown function of time $\beta(t)$. Remember that the $\beta(t)$ is responsible for the shape of the curve $\psi(t)$ via Equation (3.7). However, it is only an intermediary step in the calculation of the parameters α, σ, and τ. We do not use this $\beta(t)$ for modeling purposes, and include it here purely for interests sake. The $\beta(t)$ used in option evaluation is found by matching the futures prices and the expected future spot prices. Hence our model is forward looking in that it takes advantage of the beliefs of traders in the energy market. The corresponding pictures for crude oil and electricity are located in Appendix A.

Tables (4.8), (4.9), and (4.10) show the minimum variance estimators $\widehat{\alpha}$, $\widehat{\sigma}$, and $\widehat{\tau}$ for natural gas, crude oil, and electricity over the period 1997-1999.

Table 4.8: 2-Factor (b) Model Minimum Variance Estimators (1997)

	$\widehat{\alpha}$	$\widehat{\sigma}$	$\widehat{\tau}$
Natural Gas	0.232	0.0331	0.0240
Crude Oil	0.141	0.0174	0.0162
Electricity	0.188	0.0418	0.0377

We now perform the simulation experiment described in Subsection (4.1.4) for the case of the 2-factor (b) model. For all possible combinations of the parameter sets

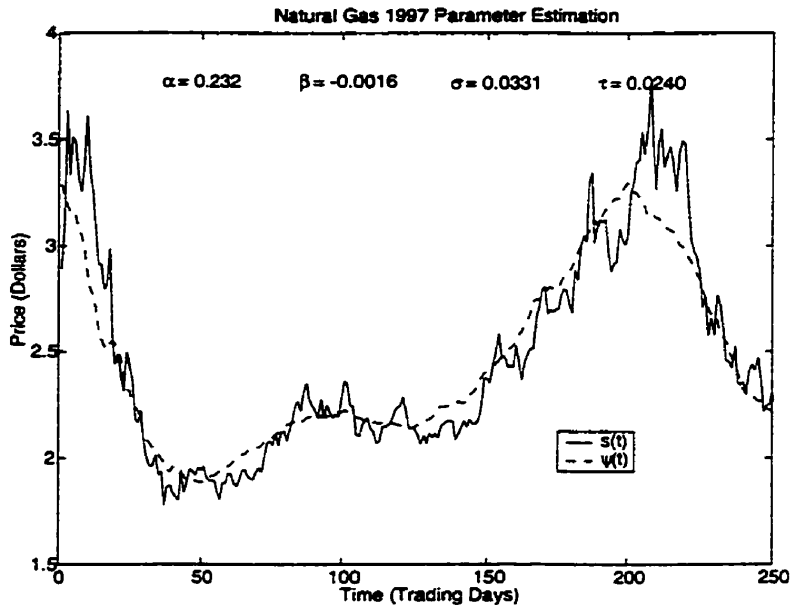

Figure 4.4: Natural Gas 1997 Parameter Estimation

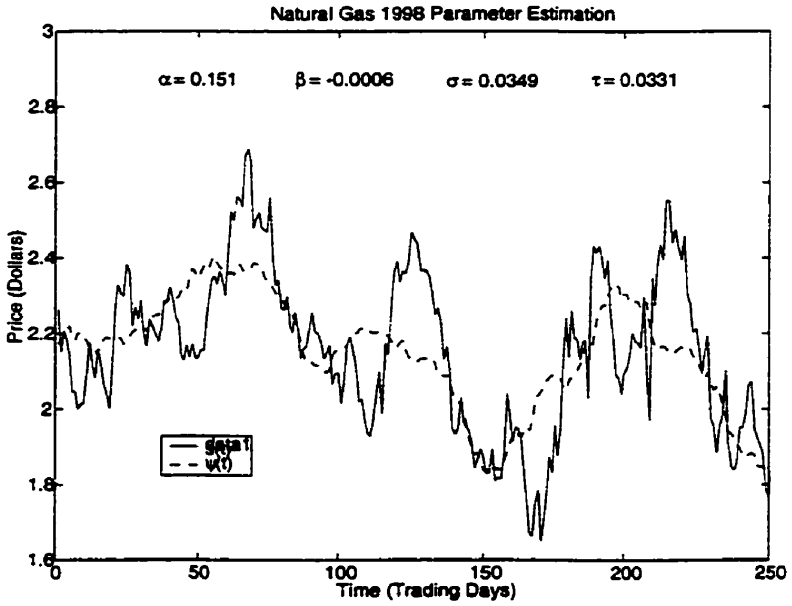

Figure 4.5: Natural Gas 1998 Parameter Estimation

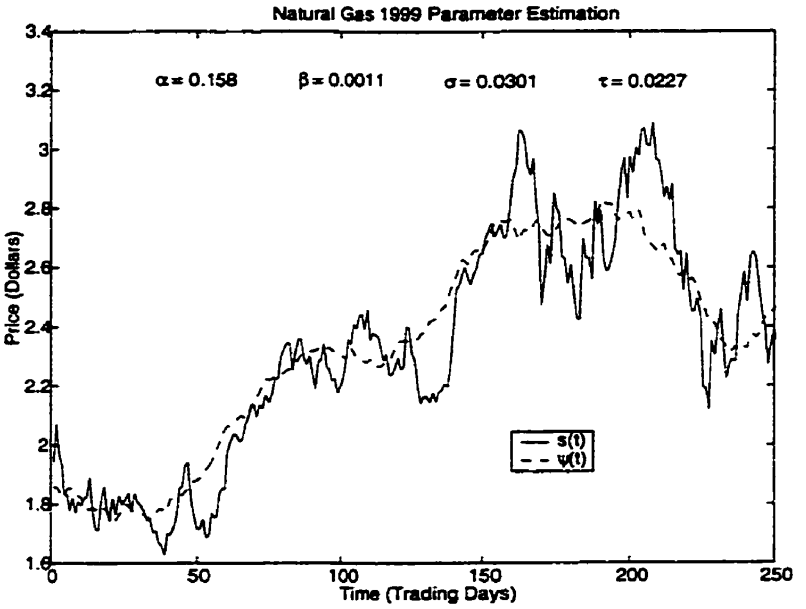

Figure 4.6: Natural Gas 1999 Parameter Estimation

$\alpha = 0.15$ and $\alpha = 0.2$, $\sigma = 0.02$, $\sigma = 0.03$, and $\sigma = 0.04$, and $\tau = 0.02$, $\tau = 0.03$, and $\tau = 0.04$, 100 sample paths consisting of 250 data points were simulated. The results are displayed in Tables (4.11) and (4.12).

Although the results vary for different combinations of the parameters, we find that on the whole they are satisfactory. Bearing in mind that we are attempting to reconstruct a hidden variable with limited information, we believe this to be a worthy achievement. According to Tables (4.11) and (4.12), the (approximate) error bars for the best estimators of the 2-factor (b) model are

$$\Delta\alpha = 0.07 \qquad \Delta\sigma = 0.003 \qquad \Delta\tau = 0.015 \tag{4.27}$$

Table 4.9: 2-Factor (b) Model Minimum Variance Estimators (1998)

	$\widehat{\alpha}$	$\widehat{\sigma}$	$\widehat{\tau}$
Natural Gas	0.151	0.0349	0.0331
Crude Oil	0.170	0.0283	0.0231
Electricity	0.172	0.0479	0.0437

Table 4.10: 2-Factor (b) Model Minimum Variance Estimators (1999)

	$\widehat{\alpha}$	$\widehat{\sigma}$	$\widehat{\tau}$
Natural Gas	0.158	0.0301	0.0227
Crude Oil	0.146	0.0212	0.0200
Electricity	0.154	0.0454	0.0397

Table 4.11: 2-Factor (b) Model Parameter Estimation Experiment ($\alpha = 0.15$)

$\alpha = 0.15$		τ		
σ		0.02	0.03	0.04
0.02	$\overline{\alpha}$	0.156	0.108	0.085
	s_α	0.034	0.030	0.023
	$\overline{\sigma}$	0.0204	0.0213	0.0223
	s_σ	0.0008	0.0010	0.0010
	$\overline{\tau}$	0.0191	0.0287	0.0391
	s_τ	0.0044	0.0067	0.0123
0.03	$\overline{\alpha}$	0.183	0.155	0.127
	s_α	0.039	0.040	0.028
	$\overline{\sigma}$	0.0299	0.0307	0.0315
	s_σ	0.0012	0.0015	0.0014
	$\overline{\tau}$	0.0245	0.0296	0.0370
	s_τ	0.0049	0.0068	0.0870
0.04	$\overline{\alpha}$	0.204	0.179	0.151
	s_α	0.038	0.038	0.033
	$\overline{\sigma}$	0.0394	0.0405	0.0407
	s_σ	0.0019	0.0021	0.0019
	$\overline{\tau}$	0.0269	0.0339	0.0408
	s_τ	0.0056	0.0073	0.0104

Table 4.12: 2-Factor (b) Model Parameter Estimation Experiment ($\alpha = 0.2$)

$\alpha = 0.2$		τ		
σ		0.02	0.03	0.04
0.02	$\overline{\alpha}$	0.161	0.109	0.082
	s_α	0.040	0.028	0.024
	$\overline{\sigma}$	0.0210	0.0221	0.0237
	s_σ	0.0009	0.0011	0.0010
	$\overline{\tau}$	0.0196	0.0297	0.0444
	s_τ	0.0049	0.0097	0.0231
0.03	$\overline{\alpha}$	0.212	0.158	0.123
	s_α	0.038	0.039	0.030
	$\overline{\sigma}$	0.0303	0.0313	0.0328
	s_σ	0.0015	0.0014	0.0015
	$\overline{\tau}$	0.0216	0.0295	0.0403
	s_τ	0.0041	0.0073	0.0096
0.04	$\overline{\alpha}$	0.226	0.192	0.158
	s_α	0.046	0.040	0.028
	$\overline{\sigma}$	0.0402	0.0412	0.0420
	s_σ	0.0020	0.0018	0.0018
	$\overline{\tau}$	0.0251	0.0325	0.0390
	s_τ	0.0054	0.0073	0.0075

CHAPTER 5
FUTURES PRICES

Traders in the energy market want to cover their positions by entering into futures contracts. This is a simple agreement between a buyer and seller that an asset will be bought and sold at a certain price at a specific time in the future. Naturally the price that is agreed upon is related to the expected value of the energy price in the future. Further the quoted futures prices in the market reflect the beliefs of an amalgamation of traders. We can use this extremely valuable information to calibrate the trinomial trees we use to model energy prices. In this chapter, we follow closely the excellent expose of futures contracts in [Hull 1999].

5.1 Futures Contracts in the Energy Market

Definition 5.16 *A **futures contract** is an agreement between traders to buy or sell an asset at a fixed price at a certain time in the future. Let $f(t,T)$ denote the **futures price** of a contract beginning at time t and ending at time T. The time T is known as the **date of maturity** of the futures contract. We often write the futures price as a function of a single variable $f(s)$: in this case we shall always assume that $t = 0$ (or $t = t_0$) and $T = s$.*

In natural gas, crude oil, and electricity markets the date of maturity is typically the first trading day of each month. In other words futures contracts are arranged on

a monthly basis. For example the February futures price is the price that traders agree now to exchange for the commodity at the beginning of February.

At the time the futures contract is arranged no money is exchanged. A futures contract costs nothing to either trader involved in the deal. This is because it is supposed to be equally valuable to both sides. The buyer gains protection from skyrocketing energy prices and the seller is guaranteed to be unaffected by plummeting energy markets. Of course it is entirely possible that one of the traders will be disappointed by the actual outcome. If the energy price realized at time T is high above the futures price then the seller will have missed an opportunity to make considerable profits. Alternatively if $s(T)$ the is far below $f(T)$ then the buyer will have missed the chance to save money. By reducing their mutual risk the traders have effectively removed the possibility of either of them making a killing (at the expense of the other).

This poses an obvious question: Is the futures price equal to the expected future spot price? Let us examine this idea using a financial pricing argument.

5.2 Financial Pricing Theory

Axiom 1 *All investment opportunities in energy markets have zero net present value.*

In other words it is impossible to enter into an arrangement that realizes a positive instantaneous profit with no risk of loss.

5.2.1 Risk and Return

It is an established fact that the higher the risk of an investment the higher the expected return demanded by an investor. There are also 2 types of risk in the economy: systematic and nonsystematic. Nonsystematic risk is unimportant as it can be completely diversified away. Systematic risk is the result of a correlation between the returns from a particular investment and the returns from the entire energy market. Therefore an investor requires higher returns than the risk free interest rate for bearing positive amounts of systematic risk. Alternatively an investor is prepared to accept lower expected returns than the risk free interest rate when the systematic risk in an investment is negative.

5.2.2 The Risk in a Futures Position

Consider a speculator who takes a long position in a futures contract[1] in the hope that the energy price will be above the futures price at maturity. We suppose that the speculator puts the present value of the futures price into a risk free investment at time while simultaneously taking a long futures position. The proceeds of the risk free investment are used to buy the asset on the delivery date. The asset is then immediately sold for its market price. This means that the cash flows to the speculator are

$$\text{Time 0: } -f(T)\mathrm{e}^{-rT}$$

[1]That is the speculator agrees to purchase the asset in the future at the agreed upon price.

Time T: $+s(T)$

where r is the *risk free interest rate.*

The present value of this investment is $\varphi(T)e^{-\mu T} - f(T)e^{-rT}$ where we have denoted the expected future spot price by $\varphi(t) = \overline{s}(t)$.

As this financial opportunity can have no net present value, we set this quantity to zero and solve for the futures price

$$f(T) = \varphi(T)\, e^{(r-\mu)T} \tag{5.1}$$

or equivalently when $t \neq 0$ we have

$$f(t,T) = \varphi(T)\, e^{(r-\mu)(T-t)}$$

where μ is the discount rate appropriate for the investment[2]. The number μ depends on the systematic risk of the investment. Let ρ be the correlation between $s(T)$ and the level of the energy market. Then there are 3 cases to consider:

(1) $\rho < 0$: $\mu < r$ and $f(T) > \varphi(T)$ (*Contango*)

(2) $\rho = 0$: $\mu = r$ and $f(T) = \varphi(T)$

(3) $\rho > 0$: $\mu > r$ and $f(T) < \varphi(T)$ (*Normal Backwardation*)

5.3 Statistical Analysis

We now explain the method of estimating the expected return μ. This is the final piece in the puzzle of our energy price model. Once the best estimator $\widehat{\mu}$ has been obtained

[2]That is μ is the expected return required by investors on the investment.

we will have at our disposal all the necessary tools to price options. For the sake of simplicity, we will assume that μ is a constant parameter that does not depend on the start and end times t and T of the futures contract[3].

5.3.1 Hypothesis Testing

The expected value $\varphi(T)$ can be taken to be the realization of the energy price $s(T)$ at the date of maturity. The futures prices $f(t,T)$ will be compared to the values of $\varphi(t)$. In essence we are observing the differences between what investors believed would happen and what actually occurred in the energy market.

It is convenient to define a normalized parameter by $\lambda = \mu - r$[4]. Then

$$\lambda = \frac{1}{(T-t)} \ln[\frac{\varphi(T)}{f(t,T)}] \tag{5.2}$$

Let $\{\varphi_j\}$ and $\{f_{ij}\}$ be sample sets of futures prices and realized energy prices at start and end times $\{(t_i, T_j)\}$. These can be used in conjunction with Equation (5.2) to obtain a sample set of numbers $\{\lambda_{ij}\}$. It is then possible to do a statistical analysis on $\{\lambda_{ij}\}$ to determine which of the 3 possible situations is realized for a particular energy price process: contango ($\lambda < 0$), no systematic risk ($\lambda = 0$), or normal backwardation ($\lambda > 0$).

[3]In reality the expected return μ is probably a (seasonal) function of both t and T but this issue will not be addressed in this thesis.

[4]It is customary to call $\frac{\mu - r}{\sigma}$ the market price of risk where σ is the volatility of $s(t)$.

We perform a student t-test. For a thorough description of this procedure, we refer to [Weiss 1999].

For the sake of simplicity we shall form a null hypothesis of no systematic risk and make no assumptions about the alternative choices.

Null Hypothesis: $\lambda = 0$

Alternative Hypothesis: $\lambda \neq 0$ (2-tailed)

Suppose that the parameter λ is normally distributed with mean 0. Then, for samples of size n, the studentized version of λ

$$t = \frac{\overline{\lambda}}{s_\lambda / \sqrt{n}}$$

has the student t distribution with $n-1$ degrees of freedom where $\overline{\lambda}$ and s_λ are the sample mean and standard deviation of λ defined by

$$\overline{\lambda} = \frac{\sum \lambda}{n}$$

and

$$s_\lambda = \sqrt{\frac{\sum \lambda^2 - (\sum \lambda)^2/n}{n-1}}$$

By elementary statistics we know that the t-test is robust to moderate violations of the normality assumption.

We shall consider samples of size $n = 101$ under a significance level $\alpha = 0.05$. Then for a particular value of the test statistic t we will accept the null hypothesis if

$t \in [-1.984, 1.984]$ and reject it otherwise. We refer to Table IV in [Weiss 1999] to see that $t_{0.025}(100) = 1.984$.

The results for natural gas, oil, and electricity are displayed in Table (5.1). The sample sets were taken from spot prices and futures prices over the period $1997 - 1999$.

Table 5.1: Null Hypothesis t-Test

	$\overline{\lambda}$	s_λ	t
Natural Gas	-0.000819	0.006	-1.408
Crude Oil	-0.000395	0.0042	-0.927
Electricity	0.000834	0.0046	1.877

In each case, there is no evidence to suggest that the futures price is not an unbiased estimator of the expected future spot price. That is, we can safely assume that on average $\lambda = 0$ for all energy commodities under consideration. Thus we assume that there is no systematic risk for the remainder of this thesis and simply put

$$\varphi(t) = f(t) \tag{5.3}$$

Equation (5.3) is the condition that, if satisfied, will render our numerical solutions risk neutral, by ensuring that all investment opportunities in the energy market, have no net present value.

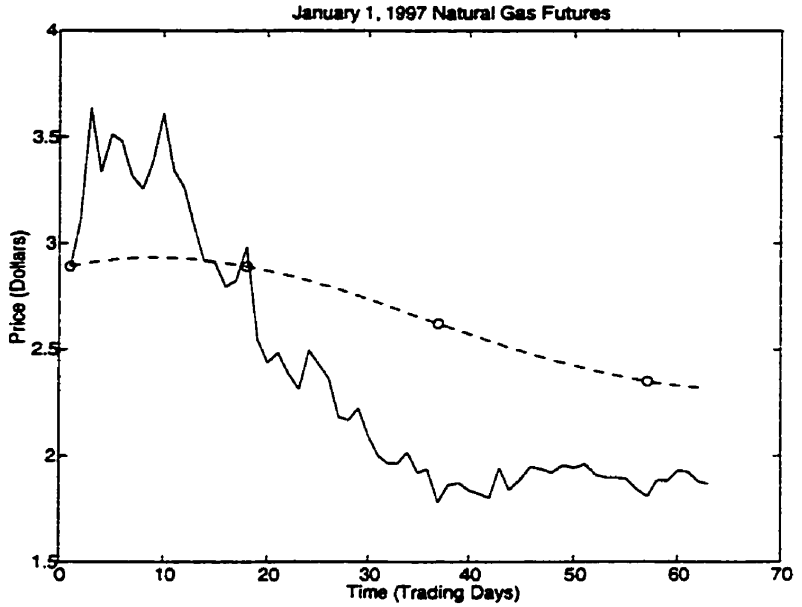

Figure 5.1: January 1, 1997 Natural Gas Futures

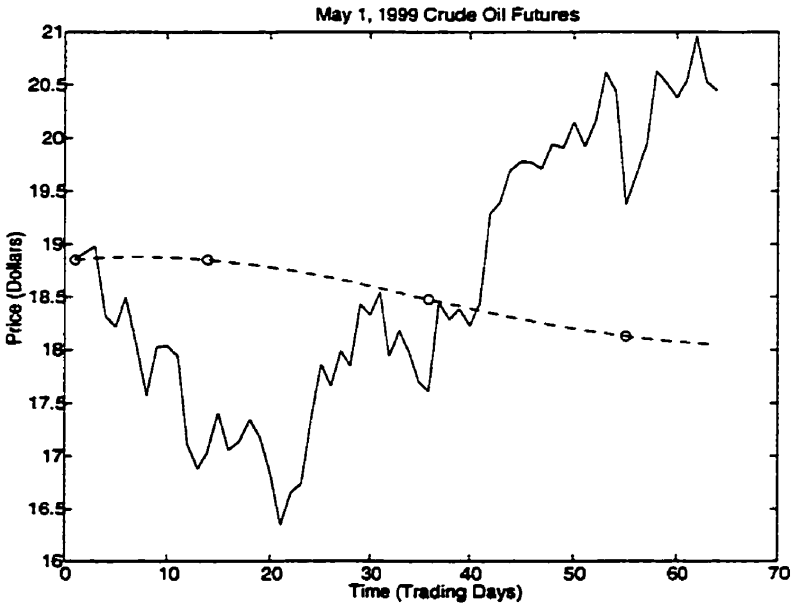

Figure 5.2: May 1, 1999 Crude Oil Futures

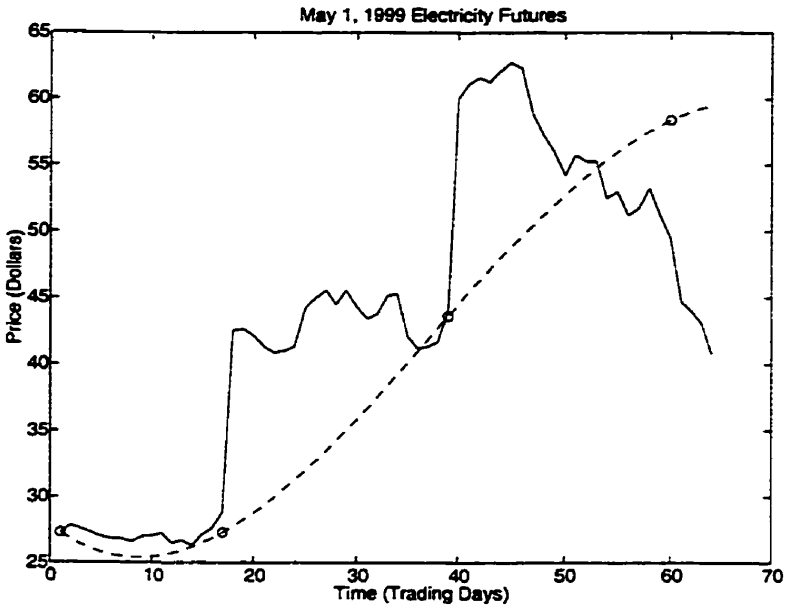

Figure 5.3: May 1, 1999 Electricity Futures

5.3.2 Futures Prices and Expected Future Spot Prices

Now we examine the relationship between futures prices and the realizations of spot prices for natural gas, crude oil, and electricity. According to what we just discussed, traders are obviously trying to predict what is actually going to happen in the future when they agree on the price in a futures contract. Figures (5.1)-(5.3) show the futures prices over a period of 3 months along with the spot prices that were later realized over this same period. In the diagrams, the solid lines depict the realized spot prices, while the circles show quoted monthly futures prices and the dashed line represents a cubic spline interpolation of missing values. Note that a circle located at time t, represent the delivery price of a futures contract entered into at time zero, and expiring at time

t.

It is clear from the pictures that although sometimes the futures prices end up being above and below the realized spot prices, on average the futures prices do an admirable job of guessing future spot prices. In a highly volatile energy market, this is probably the best that one can hope for. Refer to Appendix B to see more pictures of futures prices and realized spot prices for natural gas, crude oil, and electricity.

CHAPTER 6
NUMERICAL SOLUTIONS

Some stochastic differential equations can be solved by providing an analytic expression of the probability distribution of the random variable. It occurs often in practice that analytic solutions are impossible to arrive at. Furthermore, even when analytic solutions are available, they are unsuitable for evaluating path dependent derivative securities such as American style options. Therefore we investigate the possibility of solving stochastic differential equations by numerical methods that involve the construction of trinomial trees.

In this chapter, we extend the work of Hull & White, who have developed a trinomial tree implementation of a linear mean reverting model appropriate for interest rates. We extend and modify their procedures so that they may be applied to 1 and 2 factor nonlinear models applicable the exciting area of energy modeling. In fact it is fair to say that this thesis attempts to do for the energy industry what Hull & White accomplished for interest rate derivative securities. The following list briefly describes the main innovations offered here:

(1) Following a suggestion made in [Hull & White 1990], we develop a drift adapted method of arranging the geometry of trinomial trees that speeds convergence by minimizing errors near the median nodes of all the branches in a tree.

(2) We illuminate the notion of creating a risk neutral calibrated model by matching the futures price curve available in the energy market.

(3) We extend the method of fast trees introduced in [Hull & White 1994 (a)] by generalizing the procedure to admit random processes with nonlinear drifts. This innovation decomposes a trinomial tree into a preliminary tree that is subsequently transformed into a final tree by arranging the geometry of each branch so that the futures prices are precisely matched. This has the pleasing effect of separating the deterministic and stochastic parts of the random process. Thus each component can be dealt with one at a time instead of tackling the whole problem all at once. The fast trees are infinitely easier to understand and implement. Furthermore, they run faster (hence our name for them) and tend to converge more rapidly than ordinary trinomial trees.

(4) The nonlinear Pilipovic 2-factor models are decoupled by making substitutions similar to the ones found in [Hull & White 1994 (b)]. This completes our nonlinear analysis of the linear methods developed by Hull & White.

6.1 Trinomial Trees

We begin with a few geometric definitions.

Definition 6.17 *A **directed graph** (G, R) is a set of nodes G together with a relation R on the node set.*

Every element (a, b) in the relation R can be interpreted as a ray (directed line segment) connecting a to b. For the sake of convenience, we will often refer to a graph

(G, R) as simply G when it is unnecessary to specify the relation R.

Definition 6.18 *A **path** is a subset of rays in R of the form $\{(a, \beta_0), (\alpha_1, \beta_1), ..., (\alpha_n, b)\}$ where a is the **start node** and b is the **end node** in the path.*

The path can be interpreted as one of the possible ways of getting from a to b.

Definition 6.19 *A **circuit** is a path whose start and end nodes are identical*

Definition 6.20 *A **tree** T is a directed graph G that has no circuits.*

Definition 6.21 *A **trinomial tree** is a tree where every node has precisely 3 rays emanating from it.*

By our definition, a trinomial tree always has a unique node called the *root* with the property that there is no ray arriving at the root. The root can be thought of as the start node of all possible paths throughout the tree. For the remainder of this thesis, all trees that are mentioned are understood to be trinomial trees.

The fact that our trees have no circuits has a natural interpretation. Consider the time interval $[0, T]$. In our construction, a ray from a to b allows for the possibility of branching from a to b in the next moment of time. That is, rays in our trees point in the direction of increasing time. The existence of a circuit would then indicate the potential to travel back through time. In this thesis, we tacitly assume that reverse time travel is impossible for trading purposes!

Suppose that a random variable $s(t)$ satisfies an Ito process of the form

$$ds = a[s, \pi(t)]dt + b\,(s)\,dz \tag{6.1}$$

where $\pi(t)$ is an unknown function of time.

The parameter $\pi(t)$ will be selected so that the expected value of at every branch in the tree satisfies Equation (5.3)

$$\varphi\,(t) = f\,(t)$$

This renders the probabilities in the tree to be risk neutral by the zero net present value argument explained in Chapter (5). Hence the tree can be used to evaluate derivative securities.

In order to construct an additive trinomial tree, we require that the standard deviation be a constant number σ. To this end, it may be necessary to transform to a new random variable

$$S = \sigma \int \frac{ds}{b(s)}$$

We assume the antiderivative exists so that we can write $S = S(s)$. Furthermore, we suppose the transformation is invertible and we denote the inverse function by $s = s(S)$.

By the Ito formula

$$dS = A[S, \pi(t)]dt + \sigma\ dz \tag{6.2}$$

where the new drift is

$$A[S, \pi(t)] = \sigma\{\frac{a[s(S), \pi(t)]}{b[s(S)]} - \frac{b'[s(S)]}{2}\} \tag{6.3}$$

Note that the above framework is valid for the two cases of primary interest to us, namely $b(s) = \sigma$ (Hull & White) and $b(s) = \sigma s$ (Pilipovic), as $s(t)$ never reaches zero in finite time according to [Wilmott 1998].

6.1.1 Tree Structure

We first select the number of time steps n. Then the length of a time step is $\Delta t = \frac{T}{n}$. It is desirable that $\Delta t \in (0, 1)$. The length of a small change in S is then chosen to be

$$\Delta S = \sigma\sqrt{3\ \Delta t}$$

[Hull & White 1990] suggest that this is a good choice for ΔS in that it minimizes errors and speeds convergence.

Define (i, j) as the node where the time is $t_i = i\ \Delta t$ and the transformed random variable is $S_{ij} = \widehat{\Phi}_i + j\ \Delta S$ where $\widehat{\Phi}_i = S_{i0}$ is the position of the *median* node of the i^{th} branch in the tree. We adopt this unusual notation for the median nodes $\widehat{\Phi}_i$ as we will see later that they are approximately equal to the expected values $\Phi(t)$ of the random process $S(t)$ at the discrete times $t_i = i\ \Delta t$ for $i = 0, ..., n$. We assume that the spot price of the energy is known at time 0 and put $S_0 = S(s_0)$. We imagine that we are creating a probability distribution that will represent the future evolution of the spot

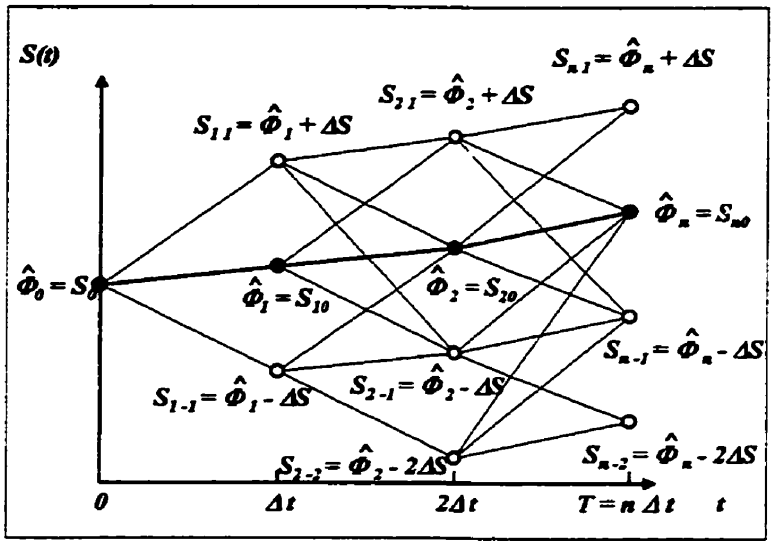

Figure 6.1: Ordinary Trinomial Tree Structure

price. Hence the median node at time 0 is $\widehat{\Phi}_0 = S_0$. Observe that the root node is labelled $(0,0)$ although we write S_0 in place of S_{00} as a matter of style. Figure (6.1) shows the geometry of an ordinary trinomial tree.

By Theorem (2.12) the mean and variance of $S(t_i + \Delta t) - S(t_i)$ are

$$\begin{aligned} M_{ij} &= A(S_{ij}, \pi_i)\Delta t \\ V &= \sigma^2 \Delta t \end{aligned} \tag{6.4}$$

at node (i,j) in the tree for S defined by Equation (6.2).

Let the drifts of the median nodes be denoted by $M_{i0} = A(\widehat{\Phi}_i, \pi_i)\Delta t$. The median nodes for $i = 1, ..., n$ are defined recursively by

$$\widehat{\Phi}_i = \widehat{\Phi}_0 + \sum_{k=0}^{i-1} M_{i0} \tag{6.5}$$

Then the median nodes are all precisely connected by their drifts as dictated by Equation (6.3). This structure makes the computer program particularly elegant. Also, we propose (without proof) that it probably speeds convergence. Why? Normally, the error of convergence associated with a node is $O(\Delta t)$ as we are using a linear approximation in the tradition of Euler. However, it is stated in [Hull & White 1990] that when the ray emanating from a node (i, j) lands directly on another node, the error of convergence associated with that node is $O(\Delta t^2)$. Furthermore, we shall demonstrate later that the median nodes are located near the expected values of the random process $S(t)$. This means that most of the discrete probability distribution is concentrated in the vicinity of the median nodes. Therefore, by minimizing the errors near the median nodes, we aim to significantly enhance the convergence rate of the trinomial tree. This leads to the phrase *drift adapted mesh*, which we will simply refer to as an ordinary trinomial tree.

We now turn our attention to the calculation of the branching probabilities. By choosing the branching probabilities so that the mean M_{ij} and variance V of the random process (6.2) are imitated at every node (i, j), we ensure that the discrete process created by our trinomial tree will approach (in a sense that will be clarified later) the continuous process as $\Delta t \to 0$.

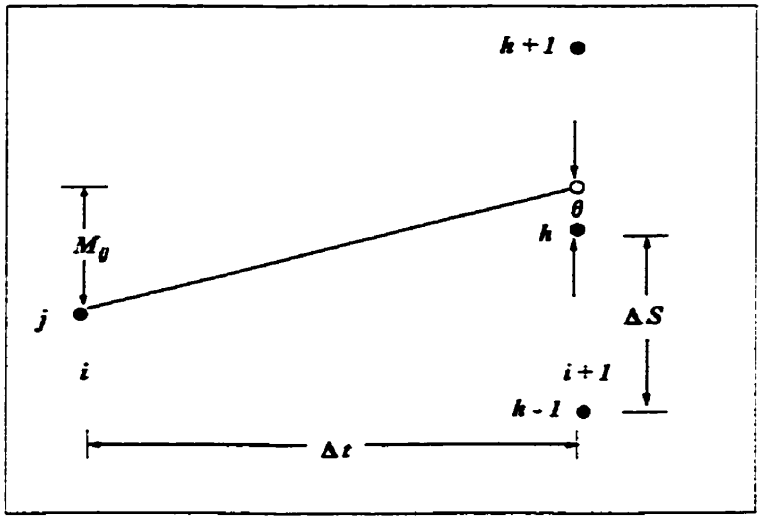

Figure 6.2: Trinomial Tree Node Branching Process

6.1.2 Branching Probabilities

Suppose we are at node (i, j). Let $(i+1, h_{ij})$ be the closest node to the drift M_{ij} bearing in mind that each branch is adapted by the drift M_{i0} of the median node $\widehat{\Phi}_i$

$$h_{ij} = \left[j + \frac{M_{ij} - M_{i0}}{\Delta S} \right] \qquad \text{(Closest Integer)} \tag{6.6}$$

Then, starting from node (i, j), it is possible to branch to one of the nodes $(i+1, h_{ij}+1)$, $(i+1, h_{ij})$, or $(i+1, h_{ij}-1)$. We shall respectively call these nodes the up, middle, and down nodes associated with node (i, j). Hence each node in the tree has three rays emanating from it, with the exception of all nodes of the form (n, j) at time $T = n\ \Delta t$, which have none (they are at the top of the tree).

Let $\theta_{ij} = M_{ij} - M_{i0} - (h_{ij} - j)\Delta S$ be the offset between the adapted drift and

this closest node. Then the positions of the up, middle, and down nodes are related to the location of their source node (i, j) by

$$\begin{aligned} S_{i+1,h_{ij}+1} &= S_{ij} + M_{i0} + (M_{ij} - M_{i0} + 1)\Delta S - \theta_{ij} \\ S_{i+1,h_{ij}} &= S_{ij} + M_{i0} + (M_{ij} - M_{i0})\Delta S - \theta_{ij} \\ S_{i+1,h_{ij}-1} &= S_{ij} + M_{i0} + (M_{ij} - M_{i0} - 1)\Delta S - \theta_{ij} \end{aligned}$$

Denote the probabilities of branching to the up, median, and down nodes by $p_{ij}^{(u)}$, $p_{ij}^{(m)}$, and $p_{ij}^{(d)}$ respectively. We will also respectively write these branching probabilities as $p_{ij}^{(1)}$, $p_{ij}^{(0)}$, and $p_{ij}^{(-1)}$, or possibly as p_u, p_m, and p_d, when it is convenient to do so. Then the linear system of equations satisfied by the branching probabilities is obtained by calculating the mean and variance of the discrete process and setting them to M_{ij} and V, together with the universal condition that all probabilities must add up to 1.

$$\begin{aligned} \left[p_{ij}^{(u)}(h_{ij} - j + 1) + p_{ij}^{(m)}(h_{ij} - j) + p_{ij}^{(d)}(h_{ij} - j - 1)\right]\Delta S &= M_{ij} \\ p_{ij}^{(u)}(\Delta S - \theta_{ij})^2 + p_{ij}^{(m)}\theta_{ij}^2 + p_{ij}^{(d)}(\Delta S + \theta_{ij})^2 &= V \\ p_{ij}^{(u)} + p_{ij}^{(m)} + p_{ij}^{(d)} &= 1 \end{aligned}$$

The solution reveals the desired probabilities

$$\begin{aligned} p_{ij}^{(u)} &= \frac{1}{6} + \frac{1}{2}\left(\frac{\theta_{ij}^2}{\Delta S^2} + \frac{\theta_{ij}}{\Delta S}\right) \\ p_{ij}^{(m)} &= \frac{2}{3} - \frac{\theta_{ij}^2}{\Delta S^2} \\ p_{ij}^{(d)} &= \frac{1}{6} + \frac{1}{2}\left(\frac{\theta_{ij}^2}{\Delta S^2} - \frac{\theta_{ij}}{\Delta S}\right) \end{aligned} \tag{6.7}$$

Figure (6.2) shows a picture of the branching probabilities at a typical node (i, j).

Theorem 6.22 *The branching probabilities (6.7) are always between* 0 *and* 1.

Proof. It is easily shown that $p_{ij}^{(u)}$ and $p_{ij}^{(d)}$ are always positive for all values of θ_{ij}. To see this let $x = \frac{\theta_{ij}}{\Delta S}$. By the quadratic formula, we have

$$\frac{1}{2}x^2 + \frac{1}{2}x + \frac{1}{6}$$

for $p_{ij}^{(u)}$. The discriminant of this parabola is evidently negative.

$$\begin{aligned} b^2 - 4ac &= (\frac{1}{2})^2 - 4(\frac{1}{2})(\frac{1}{6}) \\ &= \frac{1}{4} - \frac{1}{3} \\ &< 0 \end{aligned}$$

Hence $p_{ij}^{(u)}$ is positive for all values of x. A similar result holds for $p_{ij}^{(d)}$.

It now suffices to show that $p_{ij}^{(u)} + p_{ij}^{(d)} \leq 1$. By our method of selecting the integer h_{ij} we know that

$$|\theta_{ij}| \leq \frac{\Delta S}{2}$$

Therefore

$$\begin{aligned} p_{ij}^{(u)} + p_{ij}^{(d)} &= \frac{1}{3} + \frac{\theta_{ij}^2}{\Delta S^2} \\ &\leq \frac{1}{3} + \frac{1}{4} \\ &< 1 \end{aligned}$$

■

By our construction, the offset θ_{i0} is zero for the median node $\widehat{\Phi}_i$ of every branch in the tree, see Figure (6.2) for clarification. This means that the branching probabilities emanating from any median node are always given by Equation (6.7) to be $\frac{1}{6}$, $\frac{2}{3}$, and $\frac{1}{6}$ for moving to the up, middle, and down nodes. According to [Hull & White 1990] this means that the errors of convergence associated with the median nodes are $O(\Delta t^2)$. Remember that the median nodes are located near the expected values $\Phi(t)$ of the transformed price process $S(t)$. This means that we have succeeded in minimizing the errors of convergence in the region of the tree where most of the probability is accumulated. Hence we have developed what we call a drift adapted mesh.

6.1.3 Discrete Probability Distribution

The maximum and minimum values of j are different at each branch in the tree. It is easiest to calculate these boundary values recursively by setting $j_{\max}(0) = j_{\min}(0) = 0$ and defining

$$\begin{aligned} j_{\max}(i) &= h_{i-1,j_{\max}(i-1)} + 1 \\ j_{\min}(i) &= h_{i-1,j_{\min}(i-1)} - 1 \end{aligned} \tag{6.8}$$

for $i = 1, ..., n$. This is guaranteed to work as long as the drift function A is continuous.

Denote the probability of reaching node (i, j) by P_{ij}. These cumulative proba-

bilities are easy to calculate recursively provided we know the branching probabilities at every node in the tree - and we do by Equation (6.7).

$$\begin{aligned} P_{00} &= 1 \\ P_{ij} &= \sum_k P_{i-1,k} q[(i-1,k) \rightarrow (i,j)] \end{aligned}$$

for $i = 1, ..., n$ and $j = j_{\min}(i), ..., j_{\max}(i)$ where $q[(i-1,k) \rightarrow (i,j)]$ is the probability of branching from node $(i-1,k)$ to node (i,j). Observe that $\sum_j P_{ij} = 1$ for $i = 0, ..., n$. The probabilities P_{ij} define a discrete probability distribution that approximates the continuous solution of the stochastic differential equation (6.2).

This completely specifies the tree construction. Note that a trinomial tree is numerically efficient if the number of nodes does not grow exponentially in time. For instance, the trees constructed in the sequel are such that $j_{\max}(i)$ and $j_{\min}(i)$ are $O(n)$ for all $i = 0, ..., n$. Therefore the computer programs that construct these trinomial trees run in time $O(n^2)$. Let $Q(s,t)$ be the continuous probability distribution that is the solution of the stochastic differential equation (6.2). The discrete probabilities P_{ij} are approximations of the integrals of $Q(s,t)$ over small neighborhoods of the points (s_{ij}, t_i) where $j = j_{\min}(i), ..., j_{\max}(i)$ for $i = 0, ..., n$.

Definition 6.23 *A numerical method of solving a stochastic differential equation converges weakly for a class of functions C if*

$$E_Q[g(s,t)] = \lim_{\Delta t \rightarrow 0} E_P(g_{ij})$$

for all $g \in C$, where E_P and E_Q are the discrete and continuous expectations under the probability distributions P_{ij} and $Q(s,t)$.

Conjecture 6.24 (Trinomial Tree) *The trinomial tree converges weakly to the solution of Equation (6.1) for the purpose of evaluating derivative securities.*

[Hull & White 1990] outline a proof by establishing the equivalence of the trinomial tree with a corresponding explicit finite difference method for solving a partial differential equation describing the value of a derivative security. The finite difference method is then assuredly consistent and stable as long as the branching probabilities $p_{ij}^{(u)}$, $p_{ij}^{(m)}$, and $p_{ij}^{(d)}$ are positive - and they are by Theorem (6.22).

6.1.4 Matching Futures Prices

Once the tree is ready, the spot price at node (i,j) is given by the inverse transformation

$$\begin{aligned} s_{ij} &= s(S_{ij}) \\ &= s\left(r_i + j\Delta S\right) \end{aligned}$$

The expected value of s at time $i\ \Delta t$ is given by

$$\varphi_i = \sum_{j=j_{\min}(i)}^{j_{\max}(i)} P_{ij} s_{ij} \tag{6.9}$$

Let $f(t)$ be the futures price of the energy in the market. Then the expected future spot price is related to the futures price by Equation (5.3)

$$\varphi(t) = f(t)$$

We suppose that the unknown function of time $\pi(t)$ dictates the shape of $\varphi(t)$. In the special case when $a[s, \pi(t)]$ is a linear function of s this can be done by a theoretical calculation. By Theorem (2.11) the expectation $\varphi(t)$ is the solution of the ordinary differential equation

$$\dot{\varphi}(t) = a[\varphi(t), \pi(t)]$$

Thus, the unknown function of time is found by solving for $\pi(t)$ in the equation $\dot{f}(t) = a[f(t), \pi(t)]$ where the derivative can be approximated by a finite difference method. This is guaranteed to work as long as a is a linear function of s. Indeed this is the case for all of the energy models we go on to study later[1]. The result is a particularly elegant tree with risk neutral probabilities.

6.2 Log Normal Distribution

We now apply our numerical methods to the specific example of a log normal distribution. This simple model will provide a stepping stone to the relatively complex mean reverting models that we use to describe energy prices. Furthermore, the unique structure of a trinomial tree with constant drift and constant volatility illuminates the

[1]That is, their drifts are linear before the transformation that renders the standard deviation constant, after which they become nonlinear, as we mentioned before. It is the original drift $a[s, \pi(t)]$ - not the transformed one $A[S, \pi(t)]$ - that is used in the calculation of the unknown function of time $\pi(t)$.

relationship between the deterministic and stochastic parts of a random process. This knowledge can be exploited to create trees that are a full order of magnitude more efficient than before.

Consider a generalized log normal random variable that satisfies the stochastic differential equation (2.6)

$$dl = \beta(t)l\ dt + \tau l\ dy$$

where $\beta(t)$ is the drift parameter (taken to be an unknown function of time), and τ is the volatility.

Recall that in the standard log normal distribution the drift parameter is a constant number. We will demonstrate our procedure by allowing $\beta(t)$ to be an unknown function of time. Put $\psi(t) = \bar{l}(t)$. Evidently we can choose $\beta(t)$ so that the expected value $\psi(t)$ of the random process $l(t)$ assumes any shape we desire. We begin by doing a log transformation so that the standard deviation is constant. Substitute $L = \ln(l)$. Then

$$dL = [\beta(t) - \frac{\tau^2}{2}]dt + \tau\ dy$$

We proceed to construct a trinomial tree for L. Specify $L_0 = \ln(l_0)$. Pick the length of a small time interval Δt. Then a small change in L is $\Delta L = \tau\sqrt{3\ \Delta t}$. Denote the median nodes in the tree for L by $\widehat{\Psi}_i$. Then the median nodes are calculated

recursively by $\widehat{\Psi}_0 = L_0$ and Formula (6.5), to yield

$$\widehat{\Psi}_i = \widehat{\Psi}_0 + \sum_{k=0}^{i-1}(\beta_k - \frac{\tau^2}{2})\Delta t$$

for $i = 1, ..., n$.

Since the drift $\beta(t) - \frac{\tau^2}{2}$ does not depend on L we have $M_{ij} = M_{i0} = (\beta_i - \frac{\tau^2}{2})\Delta t$ for all i and j. By Equation (6.6) the closest node to the drift emanating from the source node (i, j) is always $(i + 1, j)$ (in other words, $h_{ij} = j$ for all i and j) and the offset of every node in the tree is $\theta_{ij} = 0$. It follows that the branching probabilities are constant throughout the tree and we write

$$\begin{pmatrix} q_u \\ q_m \\ q_d \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix} \tag{6.10}$$

The structure of the L-tree can be visualized as a diffusion centered around the median nodes $\widehat{\Psi}_i$. This suggests that it would have been conceptually simpler to build a preliminary tree representing the diffusion and then transform to the final tree by shifting the median nodes of the branches onto the expected values of $L(t)$.

The first phase is to construct a *preliminary tree* for L by setting $\beta(t) = \frac{\tau^2}{2}$ and choosing an initial value of $L_0 = 0$. This is perfectly valid: we are in effect ignoring the deterministic part and concentrating on the stochastic part of the random process. The deterministic part can be easily injected back into the final tree by shifting the central nodes of each branch to the proper level. The actual process assumed for the

preliminary tree is therefore a simple diffusion.

$$dL = \tau \, dy$$

Put $\Psi(t) = \overline{L}(t)$. Observe that the preliminary tree has the special property that $\Psi(t) = 0$. In other words, the median nodes $\widehat{\Psi}_i$ are all zero. Also, the position of node (i, j) is $L_{ij} = j \ \Delta L$ and does not depend on i. Consequently the drifts M_{ij} and the branching probabilities q_u, q_m, and q_d do not depend on i. We will capitalize on this fact in a moment when we generalize this procedure. Figure (6.3) shows the decomposition into preliminary and final trees.

The final tree is formed by shifting the median nodes of each branch onto the expected value of $L(t)$. By Theorem (2.11) this expectation is

$$\Psi_i = L_0 + \int_0^{i \ \Delta t} [\beta(t) - \frac{\tau^2}{2}] dt$$

Then the position of node (i, j) in the final tree is $L_{ij} = \Psi_i + j \ \Delta L$. Evidently the median nodes $\widehat{\Psi}_i$ are the linear approximations of the expected values Ψ_i. Therefore the trinomial tree obtained by transforming the preliminary tree into the final one is exactly the same as before. Although this trick does not really improve the discretization of a log normal random variable, it will vastly enhance the efficiency of other types of Ito processes.

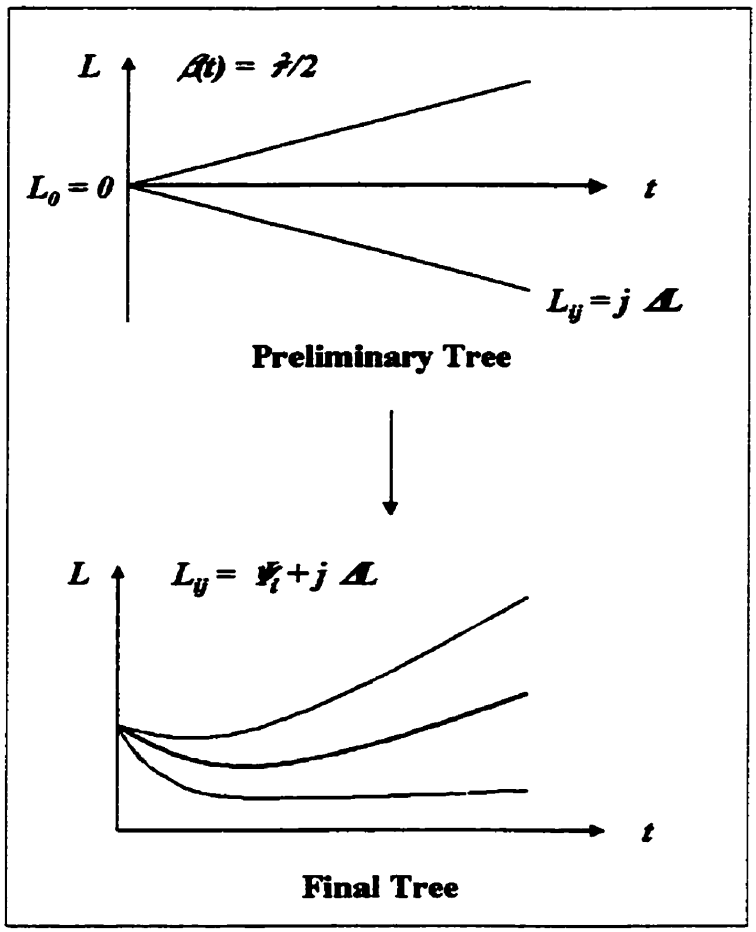

Figure 6.3: Log Normal Fast Tree Construction

6.3 Fast Trinomial Trees

We now extend the above procedure to the general case. We call this development the method of fast trinomial trees for reasons that will soon become apparent. These so-called fast trees originally appeared in [Hull & White 1994 (a)] and signaled a major breakthrough in the efficient construction of trinomial trees.

Suppose that a random variable $s(t)$ satisfies a stochastic differential equation of the form

$$ds = a[s, \pi(t)]dt + b\,(s)\,dz$$

and imagine that we have transformed to a new random variable $S(t)$ with constant standard deviation

$$dS = A[S, \pi(t)]dt + \sigma\ dz$$

We now introduce another random variable $\widehat{S}(t)$ that satisfies the Ito process

$$d\widehat{S} = A(\widehat{S}, \lambda)dt + \sigma\ dz \qquad (\widehat{S}_0 = 0) \tag{6.11}$$

Definition 6.25 *The ordinary trinomial tree for $\widehat{S}$ is called the preliminary tree for S.*

The preliminary tree turns out to have some extremely desirable properties, that make it unusually fast and easy to implement. Put $\widehat{\Phi}(t) = E[\widehat{S}(t) \mid \widehat{S}_0 = 0]$.

Then we are choosing the parameter $\pi(t)$ to be a number λ so that $\widehat{\Phi}(t) = 0$. This is accomplished by solving for the number λ in the equation

$$A(0, \lambda) = 0 \tag{6.12}$$

If a solution λ exists, we say that the random process $s(t)$ is preliminarizable. Hence our construction applies to the class of Ito processes whose transformed drifts A admit a solution of Equation (6.12).

Claim 6.26 *The expected value $\widehat{\Phi}(t)$ of the random process $\widehat{S}(t)$ is approximately zero.*

By Theorem (2.11), we have

$$\widehat{\Phi}(\Delta t) = \widehat{\Phi}_0 + \int_0^{\Delta t} E\{A[\widehat{S}(t), \lambda]\} dt$$

Now by choosing the preliminarization parameter according to Equation (6.12), we are assured that if we start at $\widehat{S}_0 = 0$, in which case $\widehat{\Phi}_0 = 0$, then over the next small time interval Δt we will arrive at $\widehat{\Phi}(\Delta t) = \mathrm{O}(\Delta t^2)$ by the Fundamental Theorem of Calculus. By resetting the expectation at time $t_1 = \Delta t$ to zero, that is, $\widehat{\Phi}(\Delta t) = 0$, we are committing an error of order $\mathrm{O}(\Delta t^2)$. Under the assumption $\widehat{\Phi}_1 = 0$, it then follows by an argument similar to the one we just made that $\widehat{\Phi}_2 = \mathrm{O}(\Delta t^2)$. By resetting the expectation at time $t_2 = 2\ \Delta t$ to zero, we commit another error of order $\mathrm{O}(\Delta t^2)$. Continuing in this fashion, we have a total error of $n\mathrm{O}(\Delta t^2) = \mathrm{O}(\Delta t)$. This heuristic

construction supports the fact that calculating λ by Equation (6.12) means that $\widehat{\Phi}(t)$ is approximately zero, as required. This is important, as the median nodes of the preliminary tree are all zero by Equation (6.5), and these median nodes will later be moved onto the expected values $\Phi(t)$ of the actual process $S(t)$, when the final tree is formed.

Definition 6.27 *The **final tree**, also called the **fast tree**, for $S(t)$ is obtained by retaining all of the relative connections and branching probabilities of the preliminary tree, and merely shifting all median nodes $\widehat{\Phi}_i$ of the preliminary tree onto the expected values Φ_i of the random process $S(t)$ at the discrete times $t_i = i\ \Delta t$ for $i = 0, ..., n$, maintaining the vertical spacing between all nodes in the same branch.*

The final tree is found by moving all branches so that their median nodes coincide with the expected values of $S(t)$. Then the position of node (i, j) in the final tree is $S_{ij} = \Phi_i + j\ \Delta S$ Note that in practice it is not necessary to solve the analytic form of the expectation $\Phi(t)$. Rather, the numbers $\Phi(t)$ are calculated by Equation (6.9) so that the futures prices $f(t)$ coincide with the expected value of $s(t)$ at every branch in the tree.

Figure (6.4) shows the geometry of a preliminary tree. Observe that the position of all nodes in the preliminary tree depend only on j (and not on i). The final tree is obtained by moving the median nodes of each branch so that the expected values of the spot prices satisfy the zero net present value condition (1). The geometry of the

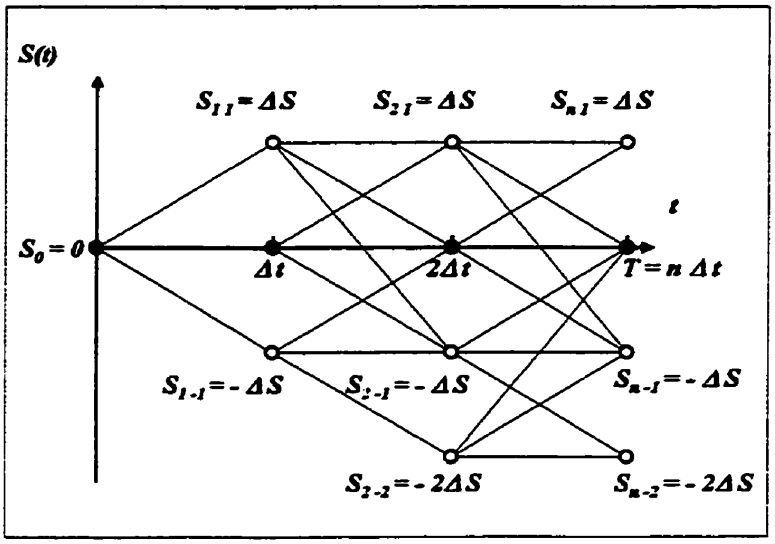

Figure 6.4: Preliminary Trinomial Tree Structure

final tree looks the same as that for an ordinary tree depicted in Figure (6.1) except that the symbol $\widehat{\Phi}$ is replaced by Φ. This is because the median nodes in the final tree resulting from the fast construction are located at precisely the expected values Φ_i by definition. This is another aspect that makes fast trees more desirable than ordinary ones.

The position of node (i, j) is $\widehat{S}_{ij} = j\ \Delta S$ and does not depend on i. Consequently the drifts M_j and the branching probabilities $p_j^{(u)}$, $p_j^{(m)}$, and $p_j^{(d)}$ do not depend on i either. This means that we only need to calculate $J_{\max} - J_{\min}$ sets of probabilities where $J_{\max} = \max_{i\in\{0,\ldots,n\}} j_{\max}(i)$ and $J_{\min} = \min_{i\in\{0,\ldots,n\}} j_{\min}(i)$. Hence the computer programs for fast trinomial trees are guaranteed to run faster than those for ordinary trees. We will do an experimental analysis of the run times of ordinary and fast trees in Chapter (8).

Conjecture 6.28 (Fast Trinomial Tree) *The fast trinomial tree converges weakly to the solution of Equation (6.1) for the purpose of evaluating derivative securities.*

[Hull & White 1994 (a)] outline a proof in the special case when $A[S, \theta(t)]$ is a linear function of S. Consider the drift of the median nodes at time $i\ \Delta t$ in the final tree

$$\Phi_{i+1} - \Phi_i = A(\Phi_i, \pi_i)\Delta t$$

Dividing both sides of this equation by Δt and taking the limit as $\Delta t \rightarrow 0$ reveals

$$\dot{\Phi}(t) = A[\Phi(t), \pi(t)]$$

and this equation is true when $A[S, \theta(t)]$ is linear in S by Theorem (2.11).

In fact the transformed drift functions we apply the fast trinomial tree method to are nonlinear. We offer an experimental proof of the conjecture for the class of Pilipovic equations by evaluating derivative securities using fast and ordinary trinomial trees and demonstrating that the results are approximately equivalent.

6.4 3 Dimensional Trees

We now introduce a method for creating a 3D trinomial tree that approximates the solution of a system of two correlated stochastic differential equations. This breakthrough idea was proposed by [Hull & White 1994 (b)].

Suppose we have a decoupled system of stochastic differential equations of the form

$$\begin{aligned} dR &= A(R,t)dt + \sigma\, dz \\ dU &= B(U,t)dt + \tau\, dy \end{aligned}$$

where the Brownian motions $z(t)$ and $y(t)$ have correlation ρ. The requirement that the equations be decoupled is essential as it allows for the construction of trinomial trees for each random variable alone (as though the other one did not exist).

Suppose that we have constructed ordinary trinomial trees for the random variables $R(t)$ and $U(t)$. Denote nodes in the R-tree by (i,j) and those in the U-tree by (i,k). Also, write the branching probabilities in the R-tree by $p_{ij}^{(u)}$, $p_{ij}^{(m)}$, and $p_{ij}^{(d)}$ and those in the U-tree by $q_{ik}^{(u)}$, $q_{ik}^{(m)}$, and $q_{ik}^{(d)}$. Then it is possible to interweave the separate trees into a single 3D tree so that the desired correlation is induced. We review this construction next.

6.4.1 Induce the Correlation

Now imagine creating a 3D tree by combining the separate trees for R and U[2]. Each pair of nodes (i,j) and (i,k) give rise to a node (i,j,k) in the (R,U)-tree. Suppose that we are at node (i,j,k). Then there are 9 possible nodes that can be branched into. These are obtained by combining all possible pairs of branchings in the individual

[2]In essence, we are taking the direct product of two trinomial trees.

trinomial trees. If the random processes were independent of one another the branching probability matrix G would be calculated by simply multiplying the corresponding probabilities together

$$G = \begin{pmatrix} p_u q_d & p_u q_m & p_u q_u \\ p_m q_d & p_m q_m & p_m q_u \\ p_d q_d & p_d q_m & p_d q_u \end{pmatrix} \tag{6.13}$$

where we have neglected to indicate the dependence of G on i, j, and k for the sake of brevity.

We now examine the adjustments necessary to induce the desired correlation. If the assets are positively correlated ($\rho > 0$) then the correct branching probability matrix is obtained by adjusting the probabilities in the following way

$$G = \begin{pmatrix} p_u q_d - \varepsilon & p_u q_m - 4\varepsilon & p_u q_u + 5\varepsilon \\ p_m q_d - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_u - 4\varepsilon \\ p_d q_d + 5\varepsilon & p_d q_m - 4\varepsilon & p_d q_u - \varepsilon \end{pmatrix} \tag{6.14}$$

where $\varepsilon = \frac{\rho}{36}$.

Note that the probability adjustments add up to zero along each row and column. This ensures that the means and variances of the individual trees for S and U are not affected. The only visible effect is to induce a correlation of exactly ρ between the two assets. Furthermore, if $\rho = 1$ then by Equations (6.7) and (6.14) we would observe the

following effect

$$\lim_{\Delta t \to 0} G = \begin{pmatrix} 0 & 0 & \frac{1}{6} \\ 0 & \frac{2}{3} & 0 \\ \frac{1}{6} & 0 & 0 \end{pmatrix}$$

This is precisely the kind of behavior we would expect if the assets were perfectly positively correlated.

We now state and prove a theorem from [Hull & White 1994 (b)] in which we make use of some elementary laws of probability which can be found in any fundamental text on probability theory. For example, see [Grimmett & Stirzaker 1995].

Theorem 6.29 *If G is given by (6.14) with $\varepsilon = \frac{\rho}{36}$ positive then $\rho = corr(S, U)$.*

Proof. Let the matrix of probability adjustments be denoted by

$$\begin{aligned} \epsilon &= (\varepsilon_{mn}) \\ &= \begin{pmatrix} -\varepsilon & -4\varepsilon & 5\varepsilon \\ -4\varepsilon & 8\varepsilon & -4\varepsilon \\ 5\varepsilon & -4\varepsilon & -\varepsilon \end{pmatrix} \end{aligned}$$

Let the median nodes of the trees for R and U be denoted by $\widehat{\Phi}_i$ and $\widehat{\Theta}_i$. Also, let the closest nodes to the drifts at nodes (i, j) and (i, k) be written as h_{ij} and g_{ik}. Suppose we are at node (i, j, k) in the (R, U)-tree. Then over the next interval of time

the covariance is given by

$$\begin{aligned}
\operatorname{cov}(R,U) &= E(RU) - E(R)\,E(U) \\
&= \sum_{m=-1}^{1}\sum_{n=-1}^{1}(p_{ij}^{(m)}q_{ik}^{(n)} + \varepsilon_{mn})R_{i+1,h_{ij}+m}U_{i+1,g_{ik}+n} - E(R)\,E(U) \\
&= \sum_{m=-1}^{1}\sum_{n=-1}^{1}(p_{ij}^{(m)}q_{ik}^{(n)} + \varepsilon_{mn})[\widehat{\Phi}_{i+1} + (h_{ij}+m)\Delta R] \\
&\qquad [\widehat{\Theta}_{i+1} + (g_{ik}+n)\Delta R] - E(R)\,E(U)
\end{aligned}$$

Recall that the branching probabilities $p_{ij}^{(m)}$ and $q_{ik}^{(n)}$ add up to one while the probability adjustments ε_{mn} add up to zero along any row or column. Hence

$$\begin{aligned}
\operatorname{cov}(R,U) &= \sum_{m=-1}^{1}[\widehat{\Phi}_{i+1} + (h_{ij}+m)\Delta R]\sum_{n=-1}^{1}[\widehat{\Theta}_{i+1} + (g_{ik}+n)\Delta R] + \\
&\qquad \sum_{m=-1}^{1}\sum_{n=-1}^{1}\varepsilon_{mn}\,(h_{ij}+m)\,\Delta R\,(g_{ik}+n)\,\Delta U - E(R)\,E(U) \\
&= E(R)\,E(U) - E(R)\,E(U) + \\
&\qquad \Delta R\,\Delta U\sum_{m=-1}^{1}\sum_{n=-1}^{1}\varepsilon_{mn}\,(h_{ij}g_{ik} + h_{ij}n + mg_{ik} + mn) \\
&= \Delta R\,\Delta U\sum_{m=-1}^{1}\sum_{n=-1}^{1}\varepsilon_{mn}\,(mn) \\
&= \Delta R\,\Delta U\,(5+1+1+5)\,\varepsilon \\
&= 12\varepsilon\,\Delta R\,\Delta U
\end{aligned}$$

This clearly demonstrates that $\operatorname{cov}(R,U)$ is independent of the node (i,j,k). Consequently the correlation of R and U is also independent of our position in the tree. Indeed this is a well known property of a system of stochastic differential equations with constant standard deviations.

The standard deviations of R and U over an interval of time are respectively $\sigma\sqrt{\Delta t}$ and $\tau\sqrt{\Delta t}$. Thus the correlation we desire is

$$\begin{aligned}\rho &= \frac{\mathrm{cov}\,(R,U)}{\sigma\tau\ \Delta t} \\ &= \frac{12\varepsilon\ \Delta R\ \Delta U}{\sigma\tau\ \Delta t} \\ &= 36\varepsilon\end{aligned}$$

■

Now suppose that the assets are negatively correlated ($\rho < 0$). Then the correct matrix is given by

$$G = \begin{pmatrix} p_u q_d + 5\varepsilon & p_u q_m - 4\varepsilon & p_u q_u - \varepsilon \\ p_m q_d - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_u - 4\varepsilon \\ p_d q_d - \varepsilon & p_d q_m - 4\varepsilon & p_d q_u + 5\varepsilon \end{pmatrix} \tag{6.15}$$

where $\varepsilon = \frac{|\rho|}{36}$. The proof of this result is similar to Theorem (6.29).

Similarly if $\rho = -1$ then Equations (6.7) and (6.15) would reveal

$$\lim_{\Delta t \to 0} G = \begin{pmatrix} \frac{1}{6} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \\ 0 & 0 & \frac{1}{6} \end{pmatrix}$$

Again this satisfies our natural intuition of the way that perfectly negatively correlated assets are supposed to behave.

Equations (6.14) and (6.15) also induce the required correlation between a pair of fast trees for $R(t)$ and $U(t)$ by proofs similar to Theorem (6.29), except that the

probability matrix G ceases to depend on i, and depends only on j and k. Hence we have a nice and neat 3D tree that accurately predicts the behavior of two correlated random variables simultaneously. We are now in a position to evaluate spread options among other things.

CHAPTER 7
MEAN REVERTING MODELS

Energy spot prices fluctuate randomly over time. In fact the spot prices are conveniently modeled by systems of stochastic differential equations. [Pilipovic 1997] has proposed a number of models that might be appropriate for the unique behavior of energy commodities. As it happens, these equations are too complex to be solved analytically. However, the continuous processes can be implemented numerically as trinomial trees with great success.

We now develop algorithms for constructing ordinary and fast trinomial trees for the mean reverting models put forward in [Pilipovic 1997]. In Section (6.3) we constructed fast trinomial trees that were more elegant than the ordinary trinomial trees we developed in Section (6.1). We intend to show numerically that the fast trinomial trees can be successfully used to evaluate derivative securities. We will do this by comparing results obtained using both fast and ordinary trees and demonstrating that their differences are within acceptable tolerance levels. This will justify the use of fast trinomial trees for the purpose of option evaluation.

7.1 1-Factor Model

We begin by dissecting the simplest variation of the Pilipovic equation; namely, the 1-factor model. This will pave the way toward understanding the relatively complex 2-factor models we will study in Sections (7.2) and (7.3).

Consider the 1-factor generalized Pilipovic equation (3.2)

$$ds = \alpha\,[l(t) - s]\,dt + \sigma s\;dz$$

In the standard 1-factor Pilipovic equation l is a constant parameter. In order to more accurately capture economic trends we allow $l(t)$ to be an unknown function of time. The parameter $l(t)$ will be chosen to make the model consistent with the term structure of futures prices observed in the market. This scheme also has the additional benefit of greatly enhancing the efficiency of the numerical implementation when we switch from ordinary to fast trees.

The spot price s is the average daily price of an energy commodity such as natural gas or oil. The long term mean $l(t)$ represents the economic trend of the market. Although it is unobservable, this quantity will be chosen so that the expected value of the model coincides with the futures prices agreed upon by a conglomeration of buyers and sellers. This, along with the condition that the market price of risk is zero, will ensure that the probabilities are risk neutral. In other words, we will arrange the geometry of the tree so that the central node of each branch always corresponds to the expected value of s as dictated by popular market opinion. This leads to eminently faster tree construction and significantly more accurate option evaluation. The parameters α and σ are estimated using historical data as explained in Section (4.1).

We begin by constructing the ordinary trinomial tree and then proceed to develop the corresponding procedure for a fast trinomial tree.

7.1.1 Log Transformation

We begin by doing a log transformation so that the standard deviation in the stochastic differential equation (3.2) is constant.

$$L(t) = \ln[l(t)] \qquad S = \ln(s)$$

By the Ito formula we have

$$dS = [\alpha(e^{L(t)-S} - 1) - \frac{\sigma^2}{2}]dt + \sigma\, dz \tag{7.1}$$

7.1.2 Separation of Deterministic and Stochastic Parts

It is now convenient to substitute another random variable that will enable us to easily separate the deterministic and stochastic parts of the random process.

$$R = L(t) - S \tag{7.2}$$

Put $\phi(t) = \overline{s}(t)$, $\Phi(t) = \overline{S}(t)$, and $\omega(t) = \overline{R}(t)$. Taking the expected value of Equation (7.2) we obtain the expectation relation $L(t) = \omega(t) + \Phi(t)$.

Then by the Higher Dimensional Ito Formula[1]

$$dR = \alpha[a(t) - e^R]dt + \sigma\, dz \tag{7.3}$$

[1]Note that the actual process obtained is $dR = \alpha[a(t) - e^R]dt - \sigma\, dz$; however, rather than introducing another Brownian motion to replace $-z$ (which would complicate notation) we have simply taken advantage of the fact that the probability distribution of $-z$ is indistinguishable from that of z.

where $a(t)$ is given by

$$a(t) = 1 + \frac{\dot{L}(t) + \frac{\sigma^2}{2}}{\alpha}$$

Observe that we now have an exponentially mean reverting process. That is e^R is bound to inexorably drift toward $a(t)$ at a mean reversion strength of α. At this point we mention that the analytic form of $a(t)$ as it involves $\dot{L}(t)$ is unimportant. The unknown function of time $l(t)$ will be revealed as the numerical solution of an ordinary differential equation derived under the assumption that the futures prices $f(t)$ match the expected values of $s(t)$ at every branch in the tree.

Using Equation (4.21) with $\varphi(t)$ replaced by $f(t)$ we obtain

$$l(t) = f(t) + \frac{\dot{f}(t)}{\alpha} \tag{7.4}$$

Recall that the futures prices in the energy market are quoted on a monthly basis. Hence it is necessary to interpolate missing values of $f(t)$. This is easily done using a cubic spline. Next, finite difference formulas similar to the ones mentioned in Subsection (4.1.1) yield $\dot{f}(t)$ [2]at all discrete times $t_i = i\ \Delta t$ for $i = 0, ..., n$. Finally, the required derivative in the exponential mean reversion level $a(t)$ is

$$\dot{L}(t) = \frac{\dot{l}(t)}{l(t)} \tag{7.5}$$

Recall that an attractive feature of the trinomial tree was the fact that the drift of the median node at time $i\ \Delta t$ landed precisely on the median node of the next branch

[2]Whenever derivatives appear in any formula in this thesis, they are understood to be obtained by finite difference methods, for numerical implementation purposes.

at time $(i+1)\Delta t$. This attribute is desirable from an algorithmic point of view and we desire to capture it here. To this end, we denote the median nodes by $\widehat{\omega}_i = R_{i0}$ and their drifts by M_{i0}. The median nodes are calculate by Equation (6.5)

$$\widehat{\omega}_i = \widehat{\omega}_0 + \sum_{k=0}^{i-1} M_{i0}$$

By Theorem (2.12) the mean and variance of $R(t+\Delta t) - R(t)$ are approximately

$$M_{ij} = \alpha(a_i - \mathrm{e}^{\widehat{\omega}_i + j\,\Delta R})\Delta t \text{ and } V = \sigma^2 \Delta t$$

at node (i,j) in the trinomial tree for R, defined by Equation (7.3).

The tree construction begins by selecting an initial value for R_0. Then $\widehat{\omega}_0 = R_0$ and the median nodes are precisely connected by their drifts according to Equation (6.5). Then, the branching probabilities are calculated by Equation (6.7).

This completely specifies the ordinary tree construction. Once the tree is ready, the spot price at node (i,j) is given by the inverse transformation

$$\begin{aligned} s_{ij} &= \mathrm{e}^{L_i - (\widehat{\omega}_i + j\Delta R)} \\ &= l_i \mathrm{e}^{-(\widehat{\omega}_i + j\Delta R)} \end{aligned} \tag{7.6}$$

where the parameter $l(t)$ has been calculated by Equation (7.4) so that the probabilities are risk neutral. Thus the tree can be used to evaluate options. We summarize the construction process of the 1-factor model (3.2) for the sake of clarity.

Algorithm 7.30 (1-Factor Model Ordinary Trinomial Tree) *1. Specify the parameters α, σ, T, and n and define the futures prices $f(t)$ for $t \in [0,T]$.*

2. Transform to a new variable $R(t)$ that satisfies the stochastic differential equation (7.3).

3. Set $\Delta t = \frac{T}{n}$ and $\Delta R = \sigma\sqrt{3\ \Delta t}$.

4. Calculate the unknown function of time $l(t)$ by solving the ordinary differential equation (7.4).

5. Pick a value for s_0. Then $R_0 = \ln(\frac{l_0}{s_0})$ and the median node at time 0 is $\widehat{\omega}_0 = R_0$. Also, put $j_{\max}(0) = j_{\min}(0) = 0$.

6. For $i = 0, ..., (n-1)$

(a) For $j = j_{\min}(i), ..., j_{\max}(i)$

Calculate the branching probabilities $\mathbf{p}_{ij}$ by Equation (6.7).

End For

(b) Calculate $\widehat{\omega}_{i+1}$ by Equation (6.5).

(c) Calculate $j_{\max}(i+1)$ and $j_{\min}(i+1)$ by Equation (6.8).

End For

7. Calculate s_{ij} by Equation (7.6).

Next we consider the construction of the fast trinomial tree for the same random process (3.2).

7.1.3 Preliminary Tree

We follow the construction of Section (6.3). The first phase is to construct a preliminary tree for a noise variable $\widehat{R}$ by setting $a(t) = 1$ and choosing an initial value of $\widehat{R}_0 = 0$. This is perfectly valid: we are in effect ignoring the deterministic part and concentrating on the stochastic part of the random process. The deterministic part is then easily injected back into the final tree by shifting the central nodes of each branch to the proper level. The actual process assumed for the preliminary tree is therefore

$$d\widehat{R} = \alpha(1 - e^{\widehat{R}})dt + \sigma\, dz \tag{7.7}$$

Recall that a stochastic differential equation is made up of a deterministic part and a stochastic part. The essences of these components are captured by the mean and variance of the probability distribution that is the solution of the stochastic differential equation. By forcing our trinomial tree to match the mean and variance of the solution we guarantee that our discrete scheme will converge to the continuous process.

By Theorem (2.12) the mean and variance of $\widehat{R}(t+\Delta t) - \widehat{R}(t)$ are approximately

$$M_j = \alpha(1 - e^{j\,\Delta R})\Delta t \text{ and } V = \sigma^2 \Delta t$$

at node (i, j) in the trinomial tree for $\widehat{R}$ (7.7). Note that this is a linear approximation for the mean and variance. Figure (7.1) shows the preliminary and final trees for a mean reverting random variable.

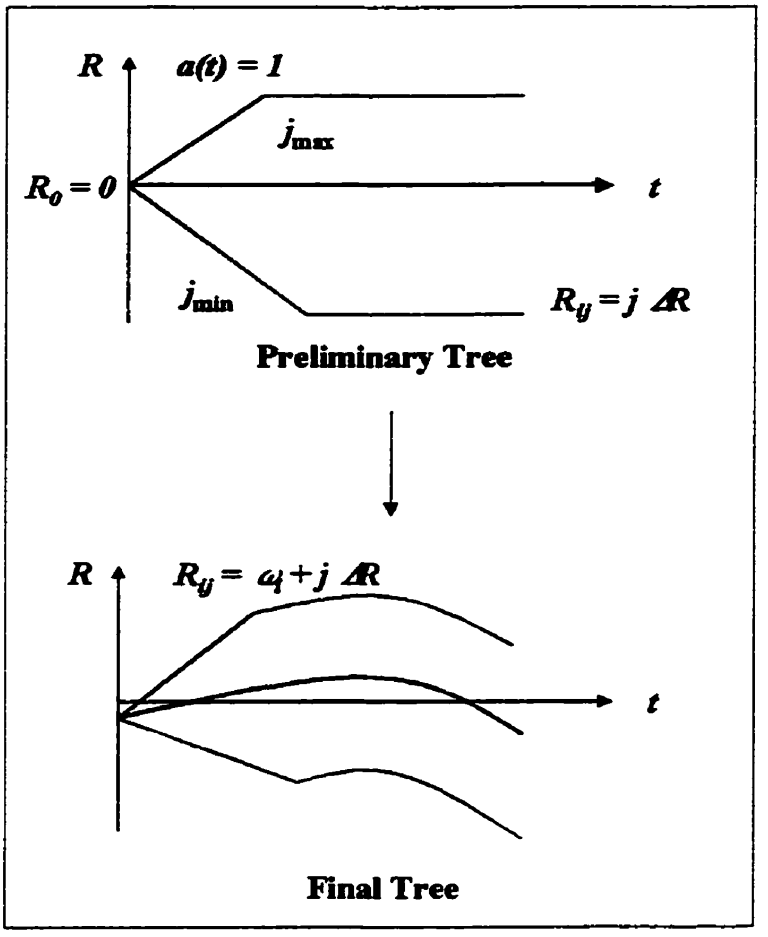

Figure 7.1: Mean Reverting Fast Tree Construction

Further, by rendering $a(t)$ constant in the preliminary tree, we have ensured that the drift M_j depends only on j and not on i. Now the median node in the branch at time $i\ \Delta t$ is assumed to be located at $\widehat{R}_{i0} = 0$, and in general, $\widehat{R}_{ij} = j\ \Delta R$ in the preliminary tree. In other words, we suppose that the expected value of $\widehat{R}(t)$ vanishes. However, this is merely an approximation and by Claim (6.26) the deviation of $\omega(t)$ from zero is negligible. We recall that this is a sufficient condition the process must satisfy for this construction to work.

7.1.4 Branching Probabilities

As it happens, following [Hull & White 1994 (a)] only three kinds of branching processes need to be considered for the mean reverting models we consider. Branching process A will dominate in the interior of the tree while branching processes B and C designate the boundary of mean reversion and ensure that the process is not allowed to drift too far away from the line of attraction. Figure (7.2) illustrates these alternatives.

Our objective now is to construct a trinomial tree that mimics the behavior of the random process. This entails deciding which of the three alternative branching processes will apply at each node. Then the branching probabilities must be calculated. This is done by matching the mean and variance of the small change in R over the next time interval Δt. Together with the well known fact the probabilities must add up to one gives us a total of three equations in the three probabilities.

The mean reversion of the random process dictates that at some point the drift

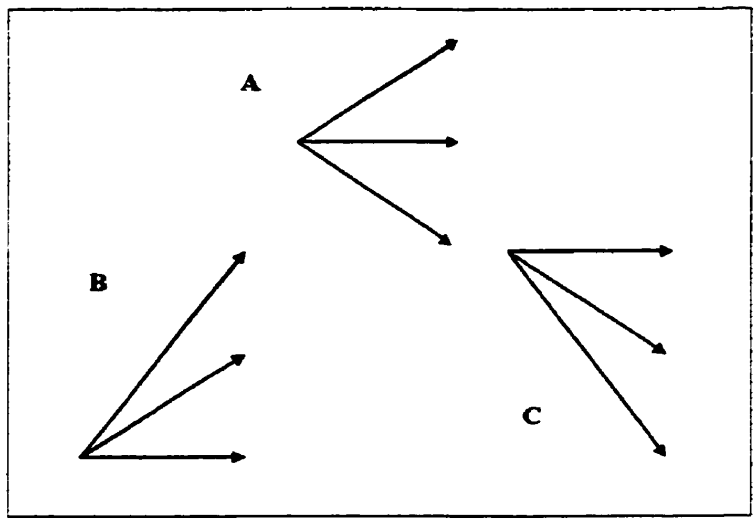

Figure 7.2: Alternative Branching Processes

of R toward 0 will cause the branching process to shift inward. From this time onward, the nodes in the tree will be forever trapped inside a band of mean reversion. Define $j_{\max}$ as the value of j where we switch from branching process A to branching process C and $j_{\min}$ as the value of j where we switch from branching process A to branching process B.

$$\begin{aligned} j_{\max} &= \left\lceil \frac{1}{\Delta R} \ln(1 + \frac{\Delta R}{2\alpha\ \Delta t}) \right\rceil \\ j_{\min} &= \left\lfloor \frac{1}{\Delta R} \ln(1 - \frac{\Delta R}{2\alpha\ \Delta t}) \right\rfloor \end{aligned} \tag{7.8}$$

Theorem 7.31 *The maximum and minimum values of j obtained by a mean reverting process obeying (7.7)*

$$d\widehat{R} = \alpha(1 - e^{\widehat{R}})dt + \sigma\ dz$$

are given by Equation (7.8).

Proof. We consider first the case when $j > 0$. In this case the drift is negative - the process is inevitably drawn back toward the line of attraction. Suppose we are at node (i, j). The switch from branching process A to C occurs when the drift is closer to node $(i+1, j-1)$ than $(i+1, j)$

$$\begin{aligned} -M_j &> \frac{\Delta R}{2} \\ \alpha(e^{j\,\Delta R} - 1)\Delta t &> \frac{\Delta R}{2} \\ j &> \frac{1}{\Delta R}\ln(1 + \frac{\Delta R}{2\alpha\ \Delta t}) \end{aligned}$$

Therefore

$$j_{\max} = \left\lceil \frac{1}{\Delta R}\ln(1 + \frac{\Delta R}{2\alpha\ \Delta t}) \right\rceil$$

Now assume that $j < 0$. Then the drift is positive and if we are at node (i, j) then the switch from branching process A to B occurs when the drift is closer to node $(i+1, j+1)$ than $(i+1, j)$

$$\begin{aligned} M_j &> \frac{\Delta R}{2} \\ \alpha(1 - e^{j\,\Delta R})\Delta t &> \frac{\Delta R}{2} \\ j &< \frac{1}{\Delta R}\ln(1 - \frac{\Delta R}{2\alpha\ \Delta t}) \end{aligned}$$

Thus

$$j_{\min} = \left\lfloor \frac{1}{\Delta R}\ln(1 - \frac{\Delta R}{2\alpha\ \Delta t}) \right\rfloor$$

■

The probabilities associated with the three possible branching processes are given by substituting the appropriate values of h in (6.7). If we are at an interior node, $h = j$ and branching process A is adopted with branching probabilities calculated by

$$\begin{aligned} p_u &= \frac{1}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} + \frac{M_j}{\Delta R}\right) \\ p_m &= \frac{2}{3} - \frac{M_j^2}{\Delta R^2} \\ p_d &= \frac{1}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} - \frac{M_j}{\Delta R}\right) \end{aligned} \tag{7.9}$$

Similarly if we are at the lower bound of the band of mean reversion, $h = j + 1$ and branching process B is adopted with branching probabilities calculated by

$$\begin{aligned} p_u &= \frac{1}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} - \frac{M_j}{\Delta R}\right) \\ p_m &= -\frac{1}{3} - \frac{M_j^2}{\Delta R^2} + 2\frac{M_j}{\Delta R} \\ p_d &= \frac{7}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} - 3\frac{M_j}{\Delta R}\right) \end{aligned} \tag{7.10}$$

Finally if we are at the upper bound of the band of mean reversion, $h = j - 1$ and branching process C is adopted with branching probabilities calculated by

$$\begin{aligned} p_u &= \frac{7}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} + 3\frac{M_j}{\Delta R}\right) \\ p_m &= -\frac{1}{3} - \frac{M_j^2}{\Delta R^2} - 2\frac{M_j}{\Delta R} \\ p_d &= \frac{1}{6} + \frac{1}{2}\left(\frac{M_j^2}{\Delta R^2} + \frac{M_j}{\Delta R}\right) \end{aligned} \tag{7.11}$$

Recall that the probabilities depend only on j. Hence the probabilities at every node in the tree are specified by calculating a total of $(j_{\max} - j_{\min} + 1)$ sets of

probabilities!

This completes the construction of the preliminary tree. It merely remains to choose the right values for $l(t)$. Then the preliminary tree will be shifted onto the final tree.

7.1.5 Final Tree

The final tree is formed by shifting the median nodes of the preliminary tree onto the expected values of the random process $R(t)$. To this end the nodes in the branch at time $i\ \Delta t$ are displaced by an amount ω_i equal to the expected value of $R(t)$ at time t_i. The position of node (i, j) in the final tree is therefore

$$R_{ij} = \omega_i + j\ \Delta R \tag{7.12}$$

Therefore, the spot price at node (i, j) is given by the inverse transformation

$$\begin{aligned} s_{ij} &= e^{L_i - (\omega_i + j\Delta R)} \\ &= e^{\Phi_i - j\Delta R} \end{aligned} \tag{7.13}$$

We assumed that $\omega(t)$ was (effectively) zero for the purpose of constructing the preliminary R-tree, but this is not (generally) true. In fact it is the solution of an ordinary integral equation involving the unknown function of time $l(t)$. The expected

value of s at time $i\,\Delta t$ is given by

$$\begin{aligned} \varphi_i &= \sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} P_{ij} s_{ij} \\ &= e^{\Phi_i} \sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} P_{ij} e^{-j\Delta R} \end{aligned} \tag{7.14}$$

where P_{ij} is the probability of reaching node (i,j). These cumulative probabilities are easily calculated as follows, provided we know the branching probability at each node in the tree.

$$\begin{aligned} P_{00} &= 1 \\ P_{ij} &= \sum_m P_{i-1\ m} q[(i-1,m) \rightarrow (i,j)] \end{aligned} \tag{7.15}$$

for $i = 1, ..., n$ and $j = \max(-i, j_{\min}), ..., \min(j_{\max}, i)$ where $q[(i-1,m) \rightarrow (i,j)]$ is the probability of branching from node $(i-1,m)$ to node (i,j). It remains to compute Φ_i by an equation given in the next subsection.

7.1.6 Matching Futures Prices

Let $f(t)$ be the futures price of the energy commodity in the market. Then the expected future spot price is related to the futures price by Equation (5.3)

$$\varphi(t) = f(t)$$

The unknown parameter Φ_i is found by setting the futures price equal to the

expected future spot price in Equation (7.14) and solving

$$\Phi_i = \ln\left(\frac{f_i}{\sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} P_{ij} e^{-j\Delta R}}\right) \tag{7.16}$$

This ends the construction of the fast trinomial tree for the 1-factor model (3.2), which we summarize next.

Algorithm 7.32 (1-Factor Model Fast Trinomial Tree) *1. Specify the parameters α, σ, T, and n and define the futures prices $f(t)$ for $t \in [0, T]$.*

2. Transform to a new variable $R(t)$ that satisfies the stochastic differential equation (7.3).

3. Set $\Delta t = \frac{T}{n}$ and $\Delta R = \sigma\sqrt{3\,\Delta t}$.

4. Calculate $j_{\max}$ and $j_{\min}$ by Equation (7.8).

5. For $j = j_{\min}, \ldots, j_{\max}$

Calculate the branching probabilities $\mathbf{p}_j$ by Equation (6.7).

End For

6. Set $P_{00} = 1$.

7. For $i = 0, \ldots, (n-1)$

(a) Calculate P_{ij} by Equation (7.15).

(b) Calculate Φ_i by Equation (7.16).

End For

8. Calculate s_{ij} by Equation (7.13).

7.2 2-Factor (a) Model (Two Correlated Assets)

It is often advantageous to model two correlated assets simultaneously. This enables one to evaluate such exotic financial derivatives as portfolios and spread options. We will do this by constructing separate trinomial trees for each of the two assets. Then we will combine them into a single 3D tree. The desired correlation will be induced by a neat mathematical trick In this way, the evolution of both random processes can be modelled at once.

Consider a pair of correlated 1-factor energy prices satisfying the 2-factor (a) model

$$\begin{aligned} ds &= \alpha\,[l(t) - s]\,dt + \sigma s\; dz \\ dv &= \beta\,[m(t) - v]\,dt + \tau v\; dy \end{aligned}$$

Let the correlation between the two related assets be written as $\rho = corr(z, y)$. Then we have a decoupled system of stochastic differential equations and we follow the procedure outlined in Section (6.4). Again we start with an ordinary trinomial tree and then move to the fast one.

7.2.1 Construct Separate Trees

We begin by doing log transformations on the random processes to render their standard deviations constant. The substitution for the first asset is the same as before

$$L(t) = \ln[l(t)] \qquad S = \ln(s)$$

while the substitution for the second asset is

$$M(t) = \ln[m(t)] \qquad V = \ln(v)$$

Similarly we introduce new random variables by the substitutions

$$R = L(t) - S \qquad U = M(t) - V$$

to facilitate tree construction. Then the system of stochastic differential equations for R and U is

$$\begin{aligned} dR &= \alpha[a(t) - \mathrm{e}^R]dt + \sigma\, dz \\ dU &= \beta[b(t) - \mathrm{e}^U]dt + \tau\, dy \end{aligned}$$

where

$$a(t) = 1 + \frac{\dot{L}(t) + \frac{\sigma^2}{2}}{\alpha} \qquad b(t) = 1 + \frac{\dot{M}(t) + \frac{\tau^2}{2}}{\beta}$$

We observe that the correlation between R and U is given by ρ as the Brownian motions $z(t)$ and $y(t)$ of the stochastic differential equations for R and U are the same as those for S and V.

The ordinary trees for R and U are constructed according to the procedure described in Section (7.1). These trees are formed individually, as though the other random process in the system were not present. Let the nodes in R-tree be denoted by (i, j) as before and let the nodes in the U-tree be denoted by (i, k). Note that the i index represents time for both assets. The length of a time interval Δt is the same for both trees. The length of a small change in R is $\Delta R = \sigma\sqrt{3\ \Delta t}$ and the length of a small change in U is $\Delta U = \tau\sqrt{3\ \Delta t}$. Furthermore we denote the branching probabilities in the R-tree by $p_{ij}^{(u)}$, $p_{ij}^{(m)}$, and $p_{ij}^{(d)}$ and those in the U-tree by $q_{ik}^{(u)}$, $q_{ik}^{(m)}$, and $q_{ik}^{(d)}$.

Let the futures prices of $s(t)$ and $v(t)$ be respectively written as $f\,(t)$ and $g\,(t)$. The unknown functions of time $l\,(t)$ and $m\,(t)$ are obtained by matching the futures prices $f\,(t)$ and $g\,(t)$ for each asset separately. This is done by solving the system of ordinary differential equations

$$\begin{aligned} l(t) &= f(t) + \frac{\dot{f}(t)}{\alpha} \\ m(t) &= g(t) + \frac{\dot{g}(t)}{\beta} \end{aligned}$$

7.2.2 Induce the Correlation

The complete 3D tree for R and U is created in precisely the same way as we did in Subsection (7.4.1). Each pair of nodes (i,j) and (i,k) give rise to a node (i,j,k) in the (R,U)-tree. Suppose that we are at node (i,j,k). Then there are 9 possible nodes that can be branched into. These are obtained by combining all possible pairs of branchings in the individual trinomial trees. By Theorem (6.29) it follows that for $\rho > 0$ the branching probability matrix is given by Equation (6.14)

$$G = \begin{pmatrix} p_u q_d - \varepsilon & p_u q_m - 4\varepsilon & p_u q_u + 5\varepsilon \\ p_m q_d - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_u - 4\varepsilon \\ p_d q_d + 5\varepsilon & p_d q_m - 4\varepsilon & p_d q_u - \varepsilon \end{pmatrix}$$

where $\varepsilon = \frac{\rho}{36}$ and for $\rho < 0$ we have the matrix (6.15)

$$G = \begin{pmatrix} p_u q_d + 5\varepsilon & p_u q_m - 4\varepsilon & p_u q_u - \varepsilon \\ p_m q_d - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_u - 4\varepsilon \\ p_d q_d - \varepsilon & p_d q_m - 4\varepsilon & p_d q_u + 5\varepsilon \end{pmatrix}$$

where $\varepsilon = \frac{|\rho|}{36}$.

This provides us with an ordinary 3D tree that converges to the system (4.14). We are now in a position to evaluate spread options for the purpose of authenticating the approximate tree. This will be done in Subsections (8.1.2) and (8.2.2).

Algorithm 7.33 (2-Factor (a) Model Ordinary Trinomial Tree) *1. Construct ordinary trinomial trees for s and v by Algorithm (7.30).*

2. Combine the trees and induce the correlation ρ by Equation (6.14) or (6.15).

7.2.3 Fast Trinomial Tree

The fast 3D tree follows precisely the same construction pattern as the ordinary one except that we build separate fast trinomial trees for each asset before combining them into a single tree.

Algorithm 7.34 (2-Factor (a) Model Fast Trinomial Tree) *1. Construct fast trinomial trees for s and v by Algorithm (7.32).*

2. Combine the trees and induce the correlation ρ by Equation (6.14) or (6.15).

7.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

We now move on to consider the 2-factor (b) model

$$\begin{aligned} ds &= \alpha(l-s)\,dt + \sigma s\,dz \\ dl &= \beta(t) l\,dt + \tau l\,dy \end{aligned}$$

In the standard 2-factor (b) Pilipovic equation β is a constant parameter. In order to more accurately capture economic trends, we allow $\beta(t)$ to be an unknown

function of time. The parameter $\beta(t)$ will be chosen to make the model consistent with the futures prices observed in the market. This is nearly identical to our development of the 1-factor model in Section (7.1). However, there are a few subtle differences that must be addressed. Although it is possible for the Brownian motions z and y to be correlated, we assume that they are not, for the sake of convenience.

7.3.1 Constant Standard Deviations

The first step in our operation is a log transformation that renders the standard deviations in the system of stochastic differential equations constant.

$$S = \ln(s) \qquad L = \ln(l)$$

By the Ito formula we have

$$\begin{aligned} dS &= [\alpha(e^{L-S} - 1) - \frac{\sigma^2}{2}]dt + \sigma\, dz \\ dL &= \left[\beta(t) - \frac{\tau^2}{2}\right] dt + \tau\, dy \end{aligned}$$

7.3.2 Decouple the Equations

It is now desirable to eliminate the dependence of S on L. This can be done by introducing a new random variable

$$R = L - S \tag{7.17}$$

Then by the Higher Dimensional Ito Formula it follows that

$$\begin{aligned} dR &= \alpha[A(t) - e^R]dt + \upsilon\, dw \\ dL &= \left[\beta(t) - \frac{\tau^2}{2}\right] dt + \tau\, dy \end{aligned} \tag{7.18}$$

where w is a new Brownian motion with the standard deviation of R implicitly defined by

$$\upsilon^2 = \sigma^2 + \tau^2$$

and $B(t)$ is some unknown function of time given by

$$A(t) = 1 + \frac{\beta(t) + \frac{\sigma^2 - \tau^2}{2}}{\alpha}$$

Observe that a linear combination of two Brownian motions is also a Brownian motion. Indeed by the Higher Dimensional Ito formula we have

$$\upsilon\, dw = \tau\, dy - \sigma\, dz$$

Even though S and L are uncorrelated by our assumption, the random processes R and L are positively correlated by an amount

$$\rho = \frac{\tau}{\upsilon}$$

7.3.3 Construct Separate Trees

The construction of an ordinary trinomial tree for R proceeds precisely as in Section (7.1). We let the nodes in the R-tree be denoted by (i, j), denote the length of a small

change in R by $\Delta R = v\sqrt{3\ \Delta t}$, and write the branching probabilities in the R-tree by $p_{ij}^{(u)}$, $p_{ij}^{(m)}$, and $p_{ij}^{(d)}$. The median nodes are denoted by $\widehat{\omega}_i$ and are calculated by Equation (6.5). It is interesting to note that the median nodes are not equal to the expected values ω_i (except at time 0 when $\widehat{\omega}_0 = \omega_0$). However, they are approximately by our Claim (6.26).

The random process for L is a geometric Brownian motion and the trinomial tree is easy to construct. In fact we already did the analysis of $l(t)$ in Section (6.2). Nodes in the L-tree will be written as (i,k) and a small change in L is given by $\Delta L = \tau\sqrt{3\ \Delta t}$. The probabilities in the L-tree are given by Equation (6.10)

$$\begin{pmatrix} q_u \\ q_m \\ q_d \end{pmatrix} = \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix}$$

and the position of node (i,k) is $\Psi_i + k\ \Delta L$ where

$$\Psi_i = L_0 + \sum_{k=0}^{i-1}(\beta_k - \frac{\tau^2}{2})\Delta t \tag{7.19}$$

7.3.4 Induce the Correlation

This done the preliminary trees for R and L are combined according the procedure explained in Subsection (6.4.1) to form a 3D (R,L)-tree. Since R and L are positively

correlated the branching probability matrix is given by Equation (6.14)

$$G = \begin{pmatrix} p_u q_d - \varepsilon & p_u q_m - 4\varepsilon & p_u q_u + 5\varepsilon \\ p_m q_d - 4\varepsilon & p_m q_m + 8\varepsilon & p_m q_u - 4\varepsilon \\ p_d q_d + 5\varepsilon & p_d q_m - 4\varepsilon & p_d q_u - \varepsilon \end{pmatrix}$$

where $\varepsilon = \frac{\rho}{36}$.

7.3.5 Matching Futures Prices

Once again the unknown function of time $\beta(t)$ can be theoretically calculated by solving the ordinary system of differential equations (3.7)

$$\begin{aligned} \dot{\varphi}(t) &= \alpha[\psi(t) - \varphi(t)] \\ \dot{\psi}(t) &= \beta(t)\psi(t) \end{aligned}$$

with the expected future spot price $\varphi(t)$ replaced by the futures price $f(t)$. This provides a satisfactory method of calculating the drift parameter $\beta(t)$. The solution is

$$\beta(t) = \frac{\dot{f}(t) + \frac{\ddot{f}(t)}{\alpha}}{f(t) + \frac{\dot{f}(t)}{\alpha}} \tag{7.20}$$

The spot price at node (i, j, k) is recovered by the inverse transformation

$$s_{ijk} = e^{\psi_i + k\,\Delta L - (\bar{\omega}_i + j\,\Delta R)} \tag{7.21}$$

and the tree can be used to evaluate options.

Remember that the remaining parameters α, σ, and τ are estimated by historical data using the procedure explained in Section (4.3).

Algorithm 7.35 (2-Factor (b) Model Ordinary Trinomial Tree) *1. Specify the parameters α, σ, T, and n and define the futures prices $f(t)$ for $t \in [0,T]$.*

2. Transform to the decoupled system (7.18).

3. Set $\Delta t = \frac{T}{n}$, $\Delta R = v\sqrt{3\ \Delta t}$, and $\Delta L = \tau\sqrt{3\ \Delta t}$.

4. Calculate the unknown function of time $\beta(t)$ by solving the ordinary differential equation (7.20).

5. Pick a value for s_0 and set $l_0 = s_0 + \frac{f_1 - s_0}{\alpha\ \Delta t}$. Then $R_0 = \ln(\frac{l_0}{s_0})$ and $L_0 = \ln(l_0)$ and the median nodes in the trees for R and L at time 0 are $\widehat{\omega}_0 = R_0$ and $\Psi_0 = L_0$. Also, put $j_{\max}(0) = j_{\min}(0) = 0$.

6. Set $\mathbf{q} = \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix}$.

7. For $i = 0, ..., (n-1)$

(a) For $j = j_{\min}(i), ..., j_{\max}(i)$

(i) Calculate the branching probabilities $\mathbf{p}_{ij}$ by Equation (6.7).

(ii) Calculate the matrix G_{ij} by Equation (6.14)

End For

(b) Calculate $\widehat{\omega}_{i+1}$ and Ψ_{i+1} by Equations (6.5) and (7.19).

(c) Calculate $j_{\max}(i+1)$ and $j_{\min}(i+1)$ by Equation (6.8).

End For

7. Calculate s_{ijk} by Equation (7.21).

7.3.6 Fast Trinomial Tree

The fast trinomial trees for R and L are constructed using the procedure in Section (6.3). Furthermore, the correlation is induced by adjusting the branching probability matrix according to Equation (6.14). Then the nodes in the preliminary trees for R and L are found at

$$R_{ij} = j\ \Delta R \text{ and } L_{ik} = k\ \Delta L$$

while the shifted nodes in the final trees for R and L are located at

$$R_{ij} = \omega_i + j\ \Delta R \text{ and } L_{ik} = \Psi_i + k\ \Delta L$$

It merely remains now to (implicitly) select the parameter $\beta(t)$ so that the expected value of $s(t)$ matches the futures prices $f(t)$. The spot price at node (i, j, k) is recovered by the inverse transformation

$$\begin{aligned} s_{ijk} &= e^{\psi_i + k\ \Delta L - (\omega_i + j\ \Delta R)} \\ &= e^{\Phi_i + k\ \Delta L - j\ \Delta R} \end{aligned} \tag{7.22}$$

where $\Phi(t) = \overline{S}(t)$, $\Psi(t) = \overline{L}(t)$ and $\omega(t) = \overline{R}(t)$.

Observe that we do not directly solve for $\beta(t)$. Rather we find the parameter $\Phi_i = \psi_i - \omega_i$ which is a function of $\beta(t)$. It is not necessary to solve Φ_i as a function of $\beta(t)$ in order to match the futures prices $f(t)$. This is indeed fortunate and we do not concern ourselves with the analytic form of Φ_i here.

The expected value of s at time $i\ \Delta t$ is given by

$$\begin{aligned} \varphi_i &= \sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} \sum_{k=-i}^{i} P_{ijk} s_{ijk} \\ &= e^{\Phi_i} \sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} \sum_{k=-i}^{i} P_{ijk} e^{k\ \Delta L - j\Delta R} \end{aligned} \tag{7.23}$$

where P_{ijk} is the probability of reaching node (i, j, k). These cumulative probabilities are easily calculated provided we know the branching probability at each node in the tree.

$$\begin{aligned} P_{000} &= 1 \\ P_{ijk} &= \sum_{m,n} P_{i-1\ mn} q[(i-1, m, n) \rightarrow (i, j, k)] \end{aligned} \tag{7.24}$$

for $i = 1, ..., n$, $j = \max(-i, j_{\min}), ..., \min(j_{\max}, i)$, and $k = -i, ..., i$ where $q[(i - 1, m, n) \rightarrow (i, j, k)]$ is the probability of branching from node $(i - 1, m, n)$ to node (i, j, k).

Let $f(t)$ be the futures price of the energy commodity in the market. Then the expected future spot price is related to the futures price by Equation (5.3)

$$\varphi(t) = f(t)$$

The unknown parameter Φ_i is found by setting the risk discounted futures price equal to the expected future spot price in Equation (7.23) and solving

$$\Phi_i = \ln\left(\frac{f_i}{\sum_{j=\max(i,j_{\min})}^{\min(i,j_{\max})} \sum_{k=-i}^{i} P_{ijk} e^{k\ \Delta L - j\Delta R}}\right) \tag{7.25}$$

The following algorithm summarizes the construction of the fast trinomial tree in pseudocode style.

Algorithm 7.36 (2-Factor (b) Model Fast Trinomial Tree) *1. Specify the parameters α, σ, T, and n and define the futures prices $f(t)$ for $t \in [0,T]$.*

2. Transform to the decoupled system (7.18).

3. Set $\Delta t = \frac{T}{n}$, $\Delta R = v\sqrt{3\ \Delta t}$, and $\Delta L = \tau\sqrt{3\ \Delta t}$.

4. Set $\mathbf{q} = \begin{pmatrix} \frac{1}{6} \\ \frac{2}{3} \\ \frac{1}{6} \end{pmatrix}$.

5. Calculate $j_{\max}$ and $j_{\min}$ by Equation (7.8).

6. For $j = j_{\min}, \ldots, j_{\max}$

(a) Calculate the branching probabilities $\mathbf{p}_j$ by Equation (6.7).

(b) Calculate the matrix G_j by Equation (6.14)

End For

7. Set $P_{000} = 1$.

8. For $i = 0, \ldots, (n-1)$

(a) Calculate P_{ijk} by Equation (7.24).

(b) Calculate Φ_i by Equation (7.25).

End For

8. Calculate s_{ijk} by Equation (7.22).

CHAPTER 8
ENERGY OPTIONS

Traditionally, there has been a market for options on some energy futures contracts on NYMEX. There is, however, a sizable over-the-counter market for options on actual spot prices, utilized by large energy firms and users. We propose and implement trinomial methodologies for the valuation of these options, based on mean reverting price processes. Trinomial trees provide an efficient and accurate way of pricing these options. We will examine a subset of derivative securities known colloquially as American and European style options. There are two basic kinds of these options: call and put.

Definition 8.37 *A **call (put) option** gives the holder the right to buy (sell) an asset by a certain date T for a fixed price K. The price K in the contract is known as the **strike price** and the date T is known as the **expiration date** or **maturity**. **American** options can be exercised at any time up to maturity. **European** options can be exercised only on the expiration date itself. The values of European call and put options will be denoted by $c(t,T)$ and $p(t,T)$ while the values of American call and put options will be written as $C(t,T)$ and $P(t,T)$ where t is the start time of the contract. We often write the option price as a function of a single variable s: in this case we shall always assume that $t = 0$ (or $t = t_0$) and $T = s$.*

We now describe a practical scenario for evaluating options. In each case, the

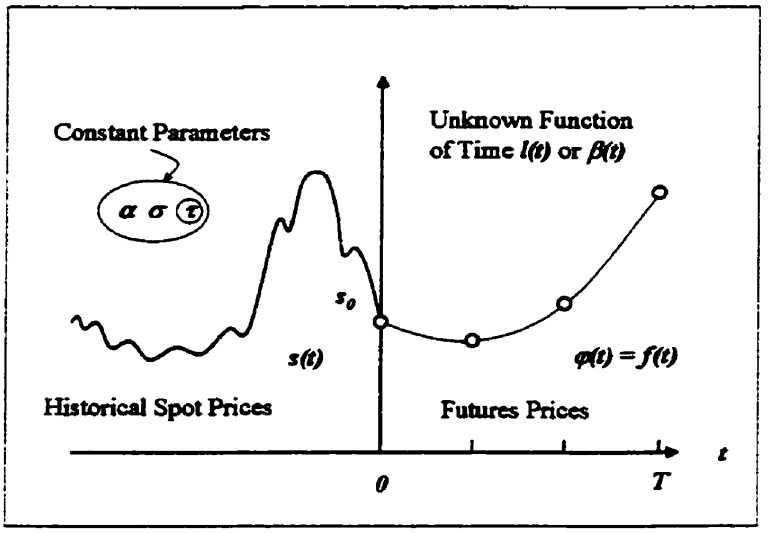

Figure 8.1: Option Evaluation Procedure

parameters α, σ, and (possibly) τ are estimated using the year of historical data immediately preceding the start date of the contract. As usual, the unknown function of time $l(t)$ or $\beta(t)$ is selected so that the geometry of the tree is consistent with the futures prices over the period of the option. Figure (8.1) shows a picture of this procedure. Also, the duration is taken to be $T = 3$ months and the strike price is $K = 0.75s_0$ where s_0 is the spot price on the start date. Remember that futures prices are only available for each monthly contract and missing values are interpolated using a cubic spline. Last but not least, we assume that the risk free interest rate is $r = 3.5\%$ for the sake of simplicity. We shall adopt this procedure for the evaluation of all options for the remainder of this thesis.

8.1 European Option Evaluation

We suppose that the asset underlying the options is an energy spot price $s(t)$. Consider a European call option on a single unit of energy. If the energy spot price $s(T)$ exceeds the strike price K at maturity, the trader can exercise the option by purchasing the energy at the strike price and immediately selling it on the market, realizing a profit of $s(T) - K$. On the other hand, if $s(T) < K$ then the option will expire unused. Evidently the payoff of a European call realized at maturity is $[s(T) - K]^+$. Similarly the payoff of a European put option at maturity is $[K - s(T)]^+$.

In a risk neutral world, the payoff is discounted to the present by the risk free interest rate r (this is the time value of money). Therefore

$$\begin{aligned} c(T) &= \mathrm{e}^{-rT}\mathrm{E}\{[s(T) - K]^+\} \\ p(T) &= \mathrm{e}^{-rT}\mathrm{E}\{[K - s(T)]^+\} \end{aligned} \tag{8.1}$$

where the expectation is taken under a risk neutral probability distribution. For a complete discussion of risk neutral option evaluation, refer to [Hull 1999].

8.1.1 1-Factor Model

Let $s(t)$ be a spot price satisfying the 1-factor generalized Pilipovic equation (3.2). Suppose that we have constructed a risk neutral trinomial tree for $s(t)$ by matching the futures prices $f(t)$ over the duration of the contract $[0, T]$. As usual, we denote the

probability of reaching node (i, j) by P_{ij}. Then

$$c(T) = e^{-rT} \sum_{j=j_{\min}}^{j_{\max}} P_{ij}(s_{ij} - K)^+ \tag{8.2}$$

$$p(T) = e^{-rT} \sum_{j=j_{\min}}^{j_{\max}} P_{ij}(K - s_{ij})^+$$

Throughout this entire section, we display equations only for fast trees. The corresponding equations for ordinary trees are almost identical except that the upper and lower bounds $j_{\min}$, $j_{\max}$, $k_{\min}$, and $k_{\max}$ depend on i.

We now compare the European call option prices revealed by ordinary and fast trinomial trees. We calculate typical natural gas, crude oil, and electricity options from 1997, 1998, and 1999. Tables (8.1), (8.2), and (8.3) show that both methods are converging to the same values for these typical options. Also, we display the errors for reference. The actual option price is approximated by executing the ordinary and fast trinomial trees for a very large number of time steps ($n = 1600$). These (pseudo) exact values are shown in brackets beside the words Fast and Ordinary in the tables below. Furthermore, the computation times are given in seconds. These times are the result of running the programs on a PC that is capable of executing 1 000 000 floating point operations (flops) per second. Also, the ratios between their CPU times are displayed as a percentage so that we can see how much faster the fast trinomial trees really are.

It would appear that on average the fast trinomial trees require only (approximately) 45% the number of flops that the ordinary ones do. Furthermore, the fast trees seem to be converging at about the same rate in that they achieve the same errors as

Table 8.1: 1-F European Call Natural Gas (January 1, 1997)

	Fast (0.1686)			Ordinary (0.1687)			CPU Ratio
n	Value	Error	Time (s)	Value	Error	Time (s)	(%)
100	0.1688	0.0002	0.06	0.1691	0.0003	0.12	45.2
200	0.1702	0.0017	0.21	0.1704	0.0017	0.47	45.2
400	0.1686	0.0000	0.86	0.1688	0.0000	1.87	46.0
800	0.1687	0.0001	4.08	0.1689	0.0001	9.16	44.7
Parameters $\alpha = 0.167$ and $\sigma = 0.0446$. Strike $K = 2.17$.							
Spot $s_0 = 2.89$. Futures $f(1) = 2.89$, $f(2) = 2.62$, and $f(3) = 2.35$.							

the ordinary trees for similar values of n. This means that the fast trinomial trees are doubly effective in the 1-factor case!

Since both types of trees use Euler methods, we expect their errors E to be $O(\Delta t)$, or equivalently $O(\frac{1}{n})$, as $\Delta t = \frac{T}{n}$. In other words, we suppose that the error is a linear function of Δt. It is easy to see this linear relationship by plotting $\log_2(E)$ as a function of $\log_2(n)$. Why? Well, because if $E = \gamma \Delta t$ for some number γ then $\log_2(E) = \log_2(\gamma T) - \log_2(n)$. For the sake of conciseness, we only consider the crude oil (May 1, 1998) data. Figure (8.2).shows the results.

Now consider the run times t_R of our programs. Evidently the space and time requirements of the 1-factor model are proportional to $(j_{\max} - j_{\min} + 1)n$ where the

Table 8.2: 1-F European Call Crude Oil (May 1, 1998)

	Fast (5.790)			Ordinary (5.789)			CPU Ratio
n	Value	Error	Time (s)	Value	Error	Time (s)	(%)
100	5.712	0.078	0.06	5.707	0.083	0.12	45.2
200	5.753	0.037	0.21	5.750	0.039	0.46	46.0
400	5.774	0.016	0.83	5.773	0.017	1.78	46.5
800	5.784	0.005	3.23	5.784	0.006	7.05	46.7
Parameters $\alpha = 0.154$ and $\sigma = 0.0209$. Strike $K = 12.10$.							
Spot $s_0 = 16.13$. Futures $f(1) = 16.13$, $f(2) = 16.09$, and $f(3) = 17.05$.							

upper and lower bounds are given by Equation (7.8)

$$\begin{aligned} j_{\max} &= \left\lceil \frac{1}{\Delta R} \ln(1 + \frac{\Delta R}{2a\ \Delta t}) \right\rceil \\ j_{\min} &= \left\lfloor \frac{1}{\Delta R} \ln(1 - \frac{\Delta R}{2a\ \Delta t}) \right\rfloor \end{aligned}$$

Recall that $\Delta R = \sigma\sqrt{3\ \Delta t}$. This fact together with the linear approximation $\ln(1+x) \sim x$ reveals that $j_{\max} \sim \frac{n}{2aT}$ and $j_{\min} \sim -\frac{n}{2aT}$. Therefore we predict that our programs run in time $O(n^2)$. To view this relationship, we plot $\log_2(t_R)$ as a function of $\log_2(n)$. The result should be a straight line of slope 2. Figure (8.3) displays these lines for the crude oil (May 1, 1998) data. The picture clearly shows that both types of trees do indeed run in time $O(n^2)$. Of course, the fast tree has a smaller constant of proportionality than its ordinary counterpart.

We now demonstrate the seasonal nature of option prices under the 1-factor

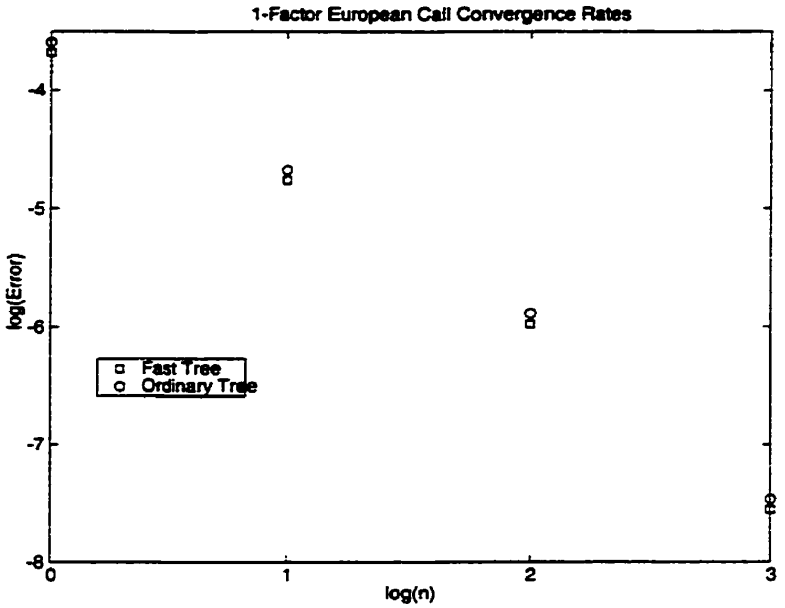

Figure 8.2: 1-Factor Model European Call Option Errors

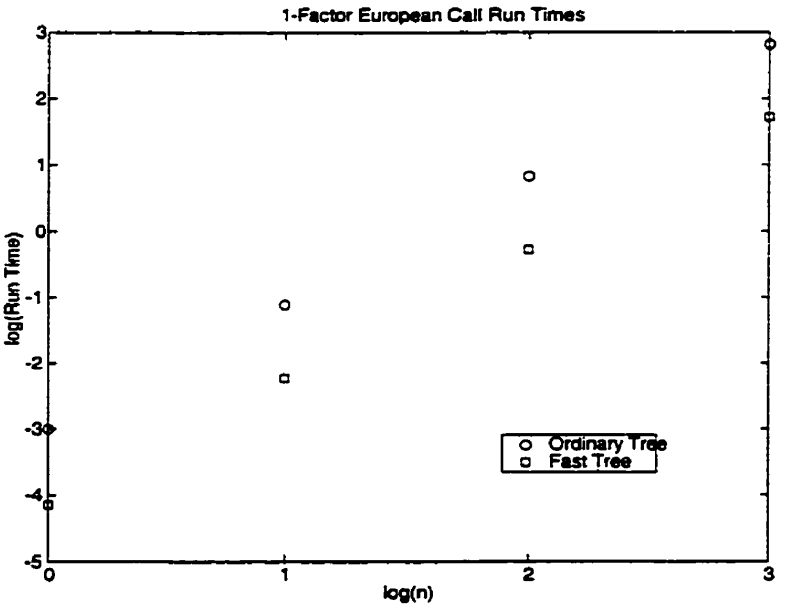

Figure 8.3: 1-Factor Model European Call Option Run Times

Table 8.3: 1-F European Call Electricity (September 1, 1999)

	Fast (7.894)			Ordinary (7.893)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	7.649	0.245	0.06	7.625	0.267	0.14	43.0
200	7.778	0.117	0.25	7.765	0.127	0.55	45.3
400	7.844	0.050	1.03	7.837	0.055	2.31	44.8
800	7.877	0.017	7.76	7.874	0.019	30.7	25.2
Parameters $\alpha = 0.142$ and $\sigma = 0.0456$. Strike $K = 24.83$.							
Spot $s_0 = 33.10$. Futures $f(1) = 33.10$, $f(2) = 29.97$, and $f(3) = 31.35$.							

model. This is done by calculating the option values at the beginning of each month in a sample year and plotting them. Although we could evaluate options on every day throughout the year, this is unnecessarily time consuming. Instead, we interpolate between the monthly values using a cubic spline. This shows the seasonal shape of the curve. Naturally we use ordinary and fast trinomial trees in these calculations for comparison. Once again we consider natural gas, crude oil, and electricity options from 1997, 1998, and 1999. To save space, we show only the crude oil 1998 option prices in Figures (8.4). The remaining pictures have been relegated to Appendix C.

The pictures clearly show the need for a time varying parameter model for the risk neutral evaluation of energy options. It is also interesting to note that ordinary and fast trees calculate option prices that are so close to one another that the difference is

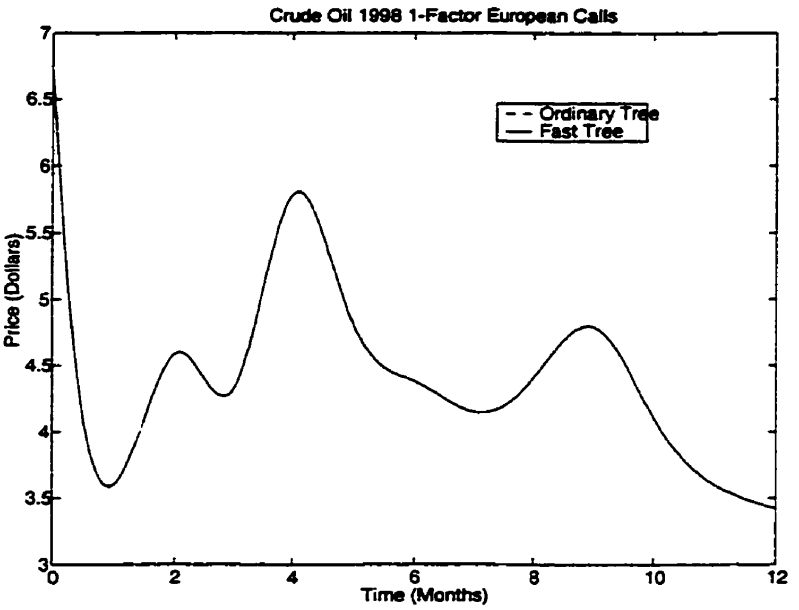

Figure 8.4: Crude Oil 1998 1-Factor Model European Call Options

not noticeable by graphing. This is strong supporting evidence for the Fast Trinomial Tree Conjecture in the case of the 1-factor model.

8.1.2 2-Factor (a) Model (2 Correlated Assets)

We continue our analysis by evaluating European spread options. If our assets are denoted by $s(t)$ and $v(t)$ and $v(t) > s(t)$ then the payoff function of a European call option at maturity is $[v(T) - s(T) - K]^+$. The European option evaluation equations are similar to (8.2) and we will not bother to derive them again although we write them

down for reference.

$$c(T) = \mathrm{e}^{-rT} \sum_{j=j_{\min}}^{j_{\max}} \sum_{k=k_{\min}}^{k_{\max}} P_{ijk}[(v_{ik} - s_{ij}) - K]^+ \tag{8.3}$$

$$p(T) = \mathrm{e}^{-rT} \sum_{j=j_{\min}}^{j_{\max}} \sum_{k=k_{\min}}^{k_{\max}} P_{ijk}[K - (v_{ik} - s_{ij})]^+$$

We proceed to evaluate spread options among all possible pairs of natural gas, crude oil, and electricity. Their correlations - like the rest of the parameters - are calculated using the year of historical data immediately prior to the option period. This is a typical scheme used by financial modelers. The actual option values are calculated using $n = 400$ and are shown in brackets for the ordinary and fast trees. Table (8.4), (8.5), and (8.6) display the results.

Table 8.4: 2-F (a) European Spread Oil-Electricity (January 1, 1999)

	Fast (0.3665)			Ordinary (0.3614)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	0.3720	0.0055	0:03	0.3671	0.0057	0:04	66.4
141	0.3690	0.0025	0:06	0.3639	0.0025	0:10	65.6
200	0.3675	0.0009	0:17	0.3624	0.0010	0:26	65.0
283	0.3666	0.0000	0:45	0.3615	0.0000	1:12	63.0
Parameters $\alpha = 0.170$, $\beta = 0.172$, $\sigma = 0.0283$, $\tau = 0.0479$, and $\rho = -0.051$.							
Strike $K = 8.59$. Spots $s_0 = 12.05$ and $v_0 = 23.50$. Futures $f(1) = 12.05$,							
$f(2) = 12.09$, $f(3) = 12.40$, $g(1) = 23.50$, $g(2) = 21.50$, and $g(3) = 20.00$.							

Table 8.5: 2-F (a) European Spread Electricity-Gas (May 1, 1997)

	Fast (9.945)			Ordinary (9.951)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	10.02	0.072	0:03	10.04	0.090	0:05	57.7
141	9.990	0.045	0:07	10.01	0.056	0:11	65.5
200	9.970	0.025	0:20	9.982	0.031	0:31	62.7
283	9.955	0.010	0:54	9.964	0.013	1:27	61.3
Parameters $\alpha = 0.176$, $\beta = 0.140$, $\sigma = 0.0410$, $\tau = 0.0444$, and $\rho = 0.388$.							
Strike $K = 13.02$. Spots $s_0 = 2.24$ and $v_0 = 19.60$. Futures $f(1) = 2.24$,							
$f(2) = 2.26$, $f(3) = 2.33$, $g(1) = 19.60$, $g(2) = 23.76$, and $g(3) = 25.73$.							

Figure (8.5) shows the linear convergence rates for the electricity and natural gas (May 1, 1997) data.

Evidently the space and time requirements of the 2-factor (a) model are proportional to $(j_{\max} - j_{\min} + 1)(k_{\max} - k_{\min} + 1)n$. By an analysis similar to the one we did in Subsection (8.1.1), this indicates that the computer programs should run in time $O(n^3)$. We can view this relationship by plotting $\log_2(t_R)$ as a function of $\log_2(n)$. The result should be a straight line of slope 3. Figure (8.6) displays these lines for the electricity and natural gas (May 1, 1997) data.

Figure (8.7) shows the seasonal shape of the option prices for the correlated pair electricity and natural gas in the year 1997. The corresponding pictures for the

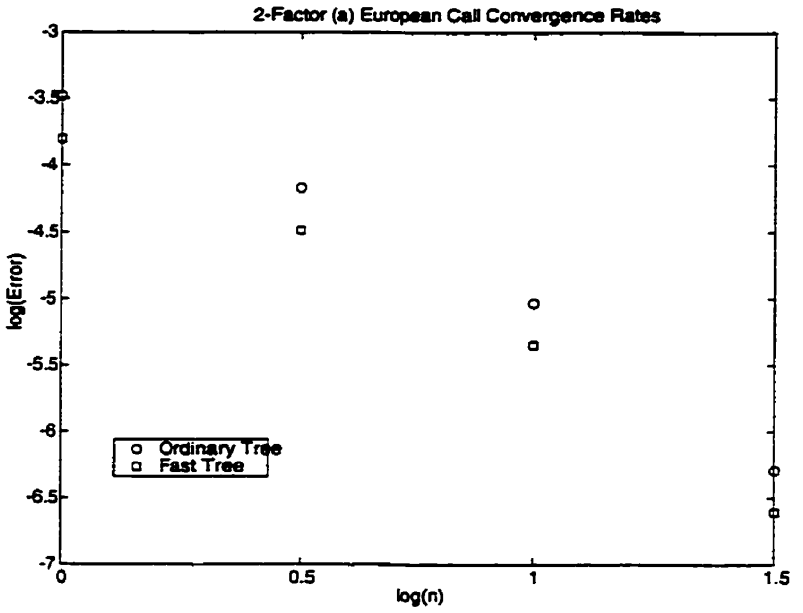

Figure 8.5: 2-Factor (a) Model European Call Errors

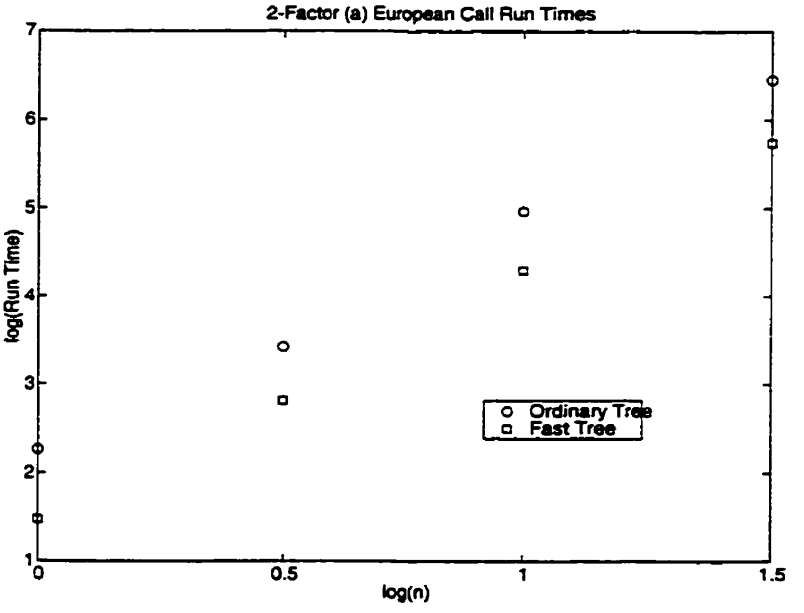

Figure 8.6: 2-Factor (a) Model European Call Option Run Times

Table 8.6: 2-F (a) European Spread Gas-Oil (September 1, 1998)

	Fast (3.438)			Ordinary (3.438)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	3.424	0.013	0:03	3.425	0.0013	0:04	66.3
141	3.430	0.008	0:07	3.430	0.008	0:11	65.5
200	3.433	0.005	0:18	3.433	0.004	0:29	62.5
283	3.436	0.002	0:49	3.436	0.002	1:16	64.5
Parameters $\alpha = 0.140$, $\beta = 0.175$, $\sigma = 0.0311$, $\tau = 0.0257$, and $\rho = 0.065$.							
Strike $K = 8.69$. Spots $s_0 = 1.75$ and $v_0 = 13.34$. Futures $f(1) = 1.75$,							
$f(2) = 1.98$, $f(3) = 2.25$, $g(1) = 13.34$, $g(2) = 13.59$, and $g(3) = 14.18$.							

remaining energy pairs may be found in Appendix C. The pictures show that ordinary and fast trees reveal virtually identical option values in the 2-factor (a) model. Again this supports the Fast Trinomial Tree Conjecture.

8.1.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

The European option prices for the 2-factor (b) model are

$$c(T) = e^{-rT} \sum_{j=j_{\min}}^{j_{\max}} \sum_{k=-n}^{n} P_{ijk}(s_{ijk} - K)^+ \tag{8.4}$$

$$p(T) = e^{-rT} \sum_{j=j_{\min}}^{j_{\max}} \sum_{k=k_{\min}}^{k_{\max}} P_{ijk}(K - s_{ijk})^+$$

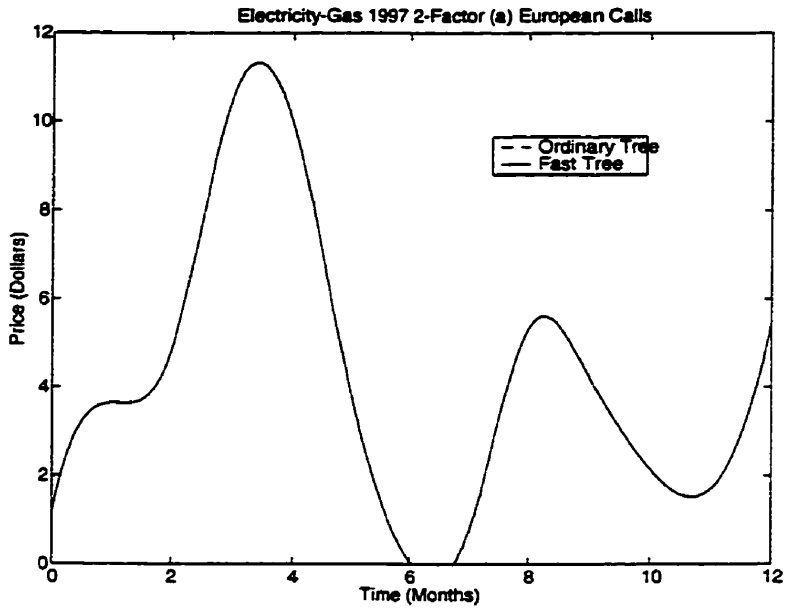

Figure 8.7: Electricity-Gas 1997 2-Factor (a) European Call Options

We evaluate typical natural gas, crude oil, and electricity options from 1997, 1998, and 1999. The actual option values are calculated using $n = 200$ and are shown in brackets for the ordinary and fast trees. The results are shown in Tables (8.7), (8.8), and (8.9).

Figure (8.5) shows the linear convergence rates for the crude oil (May 1, 1998) data.

Evidently the space and time requirements of the 2-factor (b) model are proportional to $(j_{\max} - j_{\min} + 1)(2n + 1)n$. By an analysis similar to the one we did in Subsection (8.1.1), this indicates that the computer programs should run in time $O(n^3)$. We can view this relationship by plotting $\log_2(t_R)$ as a function of $\log_2(n)$. The result

Table 8.7: 2-F (b) European Call Natural Gas (January 1, 1997)

	Fast (0.2988)			Ordinary (0.4010)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	0.3020	0.0032	0:03	0.4052	0.0042	0:03	96.7
71	0.3005	0.0017	0:07	0.4033	0.0022	0:08	87.4
100	0.2997	0.0009	0:19	0.4022	0.0012	0:21	89.6
141	0.2991	0.0004	0:49	0.4015	0.0005	0:54	91.4
Parameters $\alpha = 0.167$, $\sigma = 0.0446$, and $\tau = 0.0315$. Strike $K = 2.17$.							
Spot $s_0 = 2.89$. Futures $f(1) = 2.89$, $f(2) = 2.62$, and $f(3) = 2.35$.							

should be a straight line of slope 3. Figure (8.9) displays these lines for the crude oil (May 1, 1998) data.

Figure (8.7) shows the seasonal shape of the option prices for crude oil in the year 1998. The corresponding pictures for natural gas and electricity may be found in Appendix C.

By looking at Figure (8.10), we perceive that ordinary and fast trees clearly deviate from one another in the 2-factor (b) model. Although this would seem to discredit the Fast Trinomial Tree Conjecture, bear in mind that the relative complexity of the 2-factor (b) model - as it involves a spot price $s(t)$ mean reverting to a diffusion $l(t)$ - makes it more difficult to obtain perfect results. However, we maintain that the procedure is still satisfactory from a practical point of view. In an industrial application,

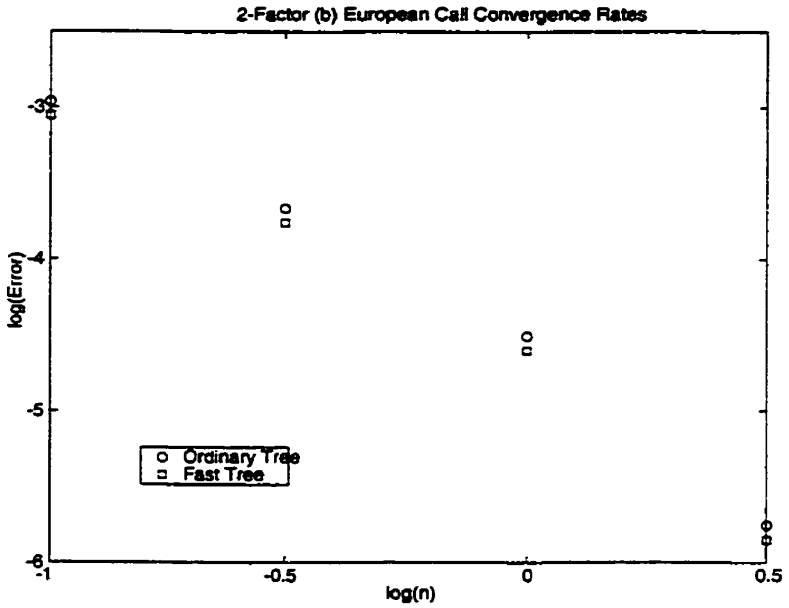

Figure 8.8: 2-Factor (b) Model European Call Errors

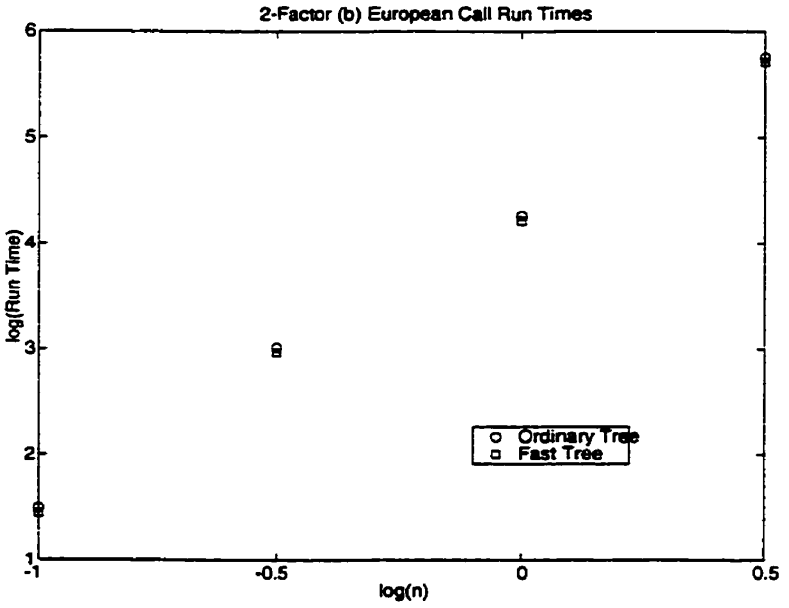

Figure 8.9: 2-Factor (b) Model European Call Optoin Run Times

Table 8.8: 2-F (b) European Call Crude Oil (May 1, 1999)

	Fast (5.754)			Ordinary (6.102)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	5.633	0.121	0:03	5.974	0.128	0:03	96.7
71	5.680	0.074	0:08	6.023	0.079	0:08	96.8
100	5.713	0.041	0:18	6.058	0.044	0:19	96.8
141	5.737	0.017	0:52	6.083	0.019	0:54	96.8
Parameters $\alpha = 0.154$, $\sigma = 0.0209$, and $\tau = 0.0176$. Strike $K = 12.10$.							
Spot $s_0 = 16.13$. Futures $f(1) = 16.13$, $f(2) = 16.09$, and $f(3) = 17.05$.							

the parameters α, σ, and τ have errors associated with them given by Equation (4.27)

$$\Delta\alpha = 0.07 \qquad \Delta\sigma = 0.003 \qquad \Delta\tau = 0.015$$

Furthermore, the highly volatile nature of the energy market renders extreme accuracy in option calculations redundant. Remember that the futures prices are in some sense the best predictors of the way spot prices will evolve. However, the futures prices - which determine the geometric shape of our trees - are still guesses, albeit good ones. Therefore the discrepancy between ordinary and fast trees is actually well within acceptable tolerance levels. Also, in the case of European option evaluation under the 2-factor (b) model, we notice that fast trees run only slightly faster than ordinary trees[1].

[1]Look at the CPU ratios in Tables (8.7), (8.8), and (8.9).

Table 8.9: 2-F (b) European Call Electricity (September 1, 1997)

	Fast (8.585)			Ordinary (11.543)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	8.280	0.306	0:03	11.157	0.386	0:03	96.7
71	8.395	0.190	0:09	11.304	0.240	0:09	96.9
100	8.478	0.108	0:22	11.409	0.134	0:22	96.8
141	8.540	0.045	1:01	11.485	0.058	1:06	92.9
Parameters $\alpha = 0.142$, $\sigma = 0.0456$, and $\tau = 0.0394$. Strike $K = 24.83$.							
Spot $s_0 = 33.10$. Futures $f(1) = 33.10$, $f(2) = 29.97$, and $f(3) = 31.35$.							

8.2 American Option Evaluation

American style options are more complicated to evaluate than their European counterparts. However, they are much more popular in the market place, and so we now analyze the effectiveness of our trees in handling them. Throughout this entire section, we develop equations only for fast trees. The corresponding equations for ordinary trees are almost identical except that the branching probabilities $\mathbf{p}$ and transition probability matrix G depend on i as well as j, and the upper and lower bounds $j_{\min}$, $j_{\max}$, $k_{\min}$, and $k_{\max}$ depend on i.

In this case, we suppose that the holder of the option is able to exercise the option at any of the times $i\,\Delta t$ for $i = 0, ..., n$. That is, at each node, the option value is calculated to be the maximum of the profit realized by immediate exercise and the

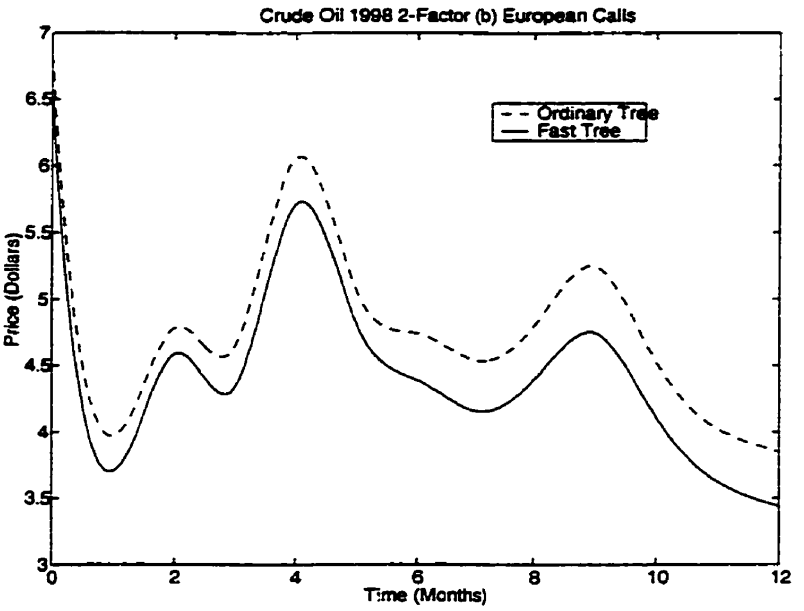

Figure 8.10: Crude Oil 1998 2-Factor (b) European Call Options

expected payoff of holding onto the option. The present value of the option is then set to the price calculated at the root $(0, 0)$.

8.2.1 1-Factor Model

Evidently, if time T is reached, it is the traders last chance to exercise the option. Hence the value of an American call or put option at node (n, j) is $(s_{nj} - K)^+$ or $(K - s_{nj})^+$ for $j = j_{\min}, ..., j_{\max}$. We proceed backward through the tree for $i = (n-1), ..., 0$ and

calculate the option value at node by the recursive formulas

$$C_{ij} = \max[(s_{ij} - K)^+, e^{-r\,\Delta t} \sum_{m=-1}^{1} p_j^{(m)} C_{i+1,h_j+m}]$$

$$P_{ij} = \max[(s_{ij} - K)^+, e^{-r\,\Delta t} \sum_{m=-1}^{1} p_j^{(m)} P_{i+1,h_j+m}]$$

for $j = \max(i, j_{\min}),\ldots,\min(i, j_{\max})$ where the notation $p_1^{(j)}$, $p_0^{(j)}$, and $p_{-1}^{(j)}$ is understood to indicate p_u, p_m, and p_d at node (i, j). Then the option prices are the values located at the root $C(T) = C_{00}$ and $P(T) = P_{00}$.

We evaluate typical natural gas, crude oil, and electricity options from 1997, 1998, and 1999. The results are shown in Tables (8.10), (8.11), and (8.12).

Table 8.10: 1-F American Call Natural Gas (January 1, 1997)

	Fast (0.9296)			Ordinary (0.9289)			CPU Ratio
n	Value	Error	Time (s)	Value	Error	Time (s)	(%)
100	0.9380	0.0084	0.10	0.9380	0.0091	0.20	52.0
200	0.9326	0.0029	0.39	0.9321	0.0033	0.76	51.7
400	0.9312	0.00016	1.59	0.9306	0.0017	3.04	52.4
800	0.9300	0.0004	7.54	0.993	0.0005	14.8	50.8
Parameters $\alpha = 0.167$ and $\sigma = 0.0446$. Strike $K = 2.17$.							
Spot $s_0 = 2.89$. Futures $f(1) = 2.89$, $f(2) = 2.62$, and $f(3) = 2.35$.							

Figure (8.11) shows the convergence rates for the crude oil (May 1, 1997) data. Again we observe the linear relationship between the errors E and the lengths of a

Table 8.11: 1-F American Call Crude Oil (May 1, 1998)

	Fast (5.825)			Ordinary (5.826)			CPU Ratio
n	Value	Error	Time (s)	Value	Error	Time (s)	(%)
100	5.758	0.067	0.10	5.757	0.069	0.20	52.0
200	5.793	0.031	0.39	5.794	0.033	0.75	52.6
400	5.811	0.014	1.53	5.812	0.015	2.89	53.0
800	5.820	0.005	6.10	5.821	0.005	11.5	53.2
Parameters $\alpha = 0.154$ and $\sigma = 0.0209$. Strike $K = 12.10$.							
Spot $s_0 = 16.13$. Futures $f(1) = 16.13$, $f(2) = 16.09$, and $f(3) = 17.05$.							

small interval of time Δt.

Figure (8.12) shows the $O(n^2)$ run times for the crude oil (May 1, 1997) data.

Figure (8.13) shows the seasonal shape of the option prices for crude oil in the year 1998. The corresponding pictures for natural gas and electricity may be found in Appendix C.

8.2.2 2-Factor (a) Model (2 Correlated Assets)

The value of an American call or put option at node (n, j, k) is $[(v_{nk} - s_{nj}) - K]^+$ or $[K - (v_{nk} - s_{nj})]^+$ for $j = j_{\min}, ..., j_{\max}$ and $k = k_{\min}, ..., k_{\max}$. The recursion equations

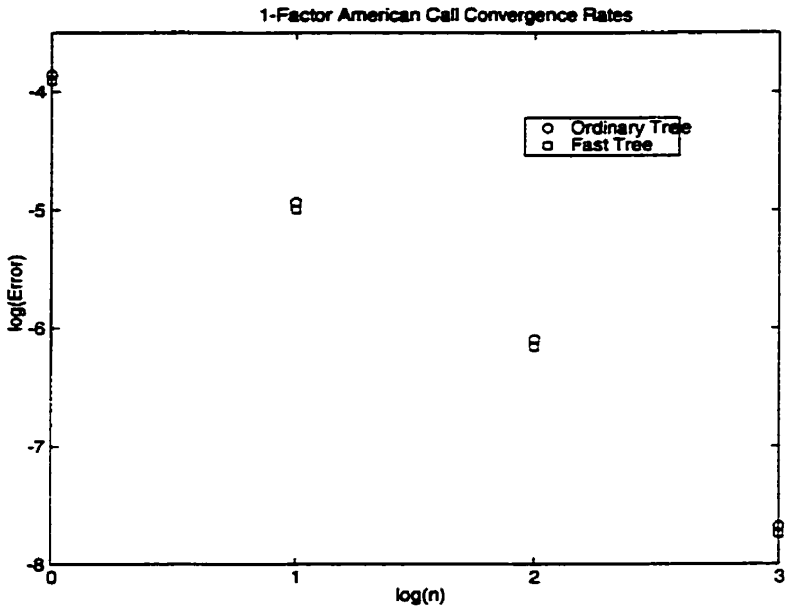

Figure 8.11: 1-Factor Model American Call Option Errors

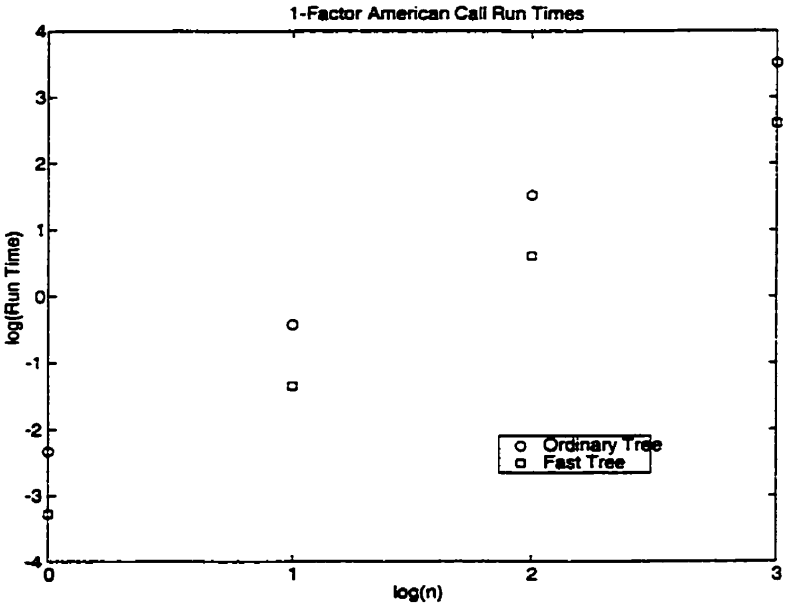

Figure 8.12: 1-Factor Model American Call Option Run Times

Table 8.12: 1-F American Call Electricity (September 1, 1999)

	Fast (0.9296)			Ordinary (0.9289)			CPU Ratio
n	Value	Error	Time (s)	Value	Error	Time (s)	(%)
100	0.9380	0.0084	0.10	0.9380	0.0091	0.20	52.0
200	0.9326	0.0029	0.39	0.9321	0.0033	0.76	51.7
400	0.9312	0.0016	1.59	0.9306	0.0017	3.04	52.4
800	0.9300	0.0004	7.54	0.9293	0.0005	14.8	50.8
Parameters $\alpha = 0.142$ and $\sigma = 0.0456$. Strike $K = 24.83$.							
Spot $s_0 = 33.10$. Futures $f(1) = 33.10$, $f(2) = 29.97$, and $f(3) = 31.35$.							

for the American option prices at node (i, j, k) are

$$C_{ijk} = \max\{[(v_{ik} - s_{ij}) - K]^+, \mathrm{e}^{-r\,\Delta t} \sum_{m=-1}^{1} \sum_{n=-1}^{1} G_j^{(mn)} C_{i+1,h_j+m,g_k+n}\}$$

$$P_{ijk} = \max\{[K - (v_{ik} - s_{ij})]^+, \mathrm{e}^{-r\,\Delta t} \sum_{m=-1}^{1} \sum_{n=-1}^{1} G_j^{(mn)} P_{i+1,h_j+m,g_k+n}\}$$

for $j = \max(i, j_{\min}),\ldots,\min(i, j_{\max})$ and $k = \max(i, k_{\min}),\ldots,\min(i, k_{\max})$ and the option prices are given by the values C_{000} and P_{000} at the root.

We evaluate all possible pairs of energy options from 1997, 1998, and 1999. The results are shown in Tables (8.13), (8.14), and (8.15).

Figure (8.14) shows the linear convergence rates for the electricity and natural gas (May 1, 1997) data.

Figure (8.15) shows the $O(n^3)$ run times for the electricity and natural gas (May

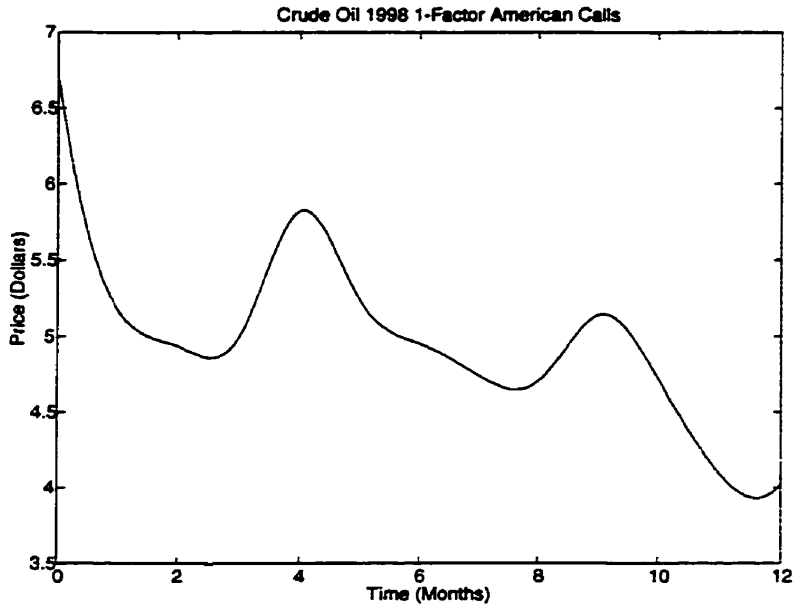

Figure 8.13: Crude Oil 1998 1-Factor Model American Call Options

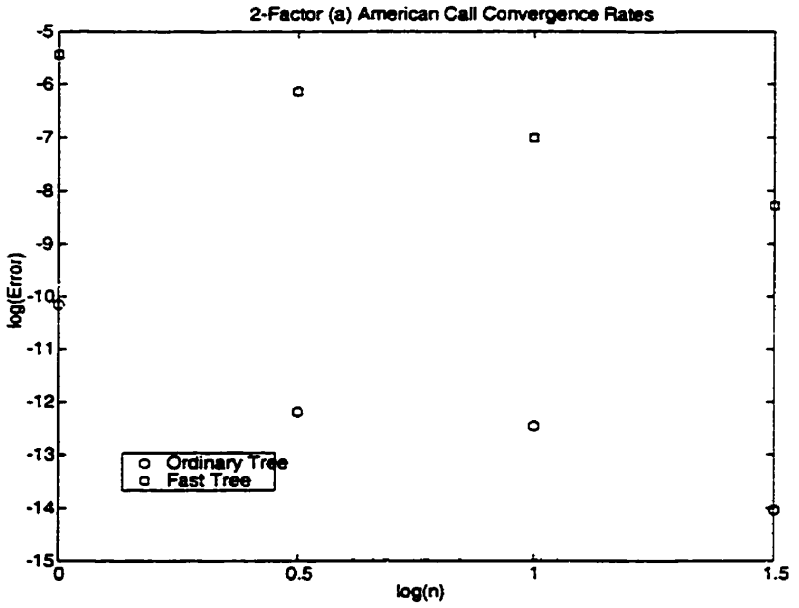

Figure 8.14: 2-Factor (a) Model American Call Errors

Table 8.13: 2-F (a) American Call Oil-Electricity (January 1, 1999)

	Fast (4.845)			Ordinary (5.301)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	4.899	0.054	0:02	5.344	0.043	0:03	66.0
141	4.878	0.032	0:05	5.328	0.027	0:07	65.0
200	4.864	0.018	0:12	5.316	0.015	0:19	64.2
283	4.853	0.008	0:33	5.307	0.006	0:53	62.1
Parameters $\alpha = 0.170$, $\beta = 0.172$, $\sigma = 0.0283$, $\tau = 0.0479$, and $\rho = -0.051$.							
Strike $K = 8.59$. Spots $s_0 = 12.05$ and $v_0 = 23.50$. Futures $f(1) = 12.05$,							
$f(2) = 12.09$, $f(3) = 12.40$, $g(1) = 23.50$, $g(2) = 21.50$, and $g(3) = 20.00$.							

1, 1997) data.

Figure (8.16) shows the seasonal shape of the option prices for the correlated pair electricity and natural gas in the year 1997. The corresponding pictures for the remaining energy pairs may be found in Appendix C.

8.2.3 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

The value of an American call or put option at node (n, j, k) is $(s_{njk} - K)^+$ or $(K - s_{njk})^+$ for $j = j_{\min}, ..., j_{\max}$ and $k = -n, ..., n$. The recursion equations for the American option

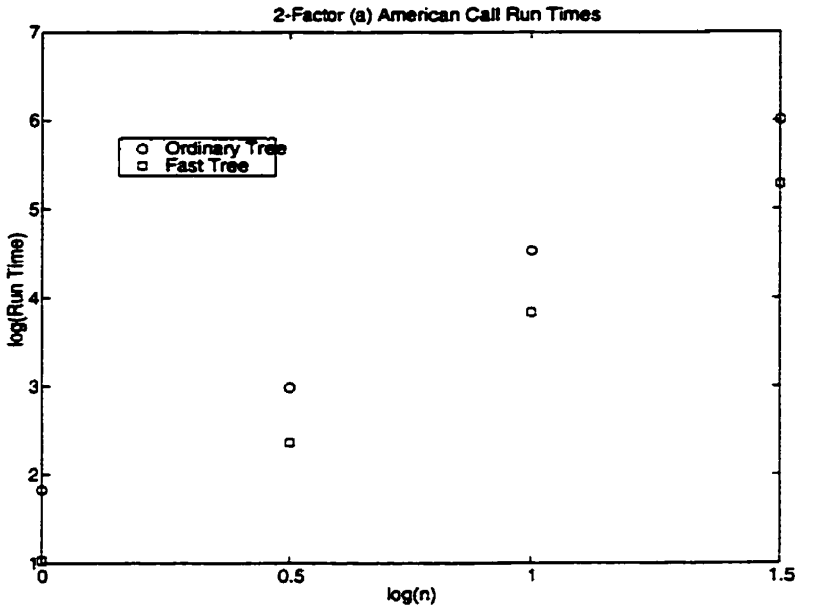

Figure 8.15: 2-Factor (a) American Call Option Run Times

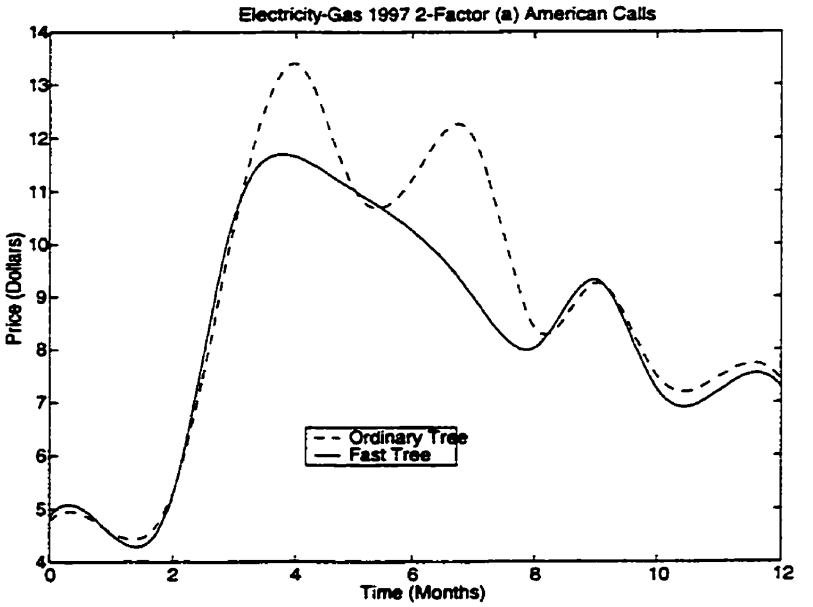

Figure 8.16: Electricity-Gas 1997 2-Factor (a) American Call Options

Table 8.14: 2-F (a) American Call Electricity-Gas (May 1, 1997)

	Fast (11.64)			Ordinary (13.40)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	11.66	0.023	0:02	13.40	0.001	0:04	57.4
141	11.65	0.014	0:05	13.40	0.000	0:08	64.9
200	11.65	0.008	0:14	13.40	0.000	0:23	61.9
283	11.64	0.003	0:39	13.40	0.000	1:04	60.4
Parameters $\alpha = 0.176$, $\beta = 0.140$, $\sigma = 0.0410$, $\tau = 0.0444$, and $\rho = 0.388$.							
Strike $K = 13.02$. Spots $s_0 = 2.24$ and $v_0 = 19.60$. Futures $f(1) = 2.24$,							
$f(2) = 2.26$, $f(3) = 2.33$, $g(1) = 19.60$, $g(2) = 23.76$, and $g(3) = 25.73$.							

prices at node (i, j, k) are

$$C_{ijk} = \max[(s_{ijk} - K)^+, \mathrm{e}^{-r\,\Delta t} \sum_{m=-1}^{1} \sum_{n=-1}^{1} G_j^{(mn)} C_{i+1,h_j+m,g_k+n}]$$

$$P_{ijk} = \max[(K - s_{ijk})^+, \mathrm{e}^{-r\,\Delta t} \sum_{m=-1}^{1} \sum_{n=-1}^{1} G_j^{(mn)} P_{i+1,h_j+m,g_k+n}]$$

for $j = \max(i, j_{\min}),\ldots,\min(i, j_{\max})$ and $k = -i,\ldots,i$ and the option prices are given by the values C_{000} and P_{000} at the root.

We evaluate typical natural gas, crude oil, and electricity options from 1997, 1998, and 1999. The results are shown in Tables (8.16), (8.17), and (8.18).

Figure (8.17) shows the linear convergence rates for the crude oil (May 1, 1998) data.

Table 8.15: 2-F (a) American Call Gas-Oil (September 1, 1998)

	Fast (3.945)			Ordinary (3.934)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
100	3.951	0.006	0:02	3.939	0.005	0:03	65.9
141	3.949	0.004	0:05	3.941	0.006	0:08	64.9
200	3.947	0.003	0:13	3.937	0.003	0:22	61.8
283	3.945	0.001	0:36	3.935	0.001	0:56	63.5
Parameters $\alpha = 0.140$, $\beta = 0.175$, $\sigma = 0.0311$, $\tau = 0.0257$, and $\rho = 0.065$.							
Strike $K = 8.69$. Spots $s_0 = 1.75$ and $v_0 = 13.34$. Futures $f(1) = 1.75$,							
$f(2) = 1.98$, $f(3) = 2.25$, $g(1) = 13.34$, $g(2) = 13.59$, and $g(3) = 14.18$.							

Figure (8.18) shows the $O(n^3)$ run times for the crude oil (May 1, 1998) data.

Figure (8.7) shows the seasonal shape of the option prices for crude oil in the year 1998. The corresponding pictures for natural gas and electricity may be found in Appendix C.

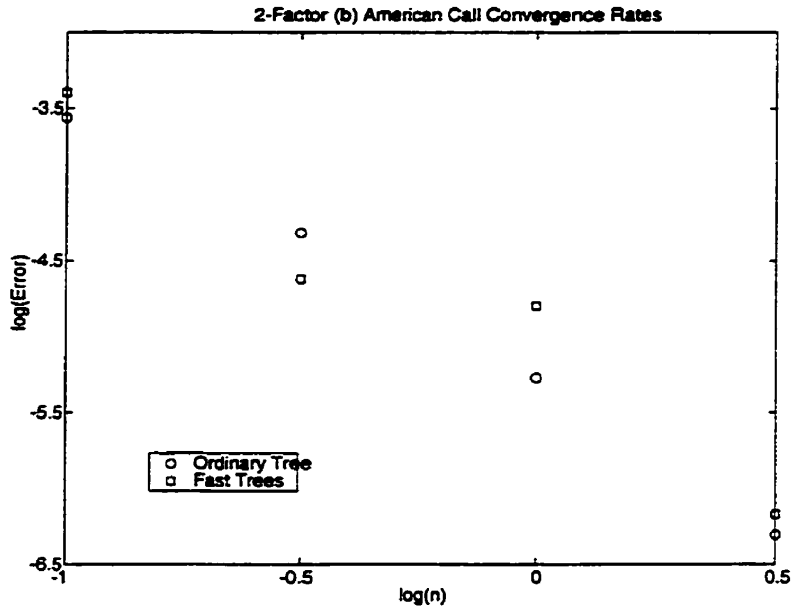

Figure 8.17: 2-Factor (b) Model American Call Errors

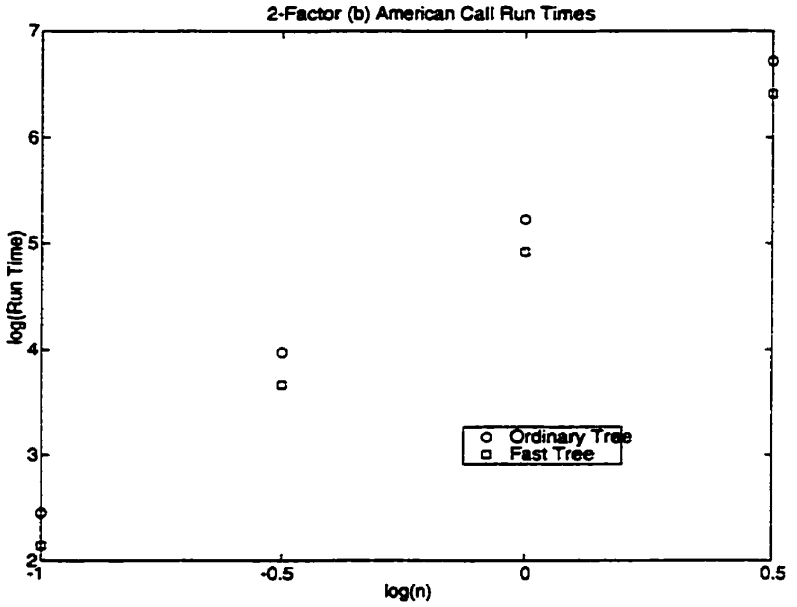

Figure 8.18: 2-Factor (b) Model American Call Option Run Times

Table 8.16: 2-F (b) American Call Natural Gas (January 1, 1997)

	Fast (0.9876)			Ordinary (1.0425)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	1.0036	0.0159	0:04	1.0594	0.0169	0:05	80.7
71	0.9962	0.0086	0:11	1.0523	0.0098	0:16	72.3
100	0.9941	0.0065	0:30	1.0476	0.0051	0:40	74.7
141	0.9895	0.0018	1:20	1.0424	0.0001	1:45	76.2
Parameters $\alpha = 0.167$, $\sigma = 0.0446$, and $\tau = 0.0315$. Strike $K = 2.17$.							
Spot $s_0 = 2.89$. Futures $f(1) = 2.89$, $f(2) = 2.62$, and $f(3) = 2.35$.							

Table 8.17: 2-F (b) American Call Crude Oil (May 1, 1999)

	Fast (5.898)			Ordinary (6.214)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	5.803	0.095	0:04	6.130	0.085	0:05	80.7
71	5.857	0.041	0:13	6.164	0.050	0:16	80.8
100	5.862	0.036	0:30	6.189	0.026	0:37	80.8
141	5.884	0.014	1:25	6.202	0.013	1:45	80.8
Parameters $\alpha = 0.154$, $\sigma = 0.0209$, and $\tau = 0.0176$. Strike $K = 12.10$.							
Spot $s_0 = 16.13$. Futures $f(1) = 16.13$, $f(2) = 16.09$, and $f(3) = 17.05$.							

Table 8.18: 2-F (b) American Call Electricity (September 1, 1997)

	Fast (12.373)			Ordinary (14.168)			CPU Ratio
n	Value	Error	Time (m:s)	Value	Error	Time (m:s)	(%)
50	12.339	0.0345	0:05	14.171	0.0035	0:06	80.7
71	12.369	0.0046	0:14	14.274	0.1064	0:17	80.9
100	12.455	0.0815	0:35	14.130	0.0372	0:44	80.7
141	12.408	0.0343	1:40	14.188	0.0200	2:09	77.4
Parameters $\alpha = 0.142$, $\sigma = 0.0456$, and $\tau = 0.0394$. Strike $K = 24.83$.							
Spot $s_0 = 33.10$. Futures $f(1) = 33.10$, $f(2) = 29.97$, and $f(3) = 31.35$.							

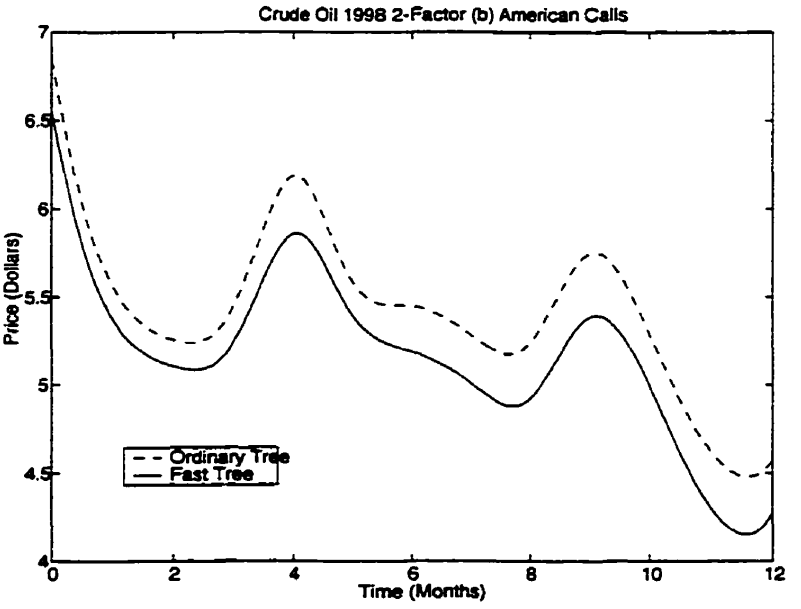

Figure 8.19: Crude Oil 1998 2-Factor (b) American Call Options

CHAPTER 9
CONCLUSION

This thesis started with a review of the elements of stochastic calculus in Chapter (2). We saw that random variables can be decomposed into a deterministic part and a stochastic part. The deterministic part is determined by the ordinary part of the differential equation while the stochastic part is the result of a Brownian motion. The deterministic part is responsible for the mean of a random process and the stochastic part reveals the standard deviation. Although this is a simple structure, stochastic differential equations can be used to model a wide range of natural phenomena, including the unpredictable fluctuations of energy spot prices.

We then defined the concept of mean reversion in Chapter (3) and explained why it was a suitable model for capturing the inherent seasonality of energy prices. We introduced the Pilipovic equations and generalized them to suit our own purposes by allowing the seasonal parameter to be an unknown function of time. We went on to explain how choosing this parameter to match futures prices has the desired effect of rendering the probabilities risk neutral; that is, all investment opportunities in the energy market, have no net present value. In this thesis, we considered three versions of the generalized Pilipovic equation:

(1) 1-Factor Model

(2) 2-Factor (a) Model (Two Correlated Assets)

(3) 2-Factor (b) Model (Single Asset with Stochastic Long Term Mean)

In Chapter (4), we developed reliable methods for estimating the parameters in the Pilipovic equations. We applied these methods to natural gas, crude oil, and electricity data from the period 1997-1999. Also, we demonstrated the correctness of our routines by simulating random sample paths with known sets of parameters and then using our programs to recover them. In all cases, the sample means were close to the original parameter values in that they were within tolerance levels established by examining the sample standard deviations. Of particular interest was the successful recovery of the volatility of the long term mean - an unobservable stochastic variable!

The fundamental relationship of futures prices and expected future spot prices was stated and proved in Chapter (5). We verified by statistical analysis that the systematic risk in the energy markets under consideration is small enough to be ignored. This allowed us to identify futures prices and expected future spot prices for the purpose of option evaluation.

In Chapter (6) we derived numerical methods for constructing what we have termed ordinary and fast trinomial trees. The fast tree splits a trinomial tree into a preliminary tree that represents the stochastic part of the random process, and shifts it branch by branch onto a final tree according to the flow of the deterministic part. This is done so that the median nodes of all the branches coincide with the corresponding futures prices. The fast trees tend to be more elegant in form and function. That is, they are easier to implement and the resulting programs run faster.

We set the stage for energy derivatives in Chapter (7) by applying our numerical methods to the mean reverting models. We wrote down concise algorithms of the implementations for the Pilipovic variations for both ordinary and fast trees.

In Chapter (8) we did a thorough numerical analysis of the ability of our programs to evaluate European and American style call options. We focused our attention on three aspects of performance:

(1) Consistency: Do ordinary and fast trees give the same results?

Table (9.1) answers this question.

Table 9.1: Ordinary and Fast Tree Option Price Accuracy

	European	American
1-Factor	Yes	Yes
2-Factor (a)	Yes	No
2-Factor (b)	No	No

In the 1-Factor model, European and American options calculated by ordinary and fast trees were virtually identical, giving strong supporting evidence to the Fast Trinomial Tree Conjecture. The 2-factor (a) model produced similar European option prices while American option values differed significantly. In the case of the 2-factor (b) model, the ordinary trees revealed option prices that were consistently higher than those for fast trees. However, in all cases the option prices calculated by fast trees were within acceptable tolerance levels. So although ordinary and fast trees do not agree

in all cases, the fast trees can be safely used for the purpose of evaluating derivative securities under all three of the models under study.

(2) Convergence: Do ordinary and fast trees have similar convergence rates?

We used an Euler method in the construction of our ordinary and fast trinomial trees. Therefore we expect their errors to be a linear functions of the length of a small time interval Δt. This linear relationship was verified numerically. Furthermore, ordinary and fast trees were demonstrated to have the same rates of convergence in that similar errors were revealed for both types of trees for equal numbers of time steps n.

(3) Speed: Are fast trees computationally faster than ordinary ones?

The CPU time ratios between fast and ordinary trees are displayed as percentages in Table (9.2).

Table 9.2: Ordinary and Fast Tree CPU Time Ratios

	European	American
1-Factor	45	50
2-Factor (a)	65	65
2-Factor (b)	95	80

In all cases, the fast trees were indeed faster than ordinary ones. The outcome was most dramatic in the 1-factor model, where the fast trees were twice as fast as their slow counterparts! The 2-factor (a) model showed a significant improvement for fast

trees, while the 2-factor (b) model speed increase for fast trees was barely satisfactory.

Bibliography

[Grimmett & Stirzaker 1995] Grimmett, G. R. and D. R. Stirzaker "Probability and Random Processes 2nd ed." *Oxford University Press*, 1995.

[Hull 1999] Hull, John C. "Options, Futures, & other Derivatives 4th ed." *Prentice Hall*, 1999.

[Hull & White 1994 (b)] Hull, John C. and Alan White "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models" *Journal of Derivatives*, Winter 1994, pp. 37-48.

[Hull & White 1994 (a)] Hull, John C. and Alan White "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models" *Journal of Derivatives*, Fall 1994, pp. 7-16.

[Hull & White 1993] Hull, John C. and Alan White "One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities" *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 2, June 1993, pp. 235-253.

[Hull & White 1990] Hull, John C. and Alan White "Valuing Derivative Securities using the Explicit Finite Difference Method" *Journal of Financial and Quantitative Analysis*, Vol. 25, No. 1, March 1990, pp. 87-100.

[Kloeden & Platen 1999] Kloeden, Peter E. and Eckhard Platen "Numerical Solution of Stochastic Differential Equations" *Springer-Verlag*, 1999.

[Lari-Lavassani, Simchi & Ware 2000] Lari-Lavassani, Ali, Mohamadreza Simchi, and Antony Ware "A Discrete Valuation of Swing Options" http://finance.math.ucalgary.ca/publications.shtml Preprint, 2000.

[Perko 1996] Perko, Lawrence "Differential Equations and Dynamical Systems 2nd ed." *Springer-Verlag*,

1996.

[Pilipovic 1997] Pilipovic, Dragana "Energy Risk: Valuing and Managing Energy Derivatives" *McGraw-Hill*, 1997.

[Weiss 1999] Weiss, Neil A. "Introductory Statistics 5th ed." *Addison Wesley Longman Inc.*, 1999.

[Wilmott 1998] Wilmott, Paul "Derivatives: The Theory and Practice of Financial Engineering" John Wiley & Sons, 1998.

APPENDIX A.
PARAMETER ESTIMATION

Here are some more pictures of parameter estimation of the 2-factor (b) model using historical data. Note that although the drift $\beta(t)$ of the long term mean is actually an unknown function of time, we display a single number β (obtained by taking the average of $\beta(t)$ over the year) in the diagrams that represents the yearly trend in the energy market. Also, bear in mind that this $\beta(t)$ is merely a side effect of the calculation of the parameters α, σ, and τ. We do not use this $\beta(t)$ for modeling purposes, and include it here purely for interests sake. The $\beta(t)$ used in option evaluation is found by matching the futures prices and the expected future spot prices. Hence our model is forward looking in that it takes advantage of the beliefs of traders in the energy market.

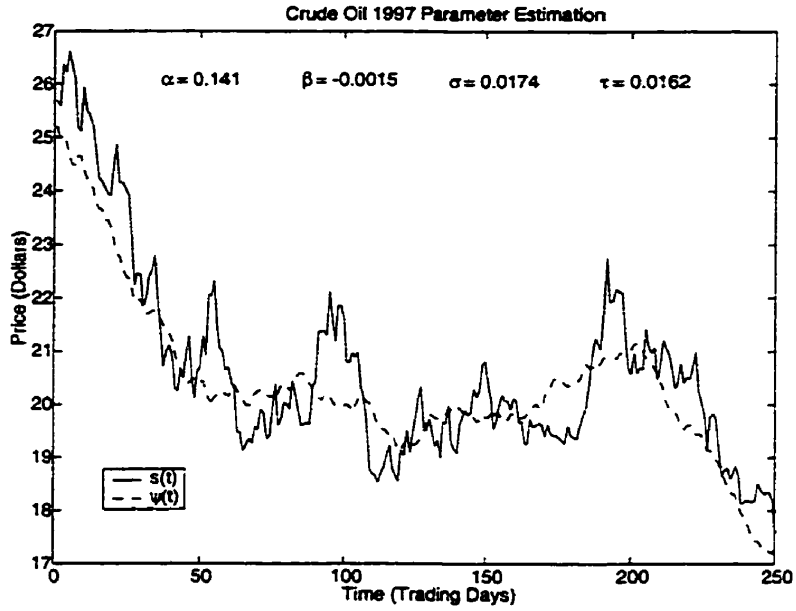

Figure A..1: Crude Oil 1997 Parameter Estimation

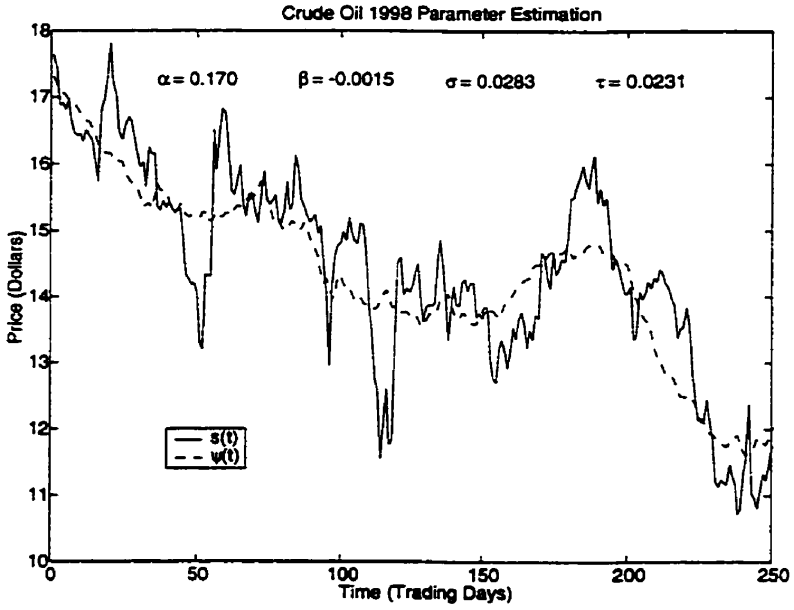

Figure A..2: Crude Oil 1998 Parameter Estimation

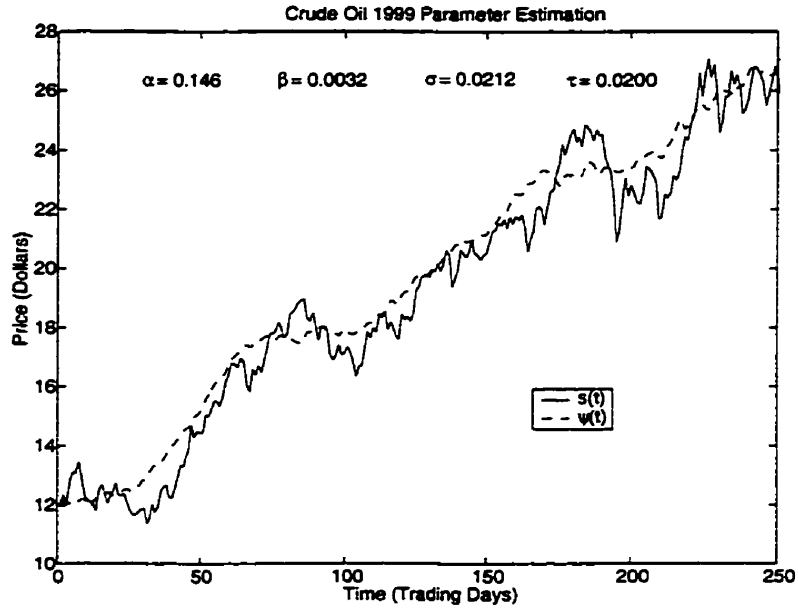

Figure A..3: Crude Oil 1999 Parameter Estimation

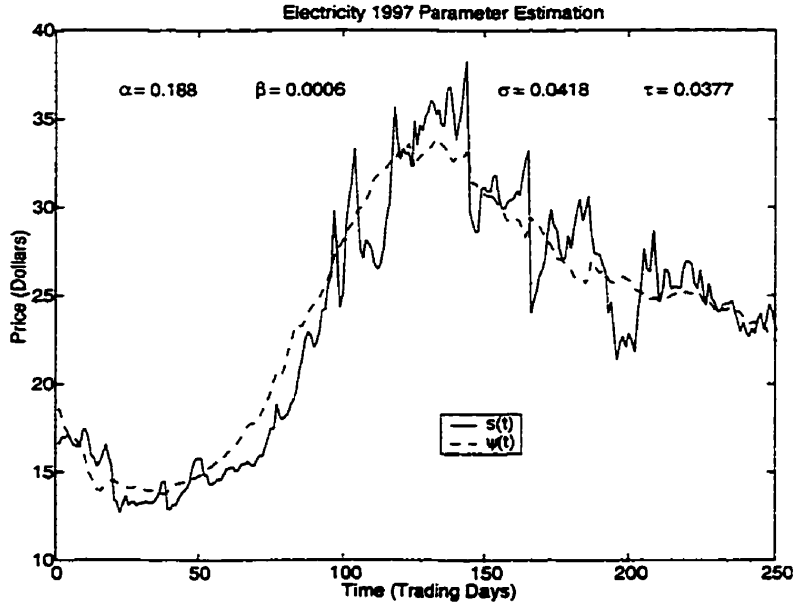

Figure A..4: Electricity 1997 Parameter Estimation

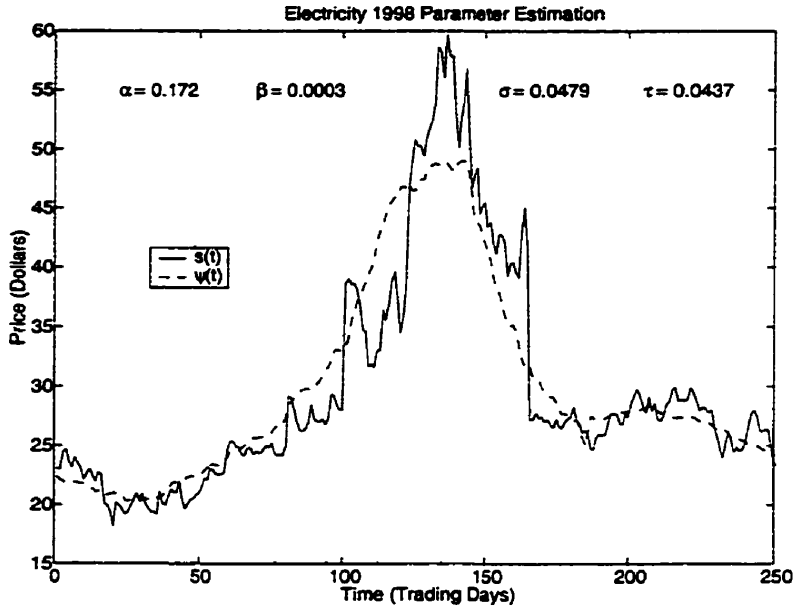

Figure A..5: Electricity 1998 Parameter Estimation

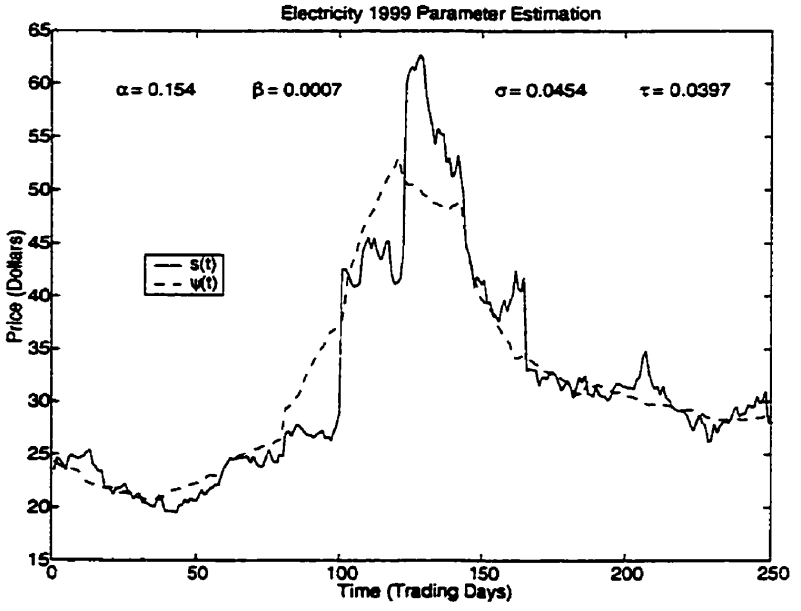

Figure A..6: Electricity 1999 Parameter Estimation

APPENDIX B.
FUTURES PRICES

Here are some more pictures of futures prices and realized spot prices for natural gas, crude oil, and electricity. Remember that in Section (5.3) we did a statistical analysis that revealed no evidence to deny the simple hypothesis that futures prices are good estimators of expected future spot prices. Figures (B..1)-(B..6) illustrate this fundamental relationship.

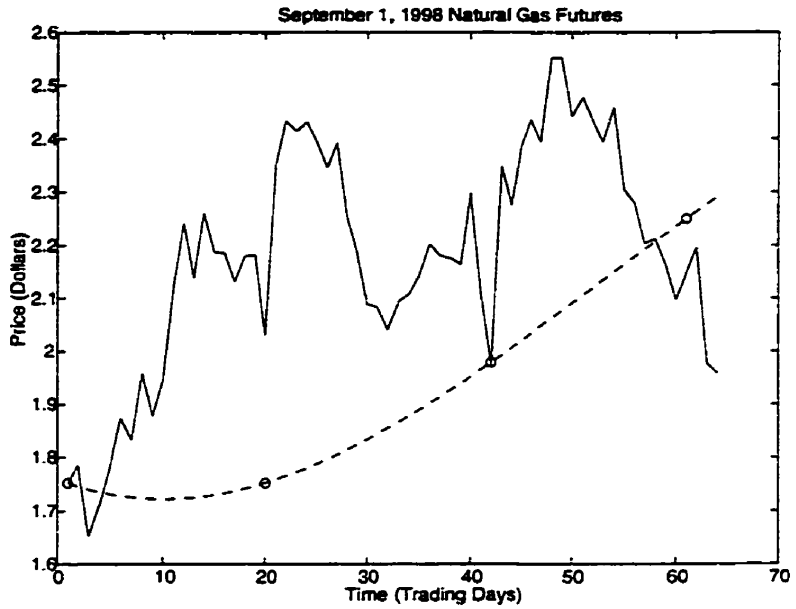

Figure B..1: September 1, 1998 Natural Gas Futures

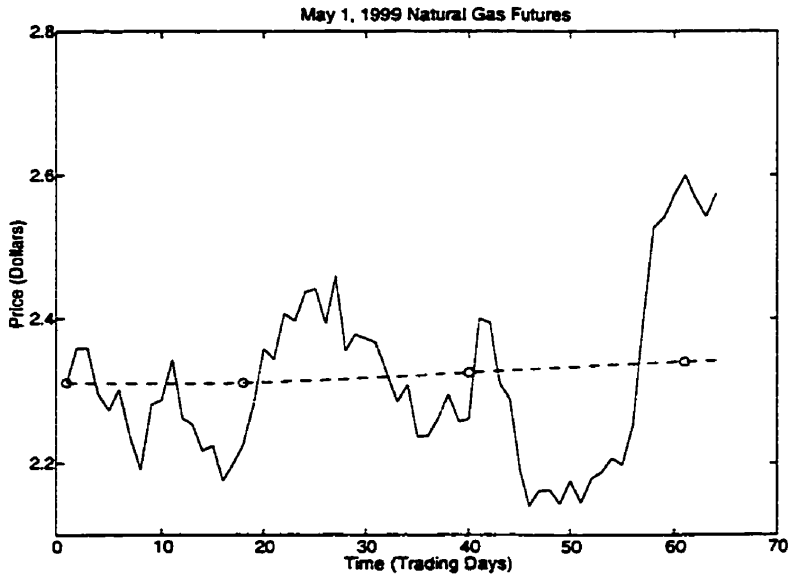

Figure B..2: May 1, 1999 Natural Gas Futures

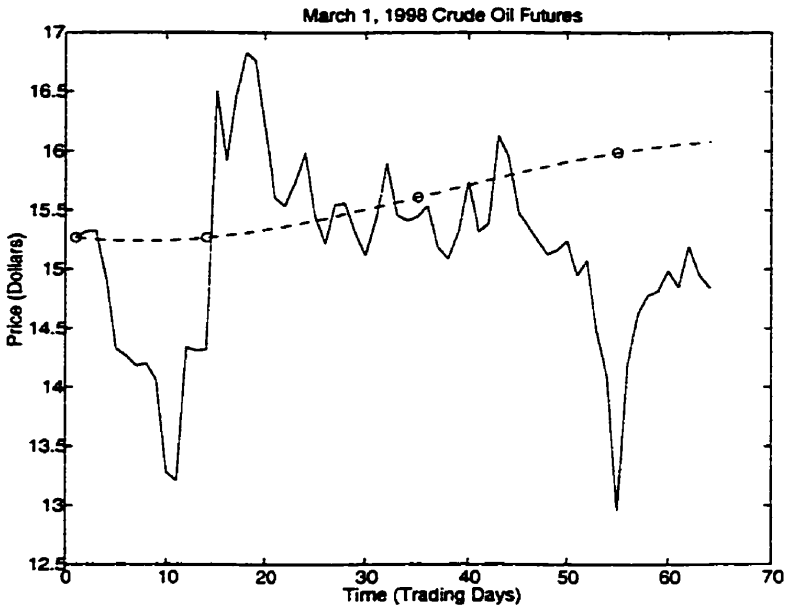

Figure B..3: March 1, 1998 Crude Oil Futures

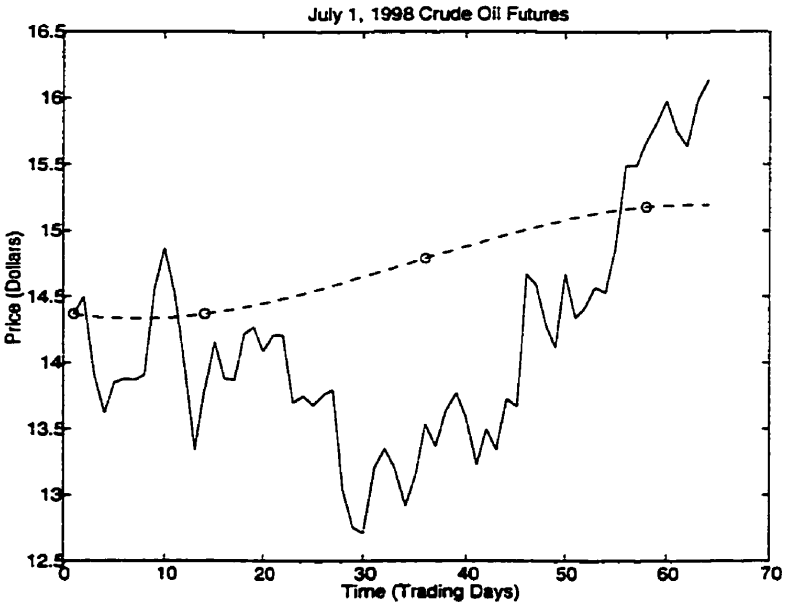

Figure B..4: July 1, 1998 Crude Oil Futures

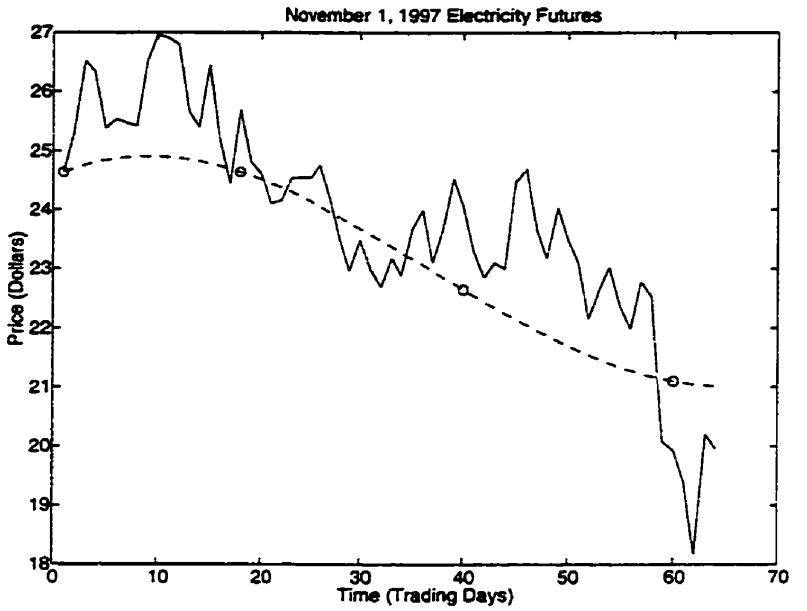

Figure B..5: November 1, 1997 Electricity Futures

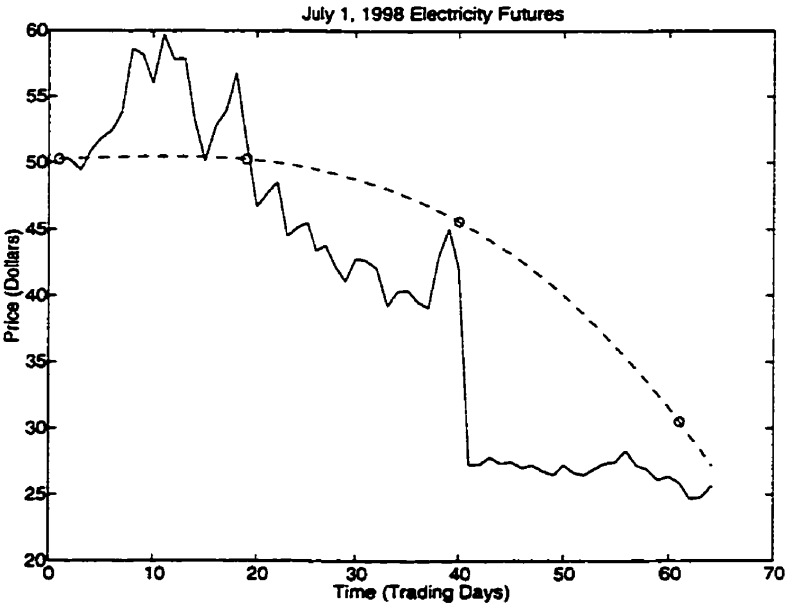

Figure B..6: July 1, 1998 Electricity Futures

APPENDIX C.
OPTION VALUES

Figures (C..1)-(C..12) show the seasonal shapes of option prices for natural gas, crude oil, and electricity over the period 1997-1999 under all three of the models under consideration. Feel free to compare the option prices revealed by the 1-factor and 2-factor (b) models for corresponding data sets. These should be similar since the model we select for the energy price process should not greatly affect the option values.

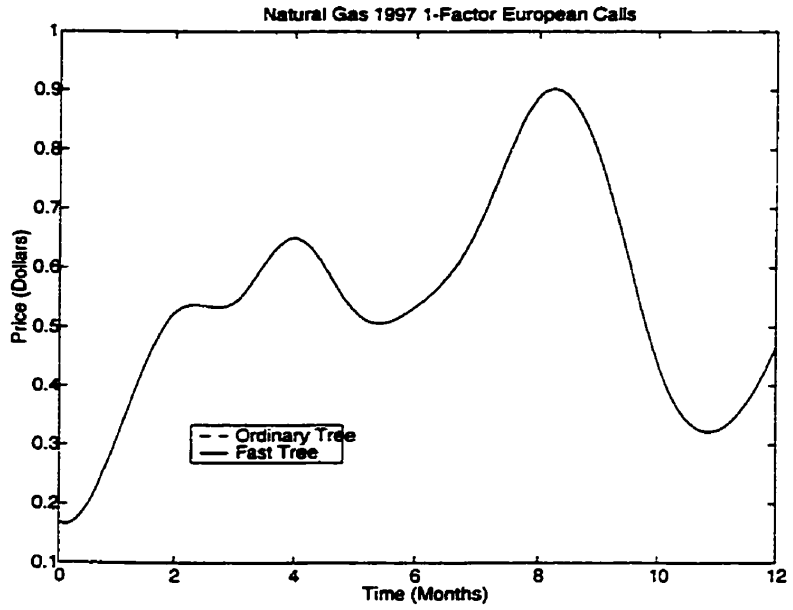

Figure C..1: Natural Gas 1997 1-Factor Model European Call Options

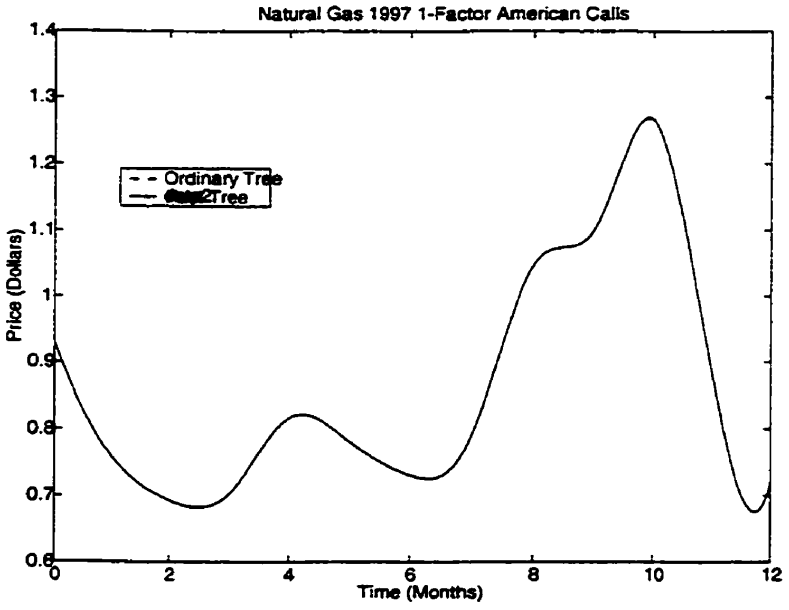

Figure C..2: Natural Gas 1-Factor Model American Call Options (1997)

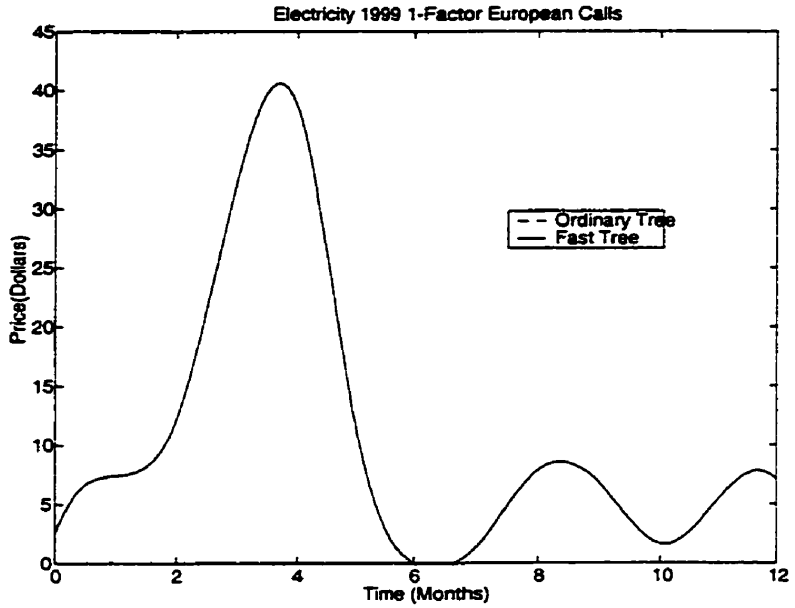

Figure C..3: Electricity 1999 1-Factor Model European Call Options

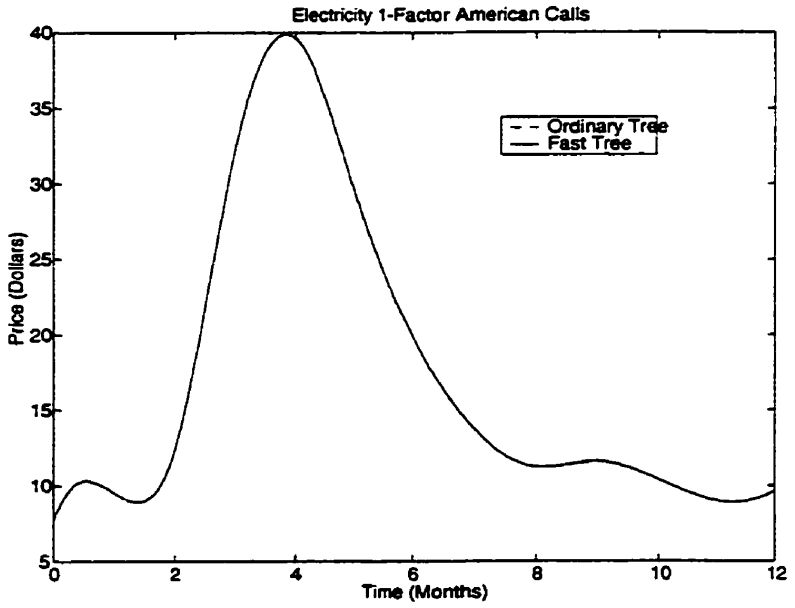

Figure C..4: Electricity 1999 1-Factor Model American Call Options

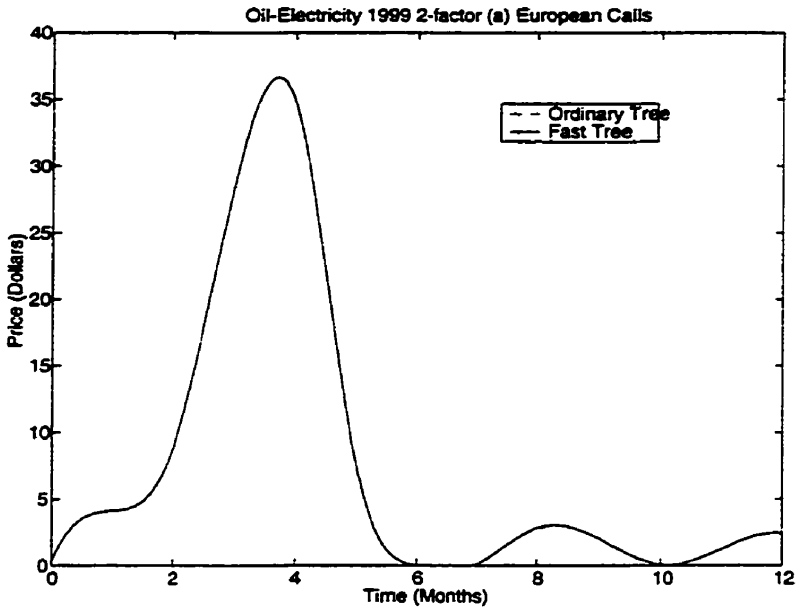

Figure C..5: Oil-Electricity 1999 2-Factor (a) Model European Call Options

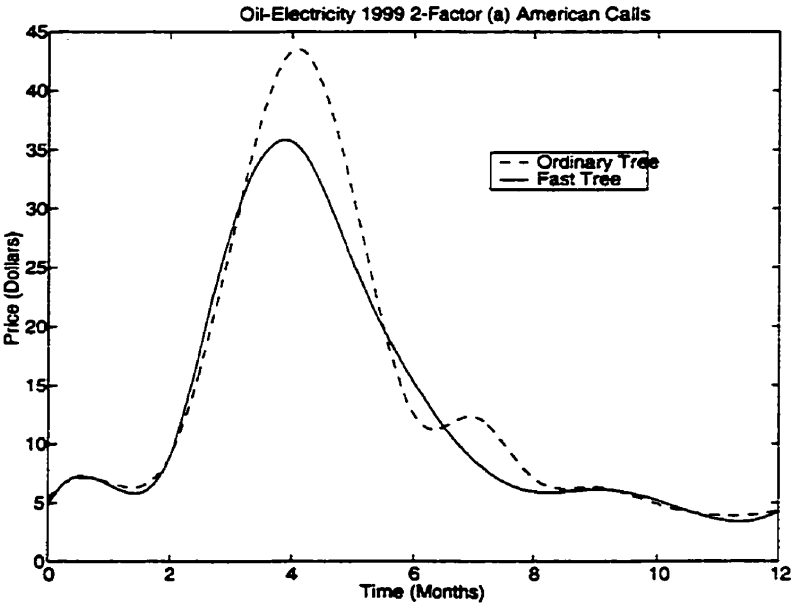

Figure C..6: Oil-Electricity 1999 2-Factor (a) American Call Options

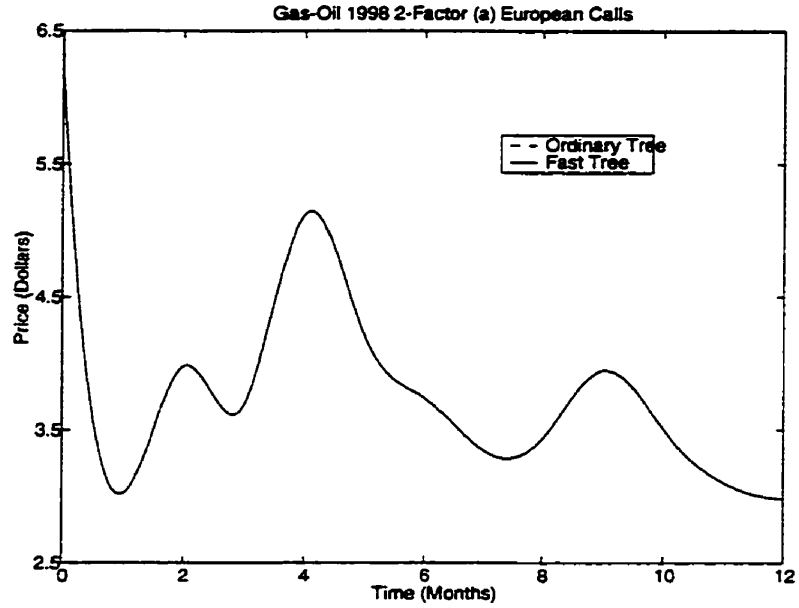

Figure C..7: Gas-Oil 1998 2-Factor (a) Model European Call Options

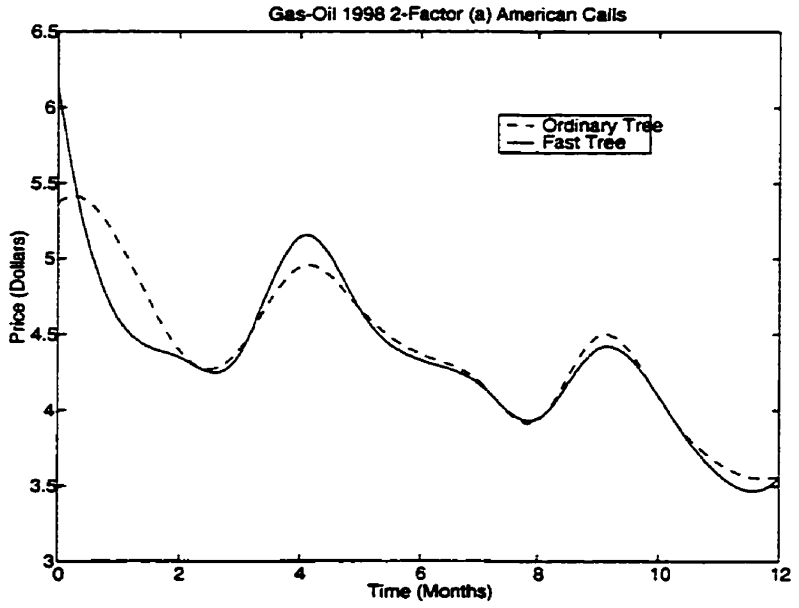

Figure C..8: Gas-Oil 1998 2-Factor (a) American Call Options

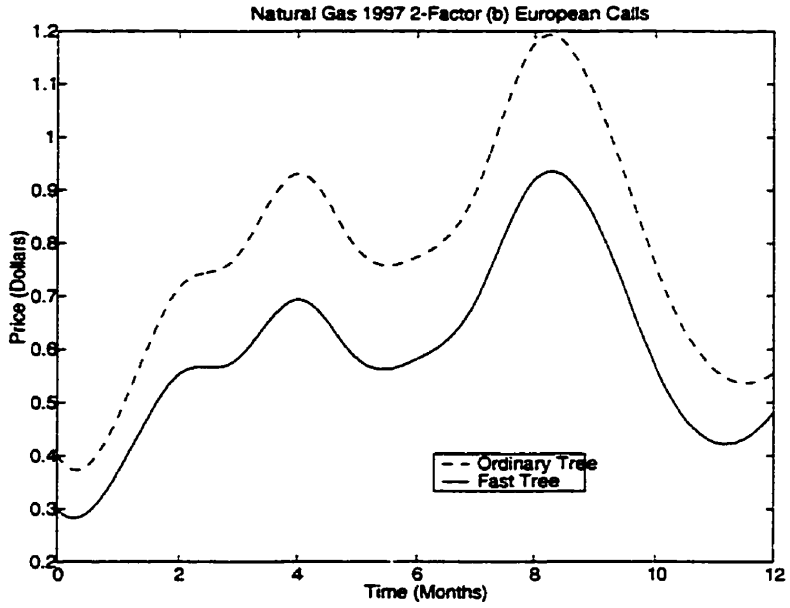

Figure C..9: Natural Gas 1997 2-Factor (b) European Call Options

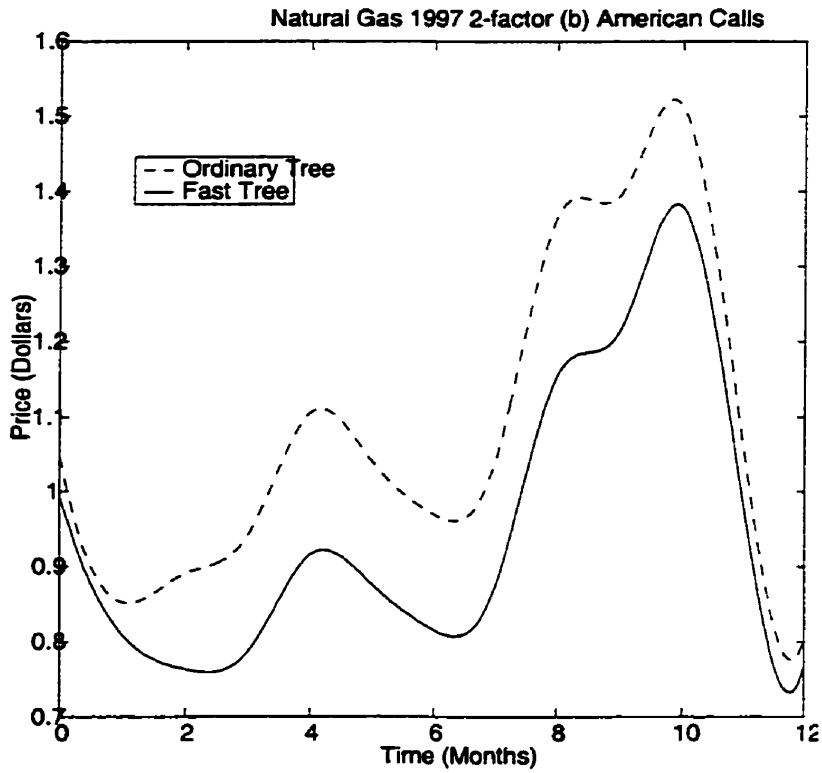

Figure C..10: Natural Gas 1997 2-Factor (b) American Call Options

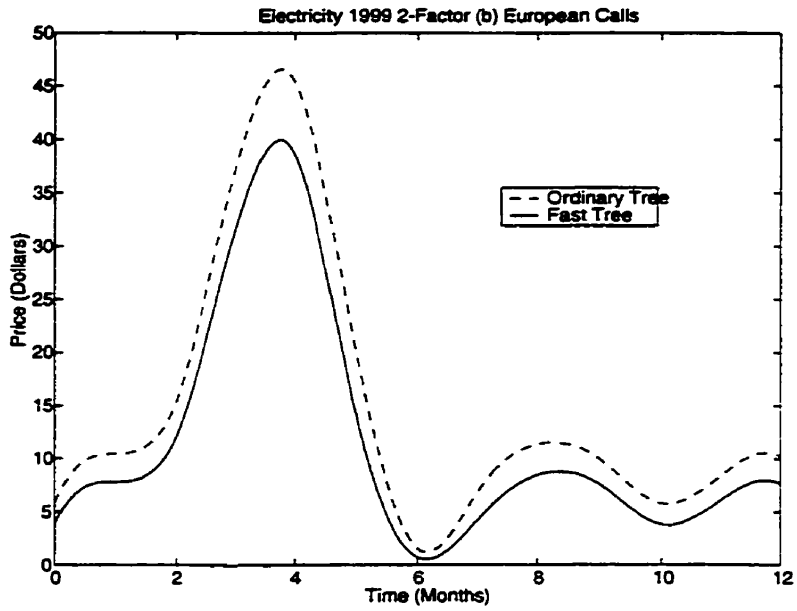

Figure C..11: Electricity 1999 2-Factor (b) European Call Options

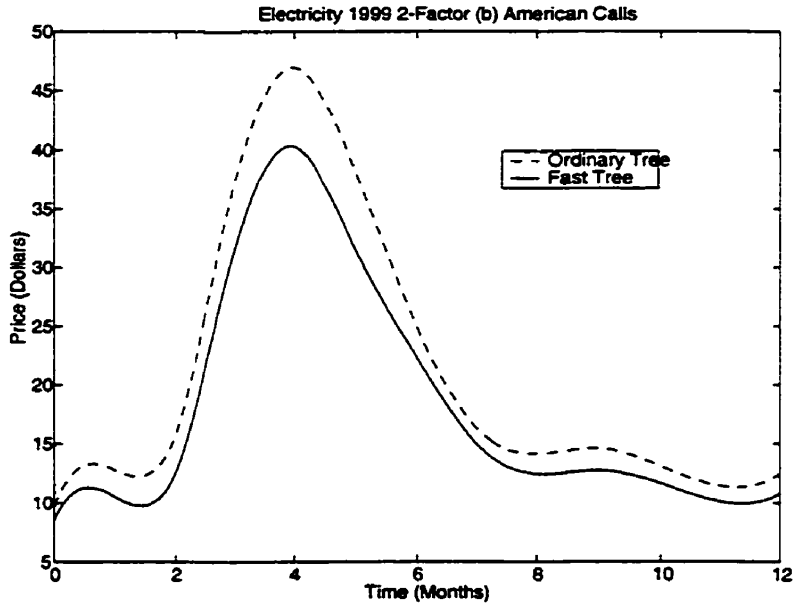

Figure C..12: Electricity 1999 2-Factor (b) American Call Options