

Binomial and Trinomial Tree Methods in Derivatives Pricing

Ettienne van Wyk

Submitted in accordance with the requirements for the degree of

Magister Scientiae

in the

**Faculty of Natural Sciences
Centre for Business Mathematics and Informatics**

at the

**North-West University
Potchefstroom 2520
South Africa**

November 2006

**Supervisor: Prof. P.J. de Jongh
Co-Supervisor: Prof. E. Maré**

Acknowledgements

Graag sou ek die volgende persone wou bedank vir hulle gewaardeerde bydrae tot hierdie verhandeling:

- Prof. P. J. de Jongh,
- Prof. E. Maré,
- Mev. A. Maré en die ander lede van die Maré gesin,
- Dr. W. Malan,
- Mnr. B. Geddes,
- Mnr. M. van Wyk (snr),
- en Mev. P. Horn.

Abstract

Tree methods for the valuation of financial derivative securities represent a recognized and well-established pricing paradigm. It has formed part of the financial engineer's "toolbox" for close on 30 years. The tree approach is multi-dimensional though: there are for example, various ways in which trees can be parametrized. Incorporating eccentricities of the financial markets like the paying of discrete dividends and volatility skews add some further complexity to the approach. A full perspective on the place of tree methods requires knowledge of the relation between the said and other pricing paradigms like numerical integration techniques and finite difference methods. Convergence properties are of definite interest to a practitioner as well. This dissertation aims to provide a general introduction to tree methods, and well by treating on the enumerated issues .

Uittreksel

Binomiaalboom metodes verteenwoordig 'n gevestigde en gerespekteerde waardasie metodologie vir finansiële afgeleide instrumente. Alhoewel die onderliggende beginsels relatief eenvoudig is, is die paradigma multi-dimensioneel, en wel ten opsigte van: die parametrisering van die algoritme, die kalibrasie na mark eksentrisiteite soos diskrete dividend betaling op sommige onderliggende bates, asook konvergensie karaktereenskappe. Verder, om 'n geheel perspektief te kry op die metodologie, is dit noodsaaklik om die verband tussen binomiaalmetodes en ander waardasie benaderings te verstaan. Hierdie verhandeling stel dit self ten doel om 'n algeme agtergrond te gee op die binomiaalwaardasie benadering, maar met die fokus op die genoemde kwessies.

“... however confusedly and meaninglessly our way may deviate from our desires,
after all it does lead us inevitably to our invisible goal.”

Stefan Zweig,
“The World of Yesterday”.

and

Adagio from the F-major string quintet by Anton Bruckner.

Contents

Acknowledgements	iii
List of symbols	xiii
List of Abbreviations	xiv
1 Introduction	1
2 Random Walk Models and Derivatives Pricing	7
2.1 Background	8
2.2 Bernoulli Experiments and Random Walks	8
2.3 Risk Neutrality and Arbitrage-Free Prices	12
2.4 Bernoulli Experiments and Arbitrage-free Derivatives	14
2.5 Summary	16
3 Tree Construction Methods	17
3.1 Background	18
3.2 Tree Calibration	18

<i>CONTENTS</i>	vii
3.2.1 Moment Matching	19
3.2.2 Partial CDF Matching	20
3.3 Well-known Tree Parameterizations	22
3.3.1 Cox, Ross and Rubinstein “Diffusion” Parameterization	22
3.3.2 Cox, Ross and Rubinstein “Jump” Parameterization	24
3.3.3 Trigeorgis’ Parameterization	25
3.3.4 Jarrow and Rudd Parameterization	27
3.3.5 Tian Parameterization	32
3.3.6 Leisen and Reimer’s Parameterization	34
3.3.7 Pricing Examples	39
3.4 Probability Measure Dichotomy	39
3.5 On Implementing Binomial Algorithms	41
3.6 Summary	42
4 Equivalence Explored	43
4.1 Background	44
4.2 The Black and Scholes Closed-Form Solution	45
4.3 Finite Difference Schemes	51
4.3.1 Equivalence to the Black and Scholes Partial Differential Equation	52
4.3.2 Equivalence to Explicit Finite Difference Schemes	54
4.4 Numerical Integration Schemes	57
4.5 Binomial and Trinomial Trees	62

4.6	Summary	66
5	Real Life Binomial Pricing	67
5.1	Background	68
5.2	The Term Structure of Riskless Return	68
5.3	Stock Borrow Costs	70
5.4	Volatility Surfaces and Implied Trees	72
5.4.1	Rubinstein Implied Trees	74
5.4.2	Derman and Kani Trees	78
5.5	The Indiscretion of Discrete Dividends	80
5.5.1	The Displaced Diffusion Model	81
5.5.2	The Accumulation Model	83
5.5.3	The Deterministic Jump Model	84
5.6	Summary	86
6	Exotic Pricing	87
6.1	Background	88
6.2	American and Bermudan Options	88
6.3	Discrete Barrier Options	91
6.4	Forward Starting Options	95
6.5	Other Exotic Pay-Offs	96
6.6	The Curse of Multi-Dimensionality	96

<i>CONTENTS</i>	ix
6.7 Summary	97
7 Numerical Issues	98
7.1 Background	99
7.2 Sense and Instability	100
7.3 Convergence	105
7.3.1 Convergence Patterns	105
7.3.2 Convergence Speed	108
7.4 Summary	112
8 Conclusion	113
Appendix - Matlab Code	121

List of Tables

3.1	Results of the moment matching procedure	20
3.2	The arbitrage bias present in Jarrow and Rudd option valuations . .	29
4.1	Numerical integration and binomial pricing	61
4.2	Numerical integration and binomial pricing ; post a discrete calculation point matching.	62
5.1	Price tree under the “Deterministic Jump” model	86
5.2	Option fair value under the different paradigms.	86
7.1	Errors in valuation - vanilla European option, at the money.	111
7.2	Errors in valuation - vanilla European option, 125% in the money. . .	111
7.3	Errors in valuation - vanilla European option, 75% moneyness.	111

List of Figures

2.1	Illustration of Bernoulli experiments and random walks	10
2.2	“Proper” and “Binomial” trees	11
3.1	Graphical view of the CDF matching process and results	21
3.2	Graphical view of the pricing bias inherent in Jarrow and Rudd trees.	30
3.3	Convergence for the respective parameterizations (vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 10	33
3.4	Convergence for the respective parameterizations (vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 12.5	35
3.5	Convergence for the respective parameterizations (Vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 7.5	39
4.1	Integration of the pay-off function with respect to the continuous or discrete (approximate) CDF.	60
4.2	Convergence diagram of numerical integration schemes versus binomial methods. (At the money, vanilla European option.	62
4.3	Convergence diagram of numerical integration schemes versus binomial methods. (75% vanilla European call option.)	63

4.4	The relationship between binomial and trinomial trees; the colored circles and bold vectors represent the trinomial tree overlay.	65
5.1	All Share 40 Implied Volatilities (January 2004)	73
5.2	Tree dynamics implied by the “Displaced Diffusion” model	82
5.3	Tree dynamics implied by the “Accumulation” model	83
5.4	Tree dynamics implied by the “Deterministic Jump” model	85
6.1	Convergence of barrier and vanilla options within a tree valuation. . .	93
7.1	The Lorenz attractor.	99
7.2	Unstable behavior observed for the “Cox et al” pricing of a vanilla European option.	101
7.3	Risk neutral probabilities implied by the “Cox et al” binomial method.	102
7.4	Tree-implied risk neutral probability distribution; evolution over six and seven time steps.	107
7.5	Wave-like convergence patterns observed for tree pricing.	108
7.6	Saw-tooth convergence patterns observed for barrier options.	109

List of symbols

The principal symbols used in the dissertation are:

$S(t), S$	: The Spot price of an asset at an arbitrary time t
K	: Strike price of a derivative
T	: Expiry of a derivative
$f(S(t), t)$	: Fair value of a derivative at time t , given an underlying value of $S(t)$
u	: “Up” factor in a binomial tree
d	: “Down” factor in a binomial tree
p, q	: Risk neutral probability of an occurrence of “Up”
μ_1, μ_2, μ_3	: First three null point moments
Q	: Stock borrow costs
σ	: Return volatility of a security
Δt	: Time interval of discrete length
dt	: Time interval of infinitesimal length
ξ	: Discrete dividends receivable
δ	: Dirac delta function
σ	: Return volatility of a security
dB_t	: Wiener process
dP_t	: Poisson process
$E^C\{\cdot \cdot\}$	: Conditional expected value under the probability measure C
$Pr.^C\{\cdot \cdot\}$	: Conditional probability under the probability measure C

List of Abbreviations

CRR : Cox, Ross and Rubinstein
JR : Jarrow and Rudd
PDE : Partial differential equation
SDE : Stochastic differential equation
BS : Black and Scholes

Chapter 1

Introduction

Cox, Ross and Rubinstein introduced the binomial tree procedure for derivatives pricing with their seminal paper “Option Pricing: A Simplified Approach”, (1979), [23].¹ After the paradigm shifting work of Black and Scholes on the pricing of derivatives securities, Cox et al (1979), proposed a novel, easy to understand and independent pricing approach. What differentiated their scheme from other finite difference schemes was a starting point other than the Black and Scholes partial differential equation: binomial trees in essence require no reference to the said. Rather, the basis of the Cox et al pricing perspective consists of an intuitive no-arbitrage reasoning, based on the ability to replicate a derivative using the underlying, and riskless assets, within in a two-outcome world, it must be added. As such the binomial tree method, although very much a numerical technique, akin to finite difference methods, finds itself in a category of its own. This procedure, 30 years after its introduction remains an invaluable part of any financial engineer’s “toolbox” in the evaluation of derivative securities. The aim of this dissertation, could be summarized as: the exposition of the fundamental ideas underlying the binomial tree procedure, its various incarnations, its consistency with other pricing paradigms, the application of binomial lattice valuations, and finally the exploration of some ideas and results concerning its convergence. But first, a general investigation of binomial methods’ history and its associated academic literature.

Since the publication of the article that introduced binomial methods, many have realized the power and elegance of the procedure; much has been done to apply, further and extend the methodology as well as to investigate its relationships with other pricing approaches. Concerning the extensions alluded to: the discretization framework

¹Some authors also acknowledge Rendleman and Barter (1979) with “Two-State Asset Pricing”, [4], as co-founders of the procedure.

proposed by Cox et al (1979), [23], leaves ample room for individual interpretation; as a consequence a number of different ways in which the tree process is parametrised have been formulated. Jarrow and Rudd, in their 1983 book “Option Pricing”, [49], for example, proposed a price tree where the outcomes are equiprobable, i.e. “up” and “down” movements are equally likely to occur from the same starting point. The reason given for the latter was an attempt towards greater accuracy. Interesting: as it turned out, the result of setting the risk neutral probability to one half was to violate exactly the principle upon which the whole binomial method is based, namely arbitrage-free pricing (this is shown in Chapter Three). Also, in violating the said condition, no advantage was gained in terms of order of convergence: Leisen and Reimer [55], established an order of convergence of one for Jarrow and Rudd trees, the same as that shown for the original Cox et al specified lattice!

Cox et al binomial trees rely on a very fundamental approximation in the derivation of tree parameters, namely that the expectation of log-returns, squared, (over an arbitrary time interval) is equal to the variance. Although not affecting the capacity to converge, it does introduce some numerical instability for certain parameter choices (further explored in Chapter Seven). Recognizing this shortcoming, Trigeorgis (1991) [86], altered the tree parameterization so as to take cognizance of the subtle difference mentioned. The result of this alteration is a procedure that produces prices almost exactly equal to the Cox et al tree, but without the instability associated with low risk free return rates and high price volatility.

As prior mentioned, the tree parameter derivation process leaves ample room for interpretation: the principal reason for this being an additional degree of freedom in the system of equations characterizing the process. Cox et al (1979) [23], used this to add the condition that “up” and “down” movements be of equal magnitude. As prior mentioned, Jarrow and Rudd (1983), [49], used the said degree of freedom to specify equiprobable price moves, whereas Trigeorgis followed Cox et al. In 1993, Tian (1993), [84], in response to the rather (or so seeming at least) arbitrary nature of the enumerated conditions, proposed that the additional degree of freedom be used to equate the third moments of the continuous and discrete price processes - instead of only the first two as done by the other systems. This yielded a binomial price process that matches the mean, variance and skewness of the continuous time- and state process posited by Black and Scholes (1973), [9]. Tian’s claim of improved pricing accuracy (following from the altered process) was unfortunately shown to have been an empty one: Diener and Diener (2004) [31], and Leisen and Reimer (1995), [55], both showed order of convergence of only one on Tian’s tree method.

A number of parameterizations for binomial trees, other than those enumerated above are in existence, for example that proposed by Van den Berg (2000) [83], as well

as the transformed process trees of Camara and Chung (2003) [18]. Of particular interest are the trees proposed by Muzzioli and Toricelli (2004), [60], where states and probabilities are derived in the context of fuzzy set logic. However, this dissertation will focus on the principal tree parameterizations as proposed by Cox et al (1979), [23], Jarrow and Rudd (1983), [49], Trigeorgis (1991), [86] and Tian (1993), [84]. In Chapter Three each of the aforementioned is discussed and well as regards the assumptions made, the derivation of tree parameters as well as the eccentricities ensuing. Some comparisons are drawn in the chapter's conclusion.

Building on the work of the above mentioned authors, Boyle (1988), with his article "A Lattice Framework for Option Pricing with Two State Variables", [12], extended binomial trees by changing the two-outcome assumption for the value of the underlying to a three-outcome one; the resulting procedure was subsequently christened trinomial trees.² The drawback of trinomial procedures is that within a three-outcome price environment, the no-arbitrage replication argument, central to binomial tree methods, no longer holds true - because of the impossibility of replicating a derivative's value in a three state world with only recourse to the underlying- and riskless assets as replication tools. Thus followed the "Babylonian exile" of trinomial trees: in name they were tree-based methods, but only as such. Traditionally the justification for trinomial trees is strictly defined in terms of equivalence to the explicit finite difference methods. In Chapter Four it is shown that the trinomial and binomial trees are indeed incarnations of the same thing, which provides added justification for the former. (A remark concerning the title of the dissertation is in order: as already mentioned, the trinomial procedure is in fact exactly equivalent to the binomial; as such it was felt that a focus on binomial trees is justified.)

There are a diverse number of different derivatives valuation methodologies in existence today: the Black and Scholes closed-form solution for vanilla European options (along with a myriad other analytical formulas for non-vanilla pay-off styles), finite difference schemes for the solution of the Black and Scholes partial differential equation, trinomial tree procedures, numerical integration schemes (that draw more on the Martingale pricing approach) and yes, binomial tree methods. In the expository article of Cox et al (1979), [23], the issue of paradigm consistency was already considered: they showed both convergence to the Black and Scholes closed-form solution for vanilla European options as well as equivalence to the Black and Scholes partial differential equation. Since then, numerous other articles have appeared that seek

²Boyle might not have been the first to consider three-jump processes in the evaluation of derivative securities, for example Parkinson (1977), [65], and Brennan and Schwartz (1978), [14] all considered this eventuality. However, the context within which the said model was presented as well as the application certainly differed from that ultimately proposed by Boyle. For the foregoing reasons it would not be unwarranted to refer to Boyle as having established the trinomial procedure for derivatives valuation.

to establish consistency links. For example, Musiela and Rutkowski (1997), [59] and Van den Berg (2000), [83], both proved convergence to the Black and Scholes closed-form solution for vanilla European options, whilst Schwartz (1977), [75], showed equivalence to the explicit finite difference solution of the Black and Scholes partial differential equation. Numerical integration schemes within a derivatives pricing context was first explored by Parkinson (1977), [65]; recently there has been renewed interest in this technique, see for example the articles by Andricopolous et al (2003), [1] and Chung and Shackleton (2004) [20]. The “Babylonian exile” prior mentioned was effectively healed when Rubinstein published an article showing equivalence between trinomial and binomial trees in 2000, [73]. Chapter Four treats on equivalence between binomial procedures and those mentioned above.

For all its elegance, binomial procedures in their original guise (as any active derivatives practitioner will know) are somewhat removed from the “real life” financial markets. For instance, riskless returns realize as a term structure, and not as a single number as assumed in almost all binomial derivations, furthermore, riskless returns are anything but deterministic. In making allowance for this, Merton (1976), [56], proposed the application of stochastic interest rates, though not in a discrete time and state environment. Accommodating a term structure of interest rates within a tree framework, although not expressly done in the expository work of Cox et al (1979), [23], is rather straight forward. Following Merton’s cue, and generalizing riskless returns to stochastic variables (within a lattice paradigm), however, requires a multi-variable modeling approach such as that proposed by Boyle (1988), [12]. As regards the cost incurred by stock borrowing - and this also extends to derivatives on currencies - expanding upon the suggestion for inclusion of proportional dividends given by Cox et al, [23], provides an easy enough solution. Geske and Shastri (1985), [35], in their comparison of different American option pricing techniques show how one would include continuous dividend yields (or equivalently stock borrow costs) in a binomial lattice framework.

Applying the principle posited by Black (1976), [8] for the valuation of options on commodity forwards and futures, namely of assuming geometric Brownian motion for the forward price, to binomial trees results in an extension that allows for the pricing of derivatives on commodity futures and forwards.³ Of particular interest in the equity markets, is the question of discrete dividends paid. It has already been alluded to that Cox et al (1979), [23] allowed for the presence of dividends, the magnitude of which depends on the level of the underlying’s value. It is well known that dividends tend to be independent of a share price’s value. Incorporating discrete dollar dividends, the size of which is independent of the share price’s value into the

³In this instance, instead of building a tree that mimics the spot price of the underlying, a forward (futures) price tree is rather built.

Black and Scholes (1979), [9] stochastic framework is an issue not yet fully decided. Chapter Five uses a basic framework proposed by Frishling (2002), [34] to explain the various approaches in existence. In terms of translating these into adjustments to the binomial procedure, once the principle has been established, is rather to the point.

The 1987 equity market correction (Black Monday, 19 October 1987), left more than just increased nervousness about the markets in October as its legacy. It is also generally regarded as the time at which volatility skews first became a feature of derivatives markets. The literature on volatility skews is quite wide - and is growing, however the authors normally coming to mind are Derman and Kani, Dupire and of course Rubinstein. (Of course the debate regarding the meaning and application of volatility skews is very much still going on.) It is interesting to note that the first (and still definitive) attempts towards explaining and accommodating the skew in derivatives valuation were based on tree methods: Derman and Kani's (1994), [29] (and Derman et al (1996), [27]) local volatility model, for all the criticism it attracts, is an elegant way of presenting skews, whilst the Rubinstein (1994) implied binomial tree, [73], certainly represent a relatively robust implied volatility procedures at this point in time. In the fifth chapter, the problem of volatility skew and its implications for binomial tree methods are treated on.

The utility of binomial procedures in the pricing of American options is quite evident from the fact that most current trading and risk management systems (Imagine, Front Arena, Riskwatch) allow for the pricing of the said by way of binomial lattices. The ease of pricing these derivatives within a binomial tree framework was already presented by Cox et al (1979), [23]; Schwartz (1977), [75], as well as Geske and Shastri (1985), [35] raised the option of pricing American options within a finite difference scheme to definite cognizance. Since then, most papers on binomial procedures, and on American option valuation takes into view either the one or the other. On the matter of convergence of American options as valued by trees, to the "actual" value, Leisen (1998), [53] provided some mathematical justification for the "rule of thumb", no-arbitrage principle first suggested by Cox et al (1979), [23], and well as relating to adjusting an option's value, within a tree for the optimality of early exercise. On the pricing of barrier options within a lattice approach: there has been a good number of publications on this, including Margrabe (1989), [54], (who investigated the issue of barrier option value convergence), Boyle and S.H. Lau (1994), [14] and of course Derman and Kani (1993), [26], and Derman et al (1995), [28]. The penultimate chapter focuses on the pricing of exotic options using the binomial procedure.

Convergence patterns, stability and order of convergence are issues of very practical concern. As might be expected, the enumerated were not addressed at length by

Cox et al in their expository work. For example, Geske and Shastri (1985), [35], were the first to point out that for certain parameter choices the Cox, Ross and Rubinstein tree resulted in unstable price approximations, although the reason for this was not alluded to. Boyle (1988), [12], and Hilliard and Schwartz (1996), [45], must have been some of the earliest authors who considered the recognition of the characteristic patterns associated with binomial procedures. The said authors, it might be added, also attempted to manipulate order of convergence by recognizing the very specific patterns and extrapolating on them; Leisen (1998) [53] acknowledged the improvement in convergence using the said method. The underlying drivers for the peculiar binomial convergence patterns, were however only seriously investigated in the late nineties. Beyond the reasons for oscillatory behavior, order of convergence and more specifically, a proper definition of the said and furthermore a systematic comparison of the convergence order of different tree construction methods was of prime concern to Diener and Diener (2003), [49] as well as Leisen and Reimer (1995), [55]. The final chapter considers these issues.

To conclude: the aim of this dissertation is to provide a general background to tree pricing methods as applied to the pricing of derivatives securities, and well by investigating the assumptions underlying, the various incarnations found as well as an investigation into some of the peculiarities exhibited by these algorithms. Very important: extensions to the paradigm - in the form of relaxing some of the underlying assumptions - are touched upon, furthermore, the relation of tree- to other pricing methods are also considered.

Chapter 2

Random Walk Models and Derivatives Pricing

Binomial tree procedures are spanned by two important concepts, namely that of a two-outcome random experiment, and a no-arbitrage argument that links the former to a derivative security's fair price. The implications of these simplistic-seeming assumptions are quite far-reaching, being amongst other things the applicability of random walk models to the underlying asset price, as well as a discrete equivalent to the Feynman-Kac result (see Chapter Four, Section Four and Goodman et al, (2004), [37]).

2.1 Background

This chapter aims towards the exposition of some of the fundamental ideas underlying binomial tree methods in derivatives pricing. More specifically, the concepts of Bernoulli trials, random walks and their properties as well as no-arbitrage reasonings and their use in linking Bernoulli variables to derivatives pricing is to be investigated. It commences with a discussion of two-outcome, or Bernoulli experiments. A generalization of these into random walks follows, whereafter some of the important and useful results pertaining are presented. The link between Bernoulli outcomes for the price of an asset underlying a derivative, and its arbitrage-free value is presented in the penultimate section by way of intuitive reasoning as well as algebraically. The chapter is concluded with a generalization of the latter into a relationship between random walk models for the asset price and the arbitrage-free value of derivatives securities.

2.2 Bernoulli Experiments and Random Walks

It could be argued that the simplest of all stochastic processes must be the Bernoulli experiment, i.e., a two-outcome or two-state random variable. In this category one finds the simple coin-toss game, where the outcome of each toss could be heads or tails exclusively, with the probability of either being constant. Another example of a Bernoulli experiment would be where the price of a financial security is categorized as either being high or low, based upon some pre-determined threshold. In the last mentioned example, although the security's price may assume any value on the positive Real line, i.e., it might be spatially continuous, a Bernoulli variable may be defined, driven by this continuous process, but which gives (from a complexity perspective) a much reduced view. The point being driven at is as follows: Bernoulli random processes can be used to simplify more complex settings.

From the preceding it should be clear that Bernoulli experiments driven by a continuous underlying process can be thought of as a two-dimensional partitioning of the primary process' cumulative density function. Should a more general k -state process be assumed, i.e., a stochastic model that allows for k possible states of the world, intuitively it should follow that as k becomes larger, i.e., as the partitioning becomes finer, the discrete view becomes progressively more similar to the underlying process. Mathematically the last mentioned would be equivalent to stating that the discrete distribution converges towards the underlying continuous process in probability. This issue is very important, for example, in testing consistency between

binomial tree based pricing methodologies and Black and Scholes pricing (this is to be discussed more in detail in Chapter Three).

Associated with Bernoulli experiments is the concept of a random walk. Consider a random variable that sums the outcome of each of n successive, identical but independent Bernoulli trials; this variable, which effectively keeps track of preceding realizations is what is referred to as a random walk. An obvious example of a random walk would be the cumulative gain or loss from a succession of coin-toss gambles. The concept of a random walk is formally introduced in the following definition.

Definition 1 (The Generalized Random Walk)

Let ϵ_i , with $i \in 0, 1, 2, \dots$ be Bernoulli random variables, defined as follows:

$$\epsilon_i = \begin{cases} \alpha, & \text{with probability } p \\ \beta, & \text{with probability } (1 - p) \end{cases}$$

The process $\{X_n = \sum_{i=0}^n \epsilon_i, n = 0, 1, 2, \dots\}$ is said to be a random walk.

Figure 2.1 presents the above in a more intuitive way; the stem-plot representing the outcomes of each of the successive two-outcome experiments, whilst the drawn line shows the random walk realization.

Random walks are well studied stochastic processes, with many interesting and useful results pertaining. Some of these are now to be introduced, the first of which relates to the probability distribution of random walk models.

Theorem 1 (Random Walk Sample Paths)

The random walk $\{X_n, n \in N\}$ has 2^n possible sample paths or realizations; the probability distribution of the process at a fixed but arbitrary time point, n can be written as follows:

$$\text{Pr.}\{X_n = (n - i)\alpha + (i)\beta\} = \binom{n}{i} p^{(n-i)} (1 - p)^i, \quad (2.1)$$

where $i = 0, 1, 2, \dots, n$.

Proof: The proof to the above follows directly from basic combinatorial concepts. Refer to Tucker (1995), [82] for an introduction to combinatorics.

Q.E.D.

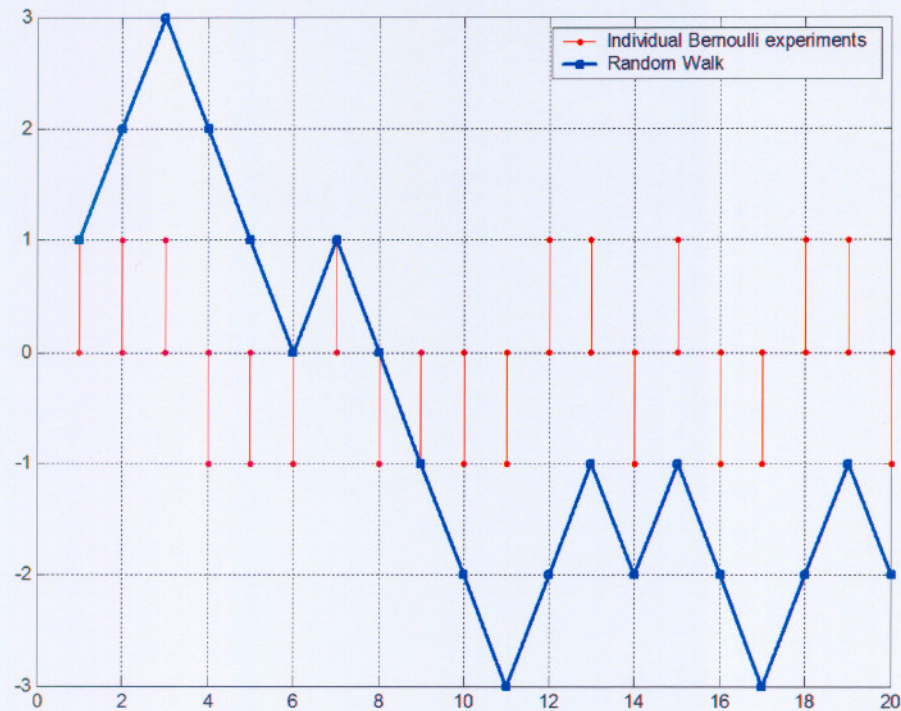


Figure 2.1: Illustration of Bernoulli experiments and random walks

From the above it should intuitively follow that many of the 2^n sample paths recombine at different nodes to imply only $(n + 1)$ terminal nodes. The implications of the recombination concept would be difficult to exaggerate - amongst other things it allows for the realistic consideration of all possible sample paths in pricing a particular derivatives contract - which is required for the sensible valuation of, for example, path-dependent options. (Providing a way to systematically discretize the entire time- and state domain of the underlying asset's process is, arguably, the principal advantage of binomial procedures.)

An important implication of the above is that the value of a random walk, when defined only in terms of the number of “up” movements, could be regarded as a binomial random variable. This is a powerful result which is used with great utility in derivatives pricing and well in the sphere of data reduction, or rather in “jumping” between points of discontinuity in the underlying asset price and derivatives value. (The last stated concept being equivalent to that used in a multi-step Monte Carlo simulation, but more on that later.)

Random walks can be visualized as sets of vertices (representing state and time realizations) connected by lines (lines representing a transition with non-zero probability). Otherwise said: one could represent the entire set of possible sample paths relating to a random walk process, by points (vertices), denoting a particular asset price realization, and connecting lines between points, denoting non-zero transition probability. Such a lattice construction, as shown in Figure 2.2 is very characteristic of binomial methods. In fact, it strongly relates to a branch of mathematics called “Graph Theory”, dealing with points and lines of connectivity. (Das (2002) [24], gives an excellent graph theoretical definition of binomial trees.)

Now, in graph theory, trees represent a class of graphs where each terminal point has only one sample path leading from the root to it. (This characteristic is also referred to as a tree having in-degree of only one.)

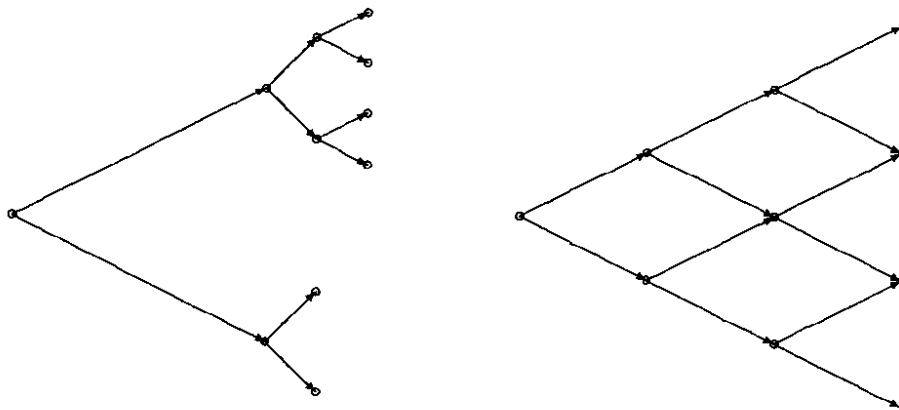


Figure 2.2: “Proper” and “Binomial” trees

As such, binomial trees as shown on the left of Figure 2.2 would be a member of the class of trees, whilst that on the right would not be: their recombining nature precluding them from the said classification. However, the name has been decided on, and as such applies. Binomial trees, it is seen, derive their name from the two-outcome nature of successive price movements, (and the consequent relevance of the binomial probability distribution), as well as the representation in terms of vertices and lines, harking back to graphs of a tree-like nature in graph theory. Based on the preceding the following could be said:

Theorem 2 *The sample paths of a random walk $\{X_n, n \in N\}$, (with no absorbing or reflecting barriers) can always be presented as a unique binomial tree with $(n+2)(n+1)\frac{1}{2}$ nodes and $(n)(n+2)$ lines.*

Proof: *The total number of nodes is simply the sum of the progression $1, 2, 3, \dots, n$, whilst the number of lines is equal to the sum of the arithmetic progression $2, 4, 6, \dots, 2n$.*
Q.E.D.

The above implies that the graph representation allows one to view an exponentially increasing number of sample paths in an efficient way such that the number of additional nodes considered for an unit-increase in the number of time steps, increases linearly only. The utility of this is to be further illustrated by way of example:

Example:

Assume an ordinary share's price in a time interval covering one hour behaves according to a Bernoulli experiment - it can either increase by one cent or decrease by one cent; the probability of an increase remaining constant through time and space. It should follow that for an observation period of one trading day, there are $2^8 = 256$ possible trajectories for the share price. Using Theorem 2 the said sample paths can be condensed in representation to 40 space-time combinations, and 80 transitions with non-zero probability. Otherwise said: all possible sample paths can be represented in a data structure consisting of 40 elements.

The principal implication of the above being that path dependent observation is feasible given the data reduction method presented. As shall be shown in Chapter Five, when considering the evaluation of path-dependent options, a particularly powerful characteristic of binomial lattice procedures is the facility of considering conditional pricing. Now, when adding a conditioning feature to pricing, but where there is no recourse to the type of data reduction above presented, the amount of computation effort expands beyond computing capacity and quite also beyond imagination as well. It could therefore be said with a good measure of justification, that the data reduction capacity of recombining binomial trees is a powerful and usefull feature. With this, the preliminary section on random walk models and their properties is concluded.

2.3 Risk Neutrality and Arbitrage-Free Prices

The saying "there is no such thing as a free lunch" must be one of the most fundamental of principles applying to derivatives pricing. In essence it states that prices

exist for all securities such that riskless profit taking is impossible. It also equates to non-exploitation, in other words, prices are set such that parties with similar information about the market are all equal in terms of the ability to calculate fair security values. This principle of no-free-lunch is generally referred to as the no-arbitrage principle; it is formally stated as:

Definition 2 (*The No-Arbitrage Principle*) *Arbitrage is the process by which profits can be made with probability one, but with no risk taken; otherwise said: arbitrage opportunities arise when trading strategies exists such that costs incurred in their execution (the cost of funding) is less than the pay-off they provide, and with complete certainty (probability one). The no-arbitrage condition in derivatives pricing therefore relates to the setting of the fair value on different securities, such that arbitrage is precluded.*

From the preceding it should follow that quoting on derivatives using prices that do not comply with the no-arbitrage principle could be a dangerous practice. A price in excess of the no-arbitrage one will leave the buyer open to exploitation, whereas a price below that of the no-arbitrage one might lead to exploitation by the market of the seller. The no-arbitrage price therefore provides an important indication of value levels for all securities.

Now, closely connected with the no-arbitrage price of a security is that of “risk neutrality”. Black and Scholes (1973): “The Pricing of Options and Corporate Liabilities”, [9], introduced the idea of “risk neutral” prices; in the logic used to construct their partial differential equation, this concept played a pivotal role. Essentially it involves the idea that a portfolio can be “immunized” against movements in a market variable over a given time interval, i.e., the “profit and loss” effect of an instantaneous change in a market variable can be neutralized. Furthermore, a “risk neutral”, or “riskless” portfolio should always provide a return equal to that of a risk less asset, such as a money market deposit (that carries no credit risk). A return in excess or shy of the last mentioned, leaves room for arbitrage - long the asset that provides the higher return and self-fund by shorting the other, thereby locking in a net return greater than zero, with probability one.

An important consequence of risk neutrality is that investor preferences and risk appetite can be ignored in calculating the arbitrage-free value of derivatives. This is quite surprising, especially in the light of early attempts to price these instruments: most models preceding that of Black and Scholes took into account variables measuring the risk aversion or preference of investors (consider for example the model postulated by Samuelson, in the (1965) paper “Rational Theory of Warrant Pricing”,

or Sprenkle (1964), "Warrant Prices as Indicators of Expectations and Preferences".) The Samuelson model depended on two variables the value of which was difficult to determine. The first being the expected rate of return for the underlying asset price (which is obviously a function of risk preference and aversion); the model also required a variable measuring the rate at which accommodation could be made for a deferred pay-out, i.e., a discount rate had to be supplied. As the measurement of perception and risk aversion and appetite is difficult if not impossible, not to mention leaving much room for subjectivity, inclusion of the said makes for unwieldy pricing. Simply said: in a framework where risk preference is taken into account, agreement on derivative prices would be hard to reach. Risk neutrality, in contrast, provides a logical framework for derivatives pricing where investor views and preferences plays no part, resulting in prices most investors can be in agreement of.

As a final remark on risk neutrality, it is worthwhile to consider that risk is represented by stochastic elements in the potential return of an investment strategy. The logic underlying risk neutral pricing, is that of removing the stochastic element in a portfolio's return completely, i.e., such a portfolio, funded by a risk less asset, should result in a completely flat profit and loss profile, implying that the arbitrage-free price of a derivative is the up-front calculation of the cost of hedging or immunizing a portfolio containing the said derivative: arbitrage-free derivative valuation models are only advanced "profit and loss" calculators.

2.4 Bernoulli Experiments and Arbitrage-free Derivatives

In their seminal paper "Option Pricing: A Simplified Approach", [23], Cox et al (1979), showed that a very close relationship exists between arbitrage-free derivatives valuations and the assumption of a two-outcome stochastic model for movements in the price of the underlying asset. In short, their argument is as follows: assume a Bernoulli model for the behavior of a derivative's underlying asset's price; further assuming arbitrage free prices for derivatives securities, and an immunization argument following from that, a two-equation, two-variable system linking the arbitrage-free price of a derivative and the two outcomes on the underlying's price results. The system solution provides a unique arbitrage free value for the derivative, unrelated to the empirical or assumed probability distribution on the underlying's value. This argument, to be referred to hereafter as the "Fundamental Result of Tree Pricing" is now presented more rigorously.

Theorem 3 (The Fundamental Result of Tree Pricing)

Assume a financial asset's price movement over a fixed but arbitrary time interval Δt modeled as a Bernoulli variable, with outcomes S_u and S_d . Furthermore, let $S, f(t)$ denote the price of the asset and a derivative thereon (respectively) at the start of the interval, whilst $f(t + \Delta t|S_u), f(t + \Delta t|S_d)$ in their turn denote the no-arbitrage value of the derivative conditional on an up and down movement in the underlying. Then, it can be stated that:

$$f(t|S) = f = \frac{(pf(t + \Delta t|S_u) + (1 - p)f(t + \Delta t|S_d))}{(1 + r\Delta t)}, \quad (2.2)$$

where r denotes the return of the riskless asset (expressed as a simple rate), and where $p = \frac{(1 + r\Delta t)S - S_d}{S_u - S_d}$. The no-arbitrage condition implies that: $S_d \leq (1 + r\Delta t)S \leq S_u$.

Proof:

Consider a portfolio consisting of a derivative, the value of which is denoted by f , furthermore, the portfolio also contains a position of D of the underlying asset, and well such that the portfolio is completely immunized against movements in the asset's price¹, i.e., the portfolio should show no change in value under either of the two outcomes assumed possible for the underlying asset's price movement. At the start and end of the observation period, the value of the portfolio can be written as: $(f + DS)$, $(f(t + \Delta t|S_d) + DS_d)$ and $(f(t + \Delta t|S_u) + DS_u)$ respectively. From the immunization assumption and last stated, the following system results:

$$(f + DS)(1 + r\Delta t) - (f(t + \Delta t|S_d) + DS_d) = 0, \quad (2.3)$$

$$(f + DS)(1 + r\Delta t) - (f(t + \Delta t|S_u) + DS_u) = 0, \quad (2.4)$$

which has the unique solution:

$$D = \frac{f(t + \Delta t|S_u) - f(t + \Delta t|S_d)}{S_u - S_d}, \quad (2.5)$$

and

$$f = \frac{(pf(t + \Delta t|S_u) + (1 - p)f(t + \Delta t|S_d))}{(1 + r\Delta t)}, \quad (2.6)$$

where $p = \frac{(1 + r\Delta t)S - d}{u - d} = \frac{R - d}{u - d}$, where $R = (1 + r\Delta t)$.

Q.E.D.

The above result is rather fascinating, and well for reasons including:

¹Refer to Leland and Rubinstein (1995), [53], for a more detailed explanation on this concept.

- In the solution, the empirical or parametric probability distribution assumed on the underlying's price movements, is absent and actually quite irrelevant.
- Due to $Sd \leq (1 + r\Delta t)S \leq Su$, the value p has the property $0 \leq p \leq 1$. Thus, it conforms to the Kolmogorov probability postulates, and one could interpret it as a probability, moreover, the derivative's value at the time period start, is interpretable as a discounted expected future value.
- The probability p , is referred to as the risk neutral probability measure. (Amongst other things it ensures that the discounted price process for the underlying is a Martingale.)

The "Fundamental Result" presented, allows one to establish a link between a Bernoulli outcome assumption on the underlying asset price movements and the no-arbitrage price for a derivative. Modeling asset price movements over a small time interval by way of Bernoulli random variables is quite acceptable, however, as the interval length increases, such a model, it might be argued would be too simplistic to realistically capture movements more likely in an actual financial market setting. However, successive changes over small time intervals could be acceptably modeled by Bernoulli variables. Furthermore, it has been seen in Section Two that such a stochastic dynamic is captured by the random walk model. Thus, assuming a random walk model for the asset price, and using the "Fundamental Result" would enable one to price a derivative in a arbitrage-free way without placing stultifying limits upon the realism of the model posited for underlying asset price dynamics.

2.5 Summary

In this chapter the concept of Bernoulli trials (and their relation to random walk models) was introduced. It was also seen that Cox et al (1979), [23], presented a link between two-outcome stochastic behavior for an asset price and a no-arbitrage price for a derivative based thereon. The latter, the "Fundamental Result of Binomial Pricing", and the assumption of random walk dynamics for the value of the underlying, yields the binomial tree method for pricing derivative securities.

Chapter 3

Tree Construction Methods

In recognition of binomial procedures' simplicity and power, quite a number of extensions and adjustments to the original parameterization posited by Cox et al (1979), [23], have been proposed. Different parameter derivations is a natural consequence of the structure of binomial lattices: the basic tenets of this pricing approach are very loosely defined; a two-outcome price process, for example, is the only requirement in so far as the discrete process definition is concerned. Ideally one would want the discrete process to inherit at least some characteristics of the continuous one. However, there are a whole number of ways in which this could be brought about. Moment matching is the preferred choice of most parameterizations, whilst some like Leisen and Reimer (1995), [55], has opted for a partial cumulative probability matching. Applying different constraints upon the discrete process, there are round five different (and considered important) tree building techniques, namely that of CRR, Jarrow and Rudd, Trigeorgis, Tian and Leisen and Reimer.

3.1 Background

In the previous chapter the concept of a random walk was introduced. Applying random walks to describe the price movement for asset prices, facilitates the calculation of arbitrage free prices. Now, given an empirical distribution for the price of an asset, how is one to ensure that the discrete model assumed within the binomial framework is consistent with it? In other words, given that an asset's price behaves according to a geometric Brownian motion, how is one to select the random walk's parameters such that convergence to the empirical distribution is guaranteed as the time interval size is reduced?

This issue can be presented in the following way as well: the "Fundamental Result of Binomial Tree Pricing", provides a way in which to derive arbitrage-free prices for derivatives given the assumption of a random walk model for a description of the underlying security's price movement, however, there are literally an uncountable number of ways in which to select the model parameters - each selection of which will yield a different arbitrage-free price for the derivative.

This chapter deals with the way in which parameter selection is to be done. The concept of calibration is introduced first; two techniques through which calibration takes place are discussed. Since the publication of Cox et al (1979), [23], - suggesting a particular parameter choice, numerous other parameter selection results have appeared. Section Three presents some of the more well-known of these parameterizations, e.g. the original one proposed by Cox, Ross and Rubinstein, the parameterization posited by Jarrow and Rudd (1983), [49], the one by Trigeorgis (1991) [86], Tian (1993), [84], and that suggested by Leisen and Reimer (1995), [86]. Section Four concerns itself with two different representations of geometric Brownian motion and the implication on tree calibration, whereas the penultimate section treats (albeit very briefly) with different ways in which binomial procedures can be implemented in a programming language.

3.2 Tree Calibration

Calibration is defined by the Oxford dictionary as "1. mark (a gauge or instrument) with a standard scale of readings. 2. compare the readings of (an instrument) with those of a standard. 3. adjust (experimental results) to take external factors into account or to allow comparison with other data."

The calibration procedure in the context used here refers to transferring characteristics from an empirical stochastic process to the process used as an approximation, and well done so as to ensure convergence. Otherwise said: calibration of one process to another (standard one) will attempt to have the approximation mimic the empirical in certain aspects. Calibration of stochastic processes generally proceeds in one of two ways. Firstly it can be done by way of moment matching, and secondly by using a partial CDF matching technique. Each of these is to be briefly investigated.

3.2.1 Moment Matching

Probability distributions are characterized by their moments, i.e., two random variables with equivalent moments are of necessity identical in distribution.¹ Furthermore, a distribution's moments give clues to certain characteristics, e.g., the first divulges information about the central tendency, the second about dispersion, the third about skewness, the fourth about the propensity for outlying realizations etc. (See for example Bain and Engelhardt (1992), [3], and Feller (1957), [33] for discussions on the moments of probability distributions.) Thus, having random variables with identical first central moments will facilitate easier comparison as regards central tendency etc. The calibration technique of "moment matching" relies upon equating moments to ensure identical behavior of two different process as regards certain measures. The idea is to be further explained by way of example:

Example:

Consider two random variables, the first of which is distributed lognormally with location and scale parameters $\mu = 2.38$ and $\sigma = 0.2$ - let it be denoted by $X(t)$. The mean, variance and skewness of this variate can be seen in the first column of Table 3.1.

Let the other variable, denoted by $Y(t)$, be a discrete random variable such that:

$$Pr.\{Y(t) = Y(0)e^{\theta(20-2i)}\} = \binom{n}{i} p^i (1-p)^{n-i},$$

where $i = 0, 1, 2, \dots, 20$. θ and p are two distributional parameters for Y , whilst the process is assumed to have starting value $Y(0)$. For a rather arbitrary parameter choice, the mean, variance and skewness of the discrete (approximation) process is shown in Column Two of Table 3.1.

¹The proper condition being: equivalence in the set of all moments; this set being countable-infinite.

<i>Moments</i>	Empirical Process	Discrete Process	Calibr. Process
Mean	11.58	10.21	11.54
Variance	5.56	4.50	5.57
Skewness	0.61	0.60	0.52

Table 3.1: Results of the moment matching procedure

Assume that the variable X represents the true or empirical distribution, with X an approximation; it would make sense to calibrate the second process so as to mimic the first in some of its characteristics. Given that the moments presented are of great importance, and it be desirable that these should match as closely as possible, then the parameters θ and p could be set such that the following system is satisfied:

$$\begin{aligned}
E\{X(t)\} &= E\{Y(t)\} \\
E\{(X(t) - E\{X(t)\})^2\} &= E\{(Y(t) - E\{Y(t)\})^2\} \\
E\{(X(t) - E\{X(t)\})^3\} &= E\{(Y(t) - E\{Y(t)\})^3\}
\end{aligned}$$

The least squares estimate of the said parameters is $\theta = 0.046$ and $p = 0.566$; the moments of the discrete distribution for a parameter choice as that just presented can be seen in the third column of Table 3.1. Consistent with what has been observed on the characterization of distributions by their moments, the calibrated discrete distribution appears to track the empirical one quite closely.

To summarize: moment matching is a technique that allows one to transfer specific properties of one probability distribution to another by the tweaking of distributional parameters. Of necessity it requires knowledge of the empirical moments, as well as a functionally tractable form for the moments of the process calibrated.

3.2.2 Partial CDF Matching

Besides their moments, probability distributions are also characterized by their Cumulative Probability Density functions, i.e., the CDF of a distribution provides a way in which distributions can be uniquely identified. Furthermore, the probability of certain events are easily specified in terms of combinations of their cumulative density. Probability properties are transferable from one random variable to another via equating the respective CDFs at pre-determined points; this is illustrated by way of an example.

Example: Building on the previous example, given that the probability estimates for the events $\{X(t) \leq 6.06\}$, $\{X(t) \leq 10\}$ and $\{X(t) \leq 16.48\}$ are of vital concern;

instead of calibrating the discrete process to match the empirical process in moments, it might be more relevant to set the parameters such that the cumulative probability functions intersect at the stated points. In Figure 3.1, the CDF resulting from the calibration process is contrasted with the original variates'. It is quite remarkable how ensuring intersection of the CDFs at set points, results in a seeming overall equivalence.

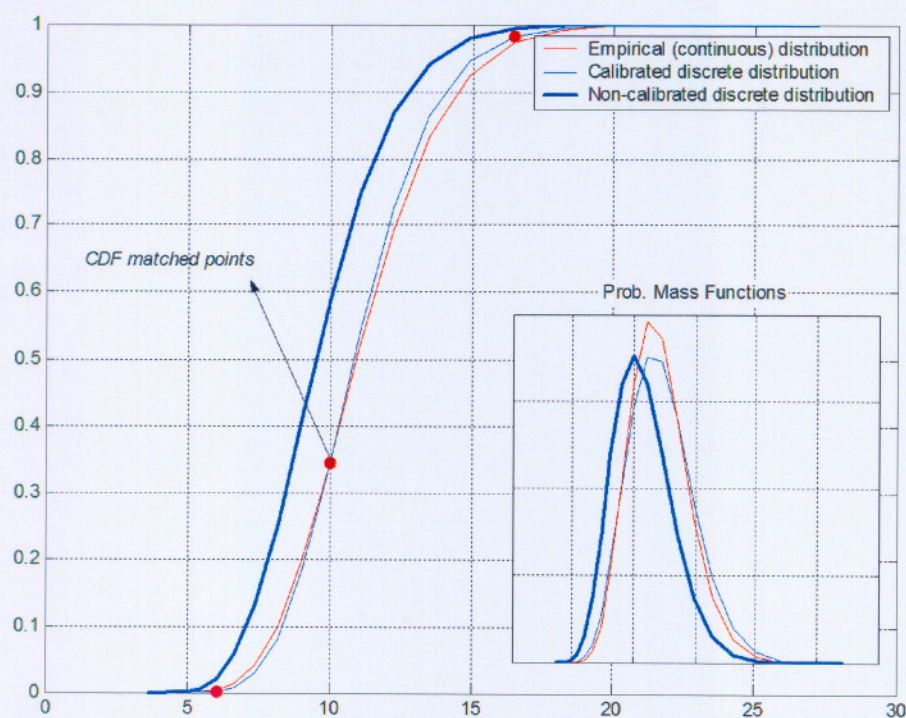


Figure 3.1: Graphical view of the CDF matching process and results

Examples of CDF matching include the many techniques that exist to approximate cumulative binomial probability mass using the Normal CDF (the rationale for this being the difficulty of calculating combinations of high order). Amongst the aforesaid would be the procedures suggested by Camp (1951), [19], Paulson (1942), [62], and Peizer and Pratt (1968), [66]. Within the realm of binomial tree construction, CDF matching is limited to the parameterizations suggested by Leisen and Reimer (1995), [55].

3.3 Well-known Tree Parameterizations

This section is to be devoted to an investigation of the most important tree parameterizations, namely the two proposed by Cox et al, the Jarrow and Rudd tree, Trigeorgis', the one posited by Tian, and finally, the Leisen and Reimer tree.

3.3.1 Cox, Ross and Rubinstein "Diffusion" Parameterization

In general "quants" parlance, binomial trees usually refer to the CRR tree parameterization, proposed in their 1979 article, [23]. Its derivation is intuitive, whilst the parameter choices are quite easy to remember. The discussion on this tree method will involve firstly the assumptions underlying it, the associated parameter derivation, and concluded with some remarks on its eccentricities.

The assumptions underlying the choice of parameters for this parameterization can be summarized as follows:

Assumptions:

- *CRR1* The underlying asset price process (the continuous time and state one) is chosen to be geometric Brownian motion, i.e., a generalized Brownian motion is assumed for the log-return process; the assumption is the same as that proposed by Black and Scholes (1973), [9]. The consequence of this assumption is that returns are assumed to be centred around a specific drift value - seen to be the riskless return, with perfect symmetry and only a limited tendency for extreme behaviour. The distribution associated with continuous return is the Gaussian. (See Neftci (1996), [59], for an introduction to geometric Brownian motion, stochastic processes and their applications in financial mathematics in general.)
- *CRR2* The discrete process is calibrated so as to have first and second order moments that approaches that of the continuous process in the limit, i.e., as the time intervals considered increase in number the first two central moments of the discrete process are to converge to those of the continuous process. This is ensured via a second order moment matching.
- *CRR3* The second moment of the discrete process is assumed to be equal to its variance.

- *CRR4* “Up” and “down” movements in the asset price are reciprocal; they are assumed to result in relative price changes of equal magnitude, but opposite direction, i.e., an “up” followed by a “down” would yield the initial price.

Parameter Derivation:

The derivation of the random walk model parameters ² proceed in the following manner:

Denote by $S(t)$ the underlying asset’s price at time t , and by q the probability of an “up” move in the price; let u represent the up-factor, and d the down-factor. Assumption CRR1 could then be expressed in terms of the following stochastic differential equation as:

$$d \ln S(t) = v dt + \sigma dB_t, \quad (3.1)$$

with $v = (r - \frac{\sigma^2}{2})$ and dB_t a Wiener process; r , following the reasoning applied in constructing the Black and Scholes partial differential equation, being equal to the continuously compounded riskless return.

Now, the first two moments of the above process are $(r - \frac{\sigma^2}{2})t$ and $\sigma^2 t$ respectively. Equating the last mentioned and the moments of the discrete process, yields:

$$q \ln \frac{S(t)u}{S(t)} + (1 - q) \ln \frac{S(t)d}{S(t)} \rightarrow vt, \quad (3.2)$$

and

$$q(\ln \frac{S(t)u}{S(t)})^2 + (1 - q)(\ln \frac{S(t)d}{S(t)})^2 \rightarrow \sigma^2 \Delta t + v^2 \Delta t^2. \quad (3.3)$$

However, Assumption CRR3 allows that the above be written as:

$$q(\ln \frac{S(t)u}{S(t)})^2 + (1 - q)(\ln \frac{S(t)d}{S(t)})^2 \rightarrow \sigma^2 \Delta t. \quad (3.4)$$

The above along with $u = \frac{1}{d}$ (Assumption CRR4) results in:

$$q = \frac{v \Delta t + \ln u}{2 \ln u}. \quad (3.5)$$

Substitution of (3.5) into (3.2) leads to:

$$(\ln u)^2 = \sigma^2 \Delta t, \quad (3.6)$$

²All the parameter derivatons in this chapter have been done from first principles; in most cases the reference articles supplied only limited algebra background for the parameter choices given. This dissertation attempts to provide full details on the algebra that underlies the various choices.

further yielding:

$$u = e^{\sigma\sqrt{\Delta t}}. \quad (3.7)$$

Finally, substitution of (3.7) into (3.5) yields:

$$q = \frac{v\Delta t}{2\sqrt{\sigma^2\Delta t}} + \frac{1}{2}. \quad (3.8)$$

Comments:

As can be seen from the solution as set out in Equations 3.7 and 3.8, the assumption given by Equation 3.4 is equivalent to ignoring the term $(v\Delta t)^2$ in equation (3.3). For small Δt , (i.e., for large n), the term $(v\Delta t)^2$ is negligible, in fact, for most choices of t, δ, σ the said term is insignificant for all values of n . However, some choices for the said parameter values causes it to become pronounced; the discretization error increasing exponentially and yielding unstable results. The penultimate chapter treats on this.

3.3.2 Cox, Ross and Rubinstein “Jump” Parameterization

The second parameterization presented by Cox et al (1979) [23], assumes that the underlying asset’s price behaves according to a jump process. Jumps are generally differentiated from “normal” movements (price moves driven by a diffusion process) in that the size of increments are constant and independent of the length of the time interval considered, whilst the probability of jumps occurring is directly proportional to the latter.

In the derivatives pricing literature jump processes are usually used as an attempt to explain the existence of the volatility skew. Perhaps the most well-known of applications of jump processes in derivatives pricing is that of Merton (1976), [56]; in the said article he derived a pricing function for vanilla European options where the underlying price process exhibits jump discontinuities. In the realm of binomial pricing, the work by Rachev and Ruschendorff (1991), [64], is insightful, as the conditions necessary for convergence to jump processes are provided. Neftci (1996), [59], provides an intuitive introduction to jump processes and their differentiation from normal ones.

Assumptions:

- *J1* The continuous underlying asset price process is assumed to be a Poisson; the discrete process exhibiting a non-stochastic trend, with jumps superimposed on this.
- *J2* The discrete and continuous processes are assumed to have identical first moments; this is ensured via a moment matching.

The discrete process following from the above assumptions can be presented mathematically as follows:

$$S(t + \Delta t) = \begin{cases} S(t)u, & \text{with probability } \lambda\Delta t \\ S(t)e^{\zeta\Delta t}, & \text{with probability } 1 - \lambda\Delta t \end{cases}$$

Parameter Derivation:

As the arrival intensity and jump size (denoted by λ and u) are known beforehand, there is only one degree of freedom left, namely in the choice of ζ . The discrete price-process discounted by the risk free rate of interest should be a Martingale, i.e., without a drift term, in order for the derivatives prices to be arbitrage free. This is the only condition applied and results in:

$$u\lambda\Delta t + e^{\zeta\Delta t}(1 - \lambda\Delta t) = R, \quad (3.9)$$

writing the above in terms of ζ yields:

$$\zeta\Delta t = \ln(R - u\lambda\Delta t) - \ln(1 - \lambda\Delta t). \quad (3.10)$$

Comments:

As prior mentioned, Merton (1996), [58], provided a formula for the risk neutral value of a vanilla European option, where a jump process is assumed for the underlying asset price dynamics. As the asset price dynamics of the discrete process converges in distribution towards the assumption made by Merton, it should follow that the price for a vanilla European option calculated using the tree parameters just derived should conform to the Merton-model based one. Cox et al (1979), [23], shows the latter to hold true. (Refer also to Hanson and Westman, [38] for an alternative approach to specifying jump-diffusion processes.)

3.3.3 Trigeorgis' Parameterization

Trigeorgis (1991), [81], proposed a parameterization similar to that of Cox et al (1979), [23], excepting the assumptions relating to the second order discrete process moment

matching the variance. As a result the Trigeorgis tree parameters differ slightly from the CRR, and as is shown in a later chapter, the instability exhibited by the CRR tree for certain parameter choices is not observed.

Assumptions:

- *TRG1* The underlying asset price process (the continuous time and -state one) is given by the geometric Brownian motion.
- *TRG2* The discrete and continuous processes are assumed to have identical first and second order moments, not only in the limit, but for any length of time interval considered.

Parameter Derivation:

The derivation of the random walk model parameters, follows amongst similar lines to the Cox et al tree, and is presented below.

Firstly consider the algebraic expression of the enumerated assumptions:

$$\ln u = -\ln d, \quad (3.11)$$

$$q \ln \frac{S(t)u}{S(t)} + (1-q) \ln \frac{S(t)d}{S(t)} = vt, \quad (3.12)$$

and

$$q \left(\ln \frac{S(t)u}{S(t)} \right)^2 + (1-q) \left(\ln \frac{S(t)d}{S(t)} \right)^2 = \sigma^2 \Delta t + v^2 \Delta t^2. \quad (3.13)$$

Substitution of (3.11) into (3.12) and (3.13) yields respectively :

$$q \ln u - (1-q) \ln u = v \Delta t, \quad (3.14)$$

and

$$q (\ln u)^2 - (1-q) (\ln u)^2 = (v \Delta t)^2 + \sigma^2 \Delta t. \quad (3.15)$$

Making q the subject of equation (3.14) yields:

$$q = \frac{v \Delta t + \ln u}{2 \ln u}. \quad (3.16)$$

Substution of (3.16) into (3.13) and some simplification results in:

$$\frac{1}{2} (v \Delta t + \ln u) \ln u + \frac{1}{2} (\ln u - v \Delta t) \ln u = \sigma^2 \Delta t + (v \Delta t)^2. \quad (3.17)$$

The above reduces to:

$$(\ln u)^2 = \sigma^2 \Delta t + (v \Delta t)^2. \quad (3.18)$$

The solution to the system earlier described can be written as:

$$u = e^{\sqrt{\sigma^2 \Delta t + (v \Delta t)^2}}, \quad (3.19)$$

$$d = e^{-\sqrt{\sigma^2 \Delta t + (v \Delta t)^2}}, \quad (3.20)$$

and

$$q = \frac{v \Delta t}{2\sqrt{\sigma^2 \Delta t + (v \Delta t)^2}} + \frac{1}{2}. \quad (3.21)$$

Comments:

The Trigeorgis tree in general, yields prices very similar to the CRR parameterization; it does not exhibit the latter's instability for particular parameter choices (as is shown in a later chapter).

3.3.4 Jarrow and Rudd Parameterization

Jarrow and Rudd, replacing the reciprocity assumption made by CRR with that of equiprobable price tree outcomes, proposed their binomial tree parameterization in the book "Option Pricing" published in 1983, [49]. This parameterization, as shall be seen, is deficient in its adherence to the no-arbitrage principle. Accommodating for this, a new parameterization appeared in Jarrow and Turnbull's, 1995 published book, "Derivatives", [48]. However, this dissertation will concern itself with the prior.

Assumptions:

- *JR1* The underlying asset price's empirical process is assumed to be geometric Brownian motion, with drift equal to the riskless rate of return.
- *JR2* The discrete and continuous processes are assumed to have identical first and second order moments (regardless of the time step size). This is ensured via a second order moment matching.
- *JR3* The risk neutral probability of an "up" movement occurring is assumed to be equal to that of a "down" occurring, implying a risk neutral probability of a half. For this reason, the model is sometimes referred to as an equal probability tree.

Parameter Derivation:

The derivation of the random walk model parameters proceed in the following manner: Assumption JR1 can be written as:

$$d \ln S(t) = v dt + \sigma dB_t, \quad (3.22)$$

with $v = (r - \frac{\sigma^2}{2})$ and dB_t a Wiener process.

The first two moments of the above process are respectively $(r - \frac{\sigma^2}{2})t$ and $\sigma^2 t$. Equating the last mentioned and the moments of the discrete process, but honoring the assumption of equiprobable outcomes (Assumption JR3) yields:

$$\frac{1}{2} \ln \frac{S(t)u}{S(t)} + \frac{1}{2} \ln \frac{S(t)d}{S(t)} = v \Delta t \quad (3.23)$$

and

$$\frac{1}{2} (\ln \frac{S(t)u}{S(t)})^2 + \frac{1}{2} (\ln \frac{S(t)d}{S(t)})^2 = \sigma^2 \Delta t + v^2 \Delta t^2. \quad (3.24)$$

Making $\ln u$ the subject of equation (3.23) and then substituting that into (3.24) yields:

$$(2v \Delta t - \ln d)^2 + (\ln d)^2 = \sigma^2 \Delta t + 2v^2 \Delta t^2, \quad (3.25)$$

implying

$$(\ln d)^2 - 2v \Delta t \ln d + 2v^2 \Delta t^2 = \sigma^2 \Delta t + v^2 \Delta t^2. \quad (3.26)$$

Completing the square and solving for $\ln d$, results in

$$d = e^{v \Delta t - \sigma \sqrt{\Delta t}}. \quad (3.27)$$

Assumption JR3 allows:

$$u = e^{v \Delta t + \sigma \sqrt{\Delta t}}. \quad (3.28)$$

Comments:

The deviation from the risk neutral probability based upon the no-arbitrage argument has important consequences, the most obvious of which is that derivatives prices implied by Jarrow and Rudd parametrised trees violate put-call parity, and thus are not arbitrage-free - this is to be illustrated in the following example.

Example:

The price of a one-year vanilla European call option struck at the money, with the risk free rate of interest at 10%, volatility at 20%, and with the underlying's spot price trading at 1000 is (according to the Black and Scholes) closed-form solution equal to 132.69; the equivalent value for a put option being 37.53. Now, as a short put option and long call option is equal in pay-off to a long vanilla forward, it should

follow that the fair value of the former be equal to that of the latter - in the absence of any arbitrage opportunities. The value of an one-year at the spot money forward contract is 95.16; this value it should be clear is equal to that of the call- less put option.

Now, pricing the said instruments within a Jarrow and Rudd tree, where the number of time steps considered vary from 5 to 50 is as shown in Table 3.2. It should be clear that the Jarrow and Rudd implied values for vanilla European options are in violation of the put-call parity, and therefore are not arbitrage-free. It should be clear though, that the bias is very small, and decreases very quickly as the number of time steps increase.

Instrument	Steps: 5	Steps: 15	Steps: 25	Steps: 100
Call Option	130.75	133.24	133.24	132.52
Put Option	35.62	38.09	38.18	37.36
Synthetic Fwd.	95.13	95.15	95.16	95.16
Arbitrage Bias	0.03	0.01	0.00	0.00

Table 3.2: The arbitrage bias present in Jarrow and Rudd option valuations

Refer to Figure 3.2 for a graphical view of the bias in pricing error present in Jarrow and Rudd; instead of errors oscillating around zero, as is the case with the Trigeorgis tree, there is a clear and consistent over-estimation of the call option price. The convergence of the risk neutral probability to one half is also shown in Figure 3.2.

The bias, it should be noted is an increasing function of the away from the moneyness, that is in evaluating options, the bias grows in accordance with the distance away from the spot value. (The observation that Jarrow and Rudd tree implied derivatives values are not arbitrage-free, is due to Prof. Eben Maré; the proof was constructed on first principles.)

Theorem 4 *Derivatives prices as implied by the 1983 Jarrow and Rudd binomial lattice tree are not arbitrage-free.*

Proof: *The proof to the theorem follows by showing that the Jarrow and Rudd price of vanilla European options violates put-call parity. Now, a long call and short put (with the same expiry and strike) is equivalent to a forward contract on the underlying with strike equal to that of the options. Therefore, preserving the no-arbitrage condition, it should follow that the risk neutral value of a long call, short put should equal the*

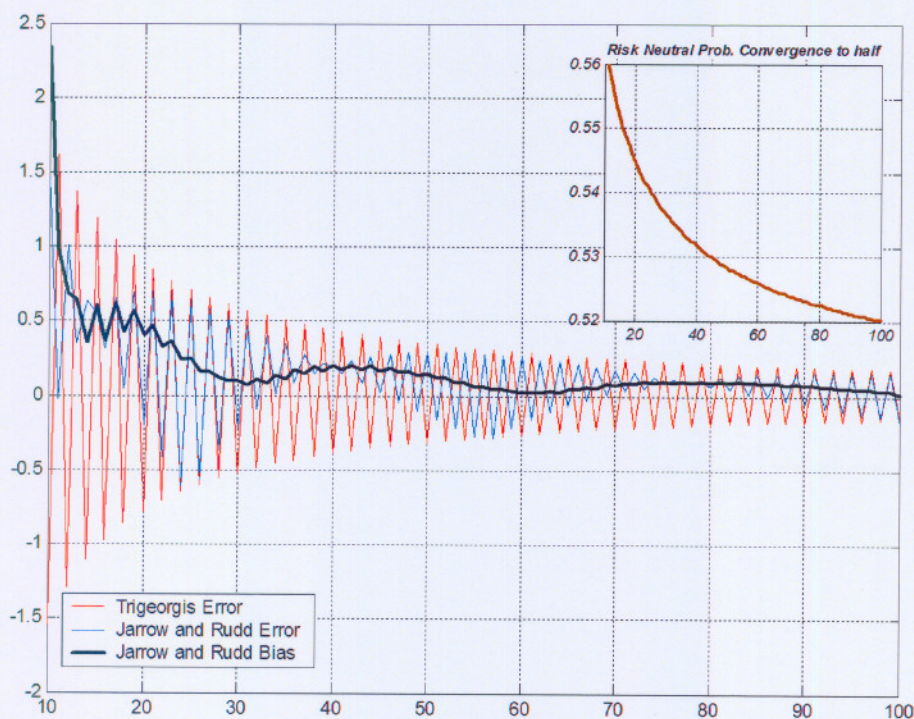


Figure 3.2: Graphical view of the pricing bias inherent in Jarrow and Rudd trees.

risk neutral value of a forward contract, i.e.

$$c(0) - p(0) = (F(0, t) - K)e^{-rt}, \quad (3.29)$$

where $c(0), p(0)$ denote the value of a call and put at time zero, and $F(0, t)$ denotes the forward price at time t of an asset, K is the strike price.

The price of a vanilla European call implied by the Jarrow and Rudd tree parameterization can be written as:

$$c(0) = \left\{ \sum_{j=a}^n \binom{n}{j} \frac{1}{2}^n u^{n-j} d^j S(0) - K \sum_{j=a}^n \binom{n}{j} \frac{1}{2}^n \right\} e^{-rt}, \quad (3.30)$$

whilst a put's price can be presented as:

$$p(0) = \left\{ K \sum_{j=0}^{a-1} \binom{n}{j} \frac{1}{2}^n - \sum_{j=0}^{a-1} \binom{n}{j} \frac{1}{2}^n u^{n-j} d^j S(0) \right\} e^{-rt}, \quad (3.31)$$

where $a = \lceil \ln Ku - \ln S(0)d^{n+1} \rceil$; $\lceil \cdot \rceil$ denoting the integer ceiling function.

Thus:

$$c(0) - p(0) = \left\{ \sum_{j=0}^n \binom{n}{j} \frac{1}{2} u^{n-j} d^j S(0) - K \sum_{j=0}^n \binom{n}{j} \frac{1}{2} \right\} e^{-rt}, \quad (3.32)$$

from the Binomial Theorem it follows that the following holds:

$$c(0) - p(0) = \left\{ S(0) \left(\frac{1}{2}u + \frac{1}{2}d \right)^n - K \right\} e^{-rt}. \quad (3.33)$$

Substituting for u and d , yields:

$$c(0) - p(0) = \left\{ S(0) e^{(r - \frac{1}{2}\sigma^2)t} \left(\cosh \sigma \sqrt{\frac{t}{n}} \right)^n - K \right\} e^{-rt}. \quad (3.34)$$

Now, if one keeps in mind that $F(0, t) = S(0)e^{rt}$, and that

$$e^{-\frac{1}{2}\sigma^2 t} \neq \left(\cosh \sigma \sqrt{\frac{t}{n}} \right)^n, \quad (3.35)$$

It follows that (3.34) cannot imply (3.29), proving that the put-call parity relationship is violated by the parameter choices implied by Jarrow and Rudd's derivation.

Q.E.D.

In addition to the above, it should be added that :

Theorem 5 *Although violating the put-call parity for time interval lengths of significant magnitude, in the limit (as the number of time steps increase to infinity), Jarrow and Rudd implied prices for vanilla European options do honor the said parity.*

Proof: Application of Taylor's expansion allows one to write

$$\left(\cosh \sigma \sqrt{\frac{t}{n}} \right)^n = \left\{ 1 + \sigma^2 \frac{t}{n} + \frac{d(n)}{n} \right\}^n, \quad (3.36)$$

where $d(n) \rightarrow 0$ as $n \rightarrow \infty$. Thus:

$$\left(\cosh \sigma \sqrt{\frac{t}{n}} \right)^n \rightarrow e^{\sigma^2 t}, \quad (3.37)$$

for $n \rightarrow \infty$. Implying that:

$$\lim_{n \rightarrow \infty} c(0) - p(0) = (F(0, t) - K) e^{-rt}. \quad (3.38)$$

The preceding shows that put-call parity is honored in the limit.

Q.E.D.

To summarize: the above shows that a bias is present in Jarrow and Rudd vanilla European options; call options are under-priced, whilst the opposite holds true for put options. However, as the risk neutral probability given by the “Fundamental Result of Binomial Pricing”, tends towards one half as the number of time steps go to infinity, it was shown that Jarrow and Rudd option prices are arbitrage-free in the limit.

3.3.5 Tian Parameterization

Tian (1993) being “A Modified Lattice Approach to Option Pricing”, [84], proposed the calibration of a binomial tree model for an asset price to the empirical process utilizing not second order moment matching posited by all authors preceding him, but a third order moment-matching. His theory was that the skewness of the discrete process plays a significant part in convergence, thus the obvious advantage of having it matched to the empirical process - which itself is skewed. Tian (1998), [80], extended the argument to include kurtosis in the calibration process, albeit with application to a three-outcome (trinomial) tree process.

Assumptions:

- *Tian1* The underlying asset price’s continuous process is assumed to be geometric Brownian motion.
- *Tian2* The discrete and continuous processes are assumed to have identical first, second and third order moments, regardless of the time step size. Thus, the discrete and continuous (empirical) processes are identical in their mean, dispersion and skewness; this is ensured via a third order moment matching.

Parameter Derivation:

Let μ_1, μ_2, μ_3 denote the first three central moments of the empirical process to which the random walk is to be calibrated. A third order moment matching results in the following system of equations:

$$pu + (1 - p)d = \mu_1, \quad (3.39)$$

$$pu^2 + (1 - p)d^2 = \mu_2, \quad (3.40)$$

and

$$pu^3 + (1 - p)d^3 = \mu_3. \quad (3.41)$$

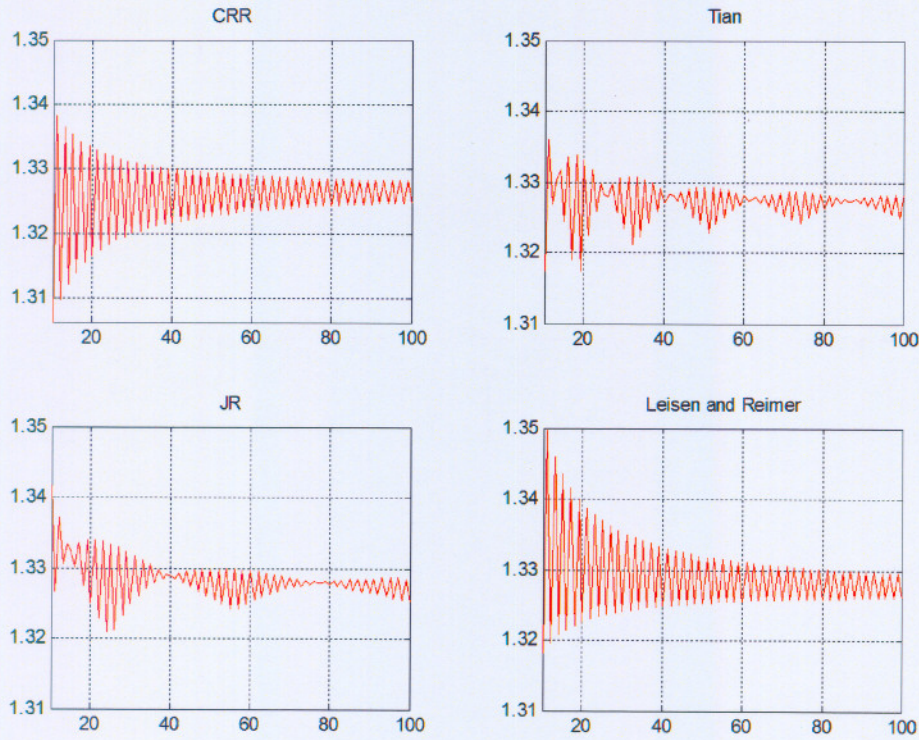


Figure 3.3: Convergence for the respective parameterizations (vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 10)

From (3.39) it follows that:

$$pu = \mu_1 - (1 - p)d, \quad (3.42)$$

and

$$p = \frac{\mu_1 - d}{u - d}. \quad (3.43)$$

Equation (3.40) can be rewritten as:

$$(pu)(u) = u(\mu_1 - (1 - p)d) = \mu_2 - (1 - p)d^2, \quad (3.44)$$

implying that

$$\mu_1(u) + (1 - p)d(d - u) = \mu_2. \quad (3.45)$$

Substitution of (3.43) into (3.45) and simplifying yields:

$$u = \frac{\mu_2 - \mu_1(d)}{\mu_1 - d}. \quad (3.46)$$

Using (3.46), equation (3.41) can be rewritten as:

$$(pu^2)(u) = u(\mu_2 - (1 - p)d^2) = \mu_3 - (1 - p)d^3. \quad (3.47)$$

Some simplifying yields

$$u\mu_2 + (1 - p)d^2(d - u) = \mu_3. \quad (3.48)$$

Substitution of (3.40) and (3.43) into the above then results in:

$$(\mu_2 - d^2)(\mu_2 - \mu_1 d) = (\mu_3 - \mu_1 d^2)(\mu_1 - d). \quad (3.49)$$

Reducing in its turn to:

$$(\mu_2 - \mu_1^2)d^2 + (\mu_2\mu_1 - \mu_3d + \mu_2^2 = \mu_1\mu_3. \quad (3.50)$$

Completing the square in the above equation, and back substitution then yields the parameter values as:

$$u = \frac{MV}{2} \{(V + 1) + \sqrt{V^2 + 2V - 3}\}, \quad (3.51)$$

and

$$d = \frac{MV}{2} \{(V + 1) - \sqrt{V^2 + 2V - 3}\}, \quad (3.52)$$

with M, V such that $\mu_1 = M$, $\mu_2 = M^2V$ and $\mu_3 = M^3V^3$.

Comments:

The Gaussian probability distribution is characterized by its first two moments, the same applies to the binomial probability distribution, implying that no “new” information is added by a third order moment matching.

Thus, contrary to Tian’s hypothesis, it should be expected that the convergence qualities of the random walk model is not enhanced by increasing the order of the moments matched. This is indeed shown in Chapter Seven.

3.3.6 Leisen and Reimer’s Parameterization

Leisen and Reimer (1995), “Binomial Models for Option Valuation - Examining and Improving Convergence”, [55], posited a parameterization of binomial trees that differs significantly from previous models. Whereas other parameterizations invariably

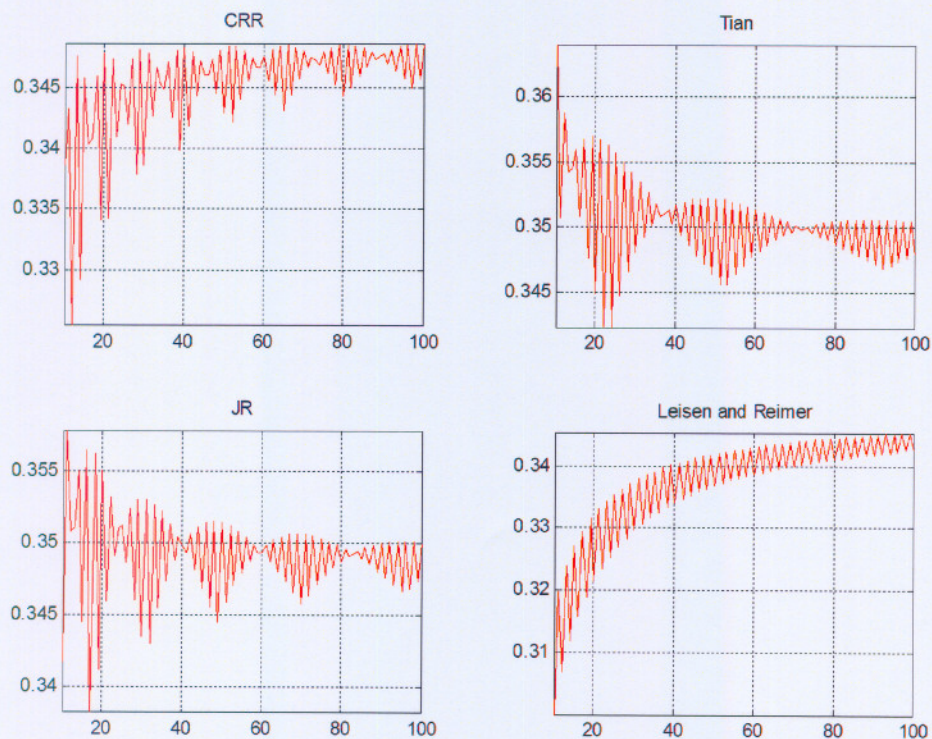


Figure 3.4: Convergence for the respective parameterizations (vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 12.5)

involve moment matchings in the derivation of parameter values, Leisen and Reimer's proposed a first order moment matching coupled with a partial CDF matching. The fundamental idea was that the CDF matching would result in the discrete distribution being forced to faster convergence towards the empirical, implying improved convergence. Fundamental to the parameterization approach is the idea of continuity correction. A short digression to explore this concept is now to follow.

Continuity Correction:

Continuity corrections are procedures applied to enhance the approximation qualities of a discrete random variable vis-à-vis a continuous one. Obviously the inverse also applies. Much has been written on enhancing the approximation qualities of the normal distribution when applied to others. The idea being that in cases where convergence in probability of a random variable is towards the normal, the last mentioned can be used to approximate CDF values, but accommodating for discretization errors by using a continuity correction. An interesting point to note: the Normal distribution first came about as an approximation for the cumulative probability mass

of the binomial distribution. This was first proposed by Abraham De-Moivre (1733), "The doctrine of chances".

The assumptions underlying the Leisen and Reimer tree are:

Assumptions:

- *LR1* The underlying asset price process is assumed to follow geometric Brownian motion, i.e., the log-return process is assumed to behave according to generalized Brownian motion.
- *LR2* The discrete and continuous processes are assumed to have identical first moments.
- *LR3* The CDFs of the discrete and continuous processes are assumed to intersect at the probability of the event that the asset price at maturity is in excess of the option strike.
- *LR4* The final assumption is that of equivalence in the expected asset price at maturity, but conditional on the option expiring in the money.

Parameter Derivation:

The empirical process followed by the variable is assumed to be $d \ln S = (r - d_y - \frac{1}{2}\sigma^2)dt + \sigma dB_t$, whilst the discrete process is as assumed in prior derivations. Assumptions ii. to iv. are equivalent to stating:

$$E^D\{S(t + \Delta t)|S(t)\} = E^C\{S(t + \Delta t)|S(t)\}, \quad (3.53)$$

$$E^D\{S(T)|S(T) > K\} = E^C\{S(T)|S(T) > K\}, \quad (3.54)$$

$$Pr.^D\{S(T)|S(T) > K\} = Pr.^C\{S(T)|S(T) > K\}, \quad (3.55)$$

with $E^D\{.\}$, $E^C\{.\}$ referring to the expected values of the discrete and continuous processes respectively; the same notational logic applying to the probability reference $Pr.^D\{.\}$. Three lemmata of relevance in this derivation are now presented:

Lemma 1 Consider the process $d \ln S = (r - d_y - \frac{1}{2}\sigma^2)dt + \sigma dB_t$. Then:

$$E\{S(T)|S(T) > K\} = S(0)e^{(r-d_y)T}N(d_1), \quad (3.56)$$

and

$$Pr.\{S(T) > K\} = N(d_2), \quad (3.57)$$

where,

$$d_1 = \frac{\ln S(0) - \ln K + (r - d_y + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}},$$

and

$$d_2 = d_1 - \sigma\sqrt{T}.$$

Proof: These are standard results that apply to lognormally distributed random variables. See Feller (1957), [79] for a full proof.

Q.E.D

The second lemma is the discrete or random walk equivalent of the above.

Lemma 2 Consider the process $d \ln S = (r - d_y - \frac{1}{2}\sigma^2)dt + \sigma dB_t$. Then:

$$E\{S(T)|S(T) > K\} = S(0) \sum_{j=a}^n \binom{n}{j} \left(\frac{up}{r}\right)^{n-j} \left(\frac{(1-p)d}{r}\right)^j. \quad (3.58)$$

$$Pr.\{S(T) > K\} = \sum_{j=a}^n \binom{n}{j} p^{n-j} (1-p)^j. \quad (3.59)$$

where, $a = \lceil \ln Ku - \ln S(0) \rceil$

Proof: These are standard results that apply to random walks. See Feller (1957), [79] for a full proof.

Q.E.D

Lemma 3 (Camp-Paulson inversion)

If $\text{Binomial}(k, n, p) = N(z)$, with k corresponding to the event z , then:

$$p = \left(\frac{b}{a}\right)^2 \left(\frac{[9a-1][9b-1] + 3z[a(9b-1)^2 + b(9a-1)^2 - 9abz^2]^{\frac{1}{2}}}{[9b-1]^2 - 9bz^2} \right)^{\frac{1}{3}}. \quad (3.60)$$

(Peizer-Pratt inversion)

If $\text{Binomial}(k, n, p) = N(z)$, with k corresponding to the event z , then:

$$p = \frac{1}{2} \pm \left[\frac{1}{4} \left[1 - \exp\left\{-\left(\frac{z}{n + \frac{1}{3}}\right)^2 \left(n + \frac{1}{6}\right)\right\} \right] \right]^{\frac{1}{2}}. \quad (3.61)$$

Refer to Paulson (1942), [62], Peizer and Pratt (1968), [63] and Camp (1951), [19] for more detailed information (including the proofs) on the above. Now, the last two stated lemmata, inverts two continuity corrections, resulting in equations in terms of the probability p .

The assumptions set out in equations (3.53) to (3.55), along with the above lemmata, can be used to construct the following system of equations:

$$pu + (2 - p)d = R, \quad (3.62)$$

(with $R = (1 + r\Delta t) \simeq e^{r\Delta t}$.)

$$\sum_{j=a}^n \binom{n}{j} \left(\frac{up}{r}\right)^{n-j} \left(\frac{(1-p)d}{r}\right)^j = N(d_2), \quad (3.63)$$

and

$$\sum_{j=a}^n \binom{n}{j} p^{n-j} (1-p)^j = N(d_1). \quad (3.64)$$

Assumption LR2 implies that $0 \leq \left(\frac{up}{r}\right) \leq 1$ and $\left(\frac{up}{r}\right) + \left(\frac{(1-p)d}{r}\right) = 1$. Thus, the continuity correction inversions (presented in lemma 3) can be used to simplify both equations (3.63) and (3.64), resulting in the following solution set:

$$d = \frac{R - pu}{1 - p}, \quad (3.65)$$

$$u = R\left(\frac{\dot{p}}{p}\right), \quad (3.66)$$

$$p = f(d_2, \frac{n}{2}), \quad (3.67)$$

and

$$\dot{p} = f(d_1, \frac{n}{2}). \quad (3.68)$$

The last two equations referring to the selected continuity correction inversion.

Comments:

It should be noted that the tree parameterization proposed by Leisen and Reimer is dependent upon events defined in terms of vanilla European options; this results in biased pricing for other derivative securities and makes the general use of this methodology problematic. However, the principle of partial CDF matching seems to be quite powerful in improving convergence; adjusting the events at which CDFs are matched to accommodate eccentricities of other derivative pay-offs might yield results equivalent in their utility to that of Leisen and Reimer tree parameterizations for vanilla European options.

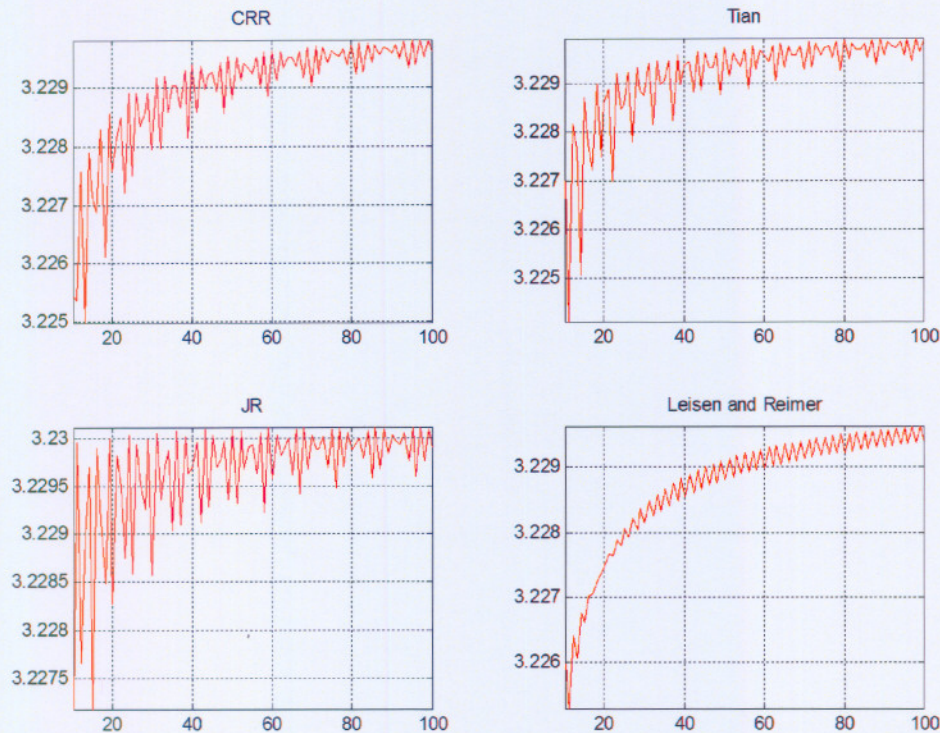


Figure 3.5: Convergence for the respective parameterizations (Vanilla European call option; one year; risk free rate = 10%; sigma = 20 %; spot = 10; strike = 7.5)

3.3.7 Pricing Examples

The convergence diagrams (price versus number of time steps) for vanilla European options, with varying degrees of moneyness are provided in figures 3.3, 3.4 and 3.3. Note the (surprisingly) varying convergence patterns for the respective parameterizations.

3.4 Probability Measure Dichotomy

The risk neutral probability has been prior defined as the probability measure that ensures the martingale property for risk free discounted asset price behavior. Harrison and Pliska (1981), [39], has shown such martingale behavior to be a necessary

condition for arbitrage-free derivatives prices. For binomial trees the risk neutral probability can always be expressed in terms of the “up” and “down” factors, and well in the following way: $p = \frac{R - d}{u - d}$.

Now, should the discrete process be calibrated towards the continuous price process, with the last mentioned already being a zero-drift process, and calibration take place through at least a first order moment matching, it would be logical to assume that the probability measure implied by the calibration should be equivalent to the risk neutral measure. Otherwise said: calibration towards a risk neutral process, should yield another risk neutral process. In the preceding parameter derivations, vis. Cox, Ross and Rubinstein’s diffusion, as well as in the Trigeorgis the measure p realized slightly different from that prior stated. Tian’s parameterization however, has the two measures equivalent.

The dichotomy alluded to above is quite interesting. Cox et al and Jarrow and Rudd’s derivations (although comprising a first order moment matching) yields a discrete process, the first moment of which is contrary to the first moment of the price process implied by the continuous, risk neutral process assumed. To expand upon this: should one assume that

$$d \ln S(t) = (r - \frac{\delta^2}{2})dt + \sigma dB_t,$$

then it follows that

$$E\{S(t + \Delta t)|S(t)\} = S(t)R.$$

Both Trigeorgis and Cox, Ross and Rubinstein calibrate towards the aforesaid return process, however they yield first order price moments of

$$E\{S(t + \Delta t)|S(t)\} = S(t)\{u\frac{v\Delta t}{2\sqrt{\sigma^2\Delta t}} + \frac{1}{2} + d(1 - \frac{v\Delta t}{2\sqrt{\sigma^2\Delta t}} + \frac{1}{2})\}.$$

The latter, though converging in the limit (as the time interval size becomes progressively smaller), is not equal to $RS(t)$.

How is this issue to be explained or resolved? Firstly note the main difference between the Tian and Trigeorgis derivations being that the first calibrates towards the price moments (i.e., towards the discrete compounded return series $dS(t) = rS(t)dt + \sigma S(t)dB_t$ ³), whilst the second performs a moment matching towards the continuous

³The “discretely compounding” continuous process implies a product normal distribution for the

return series. Now, Ito's lemma - used to construct the log-return from the discretely compounded continuous process - provides a way in which transformed diffusion process can be equated in terms of distribution (process) for continuous time, with the equation sign being defined in terms of mean square convergence. That is, should one use the lemma to derive the process of a transformed random variable, from the original non-transformed process, the second one would be applicable only in the mean square sense (limit). To explain this more symbolically:

Theorem 6 *If A is distributed like B , then from Ito's lemma the process for the transformed variable $C = f(A)$ can be derived, but only in the mean square sense.*

Thus, for larger time increments one should expect a dichotomy between the log-return and discretely compounded return processes. This dichotomy materializes in the true and risk neutral probability measures (and consequently the first price moments) differing for the Trigeorgis and Cox, Ross and Rubinstein processes. It is worthwhile noting that the "true" probability, i.e., q , converges in the limit to p . It is worthwhile to take cognizance of the presence of mention to this issue in the Cox et al (1979), [23], (as a footnote).

3.5 On Implementing Binomial Algorithms

There are various ways in which one could code the binomial tree algorithms presented in this Chapter. Higham (2002), [43], summarised some of the most general ways in which it could be done. Clewlow and Strickland (1998), [21], provides pseudo code for some of the binomial and trinomial tree methods. Fundamentally, however, it is a matter of choosing between a recursion approach (with consequently very little memory requirements), and an approach where trees are stored in matrices. The latter is especially suited to path dependent options, whereas the prior mentioned is quite efficient in terms of computation time. The examples worked out for this dissertation were based on independent (my own) Matlab implementations of the algorithms ; it is presented in the Appendix.

asset price at some future date, whilst the transformed one implies a log-normally distributed value. Binomial tree procedures effectively discretize the product normal distribution. As such, it could be argued that calibration should be towards the latter. As a final remark: the product normal distribution converges quickly to the lognormal one as the number of compounding periods increase.

3.6 Summary

In this section the essential techniques underlying the construction of binomial trees were presented. The principal tree parameterizations were also explored and well as regards the assumptions made, applicable parameter derivations and eccentricities in behaviour exhibited. An example of the latter being the violation of the no-arbitrage condition by Jarrow and Rudd implied derivatives prices. A pricing example was given in the form of convergence diagrams. Furthermore, the section investigated (albeit briefly) the implementation of binomial tree algorithms in computer programming.

Chapter 4

Equivalence Explored

If all roads lead to Rome, it should follow that some kind of equivalence relation could be defined to explain the roads' relationship to each other. In the case of derivatives pricing, most methodologies lead to risk neutral (arbitrage-free) prices. The alternative would firstly imply difficulty in trading (as agreement on the value of instruments would be hard to reach), and secondly to exploitation of some, as the more accurately priced derivatives would very much hold their own. The quantitative analyst has a whole range methodologies at his disposal to approach derivatives pricing, including: closed-form analytical formulae, solving the Black and Scholes partial differential equation (either explicitly or by way of numerical methods) and lastly, by resorting to numerical integration schemes. In this chapter the aim is to investigate the relationship between the enumerated and binomial methods of pricing.

4.1 Background

The most fundamental result (and best known in arbitrage-free derivatives pricing) is that of Black and Scholes (1973), [9]. Based upon a replication argument, that precludes arbitrage strategies, as well as some assumptions regarding the stochastic behavior of the price of the underlying, the authors derived a partial differential equation, that describes the behavior of all derivative instruments. This scheme forms the basis of many closed-form solutions; the possibility of these being due to the absence of any stochastic components from the said equation, implying that “normal” partial differential equation solution methods are applicable. Obvious examples of such closed-form solutions include vanilla European options (Black and Scholes (1973), [9], Merton (1973), [57]), forward starting options and barrier options. In the case of many derivatives for which a closed-form solution is non-existent, approximations are easily derived by tweaking the boundary conditions slightly so as to allow for solution, vis. arithmetic Asian options and options where the underlying pays fixed dollar dividends. (Refer to Geske and Shastri (1985), [35], and Hull and White (1993), [46]¹ who provided an overview of various approximation techniques used for the pricing of derivatives contracts.) In Section Two it is shown that the binomial method converges towards the closed-form solution provided by Black and Scholes for vanilla European options.

Besides solving the Black and Scholes partial differential equation explicitly, numerous algorithms used in numerical solution approximation for normal partial differential equations, are easily applied. Examples of this being the explicit finite difference scheme, the implicit finite difference scheme, Crank Nicholson’s algorithm etc. The work by Brennan and Schwartz (1978), [14], for example draws on this. Section Three comprises two equivalence investigations, the first being that the “Fundamental Result of Binomial Pricing”, Theorem 3, is in fact equal to the Black and Scholes PDE; the second part of the Section then proceeds to prove that binomial pricing is akin to applying the explicit finite difference scheme for solving the said PDE.

Shortly after the Black and Scholes (1979), [9], publication, Merton (1973), [57] provided a solution to their partial differential equation that relied on the calculation of an expectation integral. It was, however, the advent of Harris and Pliska’s (1981) work on probability measures, [41], for the underlying asset and risk neutrality, that saw the pricing of derivatives assume more of a “discounted expected pay-off” form. In fact, it has led, nowadays, to a clear distinction between partial differential equation-

¹Running slightly ahead: the latter reference, only considers binomial and trinomial tree methods, however, as trinomial methods are equivalent to the explicit finite difference PDE solution scheme, the statement made, it is felt, is legitimate.

based methods and “martingale” methods. Now, the calculation of expected values, given a continuous random variable in the expectation, yields integral type formulae. Some indefinite integrals are not tractable into a closed-form, for these, a whole array of methods have been devised to gain approximations, e.g. Simpson’s rules, the trapezoid rules, Romberg’s etc. Use of these schemes to solve the expectation is referred to as the “numerical integration” approach. In Section Four, it shall be shown that binomial methods are in fact akin to the rectangular-rule approach to solving for the discounted expected pay-off. In this respect, reference is made to the work of Boyle (1988), [12], Parkinson (1977), [65], Andricopolous et al (2003), [1], to Omberg (1987), [64], and others.

4.2 The Black and Scholes Closed-Form Solution

In general mathematics of finance parlance, Black and Scholes’ result invariably refers to the closed-form solution for the price of a vanilla European option (and not, as one could have expected to the general partial differential equation). Black and Scholes (1973), [9] provided an analytical formula with parameters the risk free rate of interest, the time to maturity of the option, the current spot price for the asset, the strike and the volatility of asset price returns, and which returns the arbitrage free value of vanilla options. As the enumerated parameters are easily observable (or estimable historically) vanilla European options are easily priced. To say that this formula revolutionized structured finance and proprietary trading, is to make a gross understatement. To the current day, the said formula forms the core of most derivatives pricing and hedging, as well as research. As such, this result could be considered key in derivatives pricing; the consistency of binomial tree pricing with the said, is explored in this section. But first, the Black and Scholes vanilla European option price result:

Theorem 7 (The Black and Scholes Closed-Form Solution)

The price of a vanilla European option, on an asset that generates no cash flows can be written as:

$$c(0) = S(0)N(d_1) - Ke^{-rT}N(d_2), \quad (4.1)$$

for a call option, and for a put option:

$$p(0) = Ke^{-rT}N(-d_2) - S(0)N(-d_1), \quad (4.2)$$

with $S(0)$, r , T , K being the current spot (cash) price of the underlying asset, the risk free rate of return, the time to expiry of the option and the strike price respectively.

The terms $N(d_1)$ and $N(d_2)$ referring to the cumulative probability density function of the Normal distribution, namely,

$$N(d_1) = \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} \exp^{-\frac{x^2}{2}} dx, \quad (4.3)$$

and d_1, d_2 defined as:

$$d_1 = \frac{\ln(S(0)) - \ln(K) + rT + \frac{1}{2}\sigma^2 T}{\sqrt{\sigma^2 T}},$$

and

$$d_2 = d_1 - \sqrt{\sigma^2 T}.$$

Proof: Refer to Black and Scholes (1973), [9] for the proof.

Q.E.D

In their original paper, Cox et al (1979) [23], showed equivalence of the binomial tree implied price for vanilla European options to the Black and Scholes one. The proof given by them is very sketchy, failing to provide rigorous mathematical backing for some of the arguments put forward - for example, nowhere is mentioned which form of the central limit theorem is used.

The Jarrow and Rudd tree, see Jarrow and Rudd (1983), [47], having the risk neutral probability set to a constant, is very simple to show consistent with the Black and Scholes result - it simply uses the very first of the central limit theorem incarnations, that of De-Moivre; for the full proof refer to Feller, (1957), [33]. Others who have contributed to showing equivalence between the Cox, Ross and Rubinstein tree implied option prices and Black and Scholes' closed-form solution include Musiela and Rutkowski (1997), [59] and Van den Berg (2000), [83]. The proof given in this dissertation is one that was constructed from first principles, and does not blindly follow the arguments from any of the prior stated articles. Furthermore, it concerns itself with the Cox, Ross and Rubinstein tree parameterization only, (see Section 3.3.1); the proof to the other parameterizations follow among similar lines, although the Jarrow and Rudd tree is considerably easier to prove consistency for.

Firstly, it is necessary to understand that the binomial tree valuation is a numerical scheme, thus, its aim is convergence to the solution. Equivalence therefore is only defined in terms of convergence. Before progressing any further, Cox et al (1979) [23] has shown that² the price of a vanilla European option, given a binomial tree scheme

²The result follows very easily from first principles as well.

is expressible in the following way:

$$c(0) = \left\{ \sum_{j=a}^n \binom{n}{j} (up)^{n-j} (d(1-p))^j S(0) - K \sum_{j=a}^n \binom{n}{j} (up)^{n-j} (d(1-p))^j \right\} e^{-rt}. \quad (4.4)$$

Where a is defined as the smallest integer greater than $\frac{\ln \frac{K}{S(0)} - n \ln d}{\ln \frac{u}{d}}$. It follows that equivalence between (4.4) and (4.1) would be proven if it is shown that:

$$\sum_{j=a}^n \binom{n}{j} (up)^{n-j} (d(1-p))^j e^{-rt} \rightarrow N(d_1), \quad (4.5)$$

and

$$\sum_{j=a}^n \binom{n}{j} (up)^{n-j} (d(1-p))^j \rightarrow N(d_2). \quad (4.6)$$

The proof for this will concern itself only with showing that the last mentioned, (4.6), holds true; (4.5)'s proof following along similar lines. The proof comprises three parts, namely:

- It should be noted that the events $\{X < a\}$ and $\{S(T) < K\}$, where $X \sim \text{Binomial}(n; p)$, are equivalent. (This being relatively self-evident.)
- After which it is necessary to show that $\frac{X - np}{\sqrt{np(1-p)}} \rightarrow^d N(0, 1)$. This is done by using the central limit theorem, however, very important: one should observe that the risk neutral probability is itself a function of the number of time steps, n , and therefore one should be careful in applying the said limit theorem.
- Finally, it shall be proven that $x = \frac{a - 1 - np}{\sqrt{np(1-p)}} \rightarrow d_2$.

We'll now progress to show that:

$$\frac{X - np}{\sqrt{np(1-p)}} \rightarrow^d N(0, 1).$$

The above relies on application of the central limit theorem. Rather than one theorem, the central limit theorem is in fact a collection of convergence results that shows that under certain conditions the sum of a number of independent realizations (from either binomial or arbitrary distributions), converges in distribution towards

the normal distribution. Development of the central limit theorem has spanned the period starting 1733; additions to it are still ongoing. Some of the greatest names in mathematical history has contributed towards it, including Laplace, Poisson, Polya and others. Refer to Cong (2003), [22], Feller (1957) [33] and [86] for a good exposition of the central limit theorem.) The oldest incarnation of the theorem is that of De-Moivre, and is now stated:

Theorem 8 (De-Moivre's Central Limit Theorem)

Let $X \sim \text{Bin}(n; p)$, and $\text{Pr}\{X = x\} = \binom{n}{x} p^x (1-p)^{n-x}$. If p is a constant function of n , then it could be said that:

$$\sqrt{np(1-p)} \text{Pr}\{X = x\} \rightarrow \frac{1}{\sqrt{2\pi}} \exp \frac{(x-np)^2}{2np(1-p)}.$$

Proof: Refer to Bain and Engelhardt (1992) [3] for the proof.

Q.E.D

The above is applicable only in the case where the parameters of the binomial distribution whose distribution is considered, are constant functions of the number of trials. This certainly holds for Jarrow and Rudd's tree parameterization, and thus, equivalence is proved by simply applying the De-Moivre central limit theorem. In the case of Cox, Ross and Rubinstein's (as well as the others) tree parameterization, as both p and a are functions of the number of time steps, n . In this situation, need arises for the central limit theorem as stated by Lyapunov; it is presented below.

Theorem 9 (Lyapunov's Central Limit Theorem)

Let $X_1, X_2, \dots, X_n$ be a sequence of independent random variables defined on the same probability space; assume that all the variables are of finite expectation and standard deviation denoted by μ_i and σ_i respectively. Define:

$$s_n^2 = \sum_{i=1}^n \sigma_i^2, \quad (4.7)$$

assume that the third moments about the mean are finite and denoted by:

$$\tau_n^3 = E\{|X_n - \mu_n|^3\}. \quad (4.8)$$

If $\lim_{n \rightarrow \infty} \frac{\tau_n}{s_n^3} = 0$ then:

$$Z_n = \frac{S_n - M_n}{s_n}, \quad (4.9)$$

converges in distribution towards the standard Normal distribution, where $S_n = \sum_{i=1}^n X_i^2$, and, $M_n = E\{S_n\}$.

Proof: Refer to Feller (1957) [33] for the proof.

Q.E.D

Now define $X = \sum_{i=1}^n \epsilon_i$ where the ϵ_i 's are independent Bernoulli random variables with probability of success p . Denote by:

$$\tau_i = E\{|\epsilon_i - E\{\epsilon_i\}|^3\}. \quad (4.10)$$

From first principles, and the fact that $0 \leq p \leq 1$, it follows that:

$$\begin{aligned} \tau_i &= (1-p)^3 p + p^3 (1-p) \\ &= \{1 - 3p + 3p^2 - p^3\}p + p^3 - p^4 \\ &= p - 3p^2 + 3p^3 - p^4 + p^3 - p^4 \\ &= p - 3p^2 + 4p^3 - 2p^4. \end{aligned}$$

Let $s_n^2 = \sum_{i=1}^n \text{Variance}\{\epsilon_i\} = \sum_{i=1}^n p(1-p) = np(1-p)$.
Then,

$$\begin{aligned} \lim_{n \uparrow \infty} \frac{\tau_i}{s_n} &= \lim_{n \uparrow \infty} \frac{p - 3p^2 + 4p^3 - 2p^4}{\sqrt{np(1-p)}} \\ &= \lim_{n \uparrow \infty} \frac{p(1-p)(-2p^2 + 2p - 1)}{\sqrt{np(1-p)}} \\ &= \lim_{n \uparrow \infty} \frac{\sqrt{p(1-p)}(-2p^2 - 2p + 1)}{\sqrt{n}}. \end{aligned}$$

Before progressing any further, an important sub-result:

Lemma 4

$$p = \frac{R - d}{u - d} \rightarrow \frac{1}{2} \text{ as } n \uparrow \infty.$$

Proof: Using the definition of the risk neutral probability, along with a substitution of the parameter choices specified by Cox, Ross and Rubinstein's tree (see Section

3.3.1), the following results

$$\begin{aligned} p &= \frac{R - d}{u - d} \\ &= \frac{e^{r\frac{t}{n}} - e^{-\sigma\sqrt{\frac{t}{n}}}}{e^{\sigma\sqrt{\frac{t}{n}}} - e^{-\sigma\sqrt{\frac{t}{n}}}}. \end{aligned}$$

A Taylor series expansion then yields:

$$\begin{aligned} &= \frac{\{1 + r\frac{t}{n} + \frac{1}{2}r\frac{t}{n} + \dots - 1 + \sigma\sqrt{\frac{t}{n}} - \frac{1}{2}\sigma^2\frac{t}{n} + \dots\}}{\{1 + \sigma\sqrt{\frac{t}{n}} + \frac{1}{2}\sigma^2\frac{t}{n} \dots - 1 + \sigma\sqrt{\frac{t}{n}} - \frac{1}{2}\sigma^2\frac{t}{n} \dots\}} \\ &= \frac{r\frac{t}{n} + \sigma\sqrt{\frac{t}{n}} + d(n)}{2\sigma\sqrt{\frac{t}{n}} + d'(n)} \\ &= \frac{r\sqrt{\frac{t}{n}} + \sigma + d(n)\sqrt{\frac{n}{t}}}{2\sigma + d'(n)\sqrt{\frac{n}{t}}}, \end{aligned}$$

with $d(n), d'(n)$ containing terms of order 1 and higher (with regards to n), it should then follow that $d(n)\sqrt{\frac{n}{t}} \rightarrow 0$ and $d'(n)\sqrt{\frac{n}{t}} \rightarrow 0$ as $n \uparrow \infty$.

Thus:

$$\lim_{n \uparrow \infty} p = \frac{1}{2}.$$

Q.E.D

If $\lim_{n \uparrow \infty} p = \frac{1}{2}$, then: $\lim_{n \uparrow \infty} \frac{\tau_i}{s_n} = 0$, so Lyapunov's condition holds for X as prior defined, and we can state that:

$$Pr.\left\{\frac{X - np}{\sqrt{np(1-p)}}\right\} \rightarrow N\left(\frac{X - np}{\sqrt{np(1-p)}}\right). \quad (4.11)$$

What still remains of the equivalence proof is to show that: $x = \frac{a - 1 - np}{\sqrt{np(1-p)}} \rightarrow -d_2$.

First note that $a - 1 = \frac{\ln \frac{K}{S(0)} - n \ln d - \epsilon \ln \frac{u}{d}}{\ln \frac{u}{d}}$ for ϵ a Real number between zero and

one. Thus,

$$\begin{aligned}
 x &= \frac{\ln \frac{K}{S(0)} - n \ln d - \epsilon \ln \frac{u}{d} - np \ln \frac{u}{d}}{\ln \frac{u}{d} \sqrt{np(1-p)}} \\
 &= \frac{\ln \frac{K}{S(0)} + n\sigma \sqrt{\frac{t}{n}} - 2\epsilon\sigma \sqrt{\frac{t}{n}} - \frac{1}{2\sigma} \{r\sqrt{nt} + \sigma n - \frac{1}{2}\sigma^2 t \sqrt{\frac{n}{t}}\} 2\sigma \sqrt{\frac{t}{n}}}{2\sigma \sqrt{\frac{t}{n}} \sqrt{np(1-p)}} \\
 &= \frac{\ln \frac{K}{S(0)} - 2\epsilon\sigma \sqrt{\frac{t}{n}} - rt - \frac{1}{2}\sigma^2 t}{2\sigma \sqrt{\frac{t}{n}} \sqrt{np(1-p)}}.
 \end{aligned}$$

Which results in:

$$\begin{aligned}
 \lim_{n \uparrow \infty} -x &= \lim_{n \uparrow \infty} \frac{\ln \frac{S(0)}{K} + rt + \frac{1}{2}\sigma^2 t - 2\epsilon\sigma \sqrt{\frac{t}{n}}}{2\sigma \sqrt{\frac{t}{n}} \sqrt{np(1-p)}} \\
 &= \frac{\ln \frac{S(0)}{K} + rt + \frac{1}{2}\sigma^2 t}{\sigma \sqrt{t}}.
 \end{aligned}$$

The above concludes the equivalence or convergence proof. ³

4.3 Finite Difference Schemes

Cox et al (1979), [23], showed as a footnote that the valuation of a derivative using the binomial method, results in solving a partial differential equation identical to the Black and Scholes partial differential equation. In other words, the authors showed that the “Fundamental Result of Binomial Pricing” is in fact the Black and Scholes

³Another possible way of conducting the above proof is by first recognizing that:

$$\begin{aligned}
 &\left\{ \sum_{j=a}^n \binom{n}{j} p^{n-j} (1-p)^j \{S(0)u^{n-j}d^j - K\} e^{-rt} \right. \\
 &\left. a \binom{n}{a} \left\{ S \int_0^{\frac{K}{R}} y^{a-1} (1-y)^{n-a} dy - \frac{K}{R} \int_0^p y^{a-1} (1-y)^{n-a} dy \right\} \right.
 \end{aligned}$$

The above result from seeing that the left hand of the equation comes about by repeated application of integration by part of the right. Now, applying the limit $n \rightarrow \infty$ to the above, one could attempt the equivalence proof as presented prior.

partial differential equation. (Note: Geske and Shastri (1985), [35], mentioned that although equivalent to explicit finite difference schemes, the binomial method is more robust as concerns stability as well as being more efficient in that convergence is better. As a matter of fact it was stated that convergence in binomial models is better by a factor of ten.) In this section the two aforementioned results will be presented and further explored.

4.3.1 Equivalence to the Black and Scholes Partial Differential Equation

The focus here is to show that the “Fundamental Result of Binomial Pricing” as presented in Theorem 3, is equivalent to the Black and Scholes partial differential equation. (The proof given here is one that was suggested by Prof. Eben Maré.)

The Black and Scholes partial differential equation, it is recalled can be written as:

$$\frac{\partial f}{\partial t} + rS \frac{\partial f}{\partial S} + \frac{1}{2} S^2 \sigma^2 \frac{\partial^2 f}{\partial S^2} = rf, \quad (4.12)$$

where f denotes the value of any derivative on S , and r the continuously compounded return on the riskless asset.

Recall the “Fundamental Result of Binomial Pricing”:

$$pf(t + \Delta t|Su) + (1 - p)f(t + \Delta t|Sd) = Rf, \quad (4.13)$$

with $p = \frac{R - d}{u - d}$, the risk neutral probability measure and $R = e^{r\Delta t}$.

Now, the first step in the proof involves considering a second order Taylor series expansion of the variables $f(t + \Delta t|Su)$ and $f(t + \Delta t|Sd)$, namely :

$$\begin{aligned} f(t + \Delta t|Su) &= f + \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} (Su - S) + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (Su - S)^2 \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \Delta t^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t \partial S} \Delta t (Su - S), \end{aligned} \quad (4.14)$$

and,

$$\begin{aligned} f(t + \Delta t|Sd) &= f + \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} (Sd - S) + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} (Sd - S)^2 \\ &\quad + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \Delta t^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t \partial S} \Delta t (Sd - S). \end{aligned} \quad (4.15)$$

Substitution of equations (4.14) and (4.15) into (4.13) (and some simplification) then yields:

$$\begin{aligned}
 f &+ \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} \{(Su - S)p + (Sd - S)(1 - p)\} \\
 &+ \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \Delta t^2 + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} \{(Su - S)^2 p + (Sd - S)^2 (1 - p)\} \\
 &+ \frac{1}{2} \frac{\partial^2 f}{\partial t \partial S} \Delta t \{(Su - S)p + (Sd - S)(1 - p)\} \\
 &= fR.
 \end{aligned} \tag{4.16}$$

By some algebra and substitution of $p = \frac{R - d}{u - d}$, one can write:

$$S\{(u - d)p + (d - 1)(1 - p)\} = S(R - 1) \tag{4.17}$$

$$(Su - S)^2 p + (Sd - S)^2 (1 - p) = S^2 \{uR - ud + Rd - 2R + 1\}. \tag{4.18}$$

Furthermore, for the CRR parameterization of the binomial tree, it is rather trivial to show that:

$$S^2 \{uR - ud + Rd - 2R + 1\} = \sigma^2 \Delta t. \tag{4.19}$$

Substitution of (4.17), (4.18) and (4.19) into (4.16) allows one to write:

$$\begin{aligned}
 &f + \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} S(R - 1) + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \Delta t^2 \\
 &+ \frac{1}{2} \frac{\partial^2 f}{\partial S^2} S^2 \sigma^2 \Delta t + \frac{1}{2} \frac{\partial^2 f}{\partial t \partial S} \Delta t S(R - 1) = fR.
 \end{aligned} \tag{4.20}$$

A second order Taylor series expansion of R results in the latter being expressed as $R = 1 + r\Delta t + \frac{1}{2}r^2\Delta t^2$. Substitution of the said expansion into (4.21) consequently results:

$$\begin{aligned}
 f &+ \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} S(r\Delta t + \frac{1}{2}r^2\Delta t^2) \\
 &+ \frac{1}{2} \frac{\partial^2 f}{\partial t^2} \Delta t^2 + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} S^2 \sigma^2 \Delta t \\
 &+ \frac{1}{2} \frac{\partial^2 f}{\partial t \partial S} \Delta t S(r\Delta t + \frac{1}{2}r^2\Delta t^2) \\
 &= f(1 + r\Delta t + \frac{1}{2}r^2\Delta t^2).
 \end{aligned} \tag{4.21}$$

Ignoring all terms in Δt of order higher than one, leaves:

$$f + \frac{\partial f}{\partial t} \Delta t + \frac{\partial f}{\partial S} S r \Delta t + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} S^2 \sigma^2 \Delta t = f(1 + r \Delta t). \quad (4.22)$$

Dividing (??) by Δt throughout, and a rearrangement of terms then yields:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial S} S r + \frac{1}{2} \frac{\partial^2 f}{\partial S^2} S^2 \sigma^2 = f r.$$

This concludes the proof that derivative pricing using the binomial tree method (at least given the CRR parameterization) is equivalent to solving the Black and Scholes PDE.

4.3.2 Equivalence to Explicit Finite Difference Schemes

Proving that the “Fundamental Result of Binomial Pricing” and the Black and Scholes partial differential equation are equivalent still leaves the issue of exactly how the solution of the partial differential equation is done via the binomial algorithm. To expand on this further, numerous numerical schemes exist that provide solution approximations to partial differential equations, for example the explicit finite difference scheme, the Crank-Nicholson method etc. The binomial pricing algorithm is in fact a solution of the Black and Scholes partial differential equation by way of the explicit finite difference method. Brennan and Schwartz (1977), [14], showed how to solve the Black and Scholes partial differential equation in an explicit finite difference framework. (The equivalence proof supplied here was constructed independently, by employing the latter and applying first principles.)

First consider the “Fundamental Result of Binomial Pricing” as given in Theorem 3:

$$f(S; t)R = pf(Su; t + \Delta t) + (1 - p)f(Sd; t + \Delta t), \quad (4.23)$$

where $R = e^{r\Delta t}$ and $p = \frac{R-d}{u-d}$, and with $f(S; t)$ the value of a derivative instrument on an asset with value $S(t)$, at time t . Self-substitution of (4.23) yields:

$$\begin{aligned} f(S; t)R^2 &= p\{pf(Suu; t + 2\Delta t) + (1 - p)f(Sud; t + 2\Delta t)\} + \\ &\quad (1 - p)\{pf(Sdu; t + 2\Delta t) + (1 - p)f(Sdd; t + 2\Delta t)\}. \end{aligned}$$

Simplification of the above then results in:

$$f(S; t)R^2 = p^2\{f(Suu; t + 2\Delta t) - 2f(Sud; t + 2\Delta t) + f(Sdd; t + 2\Delta t)\} + 2p\{f(Sud; t + 2\Delta t) - f(Sdd; t + 2\Delta t)\} + f(Sdd; t + \Delta t),$$

where $f(Sud; t + \Delta t) = f(Sdu; t + \Delta t)$ (by choice). Now, the explicit finite difference scheme proposes estimation of partial derivatives as follows:

$$\begin{aligned}\frac{\hat{\partial}f}{\partial S} &= \frac{f(Sud; t + 2\Delta t) - f(Sdd; t + 2\Delta t)}{Sud - Sdd} \\ &= \frac{f(Sud; t + 2\Delta t) - f(Sdd; t + 2\Delta t)}{Sd(u - d)} \\ \frac{\partial^2 f}{\partial S^2} &= \frac{f(Suu; t + 2\Delta t) - 2f(Sud; t + 2\Delta t) + f(Sdd; t + 2\Delta t)}{(Suu - Sud - Sdu + Sdd)} \\ &= \frac{f(Suu; t + 2\Delta t) - 2f(Sud; t + 2\Delta t) + f(Sdd; t + 2\Delta t)}{S^2(u - d)^2} \\ \frac{\hat{\partial}f}{\partial t} &= \frac{f(Sud; t + 2\Delta t) - f(S; t)}{\Delta t} = \frac{f(S; t + 2\Delta t) - f(S; t)}{2\Delta t}.\end{aligned}$$

Keeping in mind that: $p^2 = \left\{\frac{R - d}{u - d}\right\}^2$, allows for (4.24) to be written as

$$R^2 f(S; t) = S^2(R - d)^2 \frac{\partial^2 f}{\partial S^2} + 2S(dR - d^2) \frac{\hat{\partial}f}{\partial S} + f(Sdd; t + 2\Delta t).$$

Now, using a Taylor series expansion where terms of order higher than 1 are ignored yields the following

$$\begin{aligned}(R - d)^2 &= (e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}})^2 \\ &\approx (r\Delta t + \sigma\sqrt{\Delta t} - \frac{1}{2}\sigma^2\Delta t)^2 \\ &\approx \sigma^2\Delta t \\ R^2 &= e^{2r\Delta t} \\ &\approx (1 + 2r\Delta t) \\ (Rd - d^2) &= (e^{r\Delta t}e^{-\sigma\sqrt{\Delta t}} - e^{-2\sigma\sqrt{\Delta t}}) \\ &\approx (1 + r\Delta t)(1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t) - (1 - 2\sigma\sqrt{\Delta t} + \sigma^2\Delta t) \\ &\approx (r\Delta t + \sigma\sqrt{\Delta t} - \frac{1}{2}\sigma^2\Delta t).\end{aligned}$$

Lastly, a second order Taylor series expansion of u , and d and simplification yields: $(u - d) = 2\sigma\sqrt{\Delta t}$. Substitution of the various approximations into (4.24) yields

$$(1 + 2r\Delta t)f(S; t) = S^2\sigma^2\Delta t\frac{\partial^2 f}{\partial S^2} + 2S(r\Delta t + \sigma\sqrt{\Delta t} - \frac{1}{2}\sigma^2\Delta t)\frac{\partial f}{\partial S} + f(Sdd; t + 2\Delta t),$$

and

$$rf(S; t) = \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial S^2} + 2Sr\frac{\partial f}{\partial S} + \alpha, \quad (4.24)$$

where

$$\begin{aligned} \alpha &= \frac{1}{2\Delta t} \{ f(Sdd; t + 2\Delta t) - f(S; t) + 2S(\sigma\sqrt{\Delta t} - \frac{1}{2}\sigma^2\Delta t)\frac{\partial f}{\partial S} \} \\ &= \frac{1}{2\Delta t} \{ f(Sdd; t + 2\Delta t) - f(S; t) + 2S(\sigma\sqrt{\Delta t} - \frac{1}{2}\sigma^2\Delta t) \times \\ &\quad (\frac{f(Sud; t + 2\Delta t) - f(Sdd; t + 2\Delta t)}{Sd(u - d)}) \}. \end{aligned}$$

Now, using a first order Taylor series expansion, it follows that

$$\begin{aligned} Sd(u - d) &= Se^{-\sigma\sqrt{\Delta t}}(e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}) \\ &\approx S(1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t)(2\sigma\sqrt{\Delta t}) \\ &\approx S(2\sigma\sqrt{\Delta t} - \sigma^2\Delta t). \end{aligned}$$

Inclusion of the above approximation into (4.24) then results in

$$\alpha = \frac{1}{2\Delta t} \{ f(Sdd; t + 2\Delta t) - f(S; t) + f(Sud; t + 2\Delta t) - f(Sdd; t + 2\Delta t) \} = \frac{\partial f}{\partial t}. \quad (4.25)$$

Back substitution of the above into (4.24) then results in

$$rf(S; t) = \frac{1}{2}S^2\sigma^2\frac{\partial^2 f}{\partial S^2} + Sr\frac{\partial f}{\partial S} + \frac{\partial f}{\partial t}. \quad (4.26)$$

The above proves conclusively that the binomial pricing methodology is akin to solving the Black and Scholes partial differential equation via the explicit finite difference scheme.

4.4 Numerical Integration Schemes

One of the fundamental assumptions made by Black and Scholes in the derivation of their partial differential equation is that of geometric Brownian motion dynamics assumed for the underlying asset's price. This is equivalent to positing generalized Brownian motion for the continuously compounded asset price returns, which in its turn implies that cumulative returns are normally distributed in probability.

One of the most important results in the mathematics of finance is the Feynman-Kac formula. This theorem provides the solution of PDEs of type similar to the Black and Scholes PDE in terms of an expected terminal condition. More specifically it can be given as:

Theorem 10 (*Feynman-Kac Formula*)

Let μ , σ and f be bounded and smooth functions, whilst S denotes the solution to the stochastic differential equation

$$dS(t) = \mu(S, t)dt + \sigma(S, t)dB_t. \quad (4.27)$$

The solution to the Kolmogorov backward equation

$$\frac{\partial f}{\partial t} + \mu(S, t)\frac{\partial f}{\partial S} + \frac{1}{2}\sigma(S, t)^2\frac{\partial^2 f}{\partial S^2} = r(t)f,$$

with terminal condition

$$f(S, T) = \psi(S, T),$$

is given by

$$f(S, t) = E\{\psi(S, T)|S(t)\}.$$

The expectation is under the probability measure implied by (4.27).

Proof: Refer to Goodman et al (2004) [37] for a proof.

Q.E.D

Close inspection of the above shows that the Black and Scholes PDE is nothing short of the Kolmogorov backward equation of a derivative's value function - assuming that the function $\mu(S, t)$ is equal to rS . Furthermore, the geometric Brownian motion dynamics assumed for the value of the underlying's price, firstly conforms with the pre-condition of the formula, whilst implying on the other hand, log-normally distributed prices at expiry (the probability distribution function of which is tractable).

The Feynman-Kac formula therefore allows one to view the value of a derivative security, whose behaviour is described by the Black and Scholes PDE in terms of its discounted expected pay-off at expiry. Drawing upon this result, Merton (1973), [57], provided a closed-form solution to the Black and Scholes partial differential equation - refer to (4.1), whilst also initiating the “martingale” pricing approach to derivatives pricing.

It should follow that the ability to express the arbitrage free value of derivative securities in terms of expected pay-offs, makes the application of integration schemes very relevant. In the case of vanilla European options the indefinite integral is solvable, however this is not necessarily true for all options pay-offs. Definite integral solution (for non-tractable integrals) is a problem long studied; many numerical schemes exist for such solution approximation, including the simple “Rectangular rule”, the “Trapezoid rule”, Simpson’s rules etc. Parkinson (1977), [61], was the first to propose the use of these schemes in the pricing of derivatives; he considered the case of “Barrier options”. Omberg (1987), [64], expanded upon Parkinson’s work. More recently Andricopoulos et al (2003), [1], revisited this idea; they also made mention of numerical integration techniques being equivalent to a flexible multinomial grid. Multinomial grids are nothing else than generalized binomial trees.

This section will attempt to show equivalence between the binomial tree pricing method and numerical integration schemes. (The proof given here was constructed from first principles entirely.)

First recall the Riemann-Stieltjes integral definition:

Definition 3 (Riemann-Stieltjes Integral)

Let $f(x), g(x)$ be two real valued functions defined for $x \in (a, b)$. Then, the integral of $f(x)$ with respect to $g(x)$, over the said interval, is defined as:

$$\int_a^b f(x)dg(x) = \lim_{n \uparrow \infty} \sum_{i=1}^n f(x_i)\{g(x_i) - g(x_{i-1})\},$$

for $x_0 < x_1 < \dots < x_{i-1} < x_i < \dots < x_n$ a partition of $[a, b]$.

Consider once more the value of a vanilla European option as presented in Theorem 6 in the previous section. Focus on the second term of the said presentation, namely:

$$\tau = K \sum_{j=0}^{a-1} \binom{n}{j} (p)^{n-j} (1-p)^j e^{-rt}. \quad (4.28)$$

The above can also be presented as:

$$\tau = Ke^{-rt}\phi(n; p; a - 1), \quad (4.29)$$

where, $\phi(\cdot; \cdot; \cdot)$ represents the cumulative binomial probability mass function, n the number of time steps in the binomial tree scheme, p the risk neutral probability of an up move occurring, and a as defined in Subsection two.

But,

$$\phi(n; p; a - 1) = \sum_{j=0}^{a-1} \binom{n}{j} (p)^{n-j} (1-p)^j, \quad (4.30)$$

and,

$$\binom{n}{j} (p)^{n-j} (1-p)^j = \phi(n; p; j) - \phi(n; p; j - 1). \quad (4.31)$$

From the above it should be clear that:

$$\tau = \sum_{j=0}^{a-1} Ke^{-rt} \{ \phi(n; p; j) - \phi(n; p; j - 1) \}. \quad (4.32)$$

Define

$$\psi(n; p; j) = \begin{cases} \phi(n; p; j), & \text{for } j = 0, 1, \dots, n \\ 0, & \text{else} \end{cases} \quad (4.33)$$

The ψ function just explained is bounded (it always assumes values on the interval $[0, 1]$), regular and “well-behaved” over the interval $[0; n]$, thus, using the Riemann-Stieltjes integral definition, relationship (4.32) could be written as:

$$\tau = \int_0^{a-1} Ke^{-rt} d\psi(j). \quad (4.34)$$

In the previous section it was shown that:

$$\lim_{n \uparrow \infty} \phi(n; p; j) = \lim_{n \uparrow \infty} N\left(\frac{j - np}{\sqrt{np(1-p)}}\right) = N(d_2).$$

Thus, relation (4.34) could be presented as :

$$\tau = \sum_{j=0}^{a-1} Ke^{-rt} \left\{ N\left(\frac{j - np}{\sqrt{np(1-p)}}\right) - N\left(\frac{j - 1 - np}{\sqrt{np(1-p)}}\right) \right\} \quad (4.35)$$

$$= \int_0^{a-1} Ke^{-rt} dN\left(\frac{j - np}{\sqrt{np(1-p)}}\right). \quad (4.36)$$

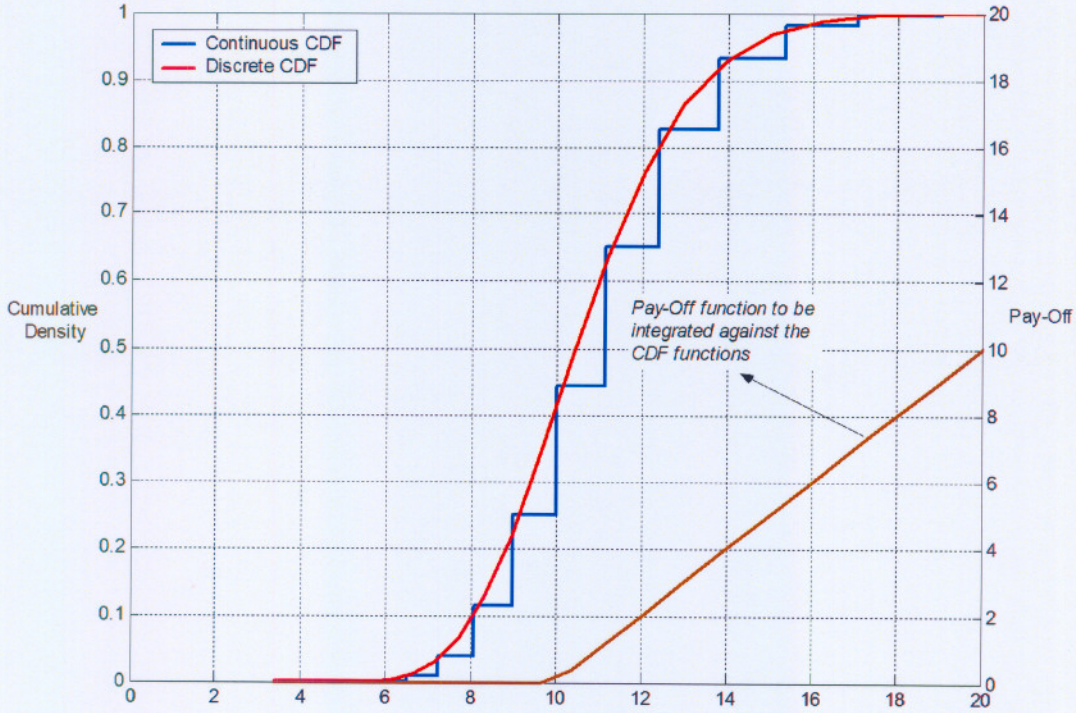


Figure 4.1: Integration of the pay-off function with respect to the continuous or discrete (approximate) CDF.

As $\lim_{n \uparrow \infty} \frac{a - 1 - np}{\sqrt{np(1-p)}} = d_2$ (proven in Subsection Two), it follows that

$$\tau = \int_0^{d_2} K e^{-rt} dN(x), \quad (4.37)$$

where $x = \frac{j - np}{\sqrt{np(1-p)}}$, for $j = 0, 1, 2, \dots, a - 1$.

Thus, calculating the binomial tree implied price of a vanilla European option is akin to solving the integral via a numerical scheme, furthermore, the scheme used is a simple rectangular one. With regards to this, two observations must be made:

- In the above proof one of the fundamental “cross-overs” made is that of approximating $N(\cdot)$ with $\phi(\cdot; \cdot; \cdot)$. There are numerous transformations, or “continuity

corrections" extant with which this could be achieved. Incorporation of such schemes could significantly add to enhancing convergence speed; Leisen and Reimer (1995), [55], for example suggested the Camp (1951), [19] and Paulson (1942) [62], and Peizer-Pratt (1968), [63], inversions with this in mind.

- Improving on the area estimation procedure, i.e., in stead of employing the simple rectangular rule, using a quadrature method, might improve on the efficiency of estimation procedures.

Example:

As an illustration of the equivalence between the binomial methods and numerical integration, consider the pricing of a vanilla European call option, at the money, expiring in one year, and where the risk free rate of interest and price volatility is respectively 10% and 20% ; it is assumed that the spot of the underlying is at 1000. The tree valuation is done according to the Cox et al. parameterization, whilst three different numerical integration methodologies are used; the true, Black and Scholes closed form solution to the valuation is 132.70. (Refer to the Table 4.1 for the results of the exercise.)

Methodology	Steps: 5	Steps: 15	Steps: 25	Steps: 100
Binomial Method	134.5	133.32	133.07	132.46
Rectangular Rule	335.6	125.37	129.45	132.86
Trapezoid Rule	335.16	125.37	129.45	132.86
Simpson's Rule	446.66	113.08	122.78	132.89

Table 4.1: Numerical integration and binomial pricing

The above table might be somewhat misleading as the number of steps represents the true number of discrete evaluation points in the case of the integration scheme, whereas the total number of price nodes evaluated under the binomial tree is much greater than the mere number of time steps. Rather equating the number of discrete calculation points, that is evaluating the integral at 21, 136, 351 and 5151 different price intervals, (corresponding to the total number of binomial discrete point observations at each of the time step values), yields results like those presented in the Table 4.2.

The convergence for binomial tree option values does appear to somewhat lag that of integration methods. Convergence patterns for the said pricing methods can be seen in Figure 4.2 and 4.3.

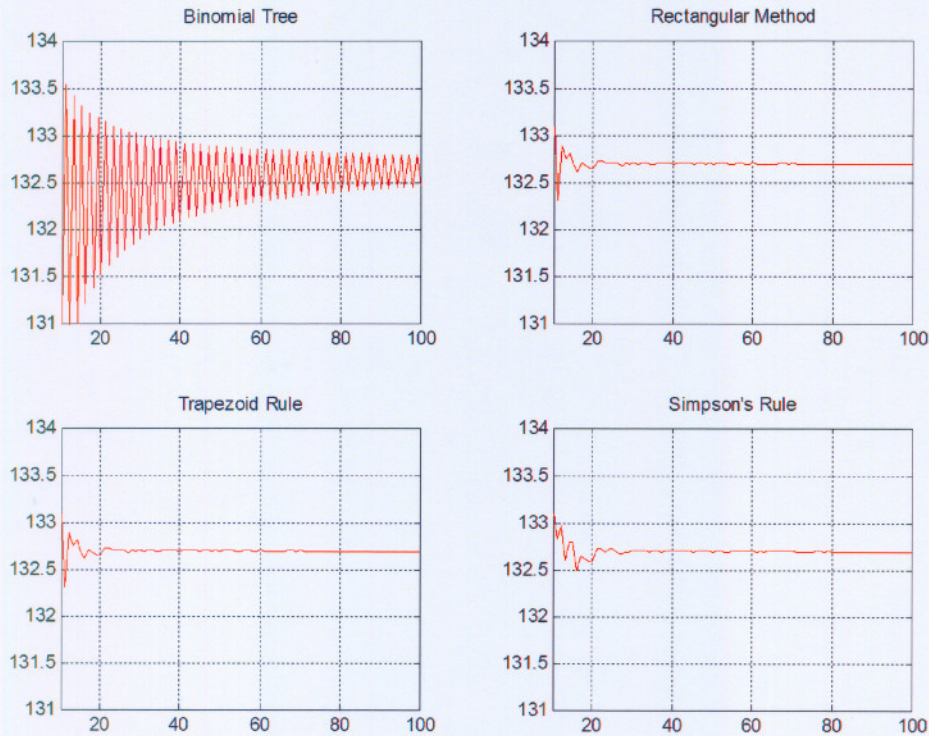


Figure 4.2: Convergence diagram of numerical integration schemes versus binomial methods. (At the money, vanilla European option.

<i>Methodology</i>	Steps: 5	Steps: 15	Steps: 25	Steps: 100
Binomial Method	134.5	133.32	133.07	132.46
Rectangular Rule	128.04	132.69	132.71	132.70
Trapezoid Rule	128.04	132.69	132.71	132.70
Simpson's Rule	139.23	132.8	132.71	132.70

Table 4.2: Numerical integration and binomial pricing ; post a discrete calculation point matching.

4.5 Binomial and Trinomial Trees

Whereas binomial trees assume that prices processes for the underlying asset can be presented by Bernoulli random variables, trinomial trees make the assumption that over small time intervals the price of an asset can assume one of three possible values. This type of tree was presented first by Boyle (1988) "A Lattice Framework

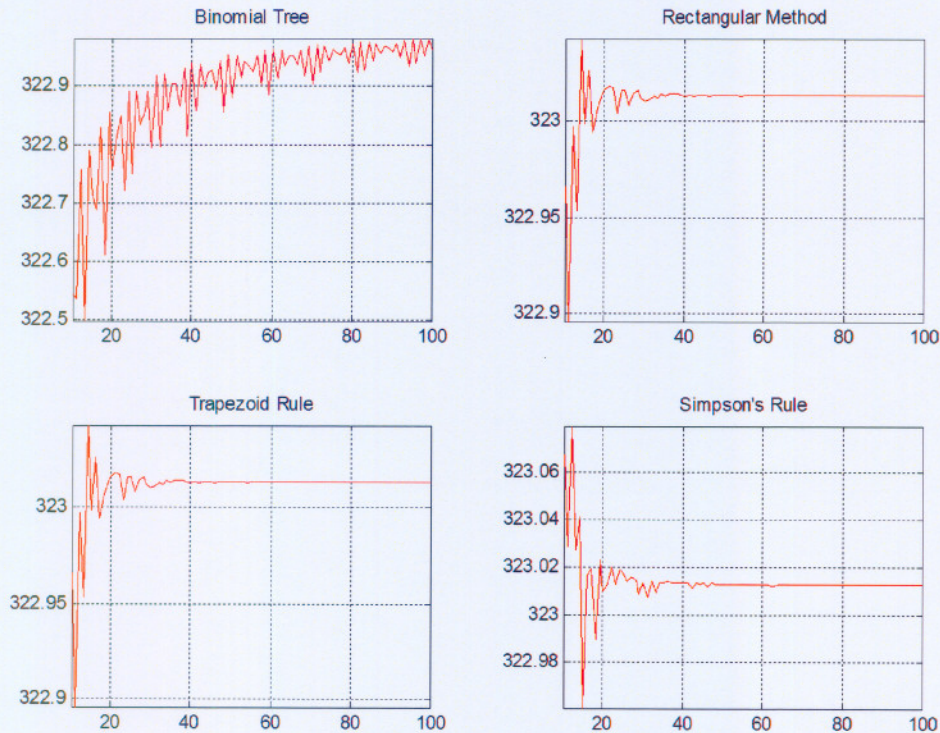


Figure 4.3: Convergence diagram of numerical integration schemes versus binomial methods. (75% vanilla European call option.)

for Option Pricing with Two State Variables, [12]. This approach although having gained wide acceptance, has the drawback that the no-arbitrage argument used by Cox et al (1979), [23], to link the variable outcomes to a fair price for the derivative no longer holds true. The reason for this being, quite simply, that one can not replicate the derivatives value for each of the three possible states using two assets (the underlying itself and a riskless asset) only. To compensate for this deficiency, the trinomial tree approach is generally regarded as an equivalent to the explicit finite difference method, and therefore is not seen as a tree method as such. Mathematically the procedure is normally presented as:

Definition 4 (Trinomial Trees)

Let $S(t)$ denote the price of some asset at time t , whilst that of a derivative on the said (conditional upon $S(t)$) is denoted by $f(t|S(t))$; the trinomial tree assumption is given as:

$$S(t + \Delta t) = \begin{cases} S(t)u, & \text{with probability } p_u \\ S(t)m, & \text{with probability } p_m \\ S(t)d, & \text{with probability } p_d \end{cases} \quad (4.38)$$

whilst the value of the derivative is given as :

$$f(t|S(t)) = \{p_u f(t + \Delta t|S(t)u) + p_m f(t + \Delta t|S(t)m) + p_d f(t + \Delta t|S(t)d)\}e^{-r\Delta t}, \quad (4.39)$$

where u, m, d are the up-, middle- and down factors applied.

The parameters u, m, d with their corresponding probabilities are normally derived by either a second order moment matching along with a stability condition, or in some cases (although rarely) by a third order matching of discrete moments to the continuous time process.

Now, because of the fact the no-arbitrage condition is not central to the derivation of most trinomial tree parameterisations, there has been somewhat of a crisis of identity associated with this methodology. How could one be sure that it is “right”? , what reason could be forwarded for its right to existence ? The answer normally given (and already alluded to), is that it can be shown that this methodology is consistent with the explicit finite difference solution of the Black and Scholes partial differential equation. In other words, it is not strictly speaking a tree based method, but rather something in between. Another issue of concern is that many trinomial parameterizations do not return arbitrage free derivatives prices, whilst numerical instability is also quite often a feature of these models: this being related to risk neutral probabilities becoming negative for some parameter choices (like volatility and the risk free returns).⁴

Rubinstein with “On the Relation Between Binomial and Trinomial Option Pricing Models” (2000), [74], provided a solution to the identity dilemma of trinomial tree procedures. He showed that these (under very loose conditions) are in fact binomial trees but regarded at every second time node only; the figure below further illustrates this:

Mathematically the equivalence can be shown as:

⁴One might question the logic of a trinomial method being unstable whilst the binomial is not ; the reason for this is that some parameterizations calibrate the probability variables independently of a binomial definition, mostly by choosing an arbitrary step size for the price. However, in the event that trinomial trees are calibrated consistent with binomial ones, the instability referred to should cease.

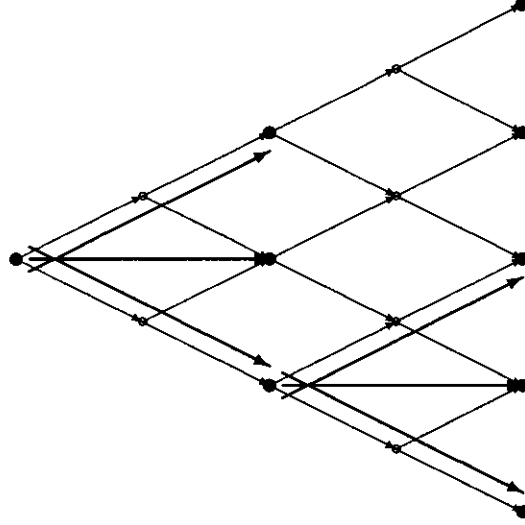


Figure 4.4: The relationship between binomial and trinomial trees; the colored circles and bold vectors represent the trinomial tree overlay.

Theorem 11 *Define a binomial price scheme as follows:*

$$S(t + \Delta t) = \begin{cases} S(t)u, & \text{with probability } p \\ S(t)d, & \text{with probability } (1 - p) \end{cases} \quad (4.40)$$

where $p = \frac{R - d}{u - d}$ so as to be consistent with the no-arbitrage condition. Were one to further define $ru = u^2$, $rm = ud$ and $rd = d^2$, with the risk neutral probabilities set equal to: $p_u = p^2$, $p_m = 2p(1 - p)$ and $p_d = (1 - p)^2$, then the a binomial consistent trinomial tree could be defined as:

$$S(t + 2\Delta t) = \begin{cases} S(t)ru, & \text{with probability } p_u \\ S(t)rm, & \text{with probability } p_m \\ S(t)rd, & \text{with probability } p_d \end{cases} \quad (4.41)$$

The above definition is very powerful in that it compresses all the information contained in a binomial scheme into a scheme comprising half the number of nodes, otherwise said: the trinomial tree will determine derivatives prices with the accuracy given by a binomial lattice but requiring only half the number of time steps. Furthermore, in contrast to many other trinomial parameterizations extant, the above definition inherits the qualities of the binomial one: arbitrage free derivatives prices as well as numerical stability (if it applies). It should be added that Tian (1993), [79] and (1998), [80], makes mention of the improved convergence properties of the trinomial tree, although he never offers an explanation for it.

4.6 Summary

This chapter focused on showing equivalence between the binomial tree method and other pricing approaches. It was proved that binomial tree implied prices for vanilla European prices converge unconditionally towards the closed-form solution given by Black and Scholes (1973), [9]. Following a proof given by Prof. E. Maré, it was furthermore shown that the “Fundamental Result of Binomial Pricing” is akin to the Black and Scholes PDE. It was also shown that solving the latter is equivalent to the explicit finite difference method for solving the said PDE.

It was furthermore shown that the relationship between binomial methods and numerical integration schemes. The chapter was concluded by a discussion of the equivalence between trinomial and binomial trees.

Chapter 5

Real Life Binomial Pricing

Models are only as good as the assumptions on which they are built; the binomial pricing model is no different. In the article first proposing this methodology (Cox et al (1979), [23]), many simplifying (and simplistic assumptions) were made. Examples of the latter include, assuming that the riskless rate of return is a single constant function over time, or that no cost is incurred in the short sale of the underlying asset, or very importantly that volatility is a constant, deterministic function both of time and the level of the current underlying stock price.

In the real financial world, most of these assumptions do not hold true. A model that has difficulty in replicating the readily available prices of derivatives in the market, might not inspire much confidence in its pricing of instruments for which no market exists. Worse, if a model predicts a different price for an asset than can be observed in the market should most certainly not be relied upon to produce accurate hedging parameters. This chapter aims towards a systematic investigation of allowances that could be made in the binomial model to accommodate some market eccentricities.

5.1 Background

In the environment originally posited by Black and Scholes (1973) [9], (and consequently by many others), riskless return rates were taken as constant, the underlying asset for derivatives were assumed to be non-cash flow generating and volatility was taken to be not only deterministic but constant across time and space. As has been mentioned in the abstract, real world finance does not behave within these constraints. Discounting of future cash flows is generally done off a curve of possible rates, volatility realizes not as a single constant value, but as a multi-dimensional stochastic function of time, space etc. Moreover, stock borrow costs, discrete dividends, storage and transport expenses are definite features of financial markets. Pricing models are almost invariably used as replicating and replication cost calculators. (Refer to Leland and Rubinstein (1995), [75], for a good discussion on this concept.) Failure to include potential replication costs could not only be disastrous to profit and loss projections, but might also open such a market participant to exploitation by other players - making nonsense of the fundamental no-arbitrage principle upon which almost all of derivatives pricing is based.

Having established the non-compromisable need to price derivatives in accordance with real world characteristics of financial markets, the question arises how this is done within a binomial pricing framework. This chapter aims towards an investigation of these. The first section deals with the term structure of risk free rates of return; the section following investigates the inclusion of stock borrow-, and other costs of hedging. The penultimate section treats on the accommodation of non-flat volatility term structures and smiles, whilst the chapter is concluded with a view on incorporating discrete cash flows generated by the underlying asset.

5.2 The Term Structure of Riskless Return

One of the key inputs of the binomial tree method is the riskless rate of return, or the return on a portfolio where the profit and loss profile shows no stochastic variability. This market variable is usually taken from market quotes of riskless asset prices, for example bonds issued by the Government or Central Bank, Treasury bills etc. Now, these instruments generally trade at yields varying with time to expiry; the main reason for this being varying inflation expectation over different time horizons. This is referred to as the term structure of interest rates (riskless returns).

The binomial model, it has been said, rests upon the an immunization argument

for a portfolio's return against stochastic changes in the underlying assets price. The absence of any stochastic component to the value of a holding, should see it (under the no-arbitrage condition), yield the riskless asset's return. Given a time-varying return on this asset type, it should follow that the yield over consecutive time intervals will not only differ, but also, given the forward value of this return, an estimation of this would be necessary. However, the no-arbitrage condition requires that rolling an investment in this asset constantly, should equal a fixed investment for the whole time horizon - allowing for the concept of forward interest rates.

Definition 5 Let r denote the instantaneous (continuously compounded) rate of riskless return. The return over an interval $[t_1; t_2]$ be written as

$$r(t_1; t_2) = \int_{t_1}^{t_2} r(s) ds.$$

This is referred to as the forward rate of return. In terms of non-instantaneous (yet continuously compounded) forward rates, the arbitrage-free, time $t = 0$, forward rate is:

$$r(t_1; t_2) = \frac{t_2 r(0; t_2) - t_1 r(0; t_1)}{t_2 - t_1}. \quad (5.1)$$

Given the above concept, making allowance for riskless return term structures is a matter of substituting the forward return over the interval $[t; t + \Delta t]$ for the flat rate normally used to calculate the "up" and "down" factors, the risk neutral probability as well as the discounting rate. To further illustrate the point, consider the Trigeorgis parameterization, and the "Fundamental Result of Binomial Pricing" in the presence of riskless return term structures:

$$f(S; t)R = pf(Su; t + \Delta t) + (1 - p)f(Sd; t + \Delta t), \quad (5.2)$$

but where we define:

$$\begin{aligned} R &= R(t; t + \Delta t) = e^{-r(t; t + \Delta t)\Delta t} \\ u &= u(t; t + \Delta t) = e^{(r(t; t + \Delta t) - \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}} \\ d &= d(t; t + \Delta t) = e^{(r(t; t + \Delta t) + \frac{1}{2}\sigma^2)\Delta t + \sigma\sqrt{\Delta t}} \\ p &= \frac{R(t; t + \Delta t) - d(t; t + \Delta t)}{u(t; t + \Delta t) - d(t; t + \Delta t)}. \end{aligned}$$

The effect of an upward sloping term structure of interest rates on the shape of a binomial tree is to introduce a marked up-trend in the tree, with the opposite holding true for downward sloping structures.

5.3 Stock Borrow Costs

The idea on which binomial pricing is based, entails the replication of a derivative's pay-off using the underlying and the riskless asset. Costs involved in short selling the asset, or potential income in lending out the same, whilst long, should therefore be taken into consideration in calculating the cost of replicating such an instrument. Securities lending is a definite feature of financial markets; lending agreements are in essence nothing else than collateralised borrowing. Ignoring the collateral required from the borrower, the costs involved in such transactions is calculated as a lease rate based on the market value of the underlying asset; this implies that costs increase or decrease in conjunction with the value of the asset. The question to be considered here is how allowance should be made within the binomial pricing methodology for stock borrow costs.

The logic of the replication paradigm assumes positions in the underlying and riskless assets of say Δ and B - refer to the "Fundamental Result of Binomial Pricing" in the Chapter Two. Given the two-outcome nature of the underlying's price behavior (over some arbitrary time interval $[t; t + \Delta t]$), it follows that the value of the replicating portfolio at the end of the interval could be calculated as either : $\Delta Su + BR$ or $\Delta Sd + BR$; taking stock borrow costs payable or receivable on the stock position to be relative to the price, and charged at say q , the replicating portfolio's value at the time interval end could be rewritten as: $\Delta Su + BR + \Delta Sq$ or $\Delta Sd + BR + \Delta Sq$. The replicating nature of this portfolio further allows one to state:

$$\Delta Su + BR + \Delta Sq = f(Su, t + \Delta t) = f_1, \quad (5.3)$$

$$\Delta Sd + BR + \Delta Sq = f(Sd, t + \Delta t) = f_2. \quad (5.4)$$

With $f(;\cdot)$ the value of the derivative being replicated. Cross-substitution of the above then yields:

$$\Delta S\{d + q\} + f_1 - \Delta S\{u + q\} = f_2, \quad (5.5)$$

$$\Delta S\{d + q - u - q\} = f_2, \quad (5.6)$$

$$\Delta = \frac{f_2 - f_1}{(d - u)S}. \quad (5.7)$$

The above result shows that the inclusion of stock borrow costs does not yield a different hedge ratio - as is expected. Back-substitution of (5.7) into (5.3) then yields:

$$BR = f_1 - S\{u + q\} \frac{f_2 - f_1}{(d - u)S} \quad (5.8)$$

$$= \frac{f_1(u - d) - (u + q)(f_1 - f_2)}{u - d} \quad (5.9)$$

$$= \frac{f_2 u - f_1 d}{u - d} + \frac{f_2 - f_1}{u - d} q. \quad (5.10)$$

Having derived the positions required to replicate the derivatives value, what remains to be seen is what the impact is on an inclusion of stock borrow costs on the risk neutral probability measure. Consider the value of the derivative at time t :

$$f = \Delta S + B \quad (5.11)$$

$$= \frac{f_2 - f_1}{(d - u)S} S + \frac{f_2 u - f_1 d}{u - d} \frac{1}{R} + \frac{f_2 - f_1}{u + d} \frac{q}{R} \quad (5.12)$$

$$= \frac{f_1 - f_2}{(u - d)} + \frac{f_2 u - f_1 d + (f_2 - f_1) q}{u - d} \frac{1}{R} \quad (5.13)$$

$$= \frac{1}{R(u - d)} \{R(f_1 - f_2) + f_2 u - f_1 d + f_2 q - f_1 q\} \quad (5.14)$$

$$= \frac{1}{R} \left\{ \frac{u - R - q}{u - d} f_2 + \frac{R - q - d}{u - d} f_1 \right\}. \quad (5.15)$$

From the above it should be clear that the inclusion of stock borrow costs results in a risk neutral probability measure changed to $p = \frac{(R - Q) - d}{u - d}$, with $Q = e^{q\Delta t}$.

The reader will recall that the "up" and "down" factors of the binomial model is set in accordance with the empirical (continuous time) process assumed for the price behavior of the underlying. It should also be noted that the inclusion of stock borrow costs calculated relative to the price of the asset changes the drift parameter in the

continuous process from $r dt$ to $(r - q)dt$. The derivation of “up” and “down” factors changes accordingly, for example the Trigeorgis tree parameterization becomes:

$$u = e^{(r-q-\frac{1}{2}\sigma^2)+\sigma\sqrt{\Delta t}} \quad (5.16)$$

$$d = e^{(r-q-\frac{1}{2}\sigma^2)-\sigma\sqrt{\Delta t}}. \quad (5.17)$$

As a final remark: stock borrow cost is equivalent to a constant dividend yield earned on an index, or to the foreign interest earned on the underlying of a currency option. Pricing currency options, or options on indices are therefore facilitated by this.

5.4 Volatility Surfaces and Implied Trees

One of the fundamental assumptions of the Black and Scholes pricing framework is that of volatility being a single constant value - independent of both time and the level of the underlying asset's price. Since the 1987 Equity market crash, however, volatility as implied by actively traded options prices have generally realized as a complex three dimensional function of both spot and the time to expiry. (Refer to Figure 5.1 for a graphical view of the volatilities as implied by options on the All Share Index 40.) This function is generally referred to as the volatility skew and smile; the implication of this being that different price trees have to be constructed for options with strike price and expiry date differing. Not only is this in direct violation of to the Black and Scholes assumption, it also appears to be counter intuitive: one stock price process should suffice to price a whole range of different options.

There are numerous explanations for the existence of the volatility smile or skew, ranging from discontinuities in the stock price process to stochastic volatility etc. Using the implied volatility surface to run a book consisting of vanilla European options is not problematic, however, the pricing of derivatives with variable or multiple strikes and expiry dates do give rise to some problems. As an illustration of this consider for example the case of an average strike option: is one to adjust the volatility over the averaging time due to the strike constantly changing ? If so, how is one to treat average price options, as these are related to average strike options by a simple parity type relation. That is, if one changes the volatility as new observations are made on the strike, is one to also change the volatility on average price options, even though the strike is constant; failure to do so would lead to arbitrage possibilities between the said contracts.

Another example of the confusion this could potentially cause, would be that of swim-

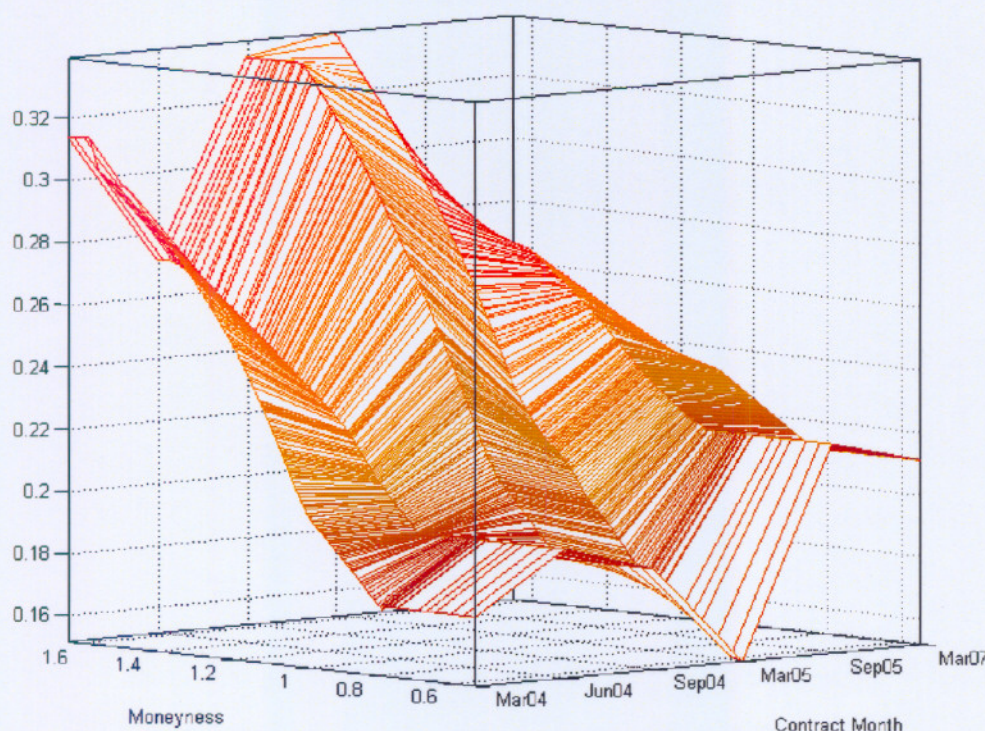


Figure 5.1: All Share 40 Implied Volatilities (January 2004)

ming¹ versus fixed volatility regimes. Otherwise said: should volatility be determined by an option's moneyness or its strike price? (An article that is considered definitive in terms of expositing the mentioned volatility regimes is Derman (1999), [25].) In the case of the former, a change in the price of the underlying would clearly cause a shift in the volatility applied. The choice of regime is not a trivial matter. Non-trivial neither in theoretical consistency, nor in potential financial impact. Gatheral (1999), "Hedging and volatility regimes", [36], showed that either of the mentioned regimes could prove itself optimal (or sub-optimal depending on the view) under certain market conditions, for example, a swimming regime tends to perform better in a whipsaw² environment, whereas fixed regimes are better suited to trending markets.

Given the ambiguity in terms of volatility choice that is a consequence of the implied

¹Swimming volatility regimes refer to sticky-delta volatility surfaces, whereas fixed ones are sticky-strike. As the price of the underlying security changes (and therefore the delta of the option), one would re-read the volatility off the surface in a sticky-delta regime.

²Whipsaw refers to range-bound, highly volatile behaviour in the price of a security.

volatility skew, researchers have attempted to provide a unified (that incorporates the skew) behavioural dynamic for the underlying asset's price, and thus correctly prices all observed options. To illustrate the just mentioned in terms of binomial trees: instead of having to build different price trees for options with different strike prices, the idea is to have one tree that caters for all possible options, as well as one that correctly prices all observed options for which there is an actively traded market.

The number of publications that have appeared and specifically deals with calibrating options pricing models to the observed skew is substantial. Two very important results that authors almost invariably refer to would be that of Ross (1976), "Options and Efficiency", [69], and that of Breeden and Litzenberger (1978), "Prices of State Contingent Claims Implicit in Options Prices", [13]. The Ross result, was the first to show the possibility of deriving a risk neutral probability distribution from a series of options prices. Breeden and Litzenberger, in turn showed the uniqueness of the implied distribution; they also showed that the distribution is derived by using the fact that an option's price, twice derived with respect to the strike is equal to the risk neutral probability of the underlying's price ending up to be equal to the strike. That is:

$$\frac{\partial^2 f}{\partial K^2} = Pr.\{S(T) = K|S(0)\}.$$

The mentioned process is however, dependent on fitting a proper function on observed options prices. Along a similar vein, Rubinstein (1998), [72], proposed constructing binomial trees consistent with an adjusted log-normal distribution for the underlying, and well where adjustment is done by an application of Edgeworth expansions (a probabilistic equivalent of Taylor's theorem). Cakici and Foster (2002), [17], on the other hand, proposed calibrating binomial price trees to historical distribution; this was done as an attempt at explaining the volatility skew.

Attempts at building binomial or trinomial tree processes (fully calibrated to observed options prices) abound, however the three that stand out as seminal are: Rubinstein (1994), [73], Derman and Kani (1996), [27], and Dupire (1994), [32]. The first two of these will be briefly discussed below.

5.4.1 Rubinstein Implied Trees

Rubinstein, in "Implied Binomial Trees", published 1994, [73], attempted to construct binomial trees calibrated to observed options price, by assuming that transition probabilities are functions both of time and the price level of the underlying. His method whilst yielding relatively stable results (that are always consistent with

the no-arbitrage principle, though in frequent violation of the Kolmogorov probability postulates), has the draw-back that only options with similar expiry dates are accommodated by the algorithm.

The process consists of first constructing a standard binomial price tree. From the prices of observed option price a state price vector is derived, the state prices corresponding to the probability of the asset price ending at the respective terminal node points. The state price vector, obviously being identical to the terminal price distribution. Now, from the terminal state price vector, through backward induction, the transition probabilities are derived, resulting in a full set of probability matrices. The transition probabilities are easily converted into local volatility values. Mathematically the algorithm can be presented as follows:

Denote by $S_j(t_i)$ the j^{th} price node value on the tree at time t_i , where $j = 0, 1, \dots, n$ and with $t_0, t_2, \dots, t_n = T$ a partitioning of the time interval $[0, T]$. Furthermore let $v_j^k(t_i)$ be the value of the k^{th} option (of $n - 1$ for which values are observed in the market) on the asset S , at time t_i and conditional upon $S_j(t_i)$; it is worthwhile to add that $v_0^k(t_0)$ represents the market observed value of the said option at the evaluation time.

Also define $\Psi_j(t_i)$ as the probability of ending at the j^{th} price node at time t_i (denoted by $S_{(t_i, j)}$), that is :

$$\Psi_j(t_i) = Pr.\{S(t_i) = S_j(t_i)\}. \quad (5.18)$$

Lastly, let R be the future value, at time t_i of 1 invested in the riskless asset at time t_0 .

Using the fact that the fair value of a derivative is equal to its discounted expected pay-off at expiry, (under the risk neutral probability distribution), it follows that :

$$v_0^k(t_0) = R(t_n) \sum_{j=0}^n v_j(t_n) \Psi_j(t_n) \text{ for } k = 0, 1, 2, \dots, n - 1. \quad (5.19)$$

The above represents a system of $n - 1$ equations in $n + 1$ variables; a unique solution could only be derived in the event that two additional conditions are added. Now, for the terminal probability distribution to comply with the second of the Kolmogorov probability postulates³, it is necessary to impose the condition :

$$\sum_{j=0}^n \Psi_j(t_i) = 1 \text{ for all } i = 0, 1, 2, \dots, n. \quad (5.20)$$

³Kolmogorov's postulates:

- a. The probability of an event occurring is always greater than zero.
- b. The sum of the probability of events occurring that partitions the whole sample space is one.

For the no-arbitrage condition to hold, the implied probabilities have to correctly price the underlying variable, i.e. the discounted expected value of the underlying has to be equal to the spot price observed; algebraically, this condition could be presented as :

$$R(t_n) \sum_{j=0}^n \Psi_j(t_i) S_j(t_i) = S_0(t_0) \text{ for all } i = 0, 1, 2, \dots, n. \quad (5.21)$$

Incorporating the two conditions just given into the already constructed equations, yields an $(n+1) \times (n+1)$ linear system, the solution of which yields the implied risk neutral probability distribution.

A binomial tree consists properly of price node vectors for each time node, and the probability of transition between consecutive nodes. Thus, in order to arrive at a proper implied tree structure, once the implied terminal probability distribution is derived, the transition probability vectors have to be induced. The way this induction is done, comprises the main differentiation between Rubinstein's and other implied tree methodologies. Rubinstein, makes the fundamental assumption that all price paths ending in the same terminal node, carries equal probability weight. Seeing that the trees considered have recombining nodes, the aforementioned assumption could also be explained as: the probability of an "up" move followed by a "down" move is assumed equivalent to a "down" move, followed by an "up".

Once the terminal probability distribution is derived, the cumulative distribution, just prior to the terminal is induced, after which the second to last distribution is calculated, and so forth. This results in not only collection of cumulative probability distributions for each of the time nodes, but also, by virtue of the induction method followed, the transition probabilities at each point. Mathematically, this algorithm could be presented in the following format:

Using the same notational arrangement made prior, but adding $\psi_j(t_i)$, denoting the conditional probability of transition from the $j-1^{th}$ node at time t_{i-1} to the j^{th} node at time t_i , that is to say :

$$\psi_j(t_i) = Pr.\{S(t_i) = S_j(t_i) | S(t_{i-1}) = S_{j-1}(t_{i-1})\}. \quad (5.22)$$

-
- c. The probability of the union of disjoint events is equal to the sum of the individual probabilities.

(Refer to Bain and Engelhardt (1992), [3] for a more detailed treatment on this.)

The assumptions and notational arrangements made, viewed in the context of the law of total probability ⁴, yields the following system of equations:

$$\Psi_i(t_i) = \Psi_{i-1}(t_{i-1})\psi_i(t_i) \quad (5.23)$$

$$\frac{1}{2}\Psi_{i-1}(t_i) = \Psi_{i-1}(t_{i-1})(1 - \psi_i(t_i)) \quad (5.24)$$

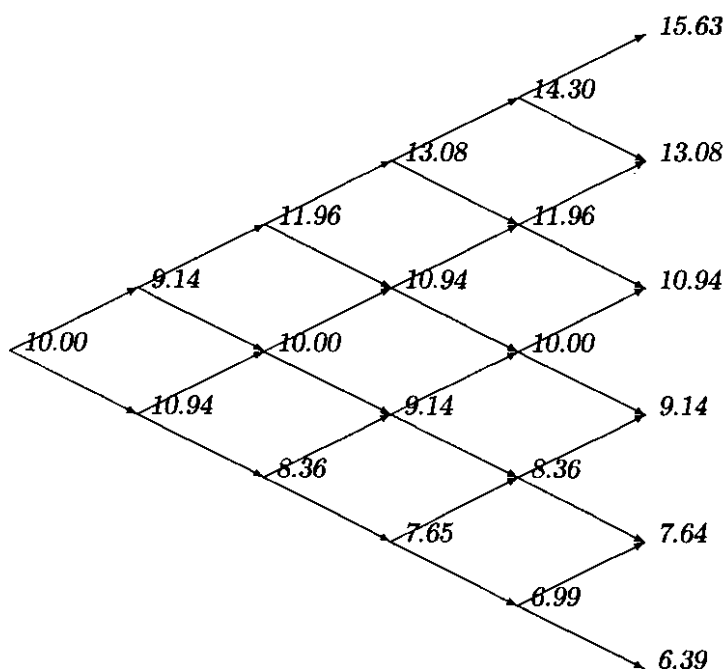
$$\frac{1}{2}\Psi_j(t_i) = \Psi_{j-1}(t_{i-1})(1 - \psi_j(t_i)) \text{ for all } 2 < j < i - 1. \quad (5.25)$$

Applying the above consecutively at each price and time node, yields the complete set of transition probability vectors, and thus also completes the implied tree. The Rubinstein implied tree is further illustrated by way of example:

Example:

Consider a market pricing in a skew in volatility, and such that one year options, with moneyness of 111%, 100%, 90% and 83% are priced at volatilities of respectively 24%, 23%, 22% and 22%.

Applying the at-the-money volatility of 23%, as well as the observed risk free rate at 10%, and spot being 10, yields a Cox et al price tree of :



⁴The probability of an event occurring is equal to the sum of the probabilities of that event and others that partition it; mathematically it could be written as: $Pr.\{A\} = \sum Pr.\{A \cap B_i\} = \sum Pr.\{A|B_i\}.Pr.\{B_i\}$

The Black and Scholes option values for two calls and puts, struck at the said money levels, and with spot at 10 and the risk free rate at 10%, along with the additional conditions placed by Rubinstein's implied tree algorithm then gives rise to the following system of equations:

$$\begin{array}{rcccccccl}
 1.00p_1 & +1.00p_2 & +1.00p_3 & +1.00p_4 & +1.00p_5 & +1.00p_6 & = & 1.00 \\
 15.64p_1 & +13.087p_2 & +10.945p_3 & +9.14p_4 & +7.65p_5 & +6.39p_6 & = & 11.05 \\
 6.64p_1 & +4.08p_2 & +1.94p_3 & +0.14p_4 & +0.00p_5 & +0.00p_6 & = & 2.09 \\
 5.64p_1 & +3.08p_2 & +0.94p_3 & +0.00p_4 & +0.00p_5 & +0.00p_6 & = & 1.43 \\
 0.00p_1 & +0.00p_2 & +0.06p_3 & +1.865p_4 & +3.35p_5 & +4.615p_6 & = & 0.85 \\
 0.00p_1 & +0.00p_2 & +1.06p_3 & +2.865p_4 & +4.35p_5 & +5.615p_6 & = & 1.39
 \end{array}$$

The least squares solution to the above is as follows:

$$\begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{pmatrix} = \begin{pmatrix} -0.024 \\ 0.489 \\ 0.066 \\ 0.917 \\ -0.961 \\ 0.515 \end{pmatrix}$$

The solution presented illustrates an unfortunate drawback of almost all implied trees (as prior alluded to), namely, implied probabilities are frequently in flagrant violation of the probability postulates; in the case where option values are close to the edges of a tree, the matrix implied by the conditions, becomes very close to singular, yielding unstable and non-sensical results. To rectify this, many solutions or adjustments have been proposed, for example applying a constrained optimization to solve for the implied probabilities.

5.4.2 Derman and Kani Trees

At round the same time Rubinstein published his article, Derman and Kani (1994), in "Riding on a Smile", [30] and also Derman and Kani ("The Local Volatility Surface", [29], proposed the local volatility surface model. The essence of their approach is the construction of a deterministic volatility function of time and current asset price. This is done by the calibration of a already constructed binomial tree to market observed option prices, and well through the risk neutral transition probabilities. Theirs was very much the same idea as Rubinstein's, however with the important difference

that it requires a full set of observed prices for each time step considered. Moreover, the Derman and Kani model is a forward induction model⁵, where the focus is very much on the transition probabilities, unlike that of Rubinstein, where cumulative probability distributions are heavily drawn on.

A general outline of the Derman and Kani approach is as follows: similar to the Rubinstein model, a standard binomial price tree is first constructed. Using the market observed price of an option, struck at the initial spot level, and expiry coinciding with the first time node, the first period transition probabilities are implied; implication of the said probabilities is done by way of equating the observed fair value of the derivative to its discounted expected pay-off at expiry. Mathematically, (using the notational arrangement made in the previous section), the said could be presented as:

$$v_0^0(t_0)R(t_1) = v_1^0(t_1)\psi_1(t_1) + v_1^0(t_1)\psi_1(t_1). \quad (5.26)$$

Once the first time period has been solved for, the transition probabilities for the second period are derived using the market observed prices ($v_0^1(t_0)$ and $v_0^2(t_0)$) for two options struck respectively at $S_0(t_1)$ and $S_1(t_1)$, and expiring at t_2 .

In order to ensure that the second period implied distribution is consistent with the first, the concept of conditional expectation is introduced.⁶ In this case, conditioning is done on the events $S(t_1) = S_0(t_1)$ and $S(t_1) = S_1(t_1)$. Denoting by $v_1^2(t_1)|S_1(t_1)$ the value of the first option at time t_1 given that the spot level is $S_1(t_1)$, the following two relationships could be written down:

$$v_0^2(t_0)R(t_2) = \psi_1(t_1)v_1^2(t_1)|S_1(t_1) + (1 - \psi_1(t_1))v_0^2(t_1)|S_0(t_1), \quad (5.27)$$

$$v_0^1(t_0)R(t_2) = \psi_1(t_1)v_1^2(t_1)|S_1(t_1) + (1 - \psi_1(t_1))v_0^2(t_1)|S_0(t_1). \quad (5.28)$$

Now, the value of either of the derivatives, (conditioned upon some spot level), in virtue of the expiry on time t_2 , could further be expanded as follows:

$$\{v_1^2(t_1)|S_1(t_1)\}R(t_2) = \psi_2(t_2)v_2^2(t_2) + (1 - \psi_2(t_2))v_1^2(t_2). \quad (5.29)$$

⁵Forward- as against backward induction models are models where temporally successive probabilities depend on prior, as against posterior derived values. Very simply put: backward induction models count backwards, whilst forward ones count in normal fashion.

⁶Symbolically yhat is: $E\{X\} = E\{X|A\}Pr.\{A\} + E\{X|\bar{A}\}Pr.\{\bar{A}\}$

Substitution of the above, into (5.28), yields an equation in $\psi_2(t_2)$ that is easily solved. Applying the same logic to the other option would result in a complete transition probability matrix for the second time step. Iterative application of the above algorithm, where, in each case, the number of market quotes used, is equal to the cardinality of the time step, and where options are chosen such that expiry coincides with the end of the time step considered, should result in a complete calibration of the binomial price tree.

It is worthwhile to note the absence of the no-arbitrage condition in this recursion; the reason for this being, the lack of an additional degree of freedom. This causes the procedure, in contrast to that of Rubinstein's, to be not necessarily consistent with the no-arbitrage principle, often resulting in numerical instability. A number of improvements have been suggested by other authors, to rectify the aforementioned, notable amongst these are Derman, Kani and Chriss (1994), in "Implied Trinomial Trees of the Volatility Smile", [27], where the trinomial tree structure proposed presents an additional degree of freedom which is used to ensure no-arbitrage compliance. Barle and Cakici (1994), "How to Grow a Smiling Tree", [10], whilst maintaining the binomial tree structure, suggested the alignment of the center node with the forward value, whilst options used are set to the forward moneyness. It is interesting to note that Dupire (1994), with "Pricing with a Smile", [32], suggested a trinomial tree calibration close to identical to that of Derman and Kani, but with a trinomial dynamic assumed. As a final remark: this methodology, whilst eliciting much criticism, might still be regarded as a standard for approaching the smile problem.

5.5 The Indiscretion of Discrete Dividends

The payment of discrete dividends by stocks is a fact of equity markets that cannot be ignored. The argument put forward by Black and Scholes in construction of their partial differential equation, however, explicitly assumes this point away. Jarrow and Rudd (1983), "Option Pricing", [49], refers to this as a weakness of the Black and Scholes model. Pricing derivatives without taking cognizance of potential dividend payments opens participants in such instruments to exploitation by the market, especially in the case where options are replicated: hedging costs might be significantly under estimated.

Many authors and almost all market participants have realized the untenability of the no-dividend assumption. In consequence, various accommodations have been suggested to remedy this shortcoming. This section aims to provide a brief outline of some of the suggested solutions, with especial reference to the impact on the binomial

lattice approach in pricing. In the literature, distinction is generally drawn between two types of dividends, namely proportional and fixed amount. The first meaning that the magnitude of dividends paid out is proportional to the price of the share on the ex-date. Incorporating this type of dividend payment into derivatives pricing is not exceedingly difficult; Cox et al (1979), [23], treated on this as well. Assuming continuous payments of this kind, sees proportional dividends being equivalent to a stock borrow cost, and could thus be treated similarly.

Assuming dividends to be paid as a proportion of the share price is not very realistic; a share's price could vary greatly from the time on which dividends have been declared until the last date to trade, without in any way affecting the magnitude of the dividend. Such dividends, where magnitude is assumed to be independent of the share price, being more realistic, is however less straight forward in their incorporation into derivatives pricing models. A further complication being that both the magnitude and payment date of dividends, being driven by a company's financial results and dividend policy are technically stochastic. Generally speaking, however, due to excellent companies research being available to writers of derivatives, dividend estimates, at least over the short to medium term tend to be very accurate. For this reason, in this section only ex-ante known (or estimated) dividends are to be considered.

Since the first publication of Black and Scholes' article, debate has been going on about the impact on a share's price dynamics of ex-ante known dividends, there being various valid ways in which this impact could be modeled. The issue of exactly how it is done, is furthermore, not a trivial one, as different approaches in some cases yield very different derivatives prices. Frishling (2002), "A Discrete Question", [72], provided a concise characterization of current approaches, namely the "Displaced Diffusion", "Accumulation" and "Deterministic Jump" models. The enumerated models are to be treated on briefly, and well with respect to their history, assumptions and lastly how they could be incorporated into the binomial lattice pricing approach.

5.5.1 The Displaced Diffusion Model

Roll (1977), in "An Analytical Valuation Formula for Unprotected American Call Options on Stocks with Known Dividends", [72], was the first to characterize dividend bearing stocks' prices into riskless and risky components. As such, he argued that only the risky component need be modeled.

The assumption fundamentally underlying the "Displaced Diffusion" model, drawing upon Roll's suggestion is that the stock price net of the present value of dividends

receivable over the life of the derivative is assumed to follow the geometric Brownian motion. This model is widely used today by market participants, in commercial and risk management systems as well as put forward in textbooks, notably that of Hull (1993), [47]. From a pricing point of view, this translates into simply adjusting the spot price by netting off the present value of dividends receivable over the derivative's life.

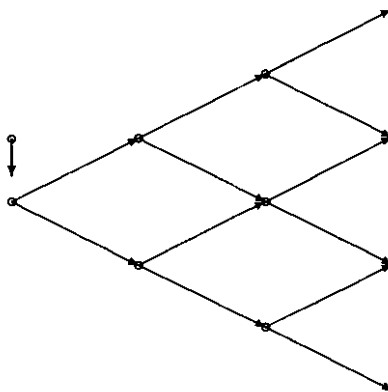


Figure 5.2: Tree dynamics implied by the “Displaced Diffusion” model

Mathematically this dynamic could be presented as:

$$d(S - \xi) = r(S - \xi)dt + \sigma(S - \xi)dB_t, \quad (5.30)$$

with ξ representing the present value (discounted using the risk free rate) of dividends receivable on the underlying over the tenor of the derivative. A passing comment on the nomenclature “Displaced Diffusion Model”: this model retains all the dynamics of the diffusion process proposed by Black and Scholes, except the point of departure for price paths are adjusted downwards by the present value of dividends; the probability density function of the price returns would therefore have the exact same moments except for the first - the same skewness, kurtosis dispersion etc. thus the “Displaced”; refer to Figure 5.2 for a graphical view of the tree dynamics implied by this model.

The “Displaced Diffusion” model is incorporated into a binomial lattice by simply

changing the “root” spot price of the tree, by deducting the present value of dividends receivable.

5.5.2 The Accumulation Model

Instead of regarding the stock price as comprising two different assets, one of which is risky, Heath and Jarrow (1988), proposed in their paper “Ex-dividend stock price behavior and arbitrage opportunities”, [41], that the stock price in its entirety be treated as risky. This effectively implies that derivatives have as underlying the stock plus all dividends received over its life.

Symbolically, the price dynamics implied by this approach is as follows:

$$d(S + \xi) = r(S + \xi)dt + \sigma(S + \xi)dB_t, \quad (5.31)$$

where ξ represents the future value of all dividends paid over the life of the derivative, with interest calculated using the risk free rate of interest. This model is easily incorporated into the Black and Scholes pricing approach by adjusting the strike price upwards by the future value (at expiry) of dividends received over the derivative’s life. That said, it should follow that no structural changes to the binomial tree method are required to make allowance for this altered price dynamics.

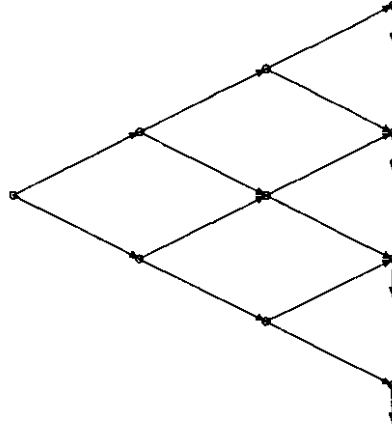


Figure 5.3: Tree dynamics implied by the “Accumulation” model

The tree dynamics consequent to the “Accumulation” model’s assumptions are illustrated in Figure 5.3.

5.5.3 The Deterministic Jump Model

The last approach to be discussed here is that posited by Wilmott (1991), [85]. The view put forward by him is that discrete dividends are in essence jumps in the price process, of which the magnitude and timing is known. As such, the dynamics of the stock price process is fundamentally altered, mathematically it could be presented as:

$$dS = rSdt + \sigma dB_t - \xi \delta(t - t_0), \quad (5.32)$$

where it is assumed that a dividend of size ξ is receivable on time t_0 , and $\delta(t - t_0)$ being the Dirac delta function, i.e.

$$\delta(t - t_0) = \lim_{h \rightarrow 0} \begin{cases} \frac{1}{h}, & \frac{1}{h} < t - t_0 < \frac{1}{h} \\ 0, & \text{else} \end{cases}$$

The Dirac delta function is sometimes viewed as the derivative of the Heaviside function:

$$H(t - t_0) = \begin{cases} 1, & 0 < t - t_0 \\ 0, & \text{else} \end{cases}$$

(Refer to Beichelt and Fatti (2002), [7] for a more detailed treatment on Dirac delta functions and their application to stochastic processes.)

The price dynamics of the asset can then be written as:

$$S(t) = \int_0^t rS(u)du + \int_0^t \sigma S(u)dB_u - \xi H(t - t_0). \quad (5.33)$$

The above clearly illustrates the “jump” effect of the dividend on the price process at time t_0 . Graphically, and well in terms of a binomial price tree, the above dynamics can be presented as in Figure 5.4.

The inclusion of the Dirac delta causes the probability distribution of the price at expiry to be non-tractable, implying that derivatives written on this asset could only be priced using numerical methods. However, the following process has been suggested as an approximation to that given in (5.33) :

$$d(S - \xi) = r(S - \xi)dt + \varsigma d(S - \xi)dB_t \quad (5.34)$$

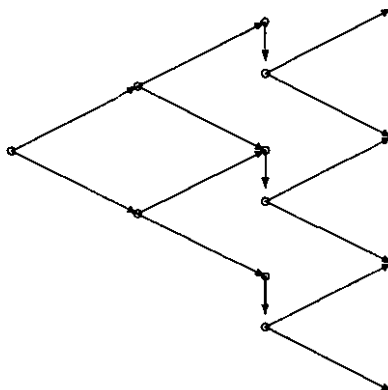


Figure 5.4: Tree dynamics implied by the “Deterministic Jump” model

where ς is chosen such that the first and second order moments of the approximation and empirical process match. More intuitively, this could be seen as compressing information on the dividends into the volatility parameter; the process generally resulting in a higher volatility.

Setting up a binomial lattice consistent with the assumed price dynamics has the drawback that nodes following a dividend payment ceases to recombine, resulting in a substantial increase in the number of nodes to be considered. However, it is fairly simple to build a tree that is consistent with the approximation process given in (5.34).

Example:

To further illustrate the enumerated approaches to incorporating discrete dividends into binomial tree pricing, consider the pricing of a one-year, at-the-money vanilla European call option. Assume an underlying of spot 10 and a volatility of 20%, but with a dividend of 1 receivable in eight month's time. Further assume a risk free return of 10%.

The price dynamics for the “Displaced Diffusion” model is easily constructed, by applying the “up” and “down” factors of 1.12 and 0.89 not to the spot but to spot net of the present value of dividends receivable, being $10 - 0.93 = 9.06$. For the “Accumulation” model, a price tree is constructed as per normal, but with the terminal

nodes adjusted downward by the dollar value of the dividend. The “Deterministic Jump” model’s dynamics can be seen in Table 5.1.

t_0	t_1	t_2	t_2 (Ex Div)	t_3
10.00	11.22	12.60	11.60	13.02
				10.33
	8.91	10.00	9.00	10.10
				8.02
		7.94	6.94	7.79
				6.18

Table 5.1: Price tree under the “Deterministic Jump” model

Traversing the respective price trees and applying the “Fundamental Result of Binomial Pricing” yields the fair price of the call options as shown in Table 5.2.

	Displaced Diff.	Accumulation	Discrete Jumps
Fair Price	0.67	0.72	0.73

Table 5.2: Option fair value under the different paradigms.

The results are in line with intuitive reasoning: the “Displaced Diffusion” model, reduces the root value of the price tree, resulting in generally lower terminal nodes and pay-off values. The “Accumulation” model in turn, effects the dividend payment only at expiry; the resulting change in price tree level is therefore only additive - instead of the multiplicative effect seen with the “Displaced” model. Finally, the “Deterministic Jump” model causes the level reduction to filter through in an additive fashion, as well as contributing to an effectively higher volatility level, resulting in an overall higher price.

5.6 Summary

Chapter Five concerned itself with the practicalities of applying binomial procedures to “real life” problems in Finance. The said problems relate to incorporating term structures on riskless returns, allowing for stock borrowing costs and importantly (if infinitely more difficult) how to build trees to be consistent with volatility skews and smiles as observed in the derivatives markets. The issue of discrete dividends was also treated on.

Chapter 6

Exotic Pricing

Path dependent options tend to be cheaper and lend themselves to hedge structures that can be created to better suit individual needs: they're more flexible. As an example consider Asian options; not only is this type of option less expensive than its European vanilla counter part, but in a market such as the South African grain futures market, which is highly volatile, price protection done with an Asian style payoff provides much better hedging cover. (A farmer who has to deliver his crop would be well advised to have the price he receives depend on the average price recorded over the whole harvesting season; failure to do this will see him exposed to much greater price volatility than he might be prepared to stomach.) Path dependent options are unfortunately, rather difficult to price and hedge, as analytical closed-form price solutions seldom exist. It is precisely the ability to easily incorporate path dependent exercise decisions that makes tree methods so attractive in this context. This section aims towards providing examples of tree based pricing for some of the more often traded exotics, namely American vanilla options, barrier options, forward starting options; in the penultimate section the problem of multi-dimensionality is discussed.

6.1 Background

At the time active trading in vanilla European options commenced on CBOT - the first formalized exchange, in modern times, that catered for options trading, trading in barrier options had already taken place for close on seven years (in the OTC market). Furthermore, most exchanges, from the inception of options trading, set contract specifications so as to accommodate for American style exercise. For neither of the options mentioned, (with the proviso that barrier events are defined in terms of discrete observation periods), does analytical, closed-form solutions exist, either for the calculation of fair value, or for hedge ratios. Apart from the aforementioned, there is a vast array of actively traded exotic pay-off derivatives, for example Asian price and -strike options, Bermudan options, very few of which are priced via analytical formulas. In fact, for some instruments like South African bond options, where the strike is defined in terms of a yield to maturity, and with American exercise style, the likelihood of an analytical method ever being derived is quite small. Now, the fact that the aforementioned derivatives have no closed form method of pricing or hedging parameter calculation, definitely does not preclude trading in them. This begs the question how market participants manage derivatives books that contain the said instruments ?

In most cases either approximations exist that provide close enough answers, or use is made of numerical techniques. In the ambit of the latter, falls finite difference methods for partial differential equation solutions, numerical integration schemes, and, binomial tree methods. What makes these methods quite powerful is the ability to incorporate active management of exercise decisions; these are extremely hard to build into analytical approaches. For this reason, tree methods and other finite difference methods form the cornerstone of almost all derivatives with some form of optimal exercise feature. In this chapter we shall attempt to illustrate the use of tree methods in the pricing of path-dependent derivatives; the options to be treated on are American (Bermudan) exercise vanilla options, Barrier options, forward starting options and Asian style options.

6.2 American and Bermudan Options

American and Bermudan options are of great academic and commercial interest, yet as previously stated no analytical solutions exist for their pricing; as a note of interest, Black and Scholes (1973) made special note in their seminal article, [9], that of the time of publication, no analytical formula existed for the value of an American

put option. One of the first attempts at pricing American options must have been that presented by Parkinson (1976), "Option Pricing: The American Put", [65], although his approach was a numerical integration one. As concerns employing finite difference methods to the valuation of these options, Brennan and Schwartz (1977), with "The Valuation of American Put Options", [15], are regarded to have been the first. Since many analytical approximations have been suggested, including that by Whaley (1984), [84], Roll (1977), [72], and what could be considered a standard technique, that of Barone-Adesi and Whaley (1987), [5]. As far as evaluating American put options by way of binomial trees are concerned, Cox et al (1979), [23], suggested it along with the exposition of the methodology; however it should be noted that this was limited to put options only, as the valuation of options where the underlying pays dividends in terms of fixed dollar amounts were not considered in their initial article. Leisen (1998), [53], evaluated the basic argument employed in binomial American option pricing, but from a dynamic optimization point of view; whilst also considering the convergence properties of binomial option pricing of American options.

American options fall into the class of early exercise type options; the buyer of the option, in each case, is allowed to decide upon the time of exercise. For American options, the time allowed for early exercise is the continuous interval spanning the inception to expiry dates, whilst for Bermudan options, the early exercise times are limited to discrete points. European options have a fixed time of exercise; Bermudan options therefore could be classed as a special case of the American, and European options, in their turn, as a special case of the Bermudan group. The optimality of early exercise is a matter of cost of carry or opportunity cost; simple relationships exist that provide the conditions for optimal exercise.¹ The early exercise feature could simply be viewed as an additional boundary condition, and well as one that

¹Early exercise is optimal only in the event that the benefit of carrying the underlying asset physically (either the long or short), exceeds the time value of the option - the differential between option value and intrinsic value. For in the money call options, exercise would be optimal, if the dividends receivable on the underlying asset are in excess of the funding cost associated. That is:

$$(S - K)(e^{r(T-t)} - 1) < De^{r(T-t)},$$

for S, K, D denoting respectively the spot price, price of exercise and dividends receivable, with $r, (T - t)$ representing the risk free rate of interest and time left to expiry. In the case of in the money put options, early exercise would result in a positive cash flow, identical to the intrinsic value of the option; the benefit of exercise is that the proceeds could be invested for the remainder of the option life, whereas no interest is receivable on the live option position. Therefore, the optimal early exercise condition could be written as:

$$(K - S)(e^{r(T-t)} - 1) - De^{r(T-t)} > p(t) - (K - S),$$

or, optimal exercise is when the interest receivable on the proceeds, net of the present value of dividends payable, exceeds the time value of the live put option.

optimizes across the time domain. Notationally, this could be presented as :

Definition 6 *Let $S(t)$ denote the spot price of an asset at an arbitrary time between inception on an option and its expiry; also denote by K and $v_{S(t)}(t)$ the option's strike price and value respectively, at time t given a spot of $S(t)$. The early exercise condition for American options is then :*

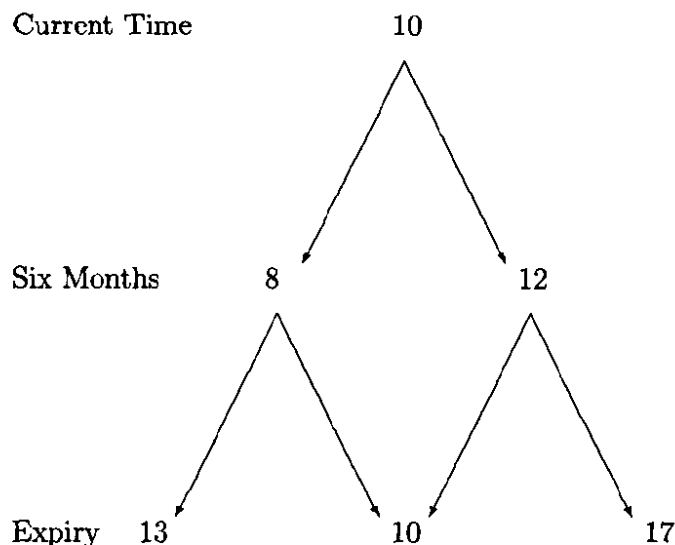
$$\begin{aligned} \max \{S(t) - K, v_{S(t)}(t)\} & \quad , \text{ for a call option,} \\ \max \{K - S(t), v_{S(t)}(t)\} & \quad , \text{ for a put option.} \end{aligned}$$

Incorporating the above into a binomial tree structure is relatively simple: at each price node, the option value is set equal to the maximum of the discounted expected value at the following node, and the immediate exercise value. The justification for this is twofold: the first being simply that this is in accordance with the no-arbitrage principle; failure of the option to reflect the higher of the intrinsic and live value would result in a option value - exercise value mispricing that is easily exploitable. The second justification for the above is a more high-minded (or sounding) one, namely the Bellman principle of dynamic optimization - see [51] for an explanation of this. Simplistically this is explicable as: the optimal value of an option, considered over all possible sample paths of the underlying and early exercise decisions, is the same as consecutively optimizing exercise decisions over small time intervals and conditioning² on the initial conditions.

To illustrate the above concepts, consider the valuation of an one year American option, struck at-the-money at 10. Assume a risk free rate of zero and equiprobable outcomes at each of the three assumed time points. Let the tree below represent the price dynamics over one year.

²Conditioning refers to calculating an expected value, or probability, but given the occurrence of some other event. In the case of early exercise features, conditioning plays a vital role, as decisions are always based on information to be had at that time. For example, in the choice between exercise or keeping an American option alive, one has to weigh the value of each, but the value is in its turn dependent upon the state of the financial asset at that time. Conditioning is easy to perform if the possible states of the variable in the future is known, a probability to each state is easily assigned, and most importantly, if the states are limited in number.

Within a binomial tree framework, valuation choices are always performed for one period at a time, implying only two states to be considered. Monte Carlo methods for example are very difficult to apply to early exercise problems, the reason being that conditioning becomes nigh on impossible given the continuous nature of both asset value realizations and the fact that it is done assuming a continuous time line. Binomial trees, whilst not denying the continuous nature of either time or the asset's value, yet provides a systematic (and process consistent) way of partitioning the respective intervals, facilitating easier conditioning.



In the absence of early exercise, the value of the option becomes quite easily $\frac{3}{4}$; however, exercise is optimal at the node (Six Months, 12), as the live option value is only $\frac{3}{2}$; the value of user-defined time of exercise is calculated as $\frac{1}{4}$.

Now, pricing Bermudan options (options where exercise is allowed at multiple discrete points) within a tree environment, is identical to that of American options, with the exception that tests for optimal exercise are performed only at points corresponding to time nodes where exercise is allowed. The requirement in this case, obviously being that time intervals where exercise is allowed be explicitly included in the tree. That is to say, should exercise on an option be allowed at times t_1, t_2 then, the tree built to cater for valuation should contain time points that correspond to times t_1, t_2 .

6.3 Discrete Barrier Options

Barrier options are options where pay-off is conditional upon whether certain barrier are breached or not for the price path over some time interval. They are broadly categorized as either knock-in or knock-out barrier the latter being options that ceases to exist in the event that a price barrier is violated by the underlying at some point in time. Knock-in options only convert into live trades when a barrier event takes place. Thus, in order to calculate the pay-off to a barrier event, knowledge on the past price behavior of the underlying is required. This type of option has been actively traded since the late 1960's; their mathematical pricing was already considered by L. Bachelier in 1904, although of course not in a risk neutral framework. Post Black and

Scholes, [9], the first pricing approach of barrier options was posited in Merton (1973) "Theory of Rational Option Pricing", [57]. In the said article, a closed form solution was provided for "down-and-out" call options, where the barrier event is defined in terms of continuous time observations (for barrier violations). The mirror principle for random walks forming the basis of the argument.³ Reiner and Rubinstein (1991), [65] extended Merton's results to all continuously monitored barrier options ; the mentioned paper is consider a classic for barrier option valuation. No analytical formulae exist for the value of discretely monitoring barrier options, however, Broadie et al (1999), [16], provided a way in which one could relate the value of the prior to continously observing barrier options.

Pricing barrier options within a finite difference approach, and more specifically within binomial trees is a problem that enjoyed much attention since the introduction of this method by Cox et al (1979), [23]. In general, the principle employed is that barrier condition violations be incorporated into a lattice valuation by simply verifying that barrier are not violated at each price-time node. In other words, the approach taken was identical to that followed for American options; the justification for this method being a simple no-arbitrage argument. However, it was realized very soon that the barrier option valuation within lattices, presents some definite challenges, as a quick glance at a convergence diagram would show (see Figure 1 below, comparing a vanilla European option's convergence with that of a knock-out barrier; the same particulars applying.)

It should be quite evident that the saw-tooth type of convergence pattern exhibited by lattice barrier option valuations, makes for difficult price inference, that is, there is a greater deal of uncertainty as to the eventual convergence value in the case of barrier options, than in the vanilla equivalent, and well because of more irregular convergence patterns. The cause for this was identified by Margrabe (1989), "Binomial Pricing of Exploding Options", [54]. More specifically, he observed that convergence is especially slow when the barrier is close to the spot value of the underlying. Of course, the reason for this is the inherent bias in lattice methods, when determining the in- or out status of barriers - similar to the overstatement of exercise values mentioned for American options. Otherwise said: whether a barrier finds itself quite substantially away from the tree node considered, or whether it is just below, because of the discrete nature of checking for violations, both would return the same result. Many authors have proposed that a solution to this might be to constrain tree nodes, such that some of the nodes are always on the same level as the barrier set; see for

³The mirror principle reduced to its most basic form is that over a specific time interval, a drift less random walk process is just as likely to end above a barrier level than below. This is easily expressed in terms of conditional probabilities that could be used to derive the entire conditional distribution.

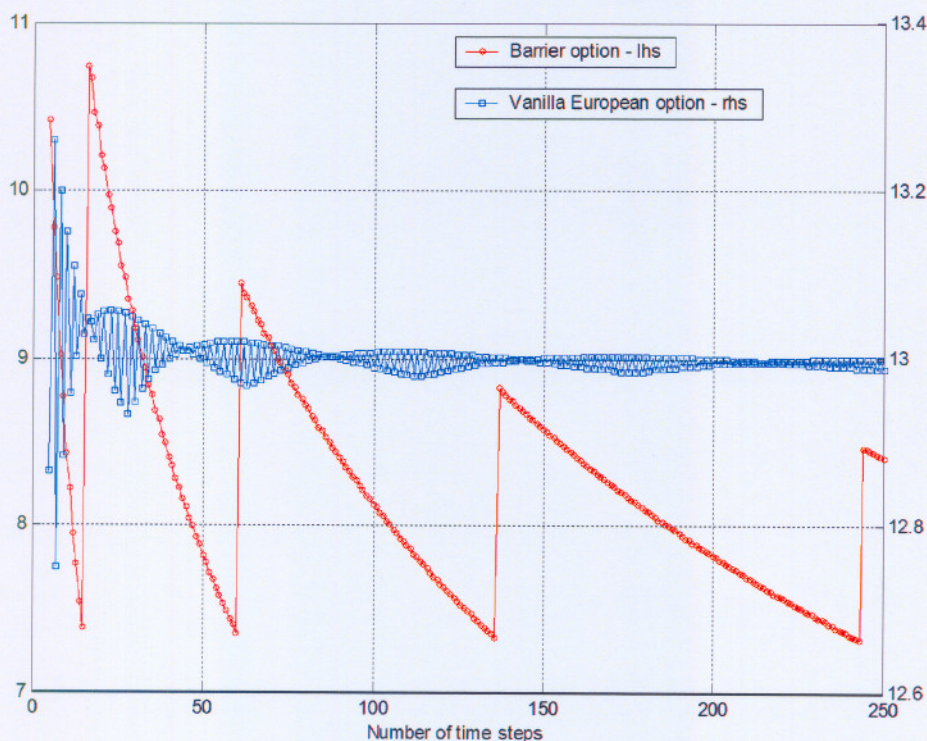


Figure 6.1: Convergence of barrier and vanilla options within a tree valuation.

example “Bumping Up Against the Barrier with the Binomial Method”, Boyle and Lau (1994), [11], as well as the seminal work by Derman and Kani (1993), “The Ins and Outs of Barrier Options”, [26]. Rogers and Stapleton (1998), [67], however, proposed altering the binomial tree probabilities to allow for the proximity problem.

Reduced to its essence, barrier option pricing within binomial trees involves the simple addition of a per node check of barrier constraints, that is the option value is derived from subsequent price nodes via the “Fundamental Result of Binomial Pricing”, but checked against the barrier conditions. In the event of a knock-out option, the value is overridden with zero if the price node falls outside of the knock-out area, whereas, the value is set to zero at all nodes where the price is outside of the knock-in area, in the case of knock-in options. Algebraically, the above could be presented as:

Definition 7 Let $S_{t_j,j}$ denote the price of the underlying at time t_j and at the j^{th} node in the tree, and $f(t_j|S_{t_j,j})$ the value of a barrier option with the barrier set as a region B ; the value being of course conditional upon the price $S_{t_i,j}$. Then, given a risk

free rate of r and the risk neutral probability of p , the value of a down-and-out option considered at the i^{th} time-interval and price level, could be written as :

$$f(t_i|S_{t_i,i}) = \begin{cases} \{pf(t_{i+1}|S_{t_{i+1},i+1}) + (1-p)f(t_{i+1}|S_{t_{i+1},i-1})\}e^{r\Delta t_i}, & S_{t_i} \notin B \\ 0, & \text{else} \end{cases}$$

the equivalent value for a down-and-in option would be :

$$f(t_i|S_{t_i,i}) = \begin{cases} 0, & S_{t_i} \in B \\ \{pf(t_{i+1}|S_{t_{i+1},i+1}) + (1-p)f(t_{i+1}|S_{t_{i+1},i-1})\}e^{r\Delta t_i}, & \text{else} \end{cases}$$

Quite often, barrier features are limited to time windows only, that is, instead of options values being subject to barrier level honoring from inception to expiry, the period within which these conditions hold might be limited to the last month of the life, or, be confined to a two month period falling somewhere between trade date and maturity. These are referred to as window barrier options; this option differs in its lattice valuation approach from that already considered, only in that the conditions are made time dependent - which is easy enough. Another feature of many traded barrier options is that of rebates.⁴ This being a cash pay-out on the violation of a barrier condition. For example, knock-out barrier options with a rebate ceases to exist when the price moves within the knock-out period, however, upon termination a cash pay-out is made to the option holder. Once again, making allowance in the lattice approach for rebates is simple in the extreme. The result earlier given, could be extended to include window observations and rebates as follows :

Definition 8 Assuming the same notational arrangements made in the prior definition, but with the proviso that a rebate of Q is paid upon knock-out or failure to knock-in, and with barrier checking limited to the time interval $[\tau, \tau']$, the valuation equation then becomes :

$$f(t_i|S_{t_i,i}) = \begin{cases} Q, & t_i \in [\tau, \tau'] \text{ and } S_{t_i} \notin B \\ \{pf(t_{i+1}|S_{t_{i+1},i+1}) + (1-p)f(t_{i+1}|S_{t_{i+1},i-1})\}e^{r\Delta t_i}, & \text{else} \end{cases}$$

the equivalent value for a knock-in option would be :

$$f(t_i|S_{t_i,i}) = \begin{cases} Q, & t_i \in [\tau, \tau'] \text{ and } S_{t_i} \in B \\ \{pf(t_{i+1}|S_{t_{i+1},i+1}) + (1-p)f(t_{i+1}|S_{t_{i+1},i-1})\}e^{r\Delta t_i}, & \text{else} \end{cases}$$

⁴In the South African warrants market, the share instalments sold by Investec have embedded knock-out barrier put options with rebates.

That concludes the discussion on barrier option valuation.

6.4 Forward Starting Options

Forward starting options are options where the start- or effective date is set to be in the future; otherwise said, forward start options could be regarded as commitments to enter into options at some time in the future; the contract detail is set up-front. The principal advantage of this type of option is that the strike is normally only set (in absolute terms) at the options start date, that is, at trade date, the strike is only specified in terms relative to the spot price that would prevail at the effective date, the most obvious example being that the strike is set at the money, or at a fixed ratio of the underlying's spot price. In some cases, the strike price might be more complexly defined, for example as the spot price but with a fixed addition; interest rate options might have strike prices set at Libor plus 250. Where the strike price ceases to be defined in relative terms to spot, analytical pricing methodologies become quite deficient in accurately determining the value of forward strike options. Also, where the option that is forward starting is itself of a more complex nature, for example forward starting barrier options, need arises for some alternative pricing technique. Binomial pricing, in virtue of its ability to easily incorporate conditional decision making, is ideally suited to price this type of instrument.

The principle used when applying binomial pricing to forward starts, is that of modeling the price behavior of the underlying for the period between valuation and start times; the conditional value of the derivative is then equated to each of the terminal nodes, that is in the stead of attaching pay-off values, conditional valuations are used. Through backward induction, the conditional valuations are then easily related to the valuation date. Mathematically the technique could be presented as follows:

Definition 9 Let $S_{t_n,j}$ denote the price of the underlying at time t_n and at the j^{th} node in the tree, and $f(t_n|S_{t_n,j})$ the value of the derivative that is valued, but conditional upon the price $S_{t_n,j}$ and considered at time t_n . Then, given a risk free rate of r and the risk neutral probability of p , the value of the derivative, forward starting at time t_n , but valued at time t_0 is presented as :

$$f(t_0|S_{t_0,0}) = \sum_{j=0}^n \binom{n}{j} p^{(n-j)} (1-p)^j f(t_n|S_{t_n,j}) e^{rt_n} \quad (6.1)$$

6.5 Other Exotic Pay-Offs

In the Introduction, mention was made of Asian options. These instruments fall in the so-called multi-dimensional category (of which the following section makes mention), which makes them exceptionally difficult to price using tree methods. This should be evident from the fact that the terminal pay-off is conditional upon average prices observed prior to expiration; each sample path within a tree has a unique average price observed. In Chapter Two, one of the advantages of binomial trees was stated as the ability to condense the information contained in the various sample paths created by the random walk model - see Theorem 2. In the case of average price per sample path, that property ceases to apply. Thus, to properly value average price options, one needs to individually consider the 2^n paths implicit in the price tree. This is not feasible; as a result, at best approximations within trees are used, or other pricing methodologies altogether.

Lookback options pose a similar problem to that of Asians - each sample path is associated with a unique pay-off. As such, apart from evaluating each binomial tree path separately - and therefore defying the purpose of binomial models, one could employ either of the approximation techniques posited by Hull and White (1993), [45] or that of Rubinstein (1994), [73]. The latter proposed a Monte Carlo simulation of sample paths within the tree. Hull and White (1993), however, proposed a technique where one approximates the derivative value at each node, based on estimating a range of either the extrema, or of the average observed. The second technique, besides being very generic, and therefore applicable to a wide range of derivative pay-offs, is also better suited to American option valuations than the Rubinstein method.

Babbs (2000), [2], provides a pricing approach that accurately solves for the price of lookback options within a tree environment. It essentially involves modelling the pay-off return at each node, rather than modelling the underlying asset's price. As such, use is made of a truncated tree, with either the top or bottom part trimmed and replaced with a reflecting barrier.

6.6 The Curse of Multi-Dimensionality

One could be forgiven for assuming from what has been presented so far about binomial trees, that this pricing methodology is capable of providing the financial engineer with everything he needs, bar perhaps world peace. The truth however, seldom is a simple affair: binomial trees and their application to derivative pricing being no ex-

ception to this. The principal limitation of these methods from a pricing perspective is that of difficulty in incorporating more than one dimension to the modeling of the value of the underlying. That is, derivatives where the underlying is defined in terms of more than one asset's value in the calculation of the pay-off. An example of this would be an option on a basket of underlying assets; this of course also includes Asian options on one underlying only, as these could be expressed as a basket of different assets as well. As each underlying has to be modeled separately, as well as co-jointly, the addition of each variable causes the number of nodes considered to increase exponentially; an illustration of this is now given. At the $i - 1^{th}$ time node on a binomial tree, there are normally i different states modeled for the underlying's price. Adding one additional underlying, results in two different trees, each with i nodes at the said time point; these are joined in a tree having $\binom{2i}{2}$ different nodes at time t_i ; the equivalent number for a three-dimensional underlying being $\binom{3i}{3}$; only regarding the first time point, the sequence of number of nodes considered across increasing dimensions is easily calculated from the preceding as: 2, 6, 20, 70, 252, 924, Keeping in mind that the enumerated sequence still only provides the number of nodes at the second time node, and that normally, time intervals number closer to 200 or more, the problem that dimensionality causes, should be self-evident.

6.7 Summary

This chapter explored the pricing of exotic derivatives using a binomial tree approach. Exotic options are basically any derivative save vanilla European options, futures and some types of forward contracts. However, the discussion was limited to a small subset of the mentioned only: American-, Bermudan- and forward-start options. The chapter was concluded with brief notes on the evaluation of Lookback and Asian options, as well as some remarks on the effect of multi-dimensionality on the pricing approach.

Chapter 7

Numerical Issues

In the field of numerical analysis, an approximation's accuracy and stability is of considerable importance, as is the optimality of the solution algorithm employed. Methods where small approximation errors accumulate and then “explode”, so as to give unexpected results are generally considered unstable and for obvious reasons are somewhat undesirable. Convergence on the other hand has to do with the optimality of a numerical solution; more optimal solutions converging faster than ones of lesser optimality. As binomial tree pricing finds itself well in the ambit of numerical methods, it follows that tree convergence and stability are important matters to be considered.

7.1 Background

Stability, or rather instability, (and well within the more heuristic - also read loose - sense in which it is treated in this dissertation), could broadly be defined as unexpected behavior.¹ Arguably the most extreme example of unstable behavior is that of so-called pseudo-stochastic functions, like the myriad of pseudo-random numbers in existence. These functions, although being deterministic (that is given the initial conditions, the value at some point in time can be known with full certainty), are so sensitive to small deviations in the starting values, that perturbations result in stochastic-like behavior. The latter is the principle underlying “Chaos Theory” - refer to Figure 7.1 for a graphical representation of the so-called Lorenz attractor, a classical example the prior mentioned functions.

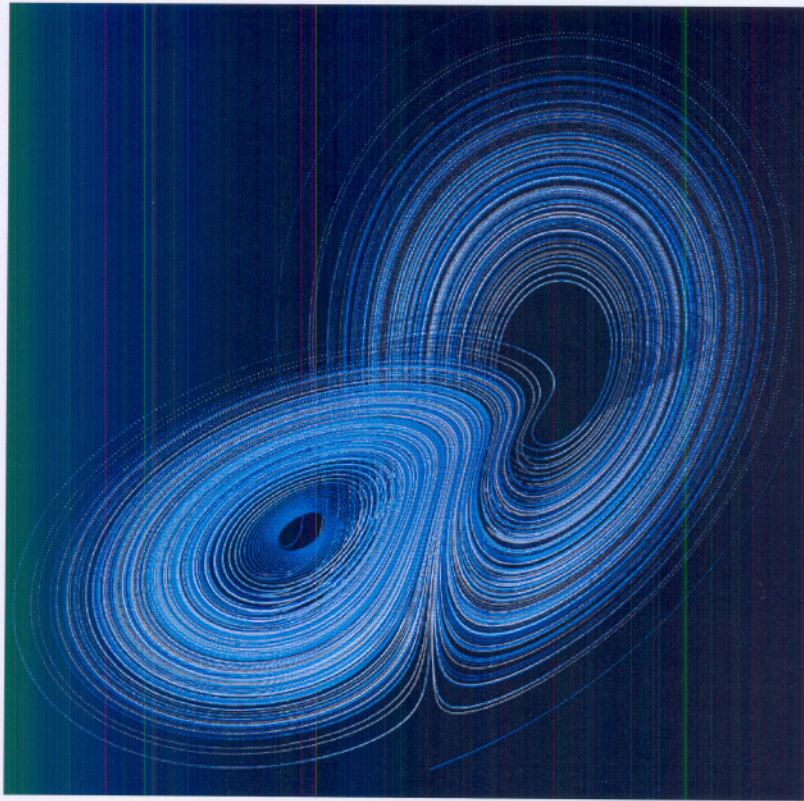


Figure 7.1: The Lorenz attractor.

Approximation errors in numerical methods are inevitable - by definition; proxy so-

¹The field of numerical analysis applies rigorous definitions to this concept, for the purposes of this discussion focus will be on a more intuitive approach.

lutions, ideally, should be well-behaved over a range of possible initial conditions; the absence of such regularity, would result in extreme difficulty when deciding upon a suitable solution value. All explicit finite difference methods are conditionally stable, that is stability is dependent upon the choice of function parameters. Binomial methods are similar to explicit finite difference methods, however, most of them display a pronounced robustness in stability. This is a point Geske and Shastri (1985), [35], made special reference to. Otherwise said: it appears as if the binomial method is more allowing of discretization error build-up than the explicit method. This section aims to investigate the robustness of binomial tree methods.

Convergence or rather claims to improvements thereon could be said to have been the principal driver in most publications on binomial trees since that of Cox et al (1979). Consider for example the subtle claim by Tian (1993), [84], “..the results indicate that the modified (*their*) binomial procedure is more accurate than the CRR (*Cox et al*)” and less subtle, the one by Leisen and Reimer (1996), [55], “Tian proposed binomial and trinomial models..... Unfortunately this interesting contribution lacks to support the ideas by mathematical arguments...”. Not only is convergence speed a point of interest, but also the convergence patterns produced by different parameterizations for different pay-off styles. The aim of the second part of this chapter is to provide a brief characterization of convergence as relating to binomial methods.

7.2 Sense and Instability

Any practitioner of derivatives pricing knows that under certain parameter choices, binomial trees, and specifically ones parametrised in accordance with the Cox, Ross and Rubinstein method, behaves in a perplexing (if not vexing) fashion. As an illustration of this, consider the valuation of a three year vanilla European option, at the money, where the underlying behaves with an annualized volatility of 5%, and with the risk free rate of interest at 25%; the graph below shows the convergence diagram.

It should be obvious that for the number of time steps less than 30 the pricing methodology becomes misbehaved - to put it rather mildly. Further investigation of this yields that for the said region, risk neutral probabilities are in flagrant violation of the Kolmogorov probability postulates, being in excess of one. Many an author considering the stability of finite difference methods in derivatives pricing have observed on unstable behaviour in derivatives pricing being frequently accompanied by such violations; see for example Diener and Diener (2004), [31], as well as Hull (1998), [47]. The blow-up behavior generally follows when the risk free rate of interest is sig-

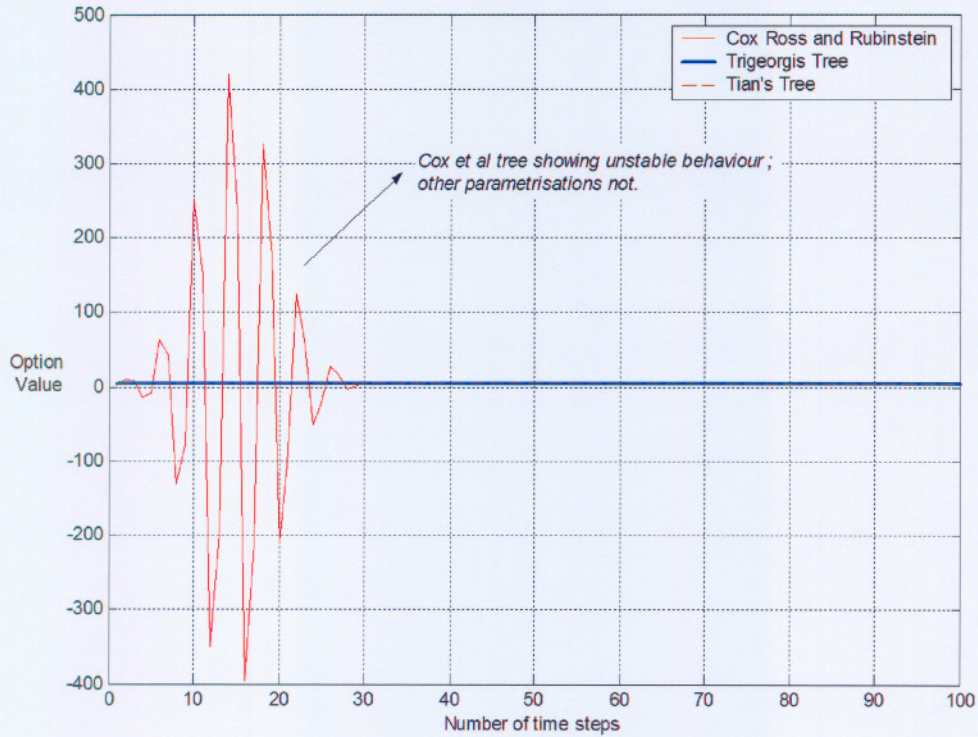


Figure 7.2: Unstable behavior observed for the “Cox et al” pricing of a vanilla European option.

nificantly higher than the return volatility (both parameters scaled for the time to expiry). Mathematically, this could be expressed as :

$$r \frac{t}{n} > \sigma \sqrt{\frac{t}{n}}.$$

To see why the above results in option values exploding, consider the price of a vanilla European call option written in terms of the binomial lattice probabilities:

$$c(0) = \left\{ \sum_{j=0}^n \binom{n}{j} p^{n-j} (1-p)^j \max\{S(0)u^{n-j}d^j - K\} \right\} e^{-rt}. \quad (7.1)$$

As p is scaled exponentially by the number of time steps, n , it should follow that for $p > 0$ the pricing formula presented could result in very high, or even negative numbers. In order to present the preceding in more concise mathematical terms, one first needs a definition of stable behavior. In the introduction a very vague, heuristic

definition of stability was given, that is of a function being well-behaved over a range of starting values. However something with a little more concision could be :

Definition 10 Let $\{f_i(x)\}_0^n$ denote a sequence of functions approximating a function $f(x)$, with $x \in A$. The approximation technique is said to be unconditionally stable on A if and only if the aforementioned function sequence converges uniformly to $f(x)$.

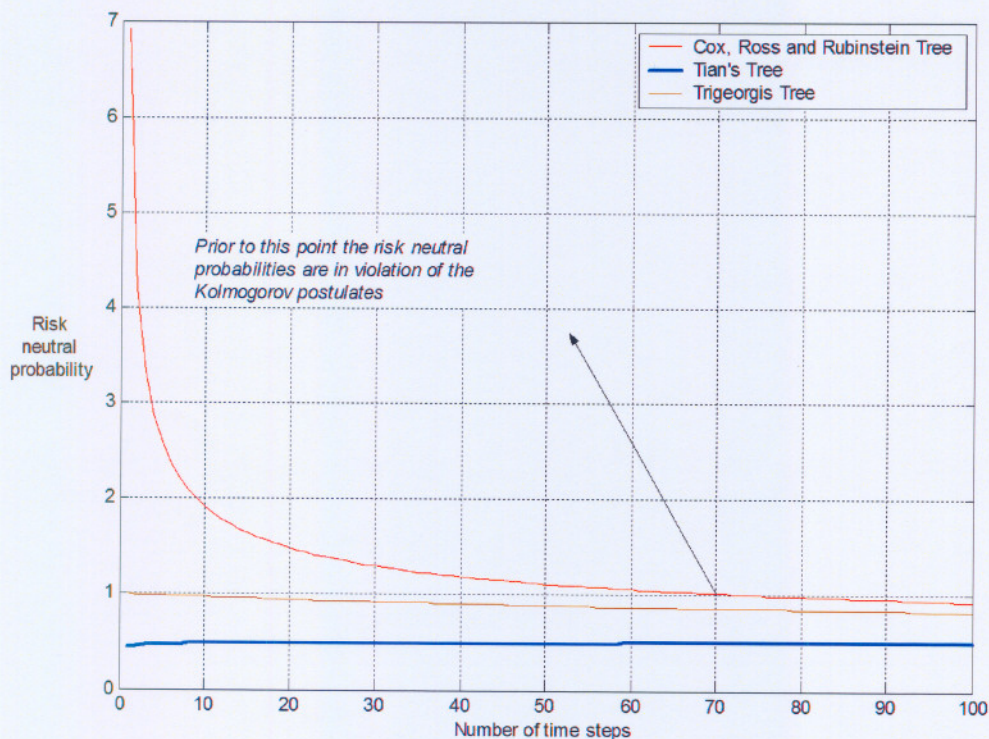


Figure 7.3: Risk neutral probabilities implied by the “Cox et al” binomial method.

The above follows because of the fact that numerical approximations can invariably be expressed as function sequences, but defined on some domain of input parameters. (Intuitively the definition is equivalent to saying that regardless of the input parameters used, all approximation functions attain a certain accuracy after an arbitrary number of iterations - time steps. It is simply a more mathematical way of saying that in terms of errors, there are no surprises to be expected after a certain number of pricing iterations.)² Now, in terms of the given definition one could make the following claim:

²A remark is in order concerning the difference between instability and non-convergence: one can

Theorem 12 *Cox, Ross and Rubinstein trees are not unconditionally stable: for certain parameter choices the valuation solution lapses into instability.*

Proof: *The proof consists of showing that the CRR tree implied value for a vanilla European call option does not converge uniformly towards the Black and Scholes one, and well on the domain $A = [0, \infty)$ of x , where the latter is defined as the ratio $x = \frac{r\sqrt{t}}{\sigma}$.*

Denote by the sequence of functions $(c_n(x))$, with $n = 1, 2, 3, \dots$, the CRR tree implied values of a vanilla European call option (for different time step values), with strike K , time to expiry t , with the spot at evaluation time S and the risk neutral probability as $p(x)$. (As the proof does not depend on the interaction between the up and down factors - denoted by u and d - for in the spot money options, the simplifying assumption is made that these are constant functions of x). Now:

$$c_n(x) = \left\{ \sum_{j=0}^n \binom{n}{j} p(x)^{n-j} (1-p(x))^j \max\{Su^{n-j}d^j - K, 0\} e^{-rt} \right\}, \quad (7.2)$$

and

$$p(x) = \frac{1}{2} - \frac{1}{2} \frac{x}{n}. \quad (7.3)$$

Substitution of (7.3) into (7.2) then results in :

$$c_n(x) = \left\{ \sum_{j=0}^n \binom{n}{j} \left(\frac{1}{2} - \frac{1}{2} \frac{x}{n}\right)^{n-j} \left(-\frac{1}{2} + \frac{1}{2} \frac{x}{n}\right)^j \max\{Su^{n-j}d^j - K, 0\} e^{-rt} \right\}. \quad (7.4)$$

Let BS denote the Black and Scholes value (a constant) of the option. Note the following lemma:

Lemma 5 *The sequence $(c_n(x))$, with $x \in A$ does not converge uniformly to $c(x)$ on A , if and only if for some ϵ there exists subsequences (c_{n_ℓ}) of (c_n) and x_ℓ in A such that :*

$$|(c_{n_\ell}(x_\ell)) - c(x_\ell)| \geq \epsilon$$

for all $\ell = 1, 2, 3, \dots$. (The above follows easily from the negation of the definition of uniform convergence, see Bartle and Sherbert, [4].)

have convergence and instability. Thus, a numerical method can converge towards the true solution, and be unstable at a certain point in the iteration; instability does not preclude convergence. Also, it was proven in Section Three, beyond any doubt, that Cox, Ross and Rubinstein trees converge unconditionally towards the Black and Scholes consistent price for derivatives.

If one were to set $x_\ell = 10\sqrt{\ell}$ and $n_\ell = \ell$, then :

$$c_{n_\ell}(x_\ell) = \left\{ \sum_{j=0}^{\ell} \binom{\ell}{j} \left(-\frac{9}{2}\right)^{\ell-j} \left(\frac{9}{2}\right)^j \max\{Su^{\ell-j}d^j - K, 0\} e^{-rt} \right\} \quad (7.5)$$

the above is an unbounded function of ℓ , so that :

$$\lim_{\ell \rightarrow \infty} |(c_{n_\ell}(x_\ell)) - BS| = \infty$$

Thus, it can be said that the sequence $c_n(x)$, although always converging to BS , does not do so uniformly on the domain of possible choices for x . In other words, the Cox, Ross and Rubinstein tree implied values for vanilla European call options are not uniformly convergent, and so follows the more general statement earlier made.

Q.E.D

Now, other tree parameterizations, interesting enough do not behave similar to the Cox et al. in the example earlier presented - refer to the convergence diagram of the Trigeorgis and Tian methods (for the same inputs). Were one to generalize on this, it would appear as if instability might be a problem related to certain parameterizations only, however could this be proven? The principal difference between the Cox et al and other parameterizations is that the former makes the simplifying assumption that the second moment of log returns is equal to the variance, i.e.,

$$E(X^2) = V(X). \quad (7.6)$$

This is of course not true in general (only in the event where the distribution is driftless does the above apply). The proper relationship is rather:

$$E(X^2) = V(X) + E(X)^2. \quad (7.7)$$

The mentioned approximation results in the absence of the mean return in the definition of the “up” and “down” factors - characteristic and eccentric to the Cox et al definition. This in turn causes the risk neutral probability,

$$p(x) = \frac{R - d}{u - d},$$

to assume values beyond the interval $[0, 1]$. In other words with the “up” and “down” factors as defined by Cox et al, the function $p(x)$ ceases to be a bounded one, ultimately causing a “blow-up” of derivatives valuations in some cases.

7.3 Convergence

Binomial tree methods display very characteristic and interesting convergence patterns. As concerns the literature on binomial convergence: on characterizing the patterns, four authors (in particular) have made contributions to this issue, namely Boyle (1988), [12], Hilliard and Schwartz (1996), [45], Leisen and Reimer (1996), [55], and finally Diener and Diener (2004), [31]. The first two, although pointing to the peculiarity of oscillations, cusp points etc. observed in binomial diagrams, and attempting to employ these towards improving the order of convergence failed to provide proper explanations of the patterns. The third contribution mentioned, being “Binomial Models for Option Valuation - Examining and Improving Convergence”, [55], (which might be considered definitive as far as the characterization of binomial pricing convergence is concerned), Leisen and Reimer makes a number of very interesting observations, whilst also defining a measure of convergence speed, which is based on the concept of bounding sequences, whereafter various tree parameterizations are compared in terms of order of convergence. The paper is concluded with a novel tree parameterization, which is shown to be at least twice as efficient as the others. Unfortunately, the proposed method is limited in its use to vanilla European options only: for other pay-off styles, it results in a strong bias in pricing results.

³ The last contribution mentioned here is “Asymptotics of the Price Oscillations of a European Call Option in a Tree Model”, by Diener and Diener (2004), [31]. This very technical article, focuses mainly on describing the wave-like, oscillatory binomial convergence patterns by way of Laplace integrals; the second focus being on using the said integral approximations to establish the speed of convergence for binomial methods.

In this section, it is aimed to briefly touch upon the types of convergence patterns as well as the causes of their eccentricities; as the speed of convergence is of great importance to the derivatives pricing practitioner, this subject is also investigated.

7.3.1 Convergence Patterns

As a general rule, convergence diagrams for binomial trees display three distinct forms, namely: regular oscillation (as found with Cox et al tree values for at the money options), wave-like oscillations, applicable to at the money European values as

³It is not entirely clear, in fact, that for alternative pay-off styles, their approximation indeed converges to the Black and Scholes partial differential equation consistent value. As an example of that consider the value of a barrier option, as derived by the methodology they proposed, as against the “true” value.

implied by other parameterizations, as well as away from the money options with Cox et al. The last general type of pattern is the saw-tooth like behavior so characteristic of binomial valuations of barrier options.

Regular Oscillations

This pattern is characteristic of an alternative over- and under estimation of the option value; both Diener and Diener (2004), [31], and Leisen and Reimer (1996), [55], ascribe this to the behavior of the variable a , being defined as

$$a = \lceil \frac{\ln \frac{K}{S} - n \ln d}{\ln u - \ln d} \rceil.$$

The said variable in effect defines the lower bound of the nodes considered for inclusion in the expected pay-off at the derivatives termination. This is more expressly seen in considering the following equivalence:

$$\sum_{j=0}^n \binom{n}{j} p^{n-j} (1-p)^j \max\{S(0)u^{n-j}d^j - K, 0\}e^{-rt} =$$

$$\sum_{j=a}^n \binom{n}{j} p^{n-j} (1-p)^j S(0)u^{n-j}d^j e^{-rt}$$

The variable a has a step-like behavior on n , the number of time steps. What this amounts to is a sort of out-of-sync movement between the risk neutral probability distribution and the event defined on it, see the figure below for a graphical description of this.

The probability mass function oscillates, but the behavior of a ensures that the same cut-of point is (within the oscillation) applied to determining the cumulative risk neutral probability of exercise. The result of this two-speed behavior being the characteristic alternating convergence patterns.

Wave-like Convergence

Depending on the moneyness of an option, the behavior of a could deviate from the regular step functions seen in the above example, for example for a moneyness of 120%, irregular two and three step functions result; the impact of this on the convergence is the addition of waves and cusp point when viewed in terms of the convergence diagram. Refer to the figure below for an example of this type of behavior.

Besides away from the money options under the Cox et al trees, other tree parameterizations, as a rule, show similar wave-like behavior even for at the money options. Tian's tree parametrisation for example, always yields decreasing trend, wave-like behavior in terms of convergence. A reason for the type of convergence alluded to

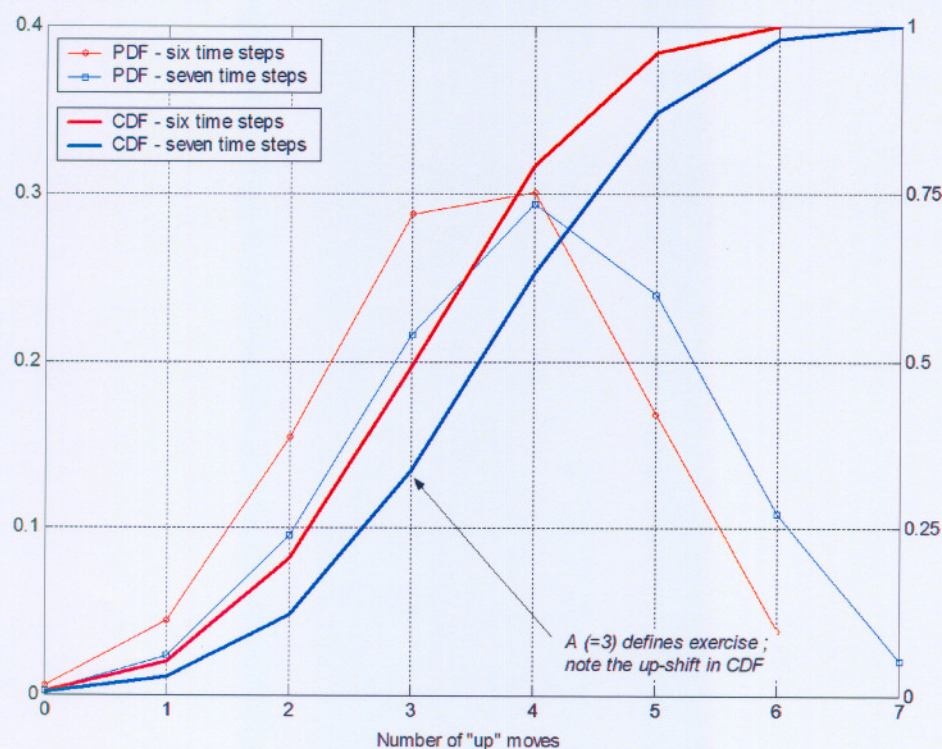


Figure 7.4: Tree-implied risk neutral probability distribution; evolution over six and seven time steps.

here, is the more complex behavior associated with the function a .

Saw-Tooth Convergence

Barrier options, at least the binomial procedure implied values for these instruments, have very characteristic saw-tooth like behavior patterns. Since the first attempts at valuation of these instruments within the tree paradigm, the saw-tooth behavior was remarked on. The reason for this relates to the positioning of nodes relative to the barrier level. In trees where some nodes are always close to the barrier level, the values are close to the "true" one and vice versa. Standard tree methods are wont to return price evolutions ever increasingly closer to the barrier level, and then suddenly reverts back to price trees where the node closest to the barrier level is quite far away, relative to the trees that preceded it. The effect is a jump to overestimation of the option value.

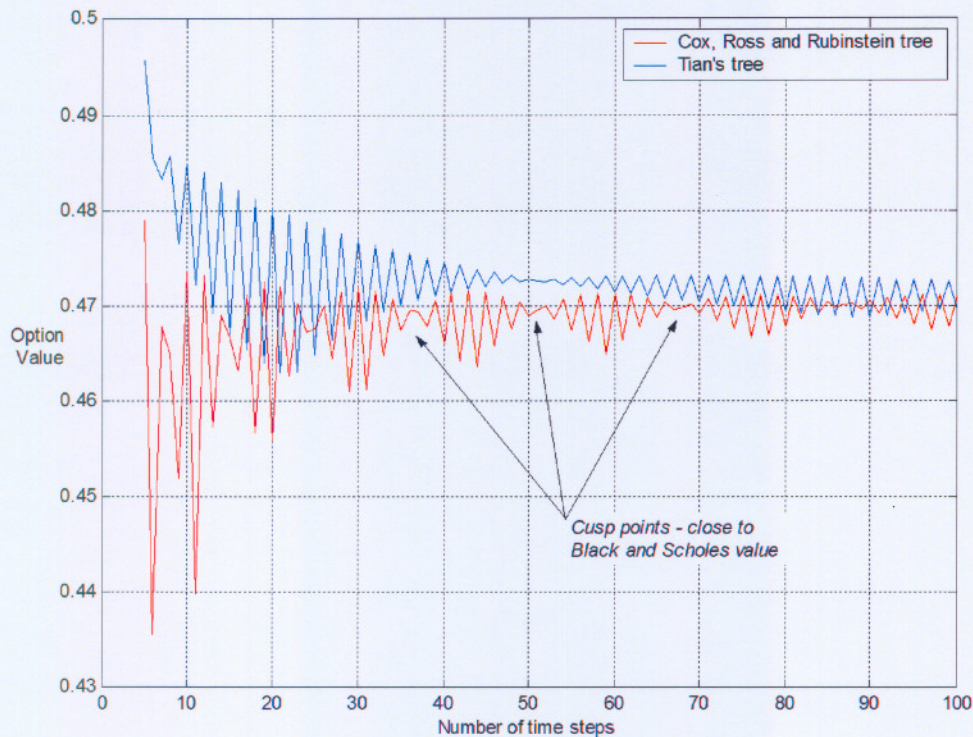


Figure 7.5: Wave-like convergence patterns observed for tree pricing.

A number of authors, (see, for example, Haug), have suggested that in the valuation of barrier options within a tree environment, use be made only of a number of time steps that finds itself in a pre-determined sequence, from our example above, such a sequence should follow as 15, 60, 136, 253, Others, like Derman and Kani (1995a) and (1995b), [26] and [28], have suggested tree constructions where nodes are constrained so as always to yield values on the barrier level; the consequence being much improved convergence.⁴

7.3.2 Convergence Speed

Many claims have been made regarding the speed of convergence of various binomial approximation techniques. Superficially considered, this seems to be valid, refer for

⁴This is quite easily achieved by using the additional degree of freedom in choosing the tree parameters to center the tree around the barrier level.

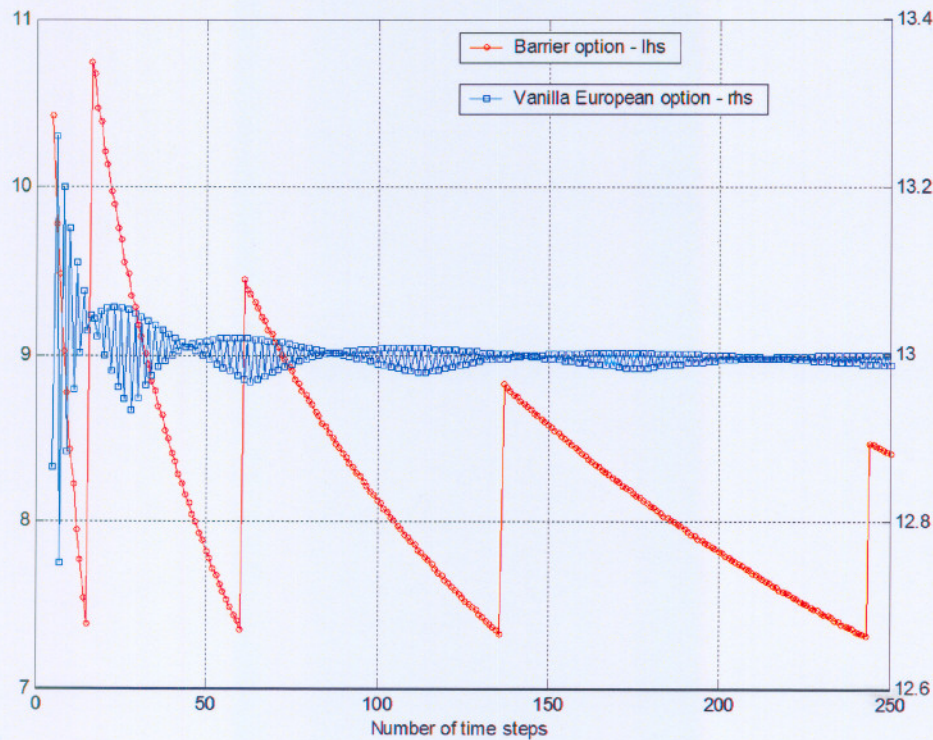


Figure 7.6: Saw-tooth convergence patterns observed for barrier options.

example to the figure below which shows convergence for vanilla European option values under various parameterizations. It does appear as if some, notably the trinomial method, perform better than the others. It is the aim of this section to shed some light on this matter, and well by reference (in the main) to work done by Diener and Diener (2004), [31].

When considering order of convergence most publications define it in terms of the landau order notation.⁵

Order of convergence of k normally implies that the error terms of the particular approximation technique are $o(\frac{1}{n^k})$. Intuitively, this means that errors diminish faster than the sequence of functions $\frac{1}{n^k}$. Now, drawing upon the work of Diener and Diener (2004), [31], the following theorem is presented:

Theorem 13 *All binomial models with parameterization such that the expected re-*

⁵A function is called $o(h)$ if $\lim_{h \rightarrow 0} \frac{g(h)-h}{h} = 0$.

turn of the underlying is set equal to the drift term of the continuous time process, converges to the Black and Scholes closed form solution by order of one; that is the error terms are $o(\frac{1}{n})$. For parameterizations where the stated condition is violated, the order of convergence is one half. (This theorem restricts itself to vanilla European options.)

Proof: Diener and Diener has proven the above by first expressing the value of a European vanilla option in terms of an incomplete Beta function, i.e.

$$\left\{ \sum_{j=a}^n \binom{n}{j} p^{n-j} (1-p)^j \{S(0)u^{n-j}d^j - K\} e^{-rt} \right\} = \\ a \binom{n}{a} \left\{ S \int_0^{\frac{pu}{R}} y^{a-1} (1-y)^{n-a} dy - \frac{K}{R} \int_0^p y^{a-1} (1-y)^{n-a} dy \right\}.$$

The above follows quite easily when considering that the said Beta function is simply an $(n-a)$ -fold integration by parts of the presented summation.

The integral form of the binomial price is then expanded asymptotically by way of Laplace extensions; with this done, one ends up with terms in $\frac{1}{n}$ of order one and higher, thus proving the theorem. For a more complete proof refer to Diener and Diener, [31], pages 277-281.

Q.E.D

It is of interest to note that Heston and Zhou (2000), [42], proved order of convergence of only one half. The reason for this is that they considered each of the two terms in equation (7.8) separately, deriving convergence speed for each, and then concluding that the overall level of convergence is the same. However, as Diener and Diener (2004) point out, ([31], page 284), the terms containing $\frac{1}{\sqrt{n}}$ cancel out, leaving only those of order one and higher.

Example:

The differences in order of convergence is quite evident when considering the evaluation of vanilla European options for various levels of moneyness. In the case of at the money, one-year options, where the underlying displays price volatility of 20%, with the riskless return at 10% and spot at 1000, Table 7.1 shows the pricing error of a call option under each of the pricing approaches - the error being expressed relative to the true price.

parameterization	Steps: 5	Steps: 15	Steps: 25	Steps: 101
Cox et al.	1.39 %	0.46 %	0.27 %	0.06 %
Tian	0.70 %	0.18 %	0.09 %	0.14 %
Jarrow and Rudd	1.46 %	0.41 %	0.48 %	0.11 %
Leisen and Reimer	4.36 %	1.60 %	0.97 %	0.24 %

Table 7.1: Errors in valuation - vanilla European option, at the money.

Changing the strike to 750, yields pricing errors as shown in Table 7.2.

parameterization	Steps: 5	Steps: 15	Steps: 25	Steps: 101
Cox et al.	0.41 %	0.09 %	0.08 %	0.01 %
Tian	0.26 %	0.04 %	0.04 %	0.01 %
Jarrow and Rudd	0.01 %	0.09 %	0.04 %	0.003 %
Leisen and Reimer	0.35 %	0.12 %	0.06 %	0.01 %

Table 7.2: Errors in valuation - vanilla European option, 125% in the money.

parameterization	Steps: 5	Steps: 15	Steps: 25	Steps: 101
Cox et al.	4.89 %	0.84 %	1.05 %	0.68 %
Tian	6.70 %	2.06 %	1.98 %	0.46 %
Jarrow and Rudd	3.59 %	1.21 %	0.71 %	0.27 %
Leisen and Reimer	15.22 %	6.032 %	3.76 %	0.97 %

Table 7.3: Errors in valuation - vanilla European option, 75% moneyness.

For 25% out of the money call options, the enumerated approximations result in valuation errors as shown in Table 7.3.

7.4 Summary

This, the final chapter concerned itself with some numerical eccentricities associated with binomial methods. To be more precise: the convergence characteristics as well as the potential for unpredictable behaviour of the various tree methods were discussed. Very importantly, a rigorous definition for order (speed) of convergence was given, and it was shown that all parameterizations converge pretty much at the same rate.

Chapter 8

Conclusion

The binomial tree method, as used in the pricing of financial derivatives caused somewhat of a revolution in the financial quantitative arena. It not only presented an easy to understand and intuitive framework within which financial instruments can be priced, but also proved general enough to be applied to a very wide and diverse range of instrument types. It is now close to 30 years since the methodology was first posited, it remains a mainstay of financial engineering. This dissertation aimed towards a providing a brief overview of the technique and well as concerns its theoretical under built, the diversity of tree parametrizations, consistency of binomial methods with other approaches as well as the application of tree evaluations in different pay-off and market assumption scenarios. In terms of the theoretical under built, an introduction to random walk models and the concept of no-arbitrage was given in the first chapter. Following the arguments put forward by Cox et al (1979), [23], it was shown that based on some very intuitive reasoning that the value of a derivative could be expressed in terms of a discounted expected value, given the assumption of Bernoulli outcomes for the underlying's price. Quite surprising was the fact that the probabilities associated with the outcomes were not explicitly defined, but rather followed from the no-arbitrage assumption and the magnitude of price movements over small time intervals. This fascinating result (termed in this document the "Fundamental Theorem of Binomial Pricing") applied to random walks provides the basic framework within which tree pricing takes place.

The following chapter dealt with the various incarnations, as it were, of binomial methods. As was mentioned in the very beginning of the said, since the seminal publication of Cox et al, numerous articles have appeared further exploring the tree pricing concept. The most well-known of these, in terms of the construction methods proposed, (and more specifically those put forward by Jarrow and Rudd, Trigeorgis,

Tian and Leisen and Reimer) were analyzed as concerns the assumptions made, the derivation of the parameters as well as some interesting characteristics resulting from the particular choices. As a case in point, it was mentioned that the Jarrow and Rudd method violates put-call parity, and therefore could not be considered a strictly arbitrage-free model.

Chapter Four treated on the consistency of binomial models with various other pricing approaches. It was shown that the price for a vanilla European option as implied by the Cox et al tree always converges to the closed-form solution provided by Black and Scholes; this was done by way of the central limit theorem, and well by first proving that the Lyapunov condition holds. Unconditional convergence (as concerns the derivative type and parameter choices) to the Black and Scholes partial derivative equation was also shown. The consistency of binomial and trinomial trees was briefly touched upon, whilst the fact that tree methods could be seen to be a type of numerical integration technique was also posited. Finally, some interesting remarks were made concerning the occurrence of a probability measure dichotomy under certain circumstances.

The application of binomial methods to real life pricing problems were investigated in Chapter Five. To be more precise, the problems associated with constructing trees consistent with term structures of risk free interest rates, pricing the cost of stock borrowing or interest (convenience yields) received on the underlying, adjusting the tree assumptions to take account of discrete dividends as well as building trees that make allowance for a term structure and skew of volatility were investigated.

Pricing exotic derivatives in a binomial tree framework was considered in Chapter Six. More specifically, American options, barrier types and forward start options were investigated. The final chapter dealt with stability and convergence of tree methods. In this chapter it was shown that Cox et al trees are only conditionally stable in behavior. The reason for instability under some circumstances was given as implied risk neutral probabilities assuming negative values. Furthermore it was shown that other tree parametrizations do not suffer similar problems, and can therefore be considered unconditionally stable. The issue of convergence, and more specifically convergence patterns were of interest in the last chapter. Mention was made of characteristic patterns in convergence (and errors); numerous references to the work of Leisen and Reimer and Diener and Diener were also made.

Bibliography

- [1] Andricopoulos A., Duck P., Newton D.P. and Widdicks M., "*Universal Option Valuation Using Quadrature*", Journal of Financial Economics, 67 (3), 2003.
- [2] Babbs S., "*Binomial valuation for look-back options*", Journal of Economic Dynamics & Control, 24 (2000) 1499 - 1525.
- [3] Bain L. and Engelhardt M., "*An Introduction to Probability Theory and its Applications*", Duxbury Press, (1992).
- [4] Bartle R. G. and Sherbert D. R., "*Introduction to Real Analysis*", Second Edition, John Wiley and Sons, 1992.
- [5] Barone-Adesi, G. and Whaley, R., "*Efficient analytic approximation of american option values*", Journal of Finance, (42) 301-320, (1987).
- [6] Bartter B. J. and Rendleman R.J., "*Two-State Asset Pricing*", Journal of Finance, December 1979, 1093-1110.
- [7] Beichelt F. E. and Fatti L. P., "*Stochastic Processes and Their Applications*", Taylor and Francis, First Edition, (2002).
- [8] Black F., "*The Pricing of Commodity Contracts*", Journal of Financial Economics, March 1976, 167-79.
- [9] Black F. and Scholes M., "*The pricing of options and corporate liabilities*", Journal of Political Economy, 81 (1973) 637-659.
- [10] Barle S. and Cakici N., "*How to grow a Smiling Tree*", The Journal of financial Engineering. 7(2), 1994, 127-146.
- [11] Boyle P. and Lau S. H., "*Bumping Up Against the Barrier with the Binomial Method*", Journal of Derivatives, 1 (Summer 1994) 6-14.
- [12] Boyle P., "*A Lattice Framework for Option Pricing with Two State Variables*", Journal of Financial and Quantitative Analysis, 23 (1), March (1988) 1-12.

- [13] Breeden D. and Litzenberger R., "*Prices of State-Contingent Claims Implicit in Options Prices*", Journal of Business, 51 (4), (1978) 621-651.
 - [14] Brennan M. J. and Schwartz S., "*Finite Difference Methods and Jump Processes Arising in the Pricing of Contingent Claims : A Synthesis*", Journal of Financial and Quantitative Analysis, September 1978, 461-474.
 - [15] Brennan M. J. and Schwartz E. S., "*The Valuation of American Put Options*", Journal of Finance, 32, May 1977, 79-93.
 - [16] Broadie M., Glasserman P. and Kou S. G., "*Connecting Discrete and Continuous Path-Dependent Options*", Finance and Stochastics, 3, 1999, 55-82.
 - [17] Cakici N. and Foster K. R., "*Implied Binomial Trees from the Historical Distribution*", Working Paper, February 2002.
 - [18] Camara A. and Chung S., "*Pricing Options with transformed-Binomial processes and their limiting distributions*", Working Paper, (2003).
 - [19] Camp B. H., "*Approximation to the Point Binomial*", Annals of Mathematical Statistics, 22, 1951.
 - [20] Chung S. and Shackleton M. B., "*Toward Option Values of Near Machine Precision using Gaussian Quadrature*", (2004).
 - [21] Clewlow L. and Strickland C., "*Implementing Derivatives Models*", John Wiley and Sons, (1998).
 - [22] Cong J. X., "*Historical Development of the Central Limit Theorem*", February 2003
 - [23] Cox J., Ross S.A., Rubinstein M., "*Option Pricing: A Simplified Approach*", Journal of Financial Economics, 7 (1979) 229-263.
 - [24] Das S. R., "*Random Lattices for Option Pricing Problems in Finance*", Working Paper, (2002).
 - [25] Derman E., "*Regimes of Volatility*", Risk, April 1999.
 - [26] Derman E. and Kani I. "*The Ins and Outs of Barrier Options*", Goldman Sachs Quantitative Strategies Research Notes, June 1993.
 - [27] Derman E., Kani I. and Chriss N., "*Implied Trinomial Trees of the Volatility Smile*", The Journal of Derivatives, 3(4) 722, 1996.
-

- [28] Derman E., Kani I., Ergener D., Bardhan I. *"Enhanced Numerical Methods for Options with Barriers"*, Goldman Sachs Quantitative Strategies Research Notes, May 1995.
 - [29] Derman E., Kani I. and Zou J. Z., *"The Local Volatility Surface"*, Goldman Sachs Quantitative Strategies Research Notes, December 1995.
 - [30] Derman E. and Kani I., *"Riding on a Smile"*, Risk, 7, 1994, 32-39.
 - [31] Diener F. and Diener M., *"Asymptotics of the Price Oscillations of a European Call Option in a Tree Model"*, Journal of Mathematical Finance, Vol. 14, No. 2 (2004) 271-293.
 - [32] Dupire B., *"Pricing with a smile,"* RISK, 7(1), 1820, (1994)
 - [33] Feller W., *"An Introduction to Probability Theory and its Applications"*, John Wiley and Sons, (1957).
 - [34] Frishling V., *"A Discrete Question"*, Risk, January 2002.
 - [35] Geske R. and Shastri K., *"Valuation by Approximation : A Comparison of Alternative Option Valuation Techniques"*, Journal of Financial and Quantitative Analysis, 20 (1), 1985, 45-71.
 - [36] Gatheral J., *"Volatility and Hedging Errors"*, Merrill Lynch (Presentation), 1999.
 - [37] Goodman J., Moon K., Szepessy A., Tempore R. and Zouraris G. *"Stochastic and Partial Differential Equations with Adapted Numerics"*, January 2004.
 - [38] Hanson F. B. and Westman J. J., *"Jump-Diffusion Stock Return Models in Finance: Stochastic Process Density with Uniform-Jump Amplitude"*, Working Paper.
 - [39] Harrison J. and Pliska S., *"Martingales and Stochastic Integrals in the Theory of Continuous Trading"*, "Stochastic Processes and their Applications", 1981.
 - [40] Haug E. G., *"Complete Guide to Option Pricing Formulas"*, 1997.
 - [41] Heath D. C. and Jarrow R. A., *"Ex-dividend stock price behavior and arbitrage opportunities"*, Journal of Business (61) 95-108, (1988).
 - [42] Heston S. and Zou G., *"On the Rate of Convergence of Discrete-Time Contingent Claims"*, Mathematics of Finance 10 (1), 53-75, (2000).
 - [43] Higham J. D., *"Nine Ways to Implement the Binomial Method for Option Valuation in Matlab"*, SIAM Review, 44 no. 4, (2002).
-

- [44] Hilliard E. and Schwartz A., "*Binomial Option Pricing under Stochastic Volatility and Correlated State Variables*", *Journal of Derivatives*, Fall 1996, 23-39.
 - [45] Hull J.C., "*Options, Futures and other Derivatives*", John Wiley and Sons, (1998).
 - [46] Hull J. C. and White A. , "*Efficient Procedures for Valuing European and American Path-Dependent Options*", *The Journal of Derivatives*, (Fall), (1993).
 - [47] Jarrow R. and Rudd A., "*Option Pricing*", Homewood, Illinois (1983).
 - [48] Jarrow R. and Turnbull S., "*Derivative Securities*", South-Western Publishing, (1995).
 - [49] Joshi M. S., "*Pricing Discretely Sampled Path-Dependent Exotic Options using Replication Methods*", Working Paper
 - [50] Just Joking, "*Bluff your way through Black's Schools*", *Journal of Financial Gymnastics*, 1 (1) 2006.
 - [51] Leisen D.P.J., "*Pricing the American put option: A detailed convergence analysis for Binomial models*", *Journal of Economic Dynamics & Control*, 22 (1998) 1419 - 1444.
 - [52] Leisen D.P.J. and Reimer M., "*Binomial Models for Option Valuation - Examining and Improving convergence*", *Journal of Applied Mathematical Finance*, 3 (1996) 319-346.
 - [53] Leland H. E. and Rubinstein M. E., "*Replicating Options with Positions in Stock and Cash*", *Financial Analysts Journal*, January-February 1995.
 - [54] Margrabe W., "*Binomial Pricing of Exploding Options*", *Topics in Money and Securities Markets*, 149, Bankers Trust Company, (1989).
 - [55] Merton R. C., "*Theory of Rational Option Pricing*", *Bell Journal of Economics and Management Science*, 4 (1), (1973), 141-183.
 - [56] Merton R. C., "*Option Pricing when Underlying Stock Returns are Discontinuous*", *Journal of Financial Economics*, January-March 1976.
 - [57] Musiela M. and Rutkowski M., "*Martingale Methods in Financial Modelling*", Springer-Verlag, 1997.
 - [58] Muzzioli S. and Torricelli C., "*A multi-period Binomial model for pricing options in a vague world*", *Journal of Economic Dynamics & Control*, 28 (2004) 861 - 887.
-

- [59] Neftci S., *"An Introduction to the Mathematics of Financial Derivatives"*, Academic Press, (1996).
 - [60] Omberg C., *"A Note on the Convergence of Binomial and Compound Option Models"*, Journal of Finance, 42, 1987.
 - [61] Parkinson M., *"Option Pricing : The American Put"*, Journal of Business, January 1977, 21-36.
 - [62] Paulson E., *"An Approximate Normalisation of the Analysis of Variance Distribution"*, Annals of Mathematical Statistics, 13, 1942, 233-235.
 - [63] Peizer D. B. and Pratt J. W., *"A Normal Approximation for Binomial, F, Beta and Other Common, Related Tail Probabilities"*, The Journal of the American Statistical Association, 63, 1968, 1416-1456.
 - [64] Rachev S.T. and Ruschendorff L., *"Models for Option Prices"*, Working Paper.
 - [65] Reiner E. and Rubinstein M. E., *"Breaking Down the Barrier"*, Risk, 4 (8), (1991), 28-35.
 - [66] Rendleman R. J. and Bartter B. J., *"The Pricing of Options on Debt Securities"*, Journal of Financial and Quantitative Analysis, 15 (1), (1980), 11-23.
 - [67] Rogers L.C.G. and Stapleton E.J., *"Fast Accurate Binomial Pricing"*, Finance and Stochastics, 2 (1998), 3-17.
 - [68] Roll R., *"An Analytical Valuation Formula for Unprotected American Call Options on Stocks with Known Dividends"*, Journal of Financial Economics, November 1977.
 - [69] Ross S., *"Options and Efficiency"*, Quarterly Journal of Economics, 90 (1), (1976), 75-89.
 - [70] Rubinstein M. E., *"Option Models"*, Unpublished Book.
 - [71] Rubinstein M. E., *"On the Relation between Binomial and Trinomial Option Pricing Models"*, UC Berkley, Research Program in Finance (Working Paper RPF-292), (2000).
 - [72] Rubinstein M. E., *"Edgeworth Binomial Trees"*, Journal of Derivatives, Spring (1998).
 - [73] Rubinstein M. E., *"Implied Binomial Trees"*, Journal of Finance, July (1994).
 - [74] Rudin W., *"Principles of Mathematical Analysis"*, Second Edition, MacGraw-Hill Book Company, 1964.
-

- [75] Schwartz E. S., *"The Valuation of Warrants : Implementing a New Approach"*, Journal of Financial Economics, January 1977, 79-93.
 - [76] Sharpe W. F., *"Investments"*, Englewood Cliffs, New Jersey, Prentice Hall, 1978.
 - [77] Stapleton R.C. and Subrahmanyam M.G., *"The Value of Options when Asset Returns are Generated by a Binomial Process"*, The Journal of Finance, December 1984, 1525-1540.
 - [78] Stulz R.M., *"Principles of Infinitesimal Stochastic and Financial Analysis"*, World Scientific, 2000.
 - [79] Tian Y., *"A Modified Lattice Approach to Option Pricing"*, Journal of Futures Markets, Vol. 13, No. 5 (1993) 563-577.
 - [80] Tian Y.S., *"A Trinomial Option Pricing Model Dependent on Skewness and Kurtosis"*, International Review of Economics and Finance, 7 (3), (1998), 3115-330.
 - [81] Trigeorgis L., *"A Log-Transformed Binomial Numerical Analysis Method for Valuing Complex Multi-Option Investments"*, Journal of Financial and Quantitative Mathematics, 3 (1991) 309-326.
 - [82] Tucker A., *"Applied Combinatorics"*, John Wiley and Sons, Third Edition, 1995.
 - [83] Van den Berg I., *"Principles of Infinitesimal Stochastic and Financial Analysis"*, World Scientific, 2000.
 - [84] Whaley R. E., *"On Valuing American Futures Options"*, Working Paper, 1984, University of Alberta.
 - [85] Wilmot P., Dewynne J. and Howison S., *"Option Pricing"*, Oxford Financial Press, 1993.
 - [86] *"Limit Distributions"*, www.statwww.epfl.ch
-

Appendix - Matlab Code

The graphs and examples included in the dissertation were constructed using Matlab code implementations (done by myself) of the various pricing approaches and algorithms. It is shown below.

```
function [a, b] = BinParameters(RFR, DeltaT, Sigma, DY, Parametrisation, CV, TTM, Spot, N)
%Returns the up and down parameters for a binomial tree

%Parametrisation is an integer with significance:
%0 - Cox Ross Rubinstein (Default), 1 - Tian Binomial, 2 - Jarrow Rudd
%3 - Trigeorgis, 4 - Leisen Reimer

if Parametrisation == 1
    M = exp(RFR * DeltaT);
    V = exp((Sigma^2) * DeltaT);
    a = 0.5 * M * V * ((V + 1) + ((V ^ 2 + 2 * V - 3)^0.5));
    b = ((M * V) ^ 2) / a;
elseif Parametrisation == 2
    v = (RFR - DY - 0.5 * (Sigma ^ 2));
    a = exp(v * DeltaT + Sigma * (DeltaT ^ 0.5));
    b = exp(v * DeltaT - Sigma * (DeltaT ^ 0.5));
elseif Parametrisation == 3
    a = exp(((Sigma^2) * DeltaT + (DeltaT * (RFR - DY - 0.5*(Sigma ^ 2))) ^ 2) ^ 0.5);
    b = 1 / a;
elseif Parametrisation == 4
% Applies the Camp-Paulson continuity correction
    R = exp(RFR * DeltaT);
    d1 = DBS1(Spot, CV, RFR, DY, TTM, Sigma);
    d2 = DBS(Spot, CV, RFR, DY, TTM, Sigma);
    a1 = ceil(N / 2); a2 = floor(a1);
    x = ((a2 / a1) ^ 2) * (((9 * a1 - 1) * (9 * a2 - 1) + 3 * d1 * ((a1 * (9 * a2 - 1) ^ 2 + ...
    a2 * (9 * a1 - 1) ^ 2 - 9 * a1 * a2 * (d1 ^ 2)) ^ 0.5)) / ((9 * a2 - 1) ^ 2 - ...
    9 * a2 * (d1 ^ 2))) ^ (1 / 3));
    y = ((a2 / a1) ^ 2) * (((9 * a1 - 1) * (9 * a2 - 1) + 3 * d2 * ((a1 * (9 * a2 - 1) ^ 2 + ...
    a2 * (9 * a1 - 1) ^ 2 - 9 * a1 * a2 * (d2 ^ 2)) ^ 0.5)) / ((9 * a2 - 1) ^ 2 - ...
    9 * a2 * (d2 ^ 2))) ^ (1 / 3));
    a = R * x / y;
    b = (R - y * a) / (1 - y);
elseif Parametrisation == 5
% Applies the Peizer-Pratt 1 continuity correction
    R = exp(RFR * DeltaT);
    d1 = DBS1(Spot, CV, RFR, DY, TTM, Sigma);
    d2 = DBS(Spot, CV, RFR, DY, TTM, Sigma);
    x = 0.5 + sign(d1) * 0.5 * ((1 - exp(-((d1 / (N * 1 / 3)) ^ 2) * (N + 1 / 6))) ^ 0.5);
    y = 0.5 + sign(d2) * 0.5 * ((1 - exp(-((d2 / (N + 1 / 3)) ^ 2) * (N * 1 / 6))) ^ 0.5);
    a = R * x / y;
    b = (R - y * a) / (1 - y);
elseif Parametrisation == 6
% Applies the Peizer-Pratt 2 continuity correction
    R = exp(RFR * DeltaT);
    d1 = DBS1(Spot, CV, RFR, DY, TTM, Sigma);
    d2 = DBS(Spot, CV, RFR, DY, TTM, Sigma);
    x = 0.5 + sign(d1) * 0.5 * ((1 - exp(-((d1 / (N + 1 / 3 + 0.1 / (N + 1))) ^ 2) * (N + 1 / 6))) ^ 0.5);
    y = 0.5 + sign(d2) * 0.5 * ((1 - exp(-((d2 / (N + 1 / 3)) ^ 2) * (N + 1 / 6))) ^ 0.5);
    a = R * x / y;
    b = (R - y * a) / (1 - y);
else
    a = exp(Sigma * (DeltaT ^ 0.5));
    b = 1 / a;
end
```

end

```
function f = BinTreeBuilder(Spot,Sigma,TimeVector,RFR,DY,DType,Parametrisation,CentralVal,DivSched)
```

%TimeVector a 2-dim matrix that explicitly define the number of timesteps
%to use for a set of time intervals

%Parametrisation is an integer with significance:
%0 - Cox Ross Rubinstein, 1 - Tian Binomial, 2 - Jarrow Rudd
%3 - Trigeorgis, 4 - Leisen Reimer

%DType a relative / yield dividend flag; 0 for a pure dividend yield, 1 for
%a proportional discrete dividend, explicitly defined discrete dividends
%are specified in the DivSched matrix

```
B(1,1) = Spot; Rho = []; DeltaT = [];
n = TimeVector(2,:) * ones(size(TimeVector,2),1) + 1;
TimeVector(2,2) = TimeVector(2,2) + 1;
PartialC = 1; dy = DY;
```

```
if DType == 1
    dy = 0;
else
    DY = 0;
end
```

```
for j=1:size(TimeVector,2)-1
    DeltaT(PartialC) = (TimeVector(1,j+1) - TimeVector(1,j)) / TimeVector(2,j+1);
    NewSC = 1;
    if j==1
        DeltaT(PartialC) = TimeVector(1,j+1) / (TimeVector(2,j+1)-1);
        NewSC = 2;
    end
    if Parametrisation >= 4
        [u, d] = BinParameters(RFR, DeltaT(PartialC), Sigma, dy, Parametrisation, CentralVal, TimeVector(1,j+1)-TimeVector(1,j), Spot, n);
    else
        [u, d] = BinParameters(RFR, DeltaT(PartialC), Sigma, dy, Parametrisation);
    end
    Rho(PartialC) = (exp(RFR*DeltaT(PartialC)) - d) / (u - d);
    for k=NewSC:TimeVector(2,j+1)
        PartialC = TimeVector(2,1:j) * ones(j,1) * k;
        DeltaT(PartialC) = DeltaT(PartialC-1);
        Rho(PartialC) = Rho(PartialC-1);
        for i=1:PartialC-1
            B(i,PartialC) = B(i,PartialC-1) * u;
        end
        B(PartialC,PartialC) = B(PartialC-1,PartialC-1) * d;
        if k > NewSC
            B(:,k) = B(:,k) * (2 - exp(DY * DeltaT(PartialC)));
        end
    end
end
end

if size(DivSched,1) > 0
    for k=1:PartialC
        for i=1:size(DivSched,1)
            if sum(DeltaT(2:1:k)) >= DivSched(i,1)
                B(:,k) = max(B(:,k) - DivSched(i,2), 0);
            end
        end
    end
end
end

f = B;
```

```
function f = VanillaTreeTraverse(B,Strike,TTM,RFR,DY,DivSched,DType,Sigma,CallOptionFlag,Parametrisation)
```

%Traverse the Tree specified by B, using the pricing pay-off logic of a
%Vanilla European Option

%TTM is a 1x2 vector with the first element being the Start Date and the second obviously the Termination Date

```
TimeSteps = size(B,2); DeltaT = TTM(end-1) / (TimeSteps-1); Omega = 1;
if CallOptionFlag == 0
    Omega = -1;
end
if DType == 1
    DY = 0;
end
if size(DivSched,1) > 0
    DS = [DivSched(:,1) - TTM(1), DivSched(:,2)];
```

```

else
    DS = [];
end

if Parametrisation == 0
    Rho = 0.5 * (1 + (RFR - 0.5 * (Sigma ^ 2)) * (DeltaT ^ 0.5) / Sigma);
elseif Parametrisation == 2
    Rho = 0.5;
else
    Rho = (exp((RFR - DY) * DeltaT) - B(2,2)/B(1,1)) / (B(1,2)/B(1,1) - B(2,2)/B(1,1));
end

for k=TimeSteps:-1:1
    for i=1:k
        if k == TimeSteps
            if B(i,k) == 0
                TempVal = Strike * CallOptionFlag;
            else
                TempVal = max(Omega*(B(i,k) - Strike),0);
            end
            V(i,k) = TempVal;
        else
            V(i,k) = (Rho * V(i,k+1) + (1-Rho) * V(i+1,k+1)) / exp(RFR * DeltaT);
        end
    end
end

f = V(1,1);

function [f, g] = VanillaOptionMI(Spot, Strike, Sigma, TTM, RFR, DY, DivSched, CallOptionFlag, NSIM, IntMethod)
% Returns the value of a Vanilla European Option where the underlying pays
% Discrete Dividends ; use is made of a Numerical Integration Procedure (Simpson's rule used by default);
% the second output argument is an upper bound on the discretisation error.

% Determine firstly the dates upon which discontinuities occur in the
% stock price process.

if isempty(DivSched) == 0
    Counter = 0; DSI = [];
    for k=1:size(DivSched,1); if DivSched(k,1) >= 0; Counter = Counter + 1; DSI(Counter, :) = DivSched(k,:); end; end
    DivSched = DSI;
    TM = CollateDates([TTM 0], DivSched, 1);
    DivSched = [0 0; TM];
else
    DivSched = [0 0; TTM 0];
end

% Build the grid upon which the conditional integration is to be based.
SpotGrid = []; NSIM = 2 * floor(NSIM / 2);
for k=2:size(DivSched,1)
    TempMean = ForwardPrice(Spot, 0, DivSched(k,1), DY, DivSched, RFR);
    TempSD = TempMean * ((exp((Sigma ^ 2) * DivSched(k,1)) * (exp((Sigma ^ 2) * DivSched(k,1))-1)) ^ 0.5);
    TempInterval(1) = max(TempMean - 3 * TempSD, 0);
    if (TempMean - 3 * TempSD) < 0; TempInterval(2) = TempMean + 4 * TempSD;
    else; TempInterval(2) = TempMean + 3 * TempSD; end
    DeltaX(k) = (TempInterval(2) - TempInterval(1)) / NSIM;
    SpotGrid(:,k) = (TempInterval(1): DeltaX(k): TempInterval(2))';
end
SpotGrid(:,1) = Spot * ones(size(SpotGrid,1),1);
ValueGrid = zeros(size(SpotGrid));
ValueGrid(:,end) = max(SpotGrid(:,end) - Strike,0);
if CallOptionFlag == 0; ValueGrid(:,end) = max(Strike - SpotGrid(:,end),0); end
% Calculate the Conditional Expectations numerically.
for k=size(DivSched,1)-1:-1:1
    DeltaT = DivSched(k+1,1) - DivSched(k,1);
    z2 = ValueGrid(:,k+1);
    if k==1; N=1; else; N = size(SpotGrid,1); end
    for j=1:N
        MU = log(max(SpotGrid(j,k)-DivSched(k,2),0)) + (RFR - DY - 0.5 * (Sigma ^ 2)) * DeltaT;
        SIGMA = Sigma * (DeltaT ^ 0.5);
        z1 = lognpdf(SpotGrid(:,k+1), MU, SIGMA);
        f = z1 .* z2;
        if IntMethod == 0
            % Simpson's rule is used.
            ValueGrid(j,k) = DeltaX(k+1) / 3 * (sum(f(1:2:end-2)) + 4 * sum(f(2:2:end-1)) + ...
            sum(f(3:2:end)));
        elseif IntMethod == 1
            % Trapezoid rule is used.
            ValueGrid(j,k) = sum((f(1:end-1) + f(2:end)) / 2 * DeltaX(k+1));
        elseif IntMethod == 2
            % Rectangular rule is used.
            ValueGrid(j,k) = sum(f(1:end) * DeltaX(k+1));
        end
    end
end
end

```

```

end

f = ValueGrid(1,1) / exp(RFR * TTM);
g = 0;

function f = IntegralEval(Funksie, FunksieArg, IntDomain, N, Method)
% Returns the definite integral value of the function <<Funksie>>, over the
% Domain <<IntDomain>>; the <<Method>> parameter has the following significance:
% 0 - Trapezoid rule (default), 1 - Gauss Quadrature,
% 2 - Simpson's rule is used, 3 - Simple (rectangular) method.

% For multivariate functions specify the value of parameters other than the
% integration variable in <<FunksieArg>>.

if nargin<4
    error('?? Not a sufficient number of input parameters')
else
    if isempty(N); N = 1000; end
    deltax = (IntDomain(2) - IntDomain(1)) / N;
    if or(nargin == 4, Method == 0)
        x = IntDomain(1):deltax:IntDomain(2);
        % z = feval(Funksie,x',FunksieArg);
        z = feval(Funksie,x');
        y = sum((z(1:end-1) + z(2:end)) / 2 * deltax);
    elseif Method == 1
        x = deltax: deltax: (IntDomain(2) - IntDomain(1));
        % z1 = feval(Funksie,(0.5 + sqrt(3) / 6) * x' + IntDomain(1),FunksieArg);
        % z2 = feval(Funksie,(0.5 - sqrt(3) / 6) * x' + IntDomain(1),FunksieArg);
        z1 = feval(Funksie,(0.5 + sqrt(3) / 6) * x' + IntDomain(1));
        z2 = feval(Funksie,(0.5 - sqrt(3) / 6) * x' + IntDomain(1));
        y = deltax / 2 * (sum(z1) + sum(z2));
    elseif Method == 2
        N = 2 * floor(N / 2);
        deltax = (IntDomain(2) - IntDomain(1)) / N;
        x = IntDomain(1): deltax: IntDomain(2);
        % z = feval(Funksie,x',FunksieArg);
        z = feval(Funksie,x');

        y = deltax / 3 * (sum(z(1:2:end-2)) + 4 * sum(z(2:2:end-1)) + ...
            sum(z(3:2:end)));
    elseif Method == 3
        x = IntDomain(1):deltax:IntDomain(2);
        z = feval(Funksie,x',FunksieArg);
        y = sum(z(1:end) * deltax);
    end
end

f = y;

function f = DES1(Spot, Strike, RFR, DY, TTM, Sigma)
%Returns d1 as specified in the BS function

f = log(Spot ./ Strike) + (RFR - DY + 0.5 * (Sigma.^ 2)) .* TTM;
f = f ./ (Sigma * (TTM.^ 0.5));

function f = DES(Spot, Strike, RFR, DY, TTM, Sigma)
%Returns d2 as specified in the BS function

f = log(Spot ./ Strike) + (RFR - DY - 0.5 * (Sigma.^ 2)) .* TTM;
f = f ./ (Sigma * (TTM.^ 0.5));

function f = BSPrice(Spot, Strike, RFR, DY, TTM, Sigma, CallOptionFlag)
%Calculates the analytical delta of a Vanilla European Option
TTM = max(TTM,0); Spot = max(Spot, 0.01);
if TTM == 0
    f = max(Spot - Strike, 0);
    if CallOptionFlag == 0
        f = max(Strike - Spot,0);
    end
else
    f = Spot .* exp(-DY .* TTM) .* normcdf(DES1(Spot, Strike, RFR, DY, TTM, Sigma));
    f = f - Strike .* exp(-RFR .* TTM) .* normcdf(DES(Spot, Strike, RFR, DY, TTM, Sigma));
    if CallOptionFlag == 0
        f = f + Strike .* exp(-RFR .* TTM) - Spot .* exp(-DY .* TTM);
    end
end

end

```