UNIVERSITY OF CYPRUS

SCHOOL OF ECONOMICS AND MANAGEMENT

DEPARTMENT OF PUBLIC AND BUSINESS ADMINISTRATION

OPTION PRICING USING NONLINEAR CONSTRAINED OPTIMIZATION ON IMPLIED NON-RECOMBINING TREES

DOCTORAL DISSERTATION

ELENI D. CONSTANTINIDE

2008
Option Pricing Using Nonlinear Constrained Optimization on Implied Non-Recombining Trees

Eleni D. Constantinide ©

A dissertation submitted to the Department of Public and Business Administration of the University of Cyprus in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Finance

Nicosia
April, 2008
Περίληψη

Σκοπός της διατριβής αυτής είναι η δημιουργία μοντέλων για εκτίμηση των παραμέτρων ενός δωδεκικού (binary) δέντρου χρησιμοποιώντας τιμές αγοράς (market prices) Ευρωπαϊκών συμβολαίων δικαιωμάτων προαίρεσης (European options). Για να εκτιμήσουμε την υποοούμενη κατανομή (implied distribution), ελαχιστοποιούμε τη διαφορά μεταξύ των τιμών της αγοράς και των τιμών που εκτιμώνται από το μοντέλο υπό περιορισμούς που διατηρούν τις πιθανότητες μετάβασης του δέντρου καλά ορισμένες και εμποδίζουν τις ευκαιρίες καυσωνικότητας (arbitrage opportunities). Το υπό μελέτη πρόβλημα είναι μη γραμμικό (μη-κυρτό) πρόβλημα βελτιστοποίησης με γραμμικούς περιορισμούς. Χρησιμοποιούμε Εξωτερική Μέθοδο Πέναλτι (Exterior Penalty Method) και για τη βελτιστοποίηση χρησιμοποιούμε τον αλγόριθμο Quasi-Newton. Η πολύτλοκη φύση του δέντρου και ο μεγάλος αριθμός περιορισμών, κάνουν το ψάξιμο για βέλτιστη λύση και την επιλογή ενός αλγόριθμου που να λειτουργεί καλά ένα πρόβλημα πρόκληση. Στο πρώτο μοντέλο (INRT), θεωρούμε ότι η μεταβλητή του συστήματος είναι το υποκείµενο αγαθό (underlying asset) σε κάθε κόμβο του δέντρου. Με αυτόν τον τρόπο παίρνουμε ευέλικτη κατανομή για το υποκείµενο αγαθό και επομένως ευέλικτη κατανομή για την τοπική διακύμανση (local volatility). Στο δεύτερο μοντέλο θεωρούμε ότι η μεταβλητή του συστήματος είναι η τοπική διακύμανση (INRT_LocalVol). Με αυτόν τον τρόπο μειώνουμε τον αριθμό των μεταβλητών στις μισές σε σχέση με το INRT μοντέλο κάνοντας το μοντέλο λιγότερο υπολογιστικά φορτισμένο. Επιπλέον, οι τοπικές διακυμάνσεις μπορούν εύκολα να ελεγχθούν με την επιμορφωμένη κατανομή για το υποκείµενο αγαθό και επομένως για την τοπική διακύμανση (local volatility). Στο τρίτο μοντέλο επιβάλουμε μη-Μαρκοβιανές υποθέσεις στο μοντέλο INRT, θεωρώντας ότι οι αποδόσεις (returns) έχουν σειριακή εξάρτηση (NMT). Αυτό το πλαίσιο δίνει ένα μοντέλο πιο πλούσιο και πιο πρακτικό από τα άλλα μοντέλα δέντρων που προσαρμόζονται στις τιμές αγοράς (implied trees) και από τα μοντέλα τιμολογίας συμβολαίων δικαιωμάτων προαίρεσης σε συνεχές χρόνο (continuous option pricing models) που υπάρχουν στην βιβλιογραφία. Ελέγχουμε τα μοντέλα χρησιμοποιώντας Ευρωπαϊκά δικαιώματα προαίρεσης αγοράς (call options) στο δείγμα FTSE 100 για το χρόνο 2003. Τα αποτελέσματα υποστηρίζουν τα προτεινόμενα μοντέλα. Τα αποτελέσματα εκτίμησης είναι ομαλά (smooth) χωρίς την παρουσία του προβλήματος υπέρ-εφαρμογής (overfitting) και οι εκτιμημένες κατανομές είναι πιο πρακτικές. Επιπλέον, το υπολογιστικό χρόνο δεν είναι μεγάλο. Για την εκτίμηση
Αμερικάνικων δικαιωμάτων προαίρεσης αγοράς (American call options) το μοντέλο INRT είναι καλύτερο από μια βελτιωμένη εκδοχή διωνυμικού δέντρου που προσαρμόζεται στις τιμές αγοράς (implied binomial tree) που προτάθηκε από τους Derman και Kani (1994) και είναι εξίσου καλό με το μοντέλο INRT_LocalVol και με την ad-hoc μέθοδο του Dumas και άλλων (1998) για ομαλοποίηση των υποοφεύγουν διακυμάνσεων (implied volatilities) Black-Scholes. Το μοντέλο NMT είναι καλύτερο από το μοντέλο INRT στην εκτίμηση Αμερικάνικων δικαιωμάτων προαίρεσης αγοράς και επίσης από την ad-hoc μέθοδο. Στα προτεινόμενα μοντέλα μπορούν να χρησιμοποιηθούν Ευρωπαϊκά συμβόλαια δικαιωμάτων προαίρεσης με την ίδια ημερομηνία εξάσκησης και με μικρές τροποποιήσεις μπορούν να χρησιμοποιηθούν συμβόλαια δικαιωμάτων προαίρεσης με διαφορετικές ημερομηνίες εξάσκησης. Τα προτεινόμενα μοντέλα παρέχουν δέντρα συνεπή με την αγορά και μπορούν να χρησιμοποιηθούν για την εκτίμηση εξωτικών και Over The Counter συμβολαίων δικαιωμάτων προαίρεσης καθώς και συμβολαίων δικαιωμάτων προαίρεσης υπό την ύπαρξη συναλλαγματικών κόστων.
Abstract

The objective of the thesis is to develop models for calibrating a non-recombining (binary) tree using option market data. In order to capture the implied distribution, we minimize the discrepancy between the observed market prices and the model values, subject to constraints that keep the probabilities well specified and prevent arbitrage opportunities. The problem under consideration is a non-convex optimization problem with linear constraints. We adopt an exterior penalty method and for the optimization we use a Quasi-Newton algorithm. Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem. In the first model (INRT), we assume that the system variable is the underlying asset at each node of the tree. In this way we obtain flexible underlying asset distribution which implies flexible local volatility distribution. In the second model we assume that the system variable is the local volatility (INRT_LocalVol). In this way, we reduce the number of variables to half, in relation to the INRT model, making the model less computationally intensive. Moreover, local volatilities can be easily controlled by imposing suitable constraints. In the third model we impose non-Markovian assumptions on the INRT model by assuming serial dependence of returns (NMT). This setup gives a richer and more realistic model than the other implied trees and continuous option pricing models in the literature. We test the models using call options data of the FTSE 100 index for the year 2003. Results strongly support our modelling approaches. Pricing results are smooth without the presence of an overfitting problem and the derived distributions are realistic. Also, the computational burden is not a major issue. The INRT model outperforms an improved version of the implied recombining (binomial) tree proposed by Derman and Kani (1994) in pricing of American calls and is equally good as the INRT_LocalVol model and the ad-hoc method of Dumas et al. (1998) of smoothing Black-Scholes implied volatilities. The NMT model outperforms the INRT model in pricing of American calls and also the ad-hoc method. The proposed models can accommodate European options with single maturities and, with minor modifications, options with multiple maturities. They can provide market-consistent trees for option replication with transaction costs and can help pricing of exotic and Over The Counter options.
Acknowledgements

While reaching at the end of the long road of my PhD studies, I feel that I have to express my deep and sincere gratitude to some people. Firstly, I would like to thank my advisor Chris Charalambous, a man that I admire for his integrity as a person and for his widespread knowledge as an academician. He was the one that push me in the road for the PhD, and supported me in difficult times. His excellent supervision and support throughout the years of my PhD studies are highly appreciated.

Next, I would like to express my gratitude and thanks to my co-advisor Spiros Martzoukos, for his support and guidance throughout my dissertation.

I am also grateful for motivation and comments to Professor Nicos Christofides of Imperial College, London.

I would also like to thank my colleague Panayiotis Andreou for useful discussions and comments on my work.

I owe a big thank you to my parents, Dimitris and Areti Constantinide for the support, understanding and encouragement that gave me all these years. Especially I would like to thank my brother Christos that was always by my side when I needed him.

Last but not least I would like to thank the many friends and the faculty at the Department of Business Administration of the University of Cyprus that have been close to me during the years of my PhD studies.

I am also grateful for financial support from a research grant on Contingent Claims from the University of Cyprus.
To my parents,
Dimitris & Areti
Constantinide
# Table of Contents

Introduction ..............................................................................................................13  
1. Implied non-recombining trees and calibration for the volatility smile ..22  
Abstract ......................................................................................................................22  
1.1. Introduction .......................................................................................................23  
1.2. The proposed methodology and initialization of the non-recombining tree ...............................................................................................27  
1.3. Risk neutrality and no-arbitrage constraints...............................................33  
1.4. The optimization algorithm ............................................................................34  
Implementation ............................................................................................................35  
1.5. Application using FTSE 100 options data ....................................................38  
1.6. Conclusions........................................................................................................50  
Appendix 1.1A: Feasibility of the initialized non-recombining tree.............52  
Appendix 1.1B: Feasibility of the initialized non-recombining tree assuming time dependent \( r_f, \delta \) and \( \Delta t \) .........................................................55  
Appendix 1.2: Formulas adjusted for time dependent \( r_f, \delta \) and \( \Delta t \).................58  
2. An Empirical Comparison between Non-Recombining and Recombining Implied Trees....................................................................................................61  
Abstract ......................................................................................................................61  
2.1. Introduction .......................................................................................................62  
2.2. Data - FTSE 100 index options........................................................................65  
2.3. Research methodology.....................................................................................68  
2.3.1 Derman and Kani (DK, 1994) Implied Binomial Tree (IBT) .................68  
Additional heuristics on the modified DK algorithm ........................................73  
2.3.2 Implied Non-Recombining Tree (INRT) ........................................................74  
2.3.3. Ad-hoc model .....................................................................................................78  
2.4. Testing the models in pricing FTSE 100 American options .....................79  
2.5. Conclusions........................................................................................................84  
3. Calibration of non-recombining implied trees for the local volatility surface ........................................................................................................86  
Abstract ......................................................................................................................86  
3.1. Introduction .......................................................................................................87  
3.2. Description of the model.................................................................................90  
3.3. Analytical formulation of the problem.........................................................97  
3.4. Data - FTSE 100 index options......................................................................103  
3.5. Calibration and testing of the model ..........................................................107  
3.6. Pricing of American options .........................................................................112  
3.7. Conclusions......................................................................................................117  
Appendix 3.1: Derivation of partial derivative \( \frac{\partial p_{i,j}}{\partial \sigma_{i,j}} \), for \( i = 1, \ldots, n - 1 \), and \( j = 1, \ldots, 2^{i-1} \) .................................................................118
Appendix 3.2: Alternative approach for computing \( \frac{\partial C_{i,j}}{\partial \sigma_{i,j}} \), \( i = 1, \ldots, n - 1, j = 1, \ldots, 2^{i-1} \) .................................................................................................................. 119

4. Option pricing on non-recombining implied trees assuming serial dependence of returns .................................................................................................................. 123

Abstract ..................................................................................................................... 123

4.1. Introduction ......................................................................................................... 124

4.2. Description of the model .................................................................................. 127

Constraints .................................................................................................................. 132

The optimization problem ......................................................................................... 133

4.3. Initialization of the model .................................................................................. 134

4.4. Analytical formulation of the problem ............................................................... 136

4.5. Data - FTSE 100 index options ......................................................................... 139

4.6. Calibration and testing of the Non-Markovian (NMT) model.......................... 143

4.7. Testing the models in pricing FTSE 100 American options ............................ 152

4.8. Conclusions ...................................................................................................... 158

Appendix 4.1: Feasibility of the initialized non-recombining tree ......................... 159

Appendix 4.2A: Derivation of \( \frac{\partial C_{i,j}}{\partial S_{i+1,2^j}} \), \( i = n - 1, \ldots, 1, j = 1, \ldots, 2^{i-1} \) .................................................. 162

Appendix 4.2B: Derivation of \( \frac{\partial C_{i,j}}{\partial S_{i+1,2^{j-1}}} \), \( i = n - 1, \ldots, 1, j = 1, \ldots, 2^{j-1} \) ........................................ 163

Summary and Conclusions .......................................................................................... 164

References ................................................................................................................... 167
Table of Figures

Figure 1.1: Non-recombining tree with 4 steps. ...................................................28
Figure 1.2: A typical triplet in a non-recombining tree. ........................................28
Figure 1.3: Smoothing of the option pay-off function at maturity. .........................32
Figure 1.4: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003. ....................................................45
Figure 1.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the full sample for $n = 6$. ......................48
Figure 1.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the restricted sample for $n = 6$. ..........49
Figure 1.6a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the restricted sample for $n = 7$. ..........50
Figure 1.6b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the restricted sample for $n = 7$. ..........50
Figure 1.A1: A typical triplet in the initialization of the non-recombining tree assuming $r, \delta$ and $\Delta t$ to be time dependent. ..............................................55
Figure 2.1: The $i^{th}$ and the $(i+1)^{th}$ time steps on a binomial tree. .................69
Figure 2.2: A typical triplet in a non-recombining (binary) tree. ..........................75
Figure 3.1: Non-recombining tree with 4 steps. .................................................90
Figure 3.2: A typical triplet on a non-recombining tree. ....................................91
Figure 3.3: Last two steps on a non-recombining (binary) tree. .........................99
Figure 3.4: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003..................................................108
Figure 3.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) for $n=6$. ........................................................................................................111
Figure 3.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) for $n=7$. ........................................................................................................111
Figure 3.5c: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the proposed model (INRT_LocalVol) for $n=7$. ........................................................................................................112
Figure 3.6: Non-recombining tree with 4 steps. .................................................127
Figure 4.2a: Neighbour-nodes of the node $(i, j)$ for $j$ even on a non-recombining tree. ........................................................................................................128
Figure 4.2b: Neighbour-nodes of the node $(i, j)$ for $j$ odd on a non-recombining tree. ........................................................................................................128
Figure 4.3: Smoothing of the option pay-off function at maturity. .......................131
Figure 4.4a: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the implied non-recombining tree (INRT) for $n=7$. .........................................................144
Figure 4.4b: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the non-Markovian model (NMT) with correlation $-2.5\%$ and for $n=7$. .........................144
Figure 4.4c: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -5% and for n=7. .................................................................145

Figure 4.4d: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -7.5% and for n=7. .................................................................145

Figure 4.4e: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -10% and for n=7. .................................................................146

Figure 4.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) model for n=7. ........................................................................................................149

Figure 4.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -2.5% and for n=7. .................................................................150

Figure 4.5c: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -5% and for n=7. .................................................................150

Figure 4.5d: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -7.5% and for n=7. .................................................................151

Figure 4.5e: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -10% and for n=7. .................................................................151

Figure 4.6: Correlation versus the root mean square error (RMSE) smoothed using cubic spline. ........................................................................................................................................153
List of Tables

Table 1.1: Results for the application of the model on the 1st trading day of each month for the year 2003. .................................................................41

Table 1.1 (continued): Results for the application of the model on the 1st trading day of each month for the year 2003. ........................................42

Table 1.1 (continued): Results for the application of the model on the 1st trading day of each month for the year 2003. ........................................43

Table 1.2: Mean and median absolute errors using our model for \( n = 6, 7, 8 \) and data from the full and the restricted sample. .........................................46

Table 1.3: Implied risk-neutral and historical statistics of the distribution of the FTSE 100 1-month log-returns. .........................................................47

Table 2.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003 which have corresponding American call option. .............67

Table 2.1b: Statistics of the FTSE 100 American (AM) style call options for the year 2003. ........................................................................................................67

Table 2.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call option. ...........................................................................................................67

Table 2.2b: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003. .................................................................................................................68

Table 2.3: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). ............................................................80

Table 2.4: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). ............................................................81

Table 2.5: Pricing errors within moneyness categories for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). ..........................................................................................82

Table 2.6: Pricing errors within maturity categories for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). ..........................................................................................83

Table 3.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003, used for the model calibration. ..................................................105

Table 3.1b: Statistics of the FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call option. ..................105

Table 3.1c: Statistics of the FTSE 100 American (AM) style call options for the year 2003. .............................................................................................................105

Table 3.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, used for the model calibration. ..................106
Table 3.2b: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call option. .................................................................106
Table 3.2c: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003. ........................................................................................................107
Table 3.3: Implied risk-neutral, and historical statistics for the distribution of the FTSE 100 1-month log-returns, using the proposed model \( \text{INRT}_\text{LocalVol} \) for \( n=7 \), and the non-recombining implied tree model for \( n=6 \) and \( n=7 \) (\( \text{INRT}(n=6), \text{INRT}(n=7) \)). ..........................................................109
Table 3.4: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree model (\( \text{INRT} \)) for \( n=6 \) and \( n=7 \), the model proposed in this chapter (\( \text{INRT}_\text{LocalVol} \)) for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\( \text{AH}_\text{Regr} \)). ....................................................................................113
Table 3.5: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree model (\( \text{INRT} \)) for \( n=6 \) and \( n=7 \), the model proposed in this chapter (\( \text{INRT}_\text{LocalVol} \)) for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\( \text{AH}_\text{Regr} \)). ....................................................................................114
Table 3.6: Pricing errors within moneyness categories for the American FTSE 100 index call options using the implied non-recombining tree model (\( \text{INRT} \)) for \( n=6 \) and \( n=7 \), the model proposed in this chapter (\( \text{INRT}_\text{LocalVol} \)) for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\( \text{AH}_\text{Regr} \)). ....................................................................................115
Table 3.7: Pricing errors within maturity categories for the American FTSE 100 index call options using the implied non-recombining tree model (\( \text{INRT} \)) for \( n=6 \) and \( n=7 \), the model proposed in this chapter (\( \text{INRT}_\text{LocalVol} \)) for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\( \text{AH}_\text{Regr} \)). ....................................................................................116
Table 4.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003. .................................................................................................................................141
Table 4.1b: Statistics of the FTSE 100 American (AM) style call options for the year 2003. .................................................................................................................................142
Table 4.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003. .................................................................................................142
Table 4.2b: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003. .................................................................................................142
Table 4.3: Implied risk-neutral, implied Non-Markovian and historical statistics for the distribution of the FTSE 100 1-month log-returns. ..........................148
Table 4.4: Pricing errors for the American FTSE 100 index call options using the Non-Markovian model for correlation -10% (\( \text{NMT}(\rho=-10\%) \)), for correlation -7.5% (\( \text{NMT}(\rho=-7.5\%) \)), for correlation -5% (\( \text{NMT}(\rho=-5\%) \)), and for correlation -2.5% (\( \text{NMT}(\rho=-2.5\%) \)), the implied non-recombining tree (\( \text{INRT} \)) model for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\( \text{AH}_\text{Regr} \)). ....................................................................................153
Table 4.5: Pricing errors for the American FTSE 100 index call options using the Non-Markovian model for correlation -10% (\( \text{NMT}(\rho=-10\%) \)), for correlation -7.5% (\( \text{NMT}(\rho=-7.5\%) \)), for correlation -5% (\( \text{NMT}(\rho=-5\%) \)),
and for correlation -2.5% \((NMT(\rho = -2.5\%))\), the implied non-recombining tree (INRT) model for \(n=7\), and the ad-hoc model of Dumas \textit{et al.} (1998) (\textit{AH_Regr}).

\begin{table}
Table 4.6: Pricing errors within moneyness categories for the American FTSE 100 index call options using the Non-Markovian model for correlation -10\% \((NMT(\rho = -10\%))\), for correlation -7.5\% \((NMT(\rho = -7.5\%))\), for correlation -5\% \((NMT(\rho = -5\%))\), and for correlation -2.5\% \((NMT(\rho = -2.5\%))\), the implied non-recombining tree (INRT) model for \(n=7\), and the ad-hoc model of Dumas \textit{et al.} (1998) (\textit{AH_Regr}).

\begin{table}
Table 4.7: Pricing errors within maturity categories for the American FTSE 100 index call options using the Non-Markovian model for correlation -10\% \((NMT(\rho = -10\%))\), for correlation -7.5\% \((NMT(\rho = -7.5\%))\), for correlation -5\% \((NMT(\rho = -5\%))\), and for correlation -2.5\% \((NMT(\rho = -2.5\%))\), the implied non-recombining tree (INRT) model for \(n=7\), and the ad-hoc model of Dumas \textit{et al.} (1998) (\textit{AH_Regr}).
Introduction

Black and Scholes (BS) introduced in 1973 their celebrated theory on option pricing. According to their theory, the underlying asset price follows a log-normal diffusion process with a constant volatility at any time and market level\(^1\). Consequently, options on the same underlying asset must be priced using the same volatility. The success of the BS framework has led traders to quote a call option’s market price in terms of whatever constant volatility makes the BS formula value equal to the market price. This volatility is called implied volatility. However, since the 1987 market crash, the market’s implied BS volatilities for index options have shown a negative relationship with strike prices, known as the volatility skew or smile. At any fixed expiration, implied volatilities vary with strike level, that is, out-of-the-money puts trade at higher implied volatilities than out-of-the-money calls. Also, for any fixed strike level implied volatilities vary with time to expiration. Often long term implied volatilities exceed short-term implied volatilities. This is known as the volatility term structure. Volatility skews and term structure are called in one name volatility smile (see, Bates, 1991, Derman and Kani, 1994, Rubinstein, 1994, Derman, Kani and Zou, 1996, Buraschi and Jackwerth, 2001, etc. for evidence on the volatility smile). These results suggest that implied volatilities could be viewed as a non flat surface the shape of which changes over the strikes and time to expiration. The smile effect causes stock return distributions to be negatively skewed with higher kurtosis (leptokurtic) than allowed by a BS log-normal distribution\(^2\) (Rubinstein and Jackwerth, 1996).

\(^1\) In mathematical terms, the evolution over an infinitesimal time \(dt\) is described by the stochastic differential equation,

\[
\frac{dS}{S} = \mu \, dt + \sigma \, dz
\]

where \(S\) is the asset level, \(\mu\) is the asset’s expected return, \(\sigma\) is the asset’s volatility and \(dz\) is a Wiener process with a mean of zero and a variance equal to \(dt\).

\(^2\) The symmetry or asymmetry of the risk neutral distribution is directly reflected in the relative prices of the out-of-the money (OTM) puts and calls. Symmetric risk-neutral distributions imply equal prices for OTM European calls and puts; skewed distributions create systematic divergences (Bates, 1991).
According to Buraschi and Jackwerth (2001) the size of the violations is such that cannot be explained just by market imperfections. These violations belie the BS theory, which assumes constant volatility (and therefore, constant implied volatility) for all options. Thus, it may be convenient to keep quoting options prices in terms of BS implied volatilities, but it is incorrect to calculate options prices using the BS formula. As Black (1975) pointed out, “One possible explanation for this misspricing pattern is that we have left something out from the formula”.

Several studies have suggested extensions of the BS model to account for the volatility smile and other empirical violations of the original model. Examples include (i) the stochastic volatility models (e.g. Scott, 1987, Wiggins, 1987, Hull and White, 1987, Melino and Turnbull, 1990, 1995, Stein and Stein, 1991, Heston, 1993), (ii) the jump-diffusion/pure jump models (e.g. Merton, 1976, Bates, 1991), and (iii) the combination of stochastic volatility and jump processes (e.g. Bates, 1996, Scott, 1997). The approach taken by these models consists of specifying the parameters of the process for the underlying asset price and for the volatility and/or jump process. Then, option prices are derived as a function of the parameters of these processes and the price of the underlying security.

However empirical evidence indicates that the non-flat implied volatility surface is not adequately explained by either stochastic volatility or jump or both (see Bakshi et al., 1997, Das and Sundaram, 1999). As researchers looked for new option pricing models that were able to explain these option prices, some tried to learn more about the stochastic process of the asset using the observed option prices. As a response, the “smile consistent” no arbitrage models or otherwise known as model calibration has emerged. Model calibration deals with the inverse problem that option pricing theory deals with. Instead of assuming a stochastic process for the underlying asset, it identifies the (unknown) stochastic process of the underlying asset given information about prices of options. Rubinstein (1994), Derman and Kani (1994) and Dupire (1994) were the first to introduce model calibration (smile}
consistent models) in a series of articles on a class of models that Rubinstein (1994) calls implied binomial trees. These implied trees are extensions of the original Cox, Ross and Rubinstein (CRR, 1979) binomial tree that assumes constant volatility.

Calibration of financial models is one of the most important issues that trading rooms must confront. A calibrated model is consistent with observed market prices, thus making it more suitable for pricing more complex or less liquid options. Consistency is achieved by extracting an implied evolution for the underlying asset price from market prices of liquid standard options. Some applications of the smile consistent models and especially of implied trees are the following:

i. To value other derivatives whose prices are not available in the market, for example standard but illiquid European style options, American style options, exotic options, Over The Counter (OTC) options with confidence that the model is valuing all options consistently with the market.

ii. Especially useful for valuing barrier options, where the probability of striking the barrier is sensitive to the shape of the smile\(^3\).

iii. To generate Monte-Carlo simulations for valuing path-dependent options\(^4\).


---

\(^3\) The theoretical value of a barrier option depends on the risk-neutral probability of the index being in-the-money at expiration but not having crossed the barrier during the option’s life time. This probability is very sensitive to volatility levels, in general, and to the volatility skew, in particular (Derman, Kani and Zou, 1996).

\(^4\) Path-dependent options contain embedded strikes at multiple market levels and are consequently sensitive to the local volatility in multiple regions (Derman, Kani and Zou, 1996).
latter class of models is more general and it nests the former class of models (Skiadopoulos, 2001). There also exist non-parametric methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.

Smile consistent deterministic volatility models are based on the assumption that the local volatility of the underlying asset is a known function of time and of the path and level of the underlying asset price\(^5\). However, they do not specify local volatility in advance, but derive it endogenously from the European option market prices. Therefore, they preserve the “pricing by no-arbitrage” property of the BS model, and the markets are complete since the option’s pay-off can be synthesized from existing assets.

Perhaps the most important contribution to the area of deterministic volatility implied trees has been put forward by Derman and Kani (DK, 1994). They were the first to show that a class of index options, which exhibit an implied volatility smile can be consistently priced using an implied binomial tree. Their model and variations of that model are very popular among practitioners. A major disadvantage of this model is the appearance of non-acceptable probability values. Barle and Cakici (1995) state that this is because of their “… strict requirement that continuous diffusion be modelled as a binomial process and on a recombining tree”. This problem can be referred to as a problem of interdependencies between nodes. Possible methods that can be used to reduce the problem of interdependencies are the calibration of trinomial (or multinomial) trees or non-recombining trees. These extra degrees of freedom allow for more flexibility in the estimation of the distribution of the underlying asset.

Trinomial trees provide a much better approximation to the continuous time process than the binomial trees for the same number of steps. However,

---

\(^5\) Local volatility is the volatility of the underlying asset at any node of the tree.
the extra degrees of freedom (additional number of nodes) require a larger number of simultaneous equations to be solved. Derman, Kani and Chriss (1996) proposed implied trinomial trees. In their model they use the additional parameters to conveniently choose the “state space” of all node prices in the tree, and let only the transition probabilities be constrained by market options prices. Chriss (1996) generalized their method for American style options.

As a response to the need of a non-recombining implied tree, in this thesis we develop models for calibrating non-recombining (binary) trees using option market data. In order to capture the implied distribution, we minimize the discrepancy between the observed market prices and the model values, subject to constraints that keep the probabilities well specified and prevent arbitrage opportunities. We built our models on non-recombining trees so as to allow the local volatility to be a function of the underlying asset and of time and to enable each node of the tree to act as an independent variable. Effectively, the problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and using methods from nonlinear constrained optimization we minimize the least squares error function on market prices. Specifically, we adopt a penalty method to handle the inequality constraints and for the optimization we use a Quasi-Newton algorithm. Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem.

The main benefit of the proposed models is their analytical structure which enables us to use efficient methods for nonlinear optimization. Although the models use a large number of variables, due to the constraints imposed and due to the fact that we use efficient methods for optimization

---

6 Other work we are aware of that uses a non-recombining tree is of Talias (2005) where for the calibration he uses genetic algorithms.
the models are not computationally intensive. Also, the extra degrees of freedom allow for more flexibility in the estimation of the underlying asset distribution. This has two implications: It prevents non-acceptable transition probability values and also helps to obtain almost perfect fit between the model and market option values. In addition, because of their binary structure, the models allow for path dependency. This creates a more natural framework for pricing path depended options and also provides a market consistent tree for option replication with transaction costs which requires non-recombining tree (see Edirisinghe et al., 1993). In contrast to Rubinstein (1994), the proposed models can be easily modified to account for European contracts with different maturities. Our models do not need any interpolation or extrapolation across strikes and time to find hypothetical options as opposed to Derman and Kani (1994). The proposed models can be easily modified to capture the observed bid/ask spreads in the market. This is very useful since the reported closing prices may not always be accurate, or may be inaccurate due to various market frictions (transaction costs, illiquidity, etc.).

The models are calibrated using FTSE 100 European call options data for the year 2003 obtained from LIFFE. The underlying asset of each call option is a future contract. Thus, on the same trading day, there are options with different underlying asset. Also, on the same trading day, options with the same underlying asset have the same maturity. Every trading day in the sample, we calibrate the above models using European call options with the same underlying asset and time to maturity. Since the models are calibrated using calls with the same expiration, the implied distribution that we find using this dataset, results from a smile the shape of which is independent of expiration time. Overall results strongly support our modeling approaches. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Also, the computational burden is not a major issue.

In the first chapter of the thesis in order to capture the implied distribution of the underlying asset, we calibrate the non-recombining tree
assuming that the system variable is the underlying asset at each node of the tree (INRT model). The proposed model is non-parametric since no restrictive assumptions are made for the underlying stochastic process. Thus, the model is flexible since it applies to arbitrary underlying asset distributions, which imply arbitrary local volatility distributions. Also, the extra degrees of freedom and the analytical structure of the model would allow us to impose smoothness constraints on the distribution of the underlying asset if required. In addition, this model provides us with an estimation of the deltas at each node of the tree.

Despite the vast literature on smile consistent models, there are only few studies that test the empirical performance of these models (see Dumas et al. 1998, Brandt and Wu, 2002, Lim and Zhi, 2002, Linaras and Skiadopoulos, 2005). The second chapter of the thesis provides an empirical comparison of the pricing performance of the INRT model proposed in the first chapter with a modified version of the implied binomial tree of Derman and Kani (DK, 1994), a very popular model among practitioners, and with the ad-hoc procedure of smoothing BS implied volatilities across strikes proposed by Dumas et al. (1998). In order to investigate the effect of the interpolation method on the pricing performance of the DK model and the ad-hoc model two interpolation methods across strikes are used, the 2nd order polynomial regression and the cubic spline. The above models are tested in pricing of American call options on the FTSE 100 for the year 2003.

In the third chapter we calibrate the non-recombining tree assuming that the system variable is the local volatility at each node of the tree (INRT_LocalVol model). In this way, we reduce the number of variables that the optimization algorithm has to deal with to half, in relation to the INRT model, making the problem less computationally intensive. Moreover, if required, we can easily control the values of the local volatilities at each node of the tree by imposing lower and upper bounds on the values of the local volatility. Knowledge of the local volatility surface is especially useful in markets with pronounced smile to measure market sentiment, to compute the
evolution of implied volatilities through time, and to price and hedge exotic options.

In the literature there is great evidence for the predictability of financial assets returns (see Fama and French, 1988, Lo and MacKinley, 1988, Jegadeesh, 1990, Jašić and Wood, 2006). However, the development of option pricing models where the returns of the underlying asset are predictable has not been investigated much yet. The only paper that we are aware of and provides a model for option pricing when stock returns are predictable is of Lo and Wang (1995). They argue that, even though the predictability of stock returns is induced by the drift of the underlying process, the drift does not enter the standard option pricing formulas. As a result, even though the market option prices are affected by predictability, standard option pricing formulas do not take account of it. In their study they make an adjustment to the BS formula (1973) that takes into account the predictability of underlying asset’s returns. They find that even small levels of predictability can be important especially for longer maturity options. They also provide several continuous-time linear diffusion processes that can capture different forms of predictability.

In the fourth chapter, we make an extension of Lo and Wang (1995) to smile-consistent option pricing models and specifically to the \textit{INRT} model proposed in the first chapter. We use the methodological framework proposed in the first chapter but now we impose non-Markovian assumptions by allowing for serial dependence on the underlying asset’s returns between two consecutive time steps (\textit{NMT} model). Thus, like the \textit{INRT} model, the \textit{NMT} model allows for a very flexible underlying asset distribution, but in addition, because of its non-Markovian nature, it gives a richer and more realistic model than the other implied trees and continuous option pricing models in the literature. By assuming serial dependence of returns, in the fourth chapter we acknowledge the potential for market inefficiencies. To the best of our knowledge, this is the first attempt in the literature of imposing
non-Markovian assumptions on the process followed by the underlying asset on an implied tree.

A note should be added on the empirical findings of the model about the implied risk-neutral mean return of the underlying asset. The expected mean return is equal to the risk-free rate \( (r_f) \) minus the dividend yield \( (\delta) \), \( r_f - \delta \). Since in our application the dividend yield equals the risk-free rate (since the underlying asset is a future contract), the implied risk-neutral return should be equal to zero. Although the formula for the computation of the transition probabilities implicitly imposes the constraint that the implied return at each node of the tree should be equal to zero, in our empirical findings the monthly implied return is around \(-0.2\%\). This is rather insignificant and is due to accumulated rounding errors. For a tree with 7 time steps, the algorithm has to solve an optimization problem of 126 variables. The computation of the implied risk-neutral log-return is done using information from the last time step of the tree, which has 64 nodes. Thus the expected log-return is the sum of the 64 path probabilities leading to the ending nodes multiplied by the log-return at each node. The large number of computations results to accumulation of rounding errors. A more accurate solution could be achieved by allowing more decimal points in the calculations.

The thesis continuous with the description of the four chapters:

1. “Implied non-recombining trees and calibration for the volatility smile”\(^7\).
2. “An empirical comparison between non-recombining and recombining implied trees”.
3. “Calibration of non-recombining implied trees for the local volatility surface”.
4. “Option pricing on non-recombining implied trees assuming serial dependence of returns”.

Finally, the thesis concludes.

\(^7\) The first chapter has been published in the journal of Quantitative Finance, 2007.
1. Implied non-recombining trees and calibration for the volatility smile

Abstract

In this chapter we capture the implied distribution from option market data using a non-recombining (binary) tree, allowing the local volatility to be a function of the underlying asset and of time. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure and use nonlinear constrained optimization to minimize the least squares error function on market prices. The proposed model can accommodate European options with single maturities and, with minor modifications, options with multiple maturities. It can provide a market-consistent tree for option replication with transaction costs (often this requires a non-recombining tree) and can help pricing of exotic and Over The Counter (OTC) options. We test our model using options data of the FTSE 100 index obtained from LIFFE. The results strongly support our modelling approach.
1.1. Introduction

Calibrating a tree, otherwise known as constructing an implied tree, means finding the stock price and/or associated probability at each node in such a way that the tree reproduces the current market prices for a set of benchmark instruments. The main benefit of calibrating a model to a set of observed option prices is that the calibrated model is consistent with today’s market prices. The calibrated model can then be used to price other more complex or less liquid securities, such as Over The Counter (OTC) options whose prices may not be available in the market.

The binomial tree is the most widely used tool in the financial pricing industry. The classic Cox-Ross-Rubinstein (CRR, 1979) binomial tree is a discretization of the Black-Scholes (BS, 1973) model since it is based on the assumption of the BS model that the underlying asset evolves according to a geometric Brownian motion with a constant volatility factor. This, however, contradicts the observed implied volatility, which suggests that volatility depends on both the strike and maturity of an option, a relationship commonly known as the volatility smile. This problem has motivated the recent literature on “smile consistent” no-arbitrage models. Consistency is achieved by extracting an implied evolution for the stock price from market prices of liquid standard options on the underlying asset. There are two classes of methodologies within this approach. Smile consistent deterministic volatility models (Rubinstein, 1994, Derman and Kani, 1994, Dupire, 1994, Barle and Kakici, 1995, Jackwerth and Rubinstein, 1996, Jackwerth, 1997, etc.); and stochastic volatility smile consistent models which allow for smile-consistent option pricing under the no-arbitrage evolution of the volatility surface (Derman and Kani, 1998, Ledoit and Santa-Clara, 1998, Britten-Jones and Neuberger, 2000, etc.). The latter class of models is more general and it nests the former class of models (Skiadopoulos, 2001). There also exist non-parametric methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution
of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.

Smile consistent deterministic volatility models are based on the assumption that the local volatility of the underlying asset is a known function of time and of the path and level of the underlying asset price. However, they do not specify local volatility in advance, but derive it endogenously from the European option prices. Therefore, they preserve the “pricing by no-arbitrage” property of the BS model, and the markets are complete since the option’s pay-off can be synthesized from existing assets.

Rubinstein (1994) finds the implied risk-neutral terminal-node probability distribution which is in the least-squares sense, closest to the lognormal subject to some constraints. The probabilities must add up to one and be non-negative. Moreover, they are calculated so that the present value of the underlying assets and all the European options calculated with these probabilities, fall between their respective bid-ask prices. This methodology allows for an arbitrary terminal-node probability distribution, but assumes that path probabilities leading to the same ending node are equal. Rubinstein’s (1994) methodology suffers from the fact that options expiring at early time steps cannot be used for the construction of the tree. Thus, options with maturity other than the maturity of the options used during the construction of the tree are not consistent with market prices.

Jackwerth (1997) introduced generalized binomial trees as an extension of Rubinstein (1994). His model allows for an arbitrary terminal-node probability distribution, but also allows path probabilities leading to the same node to take different values.

Derman and Kani (1994) constructed recombining binomial trees using a large set of option prices. For each node they need a corresponding option price with strike price equal to the previous node’s stock price and expiring at the time associated with that node. Since they have fewer option prices than required, they need to interpolate and extrapolate from given option prices.
Their trees are sensitive to the interpolation and extrapolation method and require adjustments to avoid arbitrage violations.

Barle and Cakici (1995) introduced a number of modifications which aimed to eliminate negative probabilities and improve the general stability of Derman’s and Kani’s (1994) model. Although their modified method fits the smile accurately, negative probabilities may still occur with increases in the volatility smile and interest rate. As they state, this is because of their “…strict requirement that continuous diffusion be modeled as a binomial process and on a recombining tree “. This problem can be referred to as a problem of \textit{interdependencies} between nodes.

Possible methods that can be used to reduce the problem of interdependencies are the calibration of trinomial (or multinomial) trees or non-recombining trees. These extra degrees of freedom allow for more flexibility in the estimation of the distribution of the underlying asset.

Trinomial trees provide a much better approximation to the continuous time process than the binomial trees for the same number of steps. However, the extra degrees of freedom (additional number of nodes) require a larger number of simultaneous equations to be solved. Derman, Kani and Chriss (1996) proposed implied trinomial trees. In their model they use the additional parameters to conveniently choose the “state space” of all node prices in the tree, and let only the transition probabilities be constrained by market options prices. Chriss (1996) generalized their method for American style options.

In this chapter we propose a method for calibrating a non-recombining (binary) tree, based on optimization. Specifically, we minimize the discrepancy between the observed market prices and the model values with respect to the underlying asset at each node, subject to constraints that maintain risk neutrality and prevent arbitrage opportunities. Our model is built on a non-recombining tree\textsuperscript{8} so as to allow the local volatility to be a

\textsuperscript{8} Other work we are aware of that uses a non-recombining tree is of Talias (2005) where for the calibration he uses genetic algorithms.
function of the underlying asset and of time and to enable each node of the
tree to act as an independent variable. Effectively, the problem under
consideration is a non-convex optimization problem with linear constraints.
We elaborate on the initial guess for the volatility term structure, and using
methods from nonlinear constrained optimization we minimize the least
squares error function. Specifically, we adopt a penalty method and for the
optimization we use a Quasi-Newton algorithm. Because of the combinatorial
nature of the tree and the large number of constraints, the search for an
optimum solution as well as the choice of an algorithm that performs well
becomes a very challenging problem.

Our model was created as a response for the need of a non-
recombining implied tree. The main benefit of the model is its analytical
structure which enables the use of efficient methods for nonlinear
optimization. Although the model uses a large number of variables, due to the
constraints imposed and due to the fact that we use efficient methods for
optimization the model is not computationally intensive. Also, the extra
degrees of freedom allow for more flexibility in the estimation of the
underlying asset distribution. This has two implications: It prevents non-
acceptable transition probability values and also helps to obtain almost
perfect fit between the model and market option values. In addition, because
of its binary structure, the model allows for path dependency. This creates a
more natural framework for pricing path depended options and also provides
a market consistent tree for option replication with transaction costs which
requires non-recombining tree (see Edirisinghe et al., 1993). Also, the model
provides an estimation of the Deltas at each node of the tree.

In contrast to Rubinstein (1994), the proposed model can be easily
modified to account for European contracts with different maturities. Our
model does not need any interpolation or extrapolation across strikes and
time to find hypothetical options as opposed to Derman and Kani (1994). The
proposed model can be easily modified to capture the observed bid/ask
spreads in the market. This is very useful since the reported closing prices
may not always be accurate, or may be inaccurate due to various market frictions (transaction costs, illiquidity, etc.). Finally, the extra degrees of freedom and the analytical structure of the model would allow us to impose smoothness constraints on the distribution of the underlying asset if required.

We test our model using call options data of the FTSE 100 index, for the year 2003 obtained from LIFFE. The results strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Also, the computational burden is not a major issue.

The chapter continues as follows: In section 1.2 we describe the proposed methodology and the initialization of the non-recombining tree. In section 1.3 we discuss the imposed risk neutrality and no-arbitrage constraints. In section 1.4 we describe the optimization algorithm. In section 1.5 we calibrate the model using FTSE 100 options data. Conclusions are in section 1.6. In Appendix 1.1A we prove the feasibility of the initialized tree, in Appendix 1.1B we prove the feasibility of the initialized tree taking into account that the risk-free rate, dividend yield and time step are time dependent and in Appendix 1.2 we adjust the formulas for time dependent risk free rate, dividend yield and step size.

1.2. The proposed methodology and initialization of the non-recombining tree

Our goal is to develop an arbitrage-free risk neutral model that fits the smile, is preference-free, and can be used to value options from easily observable data. In order to allow more degrees of freedom, we use a non-recombining tree. In the following section we present the proposed methodology, and describe the initialization of the tree.

The point \((i, j)\) on the tree denotes:

\(i\) : the time dimension, \(i = 1, ..., n\).
\( j \): the asset (time specific) dimension, \( j = 1, ..., 2^{i-1} \).

\( S_{i,j} \) is the value of the underlying asset at node \((i, j)\).

Figure 1.1 shows a non-recombining tree with four steps and Figure 1.2 shows a typical triplet in a non-recombining tree.

**Figure 1.1**: Non-recombining tree with 4 steps.

**Figure 1.2**: A typical triplet in a non-recombining tree.
Let $C_{\text{Mkt}}(k), \ k = 1, ..., N$ denote the market prices of $N$ European calls, with strikes $K(k)$ and single maturities $T$. Also, let $C_{\text{Mod}}(x,k), \ k = 1, ..., N$ denote the prices of the $N$ calls obtained using the model. $x$ denotes a vector containing the variables of the model which are the values of the underlying asset at each node of the tree, excluding its current value. The ideal solution is to find the values of the underlying asset (the model variables) at each node of the tree such that a perfect match is achieved between the option market prices and those predicted by the tree. However, due to market imperfections and other factors perfect matching may not always be possible. Therefore, we minimize the discrepancy between the observed market prices and the values produced by the model subject to constraints that prevent arbitrage opportunities.

We have to solve a non-convex constrained minimization problem with respect to the values of the underlying asset at each node:

$$\min_{x} \sum_{k=1}^{N} w_k f(C_{\text{Mod}}(x,k), C_{\text{Mkt}}(k))$$

(1.1)

where $f$ denotes a suitable objective function on the error between the model and market prices. We can also allow for a weight factor, $w_k$.\(^9\) In this paper we use the least squares error function which is defined as the sum of square differences between market prices and model prices produced by the tree. The method can be adjusted easily for any other objective function.

The philosophy of the initialization of the non-recombining tree is the same as that of the construction of the standard CRR binomial tree, but we adjust the formulas so that the tree does not necessarily recombine.

---

\(^9\) Weights can be related for example to the trading volume or moneyness of the options.
We denote with $u_{i,j}$ and $d_{i,j}$ the up and down factors by which the underlying asset price can move in the single time step, $\Delta t$, given that we are at node $(i,j)$. $\Delta t$, $u_{i,j}$ and $d_{i,j}$ factors are given by the following formulas\(^\text{10}\):

\[
\Delta t = \frac{T}{n-1}
\]

\[
u_{i,j} = e^{\sigma_i \sqrt{\Delta t}}
\]

\[\sigma_i = \sigma_0 e^{\lambda(i-1)\Delta t}, \lambda \in R, \ i = 1, \ldots, n-1\]

\[
d_{i,j} = e^{-\sigma_i \sqrt{\Delta t}} = \frac{1}{u_{i,j}}
\]

where $T$ is the option’s time to maturity and $\sigma_i$ is the volatility term structure at time step $i$.

We initialize the tree using the following volatility term structure:

\[
\sigma_i = \sigma_i e^{\lambda(i-1)\Delta t}, \lambda \in R, \ i = 1, \ldots, n-1
\]

where $\lambda$ is a constant parameter and $\sigma_i$ is a properly chosen initial value for the volatility. If $\lambda$ is positive, then volatility increases as we approach maturity and if $\lambda$ is negative, then volatility decreases as we approach maturity\(^\text{11}\).

In order to preserve the risk neutrality at every time step and hence obtain a feasible initial tree, we choose $\lambda$ to belong in the following interval (for proof see Appendix 1.1A):

\[
\lambda \in \left[ \frac{1}{T} \log \left( \frac{r_f - \delta}{\sigma_i} \sqrt{\Delta t} \right), +\infty \right)
\]

\(^{10}\) For simplicity, we make the assumption that the risk free rate, the dividend yield and the step size do not change across time. Formulas adjusted for time dependence can be found in Appendix 1.1B and 1.2.

\(^{11}\) Other non-monotonic functions could also be used for $\sigma_i$ but what we have tried proved adequate for our purposes.
By choosing $\lambda$ from the above interval, we allow the initial volatility to increase or decrease across time. We make several consecutive draws from interval (1.5) until we find the value of $\lambda$ that gives the “optimal” tree\textsuperscript{12}.

We denote with $S_{i,0}$ the current value of the underlying asset. The odd nodes of the tree $S_{i,j}$ are initialized using the following equation:

$$S_{i,j} = S_{i-\frac{1}{2},j-\frac{1}{2}} d_{i-\frac{1}{2},j-\frac{1}{2}}, \quad i = 2, \ldots, n, \quad j = 1, 3, \ldots, 2^{i-1} - 1 \tag{1.6a}$$

The even nodes of the tree $S_{i,j}$ are initialized using the following equation:

$$S_{i,j} = S_{i-\frac{1}{2},j-\frac{1}{2}} u_{i-\frac{1}{2},j-\frac{1}{2}}, \quad i = 2, \ldots, n, \quad j = 2, 4, \ldots, 2^{i-1} \tag{1.6b}$$

We want to point out that equations (1.3) to (1.6) are used only for initialization. Once the optimization process starts, each value of the underlying asset (except from $S_{i,0}$) acts as an independent variable in the system.

Upward transition probabilities give the probability of moving from node $(i, j)$ to node $(i+1,2j)$ whereas downward transition probabilities give the probability of moving from node $(i, j)$ to node $(i+1,2j-1)$ for $i = 1, \ldots, n-1$ and $j = 1, \ldots, 2^{i-1}$. For the upward transition probabilities $p_{i,j}$ between the various nodes of the tree we use the risk-neutral probability formula\textsuperscript{13}:

$$p_{i,j} = \frac{S_{i,j} e^{(r_{\delta} - \delta) \Delta t} - S_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}}, \quad i = 1, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \tag{1.7}$$

\textsuperscript{12} Optimal tree is the one that gives the lowest-value objective function subject to the initial constraints.

\textsuperscript{13} Probability equation (1.7) is effectively a martingale restriction (see equation (6) and relevant discussion in Longstaff, 1995). Thus the numerical implementation of the model with this probability equation is restricted to a Markovian stochastic process.
where $r_f$ denotes the annually continuously compounded riskless rate of interest and $\delta$ denotes the annually continuously compounded dividend yield. Their respective downward probability is equal to one minus the upward probability.

The call option value at the last time step is given by:

$$C_{n,j} = \max\{S_{n,j} - K, 0\}, \quad j = 1, \ldots, 2^{n-1}$$

However, the function $\max$ is non differentiable at $S_{n,j} = K$. To overcome this problem, we propose the following smoothing approximation to $C_{n,j}$:

$$\frac{C_a(n, j)}{K} = \begin{cases} 
0 & \text{for } S_{n,j} / K \leq 1 - \frac{z}{2} \\
\frac{S_{n,j}}{K} - 1 & \text{for } S_{n,j} / K \geq 1 + \frac{z}{2} \\
\frac{1}{2z}\left[\left(\frac{S_{n,j} - 1}{K}\right) + \frac{z}{2}\right]^2 & \text{for } 1 - \frac{z}{2} < S_{n,j} / K < 1 + \frac{z}{2}
\end{cases}$$

$j = 1, \ldots, 2^{n-1}$

where $z$ is a small positive constant, for example 0.01 (see Fig. 1.3).

Figure 1.3: Smoothing of the option pay-off function at maturity.
The value of the call at intermediate nodes is given by the following equation:

\[
C_{i,j} = \left( p_{i,j} C_{i+1,2j} + (1 - p_{i,j}) C_{i+1,2j-1} \right) e^{-r_i \Delta t}
\]

\[i = n - 1, ..., 1, \quad j = 1, ..., 2^{i-1}\]

(1.9b)

1.3. Risk neutrality and no-arbitrage constraints

In this section we describe the risk neutrality and no-arbitrage constraints. In order for the transition probabilities \(p_{i,j}\) defined in Eq.(1.7) to be well specified, they should take values between zero and one. This implies the following risk-neutrality constraints:

\[
S_{i,j} e^{(r_i - \delta_i) \Delta t} \leq S_{i+1,2j}
\]

\[i = 1, ..., n - 1, \quad j = 1, ..., 2^{i-1}\]

(1.10a)

\[
S_{i,j} e^{(r_i - \delta_i) \Delta t} \geq S_{i+1,2j-1}
\]

(1.10b)

Risk neutrality constraints in the non-recombining tree prevent nodes \(2j - 1\) and \(2j\) to cross, for \(i=1, ..., n\) and \(j=1, ..., 2^{i-1}\) (see Fig.1.1).

Options (puts and calls) have upper and lower bounds that do not depend on any particular assumptions on the factors that affect option prices. If the option price is above the upper bound or below the lower bound, there are profitable opportunities for arbitrageurs. To avoid such opportunities, we include the no-arbitrage constraints. Specifically, a European call with dividends should lie between the following bounds:

\[
\max\left( S_{i,i} e^{-\delta T} - Ke^{-r_i T} \right) \leq C_{mod} \leq S_{1,1}
\]

(1.11)

Also, every value of the underlying asset on the tree should be greater or equal to zero. Thus, we also impose the following constraint:
\[ S_{i,j} \geq 0, \quad i = 2, ..., n, \quad j = 1, ..., 2^{i-1} \] (1.12)

1.4. The optimization algorithm

The objective of the problem is to minimize the least squares error function of the discrepancy between the values produced by the model and the observed market prices. Thus, we have the following optimization problem:

\[
\min_x \frac{1}{2} \sum_{k=1}^{N} (C_{\text{Mod}}(x,k) - C_{\text{Mkt}}(k))^2
\] (1.13)

where \( C_{\text{Mod}}(k) \) and \( C_{\text{Mkt}}(k) \) denote the model and market price respectively of the \( k^{th} \) call, \( k = 1, ..., N \), subject to the constraints:

i) \( g_1(i, j) = S_{i,j}^e e^{(r_j - \delta)T} - S_{i+1,2^{i-2}}^M \geq 0, \quad i = 1, ..., n-1, \quad j = 1, ..., 2^{i-1} \) (1.14a)

ii) \( g_2(i, j) = S_{i+1,2^j}^M - S_{i,j}^e e^{(r_j - \delta)T} \geq 0, \quad i = 1, ..., n-1, \quad j = 1, ..., 2^{i-1} \) (1.14b)

iii) \( g_3(k) = S_{1,1} - C_{\text{Mod}}(k) \geq 0, \quad k = 1, ..., N \) (1.14c)

iv) \( g_4(k) = C_{\text{Mod}}(k) - \max(S_{1,1}^e e^{-\delta T} - K(k) e^{-r_j T}, 0) \geq 0, \quad k = 1, ..., N \) (1.14d)

v) \( g_5(i, j) = S_{i,j} \geq 0, \quad i = 2, ..., n, \quad j = 1, ..., 2^{i-1} \) (1.14e)

Since the problem under consideration is a non-convex optimization problem with linear constraints we adopt an exterior penalty method (Fiacco and McCormick, 1968) to convert the nonlinear constrained problem into a nonlinear unconstrained problem. The Exterior Penalty Objective function that we use is the following:
\[ P(x, \alpha) = \frac{1}{2} \sum_{k=1}^{N} (C_{\text{Mod}}(x, k) - C_{\text{M}k}(k))^2 \]
\[ + \alpha \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ \min(g_1(i, j), 0)^2 + \min(g_2(i, j), 0)^2 \right] \]
\[ + \frac{\alpha}{2} \sum_{k=1}^{N} \left[ \min(g_3(k), 0)^2 + \min(g_4(k), 0)^2 \right] \]
\[ + \frac{\alpha}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[ \min(g_5(i, j), 0)^2 \right] \]  
\[ (1.15) \]

The second, third and fourth terms in \( P(x, \alpha) \) give a positive contribution if and only if \( x \) is infeasible. Under mild conditions it can be proved that minimizing the above penalty function for strictly increasing sequence \( \alpha \) tending to infinity the optimum point \( x(\alpha) \) of \( P \) tends to \( x^* \), a solution of the constrained problem.

For the optimization we use a Quasi-Newton algorithm. Specifically we use the BFGS formula\(^{14} \) (Fletcher, 1987). For the procedure of Line Search in the algorithm we use the Charalambous (1992) method. To achieve the best feasible solution, i.e. the solution that gives us a feasible tree with the smallest error function we force the algorithm to draw consecutively values of \( \lambda \) from the specified interval (1.5) until the objective function is smaller than 1.E-4 and also the penalty term equals zero, i.e. we have a feasible solution.

**Implementation**

For the implementation of the optimization method, we need to calculate the partial derivatives of \( C_{\text{Mod}}(k) \) \(^{15} \) with respect to the value of the underlying asset at each node, for \( k = 1, \ldots, N \) i.e. we want to find \( \frac{\partial C(1, 1, k)}{\partial S_{i, j}} \).

---

\(^{14}\) The BFGS formula was discovered in 1970 independently by Broyden, Fletcher, Goldfarb and Shanno.

\(^{15}\) From now on we will use \( C_{1,1} \) instead of \( C_{\text{Mod}} \).
\[ i = 2, \ldots, n, \quad j = 1, \ldots, 2^{j-1} \quad \text{and} \quad k = 1, \ldots, N. \quad \text{For notational simplicity in the following, we assume that we have only one call option. For the computation of } \frac{\partial C_{1,1}}{\partial S_{i,j}}, \forall i, j \text{ we implement the following steps:}
\]

We define the triplet vector (see Fig. 1.2):

\[ S^{(i)}_{i,j} = [S_{i,j}, S_{i+1,2j}, S_{i+1,2j-1}] \quad (1.16) \]

**1st step:** Compute the partial derivatives of the risk neutral transition probabilities, \( \frac{\partial p_{i,j}}{\partial S_{i,j}}, \frac{\partial p_{i,j}}{\partial S_{i+1,2j}}, \) and \( \frac{\partial p_{i,j}}{\partial S_{i+1,2j-1}} \) for \( i = 1, \ldots, n-1, \) and \( j = 1, \ldots, 2^{i-1}. \) We summarize the derivatives in vector form (1.17).

\[
\nabla_{S^{(i)}_{i,j}} p_{i,j} \equiv \begin{bmatrix}
\frac{\partial p_{i,j}}{\partial S_{i,j}} \\
\frac{\partial p_{i,j}}{\partial S_{i+1,2j}} \\
\frac{\partial p_{i,j}}{\partial S_{i+1,2j-1}}
\end{bmatrix} = \frac{1}{e^{(rT-S^{0})N} - p_{i,j} - (1-p_{i,j})} \begin{bmatrix}
e^{(rT-S^{0})N} \\
- p_{i,j} \\
- (1-p_{i,j})
\end{bmatrix}
\]

\[ (1.17) \]

16 We do not calculate \( \frac{\partial C_{LLK}}{\partial S_{1,1}} \) since \( S_{1,1} \) is a known, fixed parameter, and thus does not take part in the optimization.
2\textsuperscript{nd} step: Compute the partial derivatives $\frac{\partial C_{i,j}}{\partial S_{i,j}}$, for $i = 2, ..., n - 1$, and $j = 1, ..., 2^{i-1}$, and $\frac{\partial C_{i,j}}{\partial S_{i+1,j}}$ and $\frac{\partial C_{i,j}}{\partial S_{i+1,j-1}}$ for $i = 1, ..., n - 1$, $j = 1, ..., 2^{i-1}$. We summarize the derivatives in vector form (1.18).

$$
\nabla_{S_{i,j}}^\alpha C_{i,j} \equiv \left[ \begin{array}{c} \frac{\partial C_{i,j}}{\partial S_{i,j}} \\ \frac{\partial C_{i,j}}{\partial S_{i+1,j}} \\ \frac{\partial C_{i,j}}{\partial S_{i+1,j-1}} \end{array} \right] = \left[ \begin{array}{c} \Delta_{i,j} \\ p_{i,j} \left( \Delta_{i+1,j} - \Delta_{i,j} e^{\alpha t} \right) e^{-\Delta t} \\ \left(1 - p_{i,j}\right) \left( \Delta_{i+1,j-1} - \Delta_{i,j} e^{\alpha t} \right) e^{-\Delta t} \end{array} \right] \tag{1.18}
$$

where

$$
\Delta_{i,j} = \frac{C_{i+1,j} - C_{i+1,j-1}}{S_{i+1,j} - S_{i+1,j-1}} e^{-\Delta t} = \frac{\partial C_{i,j}}{\partial S_{i,j}} \equiv \text{Delta Ratio} \tag{1.19}
$$

3\textsuperscript{rd} step: Compute the partial derivatives $\frac{\partial C_{n,j}}{\partial S_{n,j}}$ for $j = 1, ..., 2^{n-1}$. They are given by the following formula:

$$
\frac{\partial C_{n,j}}{\partial S_{n,j}} = \begin{cases} 
0 & \text{for } S_{n,j} \leq K\left(1 - z / 2\right) \\
1 & \text{for } S_{n,j} \geq K\left(1 + z / 2\right) \\
\frac{1}{z} \left[ \left( \frac{S_{n,j}}{K} - 1 \right) + \frac{z}{2} \right] & \text{for } K\left(1 - z / 2\right) < S_{n,j} < K\left(1 + z / 2\right) 
\end{cases} \tag{1.20}
$$
4th step: Compute the partial derivatives \( \frac{\partial C_{i,j}}{\partial S_{i,j}} \) for \( i \geq 3 \).

\[
\frac{\partial C_{i,j}}{\partial S_{i,j}} = \prod \left\{ \text{of the probabilities on the path that take us from node (1,1) to node (i-1,k)} \right\} \\
x \frac{\partial C_{i-1,k}}{\partial S_{i,j}} e^{-(i-2)\tau,\Delta t}
\]

\[
k = \begin{cases} 
  j/2 & \text{for even } j \\
  (j+1)/2 & \text{for odd } j 
\end{cases}
\]

For example,

\[
\frac{\partial C_{1,1}}{\partial S_{4,6}} = p_{1,1}(1 - p_{2,2}) \frac{\partial C_{3,3}}{\partial S_{4,6}} e^{-2\tau,\Delta t}
\]

\[
\frac{\partial C_{1,1}}{\partial S_{5,3}} = (1 - p_{1,1})(1 - p_{2,1})p_{3,1} \frac{\partial C_{4,2}}{\partial S_{5,3}} e^{-3\tau,\Delta t}
\]

1.5. Application using FTSE 100 options data

We use the daily closing prices of FTSE 100 call options of January 2003 to December 2003 as reported by LIFFE. For the risk-free rate \( r_f \), we use cubic spline interpolation for matching each option contract with a continuous interest rate that corresponds to the option’s maturity, by utilizing the 1-month to 12-month LIBOR offer rates, collected from Datastream.

\[17 \text{ FTSE 100 options are traded with expiries in March, June, September, and December. Additional serial contracts are introduced so that options trade with expiries in each of the nearest 4 months. FTSE 100 options expire on the third Friday of the expiry month. FTSE 100 options positions are marked-to-market daily based on the daily settlement price, which is determined by LIFFE and confirmed by the Clearing House.}\]
Our initial sample (for the 12 months period) consists of 99,051 observations. We adopt the following filtering rules:

i. Eliminate calls for which the call price is greater than the value of the underlying asset, i.e. $C_{Mkt} > S_{1,t}$. No observations are eliminated from this rule.

ii. Eliminate calls if the call price is less than its lower bound i.e. $C_{Mkt} < S_{1,t} e^{-rT} - Ke^{-rT}$. This rule eliminates 3,206 observations.

iii. Eliminate calls with time to maturity less than 6 calendar days, i.e. $T < 6$. This rule eliminates 3,109 observations.

iv. Eliminate calls if their closing price is less than 0.5 index points\(^{18}\). This rule eliminates 11,373 observations.

v. Eliminate calls for which the trading volume is zero (since we want highly liquid options for calibration). This rule eliminates 66,826 observations.

The final sample consists of 14,537 observations.

In the implementation, for $\sigma_i$ we use the at-the-money implied volatility given by LIFFE and for time to maturity, $T$ we use the calendar days to maturity. Also, since the underlying asset of the options on FTSE 100 is a futures contract, we make the standard assumption that the dividend yield equals the risk free rate. The model is applied every day, with $n = 6$ and also with $n = 7$. For each implementation, the options used have the same underlying asset and the same time to maturity.

The evidence for the behaviour of the futures volatility in the literature is not clear. According to Samuelson (1965) the volatility of futures price changes should increase as the delivery date nears. However, Bessembinder et al. (1996) find that the Samuelson hypothesis is not supported for options on financials futures. In order to choose the value of $\lambda$ that gives the best feasible solution we make consecutive draws from interval (1.5), which allows for both, positive and negative values of $\lambda$. The first value of $\lambda$ is that of its lower

---

\(^{18}\) FTSE 100 Index options are quoted in index points and have an assigned value of £10 per index point.
bound. However, since dividend yield equals risk free rate, instead of $|r_j - \delta|$ we set $1.E-8$. The next value of $\lambda$ equals the old plus an appropriately chosen step size.

For brevity, we present results only for the first trading day of each month of the year 2003 and only for $n = 6$ (Table 1.1). Trading Day is the trading day of each contract, Expiry is the expiration month of each contract, Asset is the value of the underlying asset at the specified trading day, $N$ is the number of contracts used for the calibration (the contracts that on the same trading day, have the same underlying asset and the same expiration day), Error is the value of the objective function, Penalty is the value of the penalty term. Ideally we want the error function and the penalty term to tend to zero. Maturity is the calendar days till the maturity of the contract, and lambda is the value of $\lambda$ that gives the best feasible solution. Also, we present results only when the number of option contracts is greater than 3, since with fewer options the distribution of the underlying asset taken will not be reliable\textsuperscript{19}.

\textsuperscript{19} In Table 1.1 we note that for the same contract (same underlying asset, same expiration) the number of contracts used in the model changes across months. That is because some contracts were removed because of the filtering rules.
<table>
<thead>
<tr>
<th>Trading Day</th>
<th>Expiry</th>
<th>Asset</th>
<th>N</th>
<th>Error</th>
<th>Penalty</th>
<th>Maturity</th>
<th>lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>01/02/2003</td>
<td>Jan-03</td>
<td>4014</td>
<td>17</td>
<td>7.933E-11</td>
<td>0</td>
<td>15</td>
<td>4.3429</td>
</tr>
<tr>
<td>01/02/2003</td>
<td>Feb-03</td>
<td>4019</td>
<td>19</td>
<td>4.2E-05</td>
<td>0</td>
<td>50</td>
<td>-1.3851</td>
</tr>
<tr>
<td>01/02/2003</td>
<td>Mar-03</td>
<td>3991</td>
<td>12</td>
<td>5.855E-12</td>
<td>0</td>
<td>78</td>
<td>-6.6823</td>
</tr>
<tr>
<td>01/02/2003</td>
<td>Jun-03</td>
<td>3995</td>
<td>11</td>
<td>2.721E-13</td>
<td>0</td>
<td>169</td>
<td>-0.3963</td>
</tr>
<tr>
<td>01/02/2003</td>
<td>Dec-03</td>
<td>3999</td>
<td>6</td>
<td>0.0208333</td>
<td>0</td>
<td>351</td>
<td>0.4096</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Feb-03</td>
<td>3675.5</td>
<td>16</td>
<td>7.254E-08</td>
<td>0</td>
<td>18</td>
<td>-3.9616</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Mar-03</td>
<td>3646</td>
<td>14</td>
<td>8.611E-11</td>
<td>0</td>
<td>46</td>
<td>-1.5180</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Apr-03</td>
<td>3644.5</td>
<td>15</td>
<td>2.722E-12</td>
<td>0</td>
<td>73</td>
<td>-6.4346</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>May-03</td>
<td>3645</td>
<td>6</td>
<td>2.355E-14</td>
<td>0</td>
<td>102</td>
<td>-6.6918</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Jun-03</td>
<td>3647</td>
<td>7</td>
<td>2.196E-12</td>
<td>0</td>
<td>137</td>
<td>-4.9425</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Sep-03</td>
<td>3640</td>
<td>5</td>
<td>1.337E-14</td>
<td>0</td>
<td>228</td>
<td>-1.1252</td>
</tr>
<tr>
<td>02/03/2003</td>
<td>Dec-03</td>
<td>3653.5</td>
<td>7</td>
<td>5.859E-11</td>
<td>0</td>
<td>319</td>
<td>-0.9666</td>
</tr>
<tr>
<td>03/03/2003</td>
<td>Mar-03</td>
<td>3657</td>
<td>16</td>
<td>6.572E-13</td>
<td>0</td>
<td>18</td>
<td>-3.9616</td>
</tr>
<tr>
<td>03/03/2003</td>
<td>Apr-03</td>
<td>3655</td>
<td>13</td>
<td>2.573E-11</td>
<td>0</td>
<td>45</td>
<td>-1.5455</td>
</tr>
<tr>
<td>03/03/2003</td>
<td>May-03</td>
<td>3655</td>
<td>9</td>
<td>2.466E-12</td>
<td>0</td>
<td>74</td>
<td>-8.3274</td>
</tr>
<tr>
<td>03/03/2003</td>
<td>Jun-03</td>
<td>3655.5</td>
<td>7</td>
<td>9.825E-14</td>
<td>0</td>
<td>109</td>
<td>-6.2312</td>
</tr>
<tr>
<td>03/03/2003</td>
<td>Sep-03</td>
<td>3645</td>
<td>9</td>
<td>1.735E-12</td>
<td>0</td>
<td>200</td>
<td>-3.3359</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>Apr-03</td>
<td>3684.5</td>
<td>16</td>
<td>4.548E-11</td>
<td>0</td>
<td>16</td>
<td>5.0218</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>May-03</td>
<td>3683.5</td>
<td>16</td>
<td>4.396E-12</td>
<td>0</td>
<td>45</td>
<td>-1.5510</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>Jun-03</td>
<td>3686.5</td>
<td>10</td>
<td>2.096E-11</td>
<td>0</td>
<td>80</td>
<td>-5.0920</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>Jul-03</td>
<td>3693</td>
<td>7</td>
<td>1.222E-11</td>
<td>0</td>
<td>108</td>
<td>-6.1873</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>Sep-03</td>
<td>3676.5</td>
<td>5</td>
<td>4.563E-11</td>
<td>0</td>
<td>171</td>
<td>-3.9295</td>
</tr>
<tr>
<td>04/01/2003</td>
<td>Mar-04</td>
<td>3667</td>
<td>5</td>
<td>1.301E-11</td>
<td>0</td>
<td>352</td>
<td>-1.8670</td>
</tr>
</tbody>
</table>

Table 1.1: Results for the application of the model on the 1st trading day of each month for the year 2003. Trading Day is the trading day of each contract, Expiry is the expiration month of each contract, Asset is the value of the underlying asset at the specified trading day, N is the number of contracts used for the calibration, Error is the value of the objective function, Penalty is the value of the penalty term, Maturity is the calendar days till the maturity of the contract and lambda is the value of λ that gives the best feasible solution.
<table>
<thead>
<tr>
<th>Trading Day</th>
<th>Expiry</th>
<th>Asset</th>
<th>N</th>
<th>Error</th>
<th>Penalty</th>
<th>Maturity</th>
<th>lambda</th>
</tr>
</thead>
<tbody>
<tr>
<td>05/01/2003</td>
<td>May-03</td>
<td>3874</td>
<td>15</td>
<td>8.178E-08</td>
<td>0</td>
<td>15</td>
<td>-4.1763</td>
</tr>
<tr>
<td>05/01/2003</td>
<td>Jun-03</td>
<td>3879</td>
<td>14</td>
<td>1.581E-11</td>
<td>0</td>
<td>50</td>
<td>-1.3588</td>
</tr>
<tr>
<td>05/01/2003</td>
<td>Jul-03</td>
<td>3885.5</td>
<td>10</td>
<td>2.535E-11</td>
<td>0</td>
<td>78</td>
<td>-8.6239</td>
</tr>
<tr>
<td>05/01/2003</td>
<td>Sep-03</td>
<td>3870.5</td>
<td>4</td>
<td>2.657E-12</td>
<td>0</td>
<td>141</td>
<td>-2.1062</td>
</tr>
<tr>
<td>05/01/2003</td>
<td>Mar-04</td>
<td>3869</td>
<td>5</td>
<td>2.448E-13</td>
<td>0</td>
<td>322</td>
<td>-1.8172</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Jun-03</td>
<td>4132</td>
<td>16</td>
<td>3.535E-09</td>
<td>0</td>
<td>18</td>
<td>25.1320</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Jul-03</td>
<td>4138.5</td>
<td>9</td>
<td>6.685E-12</td>
<td>0</td>
<td>46</td>
<td>-1.4751</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Aug-03</td>
<td>4128.5</td>
<td>9</td>
<td>9.668E-13</td>
<td>0</td>
<td>74</td>
<td>-9.0523</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Sep-03</td>
<td>4124</td>
<td>11</td>
<td>5.437E-11</td>
<td>0</td>
<td>109</td>
<td>-6.0923</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Dec-03</td>
<td>4136.5</td>
<td>9</td>
<td>3.798E-14</td>
<td>0</td>
<td>200</td>
<td>-3.2689</td>
</tr>
<tr>
<td>06/02/2003</td>
<td>Jun-04</td>
<td>4124</td>
<td>5</td>
<td>1.825E-11</td>
<td>0</td>
<td>381</td>
<td>-1.6875</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Jul-03</td>
<td>3967</td>
<td>13</td>
<td>0.0007343</td>
<td>0</td>
<td>17</td>
<td>13.5874</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Aug-03</td>
<td>3958</td>
<td>12</td>
<td>4.422E-12</td>
<td>0</td>
<td>45</td>
<td>0.1871</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Sep-03</td>
<td>3955</td>
<td>12</td>
<td>3.807E-11</td>
<td>0</td>
<td>80</td>
<td>-1.5851</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Oct-03</td>
<td>3959</td>
<td>4</td>
<td>5.92E-14</td>
<td>0</td>
<td>108</td>
<td>-6.1578</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Dec-03</td>
<td>3964</td>
<td>11</td>
<td>1.544E-12</td>
<td>0</td>
<td>171</td>
<td>-1.8451</td>
</tr>
<tr>
<td>07/01/2003</td>
<td>Mar-04</td>
<td>3956</td>
<td>7</td>
<td>1.265E-13</td>
<td>0</td>
<td>261</td>
<td>-2.4906</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Aug-03</td>
<td>4091.5</td>
<td>11</td>
<td>7.597E-12</td>
<td>0</td>
<td>14</td>
<td>1.1396</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Sep-03</td>
<td>4088.5</td>
<td>14</td>
<td>7.022E-11</td>
<td>0</td>
<td>49</td>
<td>5.6305</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Oct-03</td>
<td>4094.5</td>
<td>4</td>
<td>5.318E-11</td>
<td>0</td>
<td>77</td>
<td>-8.6243</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Nov-03</td>
<td>4096.5</td>
<td>5</td>
<td>2.43E-13</td>
<td>0</td>
<td>112</td>
<td>-5.8741</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Dec-03</td>
<td>4100.5</td>
<td>5</td>
<td>1.804E-14</td>
<td>0</td>
<td>140</td>
<td>-0.4681</td>
</tr>
<tr>
<td>08/01/2003</td>
<td>Mar-04</td>
<td>4097</td>
<td>4</td>
<td>1.013E-12</td>
<td>0</td>
<td>230</td>
<td>-1.7231</td>
</tr>
</tbody>
</table>

Table 1.1 (continued): Results for the application of the model on the 1st trading day of each month for the year 2003. Trading Day is the trading day of each contract, Expiry is the expiration month of each contract, Asset is the value of the underlying asset at the specified trading day, N is the number of contracts used for the calibration, Error is the value of the objective function, Penalty is the value of the penalty term, Maturity is the calendar days till the maturity of the contract and lambda is the value of $\lambda$ that gives the best feasible solution.
<table>
<thead>
<tr>
<th>Trading Day</th>
<th>Expiry</th>
<th>Asset</th>
<th>N</th>
<th>Error</th>
<th>Penalty</th>
<th>Maturity</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>08/01/2003</td>
<td>Jun-04</td>
<td>4098.5</td>
<td>6</td>
<td>7.7014E-12</td>
<td>0</td>
<td>321</td>
<td>0.9996</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Sep-03</td>
<td>4215</td>
<td>16</td>
<td>9.231E-11</td>
<td>0</td>
<td>18</td>
<td>-3.8161</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Oct-03</td>
<td>4222</td>
<td>9</td>
<td>4.986E-13</td>
<td>0</td>
<td>46</td>
<td>-1.4569</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Nov-03</td>
<td>4225</td>
<td>9</td>
<td>4.4323E-12</td>
<td>0</td>
<td>81</td>
<td>-8.1673</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Dec-03</td>
<td>4229</td>
<td>12</td>
<td>1.0014E-12</td>
<td>0</td>
<td>109</td>
<td>-0.6026</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Mar-04</td>
<td>4224.5</td>
<td>4</td>
<td>1.1111E-11</td>
<td>0</td>
<td>199</td>
<td>-32.5850</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Oct-03</td>
<td>4162.5</td>
<td>12</td>
<td>1.1715E-06</td>
<td>0</td>
<td>16</td>
<td>-4.3573</td>
</tr>
<tr>
<td>09/01/2003</td>
<td>Nov-03</td>
<td>4167</td>
<td>12</td>
<td>1.7641E-13</td>
<td>0</td>
<td>51</td>
<td>-1.3282</td>
</tr>
<tr>
<td>10/01/2003</td>
<td>Dec-03</td>
<td>4169.5</td>
<td>19</td>
<td>2.796E-12</td>
<td>0</td>
<td>79</td>
<td>-8.4832</td>
</tr>
<tr>
<td>10/01/2003</td>
<td>Jan-04</td>
<td>4173.5</td>
<td>5</td>
<td>5.4217E-13</td>
<td>0</td>
<td>107</td>
<td>-62.1642</td>
</tr>
<tr>
<td>10/01/2003</td>
<td>Mar-04</td>
<td>4162</td>
<td>8</td>
<td>1.1045E-11</td>
<td>0</td>
<td>169</td>
<td>-3.8885</td>
</tr>
<tr>
<td>10/01/2003</td>
<td>Jun-04</td>
<td>4171.5</td>
<td>5</td>
<td>1.0698E-13</td>
<td>0</td>
<td>260</td>
<td>-2.4977</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Nov-03</td>
<td>4330</td>
<td>12</td>
<td>9.9737E-12</td>
<td>0</td>
<td>18</td>
<td>26.2023</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Dec-03</td>
<td>4333</td>
<td>15</td>
<td>4.1511E-11</td>
<td>0</td>
<td>46</td>
<td>0.5538</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Jan-04</td>
<td>4344</td>
<td>9</td>
<td>3.2388E-12</td>
<td>0</td>
<td>74</td>
<td>-8.7229</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Feb-04</td>
<td>4354</td>
<td>7</td>
<td>2.0997E-12</td>
<td>0</td>
<td>109</td>
<td>-2.1552</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Mar-04</td>
<td>4329</td>
<td>7</td>
<td>2.6622E-13</td>
<td>0</td>
<td>136</td>
<td>-4.7931</td>
</tr>
<tr>
<td>11/03/2003</td>
<td>Jun-04</td>
<td>4343.5</td>
<td>7</td>
<td>4.4587E-13</td>
<td>0</td>
<td>227</td>
<td>-28.4200</td>
</tr>
<tr>
<td>12/01/2003</td>
<td>Dec-03</td>
<td>4415.5</td>
<td>13</td>
<td>6.9791E-11</td>
<td>0</td>
<td>18</td>
<td>15.2128</td>
</tr>
<tr>
<td>12/01/2003</td>
<td>Jan-04</td>
<td>4426</td>
<td>10</td>
<td>2.5172E-13</td>
<td>0</td>
<td>46</td>
<td>-1.4475</td>
</tr>
<tr>
<td>12/01/2003</td>
<td>Feb-04</td>
<td>4433.5</td>
<td>13</td>
<td>3.7691E-12</td>
<td>0</td>
<td>81</td>
<td>-8.1270</td>
</tr>
<tr>
<td>12/01/2003</td>
<td>Mar-04</td>
<td>4410.5</td>
<td>10</td>
<td>5.3369E-12</td>
<td>0</td>
<td>108</td>
<td>-6.0614</td>
</tr>
<tr>
<td>12/01/2003</td>
<td>Jun-04</td>
<td>4423.5</td>
<td>4</td>
<td>9.9101E-16</td>
<td>0</td>
<td>199</td>
<td>-3.2410</td>
</tr>
</tbody>
</table>

Table 1.1 (continued): Results for the application of the model on the 1st trading day of each month for the year 2003. **Trading Day** is the trading day of each contract, **Expiry** is the expiration month of each contract, **Asset** is the value of the underlying asset at the specified trading day, **N** is the number of contracts used for the calibration, **Error** is the value of the objective function, **Penalty** is the value of the penalty term, **Maturity** is the calendar days till the maturity of the contract and **$\lambda$** is the value of $\lambda$ that gives the best feasible solution.
The results obtained support our modeling approach. As we can see in Table 1.1, in all cases the solution strictly satisfies the constraints since the penalty term equals zero. Also, we see that in 67 out of 69 cases, i.e. in 97.1% of the cases the error function tends to zero with an average value of 2.34E-08. In the other 2 cases, where the error function is greater than 1.E-4, the average error is 0.01. Similar results were found for \( n = 7 \).

Even though the problem requires a constrained non-convex optimization in \( 2(2^{n-1} - 1) \) variables, the use of efficient optimization algorithms prevents the calibration of the model from becoming computationally too intensive. On average, the computational time in minutes required for each calibration had a mean (median) 1.10 (0.03) for \( n = 6 \) and 2.27 (0.08) for \( n = 7 \). The computer used for the calibration of the model had the following specifications: a Pentium 4 (3.2 GHz) CPU, Memory 1GB (RAM), and Windows XP Professional operating system. The codes were written in Matlab R2006a. The computational time needed would have decreased if the codes were written in the C/C++ language.

When models provide an exact fit there is always the concern of over-fitting. We checked the model for over-fitting by pricing options with strikes in-between those used for the optimization (calibration). Then we made plots of the call prices (market prices and estimated from the model) versus moneyness. Over-fitting was also checked using a restricted sample consisting only of options with moneyness between 0.8 and 1.1, since these options are expected to be more liquid and more accurately priced\(^\text{20}\). For brevity, we exhibit only the plots for optimizations done in the first trading day of June (middle of the year) for the two samples using a tree with \( n = 6 \). As we see, for both samples the estimated call values increase smoothly with increasing moneyness without any evidence of over-fitting (see Fig.1.4). Similar results were obtained when a tree with \( n = 7 \) was used for the calibration procedure.

\(^\text{20}\) This sub-sample has a total of 13,696 observations for the year 2003.
Figure 1.4: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003.
As a further check for over-fitting we use only part of the information to calibrate the tree and the other part to check the model using \( n = 6, 7, 8 \). Specifically, we leave out consecutively one of the \( N \) options at each time and we calibrate our model with the remaining options. In order to preserve the options’ moneyness range stable and avoid problems of extrapolation, we do not remove the options with the highest and lowest moneyness. Over-fitting is checked like before using the full and the restricted sample of options. For the calibration only cases consisting of \( N > 8 \) were used. Results for the mean and median absolute errors are given in Table 1.2. We see that the error (given an average contract size of 90 for the full and 74.4 for the restricted sample) is small and rather stable\(^{21}\).

<table>
<thead>
<tr>
<th>Model</th>
<th>Absolute Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Sample</td>
</tr>
<tr>
<td>( n = 6 )</td>
<td>Mean 1.2458</td>
</tr>
<tr>
<td></td>
<td>Median 0.9163</td>
</tr>
<tr>
<td>( n = 7 )</td>
<td>Mean 1.1375</td>
</tr>
<tr>
<td></td>
<td>Median 0.8005</td>
</tr>
<tr>
<td>( n = 8 )</td>
<td>Mean 1.1286</td>
</tr>
<tr>
<td></td>
<td>Median 0.7928</td>
</tr>
<tr>
<td>Observations</td>
<td>446</td>
</tr>
</tbody>
</table>

Table 1.2: Mean and median absolute errors using our model for \( n=6, 7, 8 \) and data from the full and the restricted sample.

Since implied volatility changes with strike and time to maturity (volatility smile) the index should have a non-lognormal distribution which implies that the log-returns will deviate from normality. In order to see how realistic is the distribution obtained from our model for year 2003, we calculate the statistics of the 1-month log-returns obtained from our model and compare them with the historical 1-month log-returns for the year 2003 and the years 2001-2005. Specifically, for each calibration

\(^{21}\) Also, we compare our model (with respect to over-fitting) with the Black-Scholes model using the Whaley (1982) approach. According to this approach we find the volatility that minimizes the sum of square differences of the Black-Scholes option prices with their corresponding market prices using nonlinear minimization. Results show that the mean (median) absolute error using this approach is 7.36 (5.94) for the full sample and 6.61 (5.60) for the restricted sample which are much higher than the errors obtained using our model for \( n=6, 7, 8 \).
(with \( n = 6 \) and \( n = 7 \)) for which the options maturity was between 28 and 32 calendar days, we calculate the first four moments (mean, variance, skewness and kurtosis). Then, in order to get a feeling for the representative statistics of 1-month log-returns we provide for each of those moments the mean and the median. The statistics for \( n = 6 \) are summarized in Table 1.3. Similar statistics were found for \( n = 7 \). Liu et al. (2005) discuss the derivations of historical, and implied real and risk-neutral distributions for the FTSE 100 index. They demonstrate that the needed adjustments to get the implied real variance, skewness and kurtosis from the implied risk-neutral ones are minimal. Thus, knowing that our implied risk-neutral moments (beyond the mean) are very close to the implied real ones, we can then compare them with the historical ones (without expecting the two distributions to be identical). As we would expect, the mean of the implied risk-neutral distribution of log-returns differs from that of the historical distribution\textsuperscript{22}. Also, as we see, both the implied risk-neutral and the historical distribution deviate from normality since they exhibit negative skewness and (mostly) excess kurtosis. This is an indication that the implied distribution is realistic.

<table>
<thead>
<tr>
<th>Implied (2003, ( n = 6 ))</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.0024</td>
<td>0.0048</td>
<td>-0.6938</td>
<td>4.5075</td>
<td>58</td>
</tr>
<tr>
<td>Median</td>
<td>-0.0013</td>
<td>0.0027</td>
<td>-0.6653</td>
<td>3.6405</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Historical</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.0106</td>
<td>0.0014</td>
<td>-0.6572</td>
<td>2.7689</td>
<td>12</td>
</tr>
<tr>
<td>2001-2005</td>
<td>-0.0021</td>
<td>0.0018</td>
<td>-1.1177</td>
<td>4.4749</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 1.3: Implied risk-neutral and historical statistics of the distribution of the FTSE 100 1-month log-returns.

In order to give further evidence for the implied distributions obtained by our model, representative implied distributions (histograms) for the 1-month log-returns in June 2003 are shown in Figures 1.5a, 1.5b (full sample) and 1.6a, 1.6b (restricted sample) for \( n = 6 \) and \( n = 7 \). To make the histograms of the implied distributions we

\textsuperscript{22} Due to rounding errors in the estimations, the implied risk-neutral mean is not equal to zero.
make use of the Pearson system of distributions as applied in Matlab. Using the first four moments of the data it is easy to find in the Pearson system the distribution that matches these moments and to generate a random sample in order to produce a histogram corresponding to the implied distribution. From the figures, it is obvious that the implied distributions have negative skewness and positive kurtosis which is consistent with historical data. These figures are representative of the vast majority of cases. Another interesting thing we observe is that distributions for $n = 6$ and $n = 7$ are practically indistinguishable for both samples.

![Figure 1.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the full sample for $n = 6$.](image)

---

23 In the Pearson system there is a family of distributions that includes a unique distribution corresponding to every valid combination of mean, standard deviation, skewness, and kurtosis.

24 Copyright 2005 The MathWorks, Inc.

25 In rare exceptions only we have implied distributions close to normal or even leptokurtic.
Figure 1.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the full sample for $n = 7$.

Figure 1.6a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the restricted sample for $n = 6$. 
1.6. Conclusions

In most options markets, the implied Black–Scholes volatilities vary with both strike and expiration, a relationship commonly known as the volatility smile. In this chapter we capture the implied distribution from option market data using a non-recombining (binary) tree allowing the local volatility to be a function of the underlying asset and of time. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and use nonlinear constrained optimization to minimize the least squares error function on market prices. Specifically we adopt a penalty method and the optimization is implemented using a Quasi-Newton algorithm. Appropriate constraints allow us to maintain risk neutrality and to prevent arbitrage opportunities. The proposed model can accommodate European options with single maturities and, with minor modifications, options with multiple maturities. Also, this
method is flexible since it applies to arbitrary underlying asset distributions, which implies arbitrary local volatility distributions. Market implied information embodied in the constructed tree can help the pricing and hedging of exotic options and of OTC options on the same underlying process. We test our model using FTSE 100 options data. The results obtained strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem, and the derived implied distributions are realistic. Also, the computational burden is not a major issue.
Appendix 1.1A: Feasibility of the initialized non-recombining tree

We initialize the tree using the following volatility term structure:

\[ \sigma_i = \sigma_1 e^{\lambda (i-1)\Delta t}, \quad \lambda \in R \text{ where } i=1, \ldots, n \]

The feasibility of the initial tree depends on the right choice of the local volatility term structure; hence to obtain a feasible initial tree we must find an interval with the appropriate values of \( \lambda \). In order to preserve the risk neutrality at every time step, the following constraints must be satisfied:

\[ S_{i,j} e^{(r_f - \delta)\Delta t} \leq S_{i+1,2j} \]  
\[ S_{i,j} e^{(r_f - \delta)\Delta t} \geq S_{i+1,2j-1} \]

Also,

\[ S_{i+1,2j} = S_{i,j} u_{i,j} = S_{i,j} e^{\sigma_j \sqrt{\Delta t}} \]   
\[ S_{i+1,2j-1} = S_{i,j} d_{i,j} = S_{i,j} e^{-\sigma_j \sqrt{\Delta t}} \]

Substituting (1.A.2a) and (1.A.2b) to (1.A.1a) and (1.A.1b) respectively we get the following inequalities:

\[ \sigma_i \geq (r_f - \delta) \sqrt{\Delta t} \]   
\[ \sigma_i \geq -(r_f - \delta) \sqrt{\Delta t} \]
Thus we have that,

$$\sigma_i \geq r_f - \delta \sqrt{\Delta t} \quad \forall i$$  \hspace{1cm} (1.A 4)

For $\lambda \geq 0$, $\sigma_i = \sigma_i e^{\lambda (i-1)\Delta t}$ is strictly increasing. Since (1.A 4) holds for every $i$ this means that

$$\min \sigma_i \geq |r_f - \delta \sqrt{\Delta t} \quad \text{or}$$

$$\sigma_1 \geq |r_f - \delta |\sqrt{\Delta t}$$ \hspace{1cm} (1.A 5)

The minimum value of $\sigma_i$ is for $i=1 (\sigma_1)$, thus (1.A 5) is independent of $\lambda$. Therefore, if $\lambda$ is positive there is no upper bound for $\lambda$.

For $\lambda < 0$, $\sigma_i = \sigma_i e^{\lambda (i-1)\Delta t}$ is strictly decreasing. Since (1.A 4) holds for every $i$ this means that

$$\min \sigma_i \geq |r_f - \delta \sqrt{\Delta t}$$

$$\sigma_n \geq |r_f - \delta \sqrt{\Delta t}$$

$$e^{\lambda (n-1)\Delta t} \geq \frac{|r_f - \delta \sqrt{\Delta t}}{\sigma_1}$$

But, $(n-1)\Delta t = T$, thus,

$$\lambda \geq \frac{1}{T} \log \left( \frac{|r_f - \delta \sqrt{\Delta t}}{\sigma_1} \right)$$ \hspace{1cm} (1.A 6)
If we allow $\lambda$ to take both negative and positive values, then $\lambda$ should belong in the interval,

$$
\lambda \in \left[ \frac{1}{T} \log \left( \frac{|r_f - \delta| \sqrt{\Delta t}}{\sigma_i} \right), +\infty \right)
$$

(1.A 7)
Appendix 1.1B: Feasibility of the initialized non-recombining tree assuming time dependent $r_f, \delta$ and $\Delta t$

We denote with $r_f(i)$ and $\delta(i)$ the risk free rate and dividend yield respectively between two consecutive time steps, i.e. between time step $i$ and $i+1$, $i=1, ..., n-1$.

![Diagram of the non-recombining tree](image-url)

**Figure 1.A1: A typical triplet in the initialization of the non-recombining tree assuming $r_f$, $\delta$ and $\Delta t$ to be time dependent.**

We initialize the tree using the following volatility term structure:

$$\sigma_i = \sigma_r e^{\lambda \sum_{j=i}^{n-1} \Delta t(j)}$$

where $\lambda \in R$ where $i=1, ..., n$

The feasibility of the initial tree depends on the right choice of the local volatility term structure; hence to obtain a feasible initial tree we must find an interval with the appropriate values of $\lambda$. In order to preserve the risk neutrality at every time step, the following constraints must be satisfied:

$$S_{i,j} e^{(r_f(i)-\delta(i))\Delta t(i)} \leq S_{i+1,2j}$$

(1.A 1a')
\[ S_{i,j} e^{(r_j(i) - \delta(i))\Delta t(i)} \geq S_{i+1,2j-1} \quad (1.A \text{ 1b}') \]

Also,

\[ S_{i+1,2j} = S_{i,j}u_{i,j} = S_{i,j}e^{\sigma_i \sqrt{\Delta t(i)}} \quad (1.A \text{ 2a}') \]

\[ S_{i+1,2j-1} = S_{i,j}d_{i,j} = S_{i,j}e^{-\sigma_i \sqrt{\Delta t(i)}} \quad (1.A \text{ 2b}') \]

Substituting (1.A 2a') and (1.A 2b') to (1.A 1a') and (1.A 1b') respectively we get the following inequalities:

\[ \sigma_i \geq (r_f(i) - \delta(i))\sqrt{\Delta t(i)} \quad (1.A \text{ 3a}') \]

\[ \sigma_i \geq -(r_f(i) - \delta(i))\sqrt{\Delta t(i)} \quad (1.A \text{ 3b}') \]

Thus we have that

\[ \sigma_i \geq |r_f(i) - \delta(i)|\sqrt{\Delta t(i)} \quad \forall i \]

(1.A 4’)

For \( \lambda \geq 0 \), \( \sigma_i = \sigma_1 e^{\frac{1}{\lambda} \sum_{j}^{\Delta t(i)} \lambda} \) is strictly increasing.

Let \( \xi_M = \max_i |r_f(i) - \delta(i)|\sqrt{\Delta t(i)} \)

Then (1.A 4’) holds for every \( i \) if

\[ \min_i \sigma_i \geq \xi \quad \text{or} \quad \sigma_1 \geq \xi \]

(1.A 5’)

The minimum value of \( \sigma_i \) is for \( i=1 \) (\( \sigma_1 \)), thus (1.A 5’) is independent of \( \lambda \). Therefore, if \( \lambda \) is positive there is no upper bound for \( \lambda \).
For \( \lambda < 0 \), \( \sigma_i = \sigma_i e^{\sum_{j=1}^{n=1} \Delta(j)} \) is strictly decreasing.

Let \( \xi_m = \min_i |r_f(i) - \delta(i)|/\Delta(i) \)

Then (1.A 4') holds for every \( i \) if

\[
\min_i \sigma_i \geq \xi_m \\
\sigma_n \geq \xi_m \\
\sigma_i e^{\sum_{j=1}^{n=1} \Delta(j)} \geq \xi_m
\]

But, \( \sum_{j=1}^{n=1} \Delta(t(j)) = T \), thus,

\[
\dot{\lambda} \geq \frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right) \quad (1.A 6')
\]

If we allow \( \lambda \) to take both negative and positive values, then \( \lambda \) should belong in the interval,

\[
\dot{\lambda} \in \left[ \frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right), +\infty \right) \quad (1.A 7')
\]
Appendix 1.2: Formulas adjusted for time dependent $r_f$, $\delta$ and $\Delta t$

We denote with $r_f(i)$ and $\delta(i)$ the risk free rate and dividend yield respectively between two consecutive time steps, i.e. between time step $i$ and $i+1$, $i = 1, \ldots, n - 1$ and with $r_f'$ and $\delta'$ we denote the risk free rate and dividend yield respectively from today till the maturity of the option, i.e. from $i = 1$ to $i = n$.

If we allow $r_f$, $\delta$ and $\Delta t$ to be time dependent the equations of the main text are replaced with the following:

\[
u_{i,j} = e^{\sigma \sqrt{\Delta t(i)}}
\]
\[
d_{i,j} = e^{-\sigma \sqrt{\Delta t(i)}} = \frac{1}{u_{i,j}}, i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]
\[
\lambda \in \left[\frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right), +\infty \right)
\]

where,

\[
\xi_m = \min_{i} \left| r_f(i) - \delta(i) \right| \sqrt{\Delta t(i)}
\]
\[
p_{i,j} = \frac{S_{i,j} e^{(r_f(i) - \delta(i)) \Delta t(i)} - S_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}}, i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]
\[
C_{i,j} = \left( p_{i,j} C_{i+1,2j} + (1 - p_{i,j}) C_{i+1,2j-1} \right) e^{-r_f(i) \Delta t(i)}
\]
\[
i = n - 1, \ldots, 1, \quad j = 1, \ldots, 2^{i-1}
\]
\[
S_{i,j} e^{(r_f(i) - \delta(i)) \Delta t(i)} \leq S_{i+1,2j}, \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]
\[
S_{i,j} e^{(r_f(i) - \delta(i)) \Delta t(i)} \geq S_{i+1,2j-1}, \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]
\[
\max \left( S_{1,1} e^{-\delta T} - K e^{-r_j T}, 0 \right) \leq C_{\text{Mod}} \leq S_{1,1} \tag{1.11'}
\]

\[
g_1(i, j) = S_{i,j} e^{(r_f(i)-\delta(i))\Delta(t)} - S_{i+1,j-1} \geq 0, \ i = 1, \ldots, n-1, \ j = 1, \ldots, 2^{i-1} \tag{1.14a'}
\]

\[
g_2(i, j) = S_{i+1,j} - S_{i,j} e^{(r_f(i)-\delta(i))\Delta(t)} \geq 0, \ i = 1, \ldots, n-1, \ j = 1, \ldots, 2^{i-1} \tag{1.14b'}
\]

\[
g_3(k) = S_{1,1} - C_{\text{Mod}}(k) \geq 0, \ k = 1, \ldots, N \tag{1.14c'}
\]

\[
g_4(k) = C_{\text{Mod}}(k) - \max(S_{1,1} e^{-\delta T} - K(k) e^{-r_j T}, 0) \geq 0, \ k = 1, \ldots, N \tag{1.14d'}
\]

\[
\nabla S_{i,j} p_{i,j} \equiv \begin{bmatrix}
\frac{\partial p_{i,j}}{\partial S_{i,j}} \\
\frac{\partial p_{i,j}}{\partial S_{i+1,j}} \\
\frac{\partial p_{i,j}}{\partial S_{i+1,j-1}}
\end{bmatrix} = \frac{1}{S_{i+1,j} - S_{i+1,j-1}} \begin{bmatrix}
e^{(r_f(i)-\delta(i))\Delta(t)} - p_{i,j} \\
- p_{i,j} \\
- (1 - p_{i,j})
\end{bmatrix} \tag{1.17'}
\]

\[
\nabla S_{i,j} C_{i,j} \equiv \begin{bmatrix}
\frac{\partial C_{i,j}}{\partial S_{i,j}} \\
\frac{\partial C_{i,j}}{\partial S_{i+1,j}} \\
\frac{\partial C_{i,j}}{\partial S_{i+1,j-1}}
\end{bmatrix} = \begin{bmatrix}
\Delta_{i,j} \\
\Delta_{i+1,j} - \Delta_{i,j} e^{\delta(i)\Delta(t)} e^{-r_f(i)\Delta(t)} \\
(1 - p_{i,j}) \Delta_{i+1,j-1} - \Delta_{i,j} e^{\delta(i)\Delta(t)} e^{-r_f(i)\Delta(t)}
\end{bmatrix} \tag{1.18'}
\]

\[
\Delta_{i,j} = \frac{C_{i+1,j} - C_{i+1,j-1}}{S_{i+1,j} - S_{i+1,j-1}} e^{-\delta(i)\Delta(t)} = \frac{\partial C_{i,j}}{\partial S_{i,j}} \equiv \text{Delta Ratio} \tag{1.19'}
\]

\[
\frac{\partial C_{1,1}}{\partial S_{1,1}} = \prod \left\{ \text{of the probabilities on the path that take us from node (1,1) to node (i-1,k)} \right\}
\]

\[
\propto - \frac{\partial C_{1,k}}{\partial S_{1,j}} e^{-\sum_{t=k}^{N} (r_{t}(h)\Delta(h))}
\]

\[
\tag{1.21'}
\]
\[ k = \begin{cases} \frac{j}{2} & \text{for even } j \\ \frac{(j+1)}{2} & \text{for odd } j \end{cases} \]
2. An Empirical Comparison between Non-Recombining and Recombining Implied Trees

Abstract

In this chapter we compare the cross-sectional pricing performance of two different implied trees. The models compared are the implied non-recombining tree (INRT) proposed in the first chapter, and a modified version of the Derman and Kani (DK, 1994) implied recombining tree, a very popular model among practitioners. The pricing performance of the two models is also compared with an ad-hoc procedure of smoothing Black-Scholes implied volatilities across strikes, proposed by Dumas et al. (1998). We use European and American call options of the FTSE 100 index for the year 2003. For each date in our sample, we calibrate the implied trees using the panel of all European style call options with the same underlying asset and time to maturity and then examine how well these models price the corresponding panel of American call options. Results show that the pricing performance of the INRT model is better than that of the DK model whereas the pricing performance of the ad-hoc procedure is not statistically different from the INRT model.
2.1. Introduction

The Black-Scholes (BS, 1973) model is based on the assumption that the underlying asset evolves according to a geometric Brownian motion with a constant volatility factor. This, however, contradicts the observed implied volatility, which suggests that volatility depends on both the strike and maturity of an option, a relationship commonly known as the volatility smile. This problem has motivated the recent literature on “smile consistent” no-arbitrage models. Consistency is achieved by extracting an implied evolution for the stock price from market prices of liquid standard options on the underlying asset. There are two classes of methodologies within this approach. Smile consistent deterministic volatility models and stochastic volatility smile consistent models which allow for smile-consistent option pricing under the no-arbitrage evolution of the volatility surface\textsuperscript{26,27}. Also, in the previous chapter, we propose a non-recombining implied tree (INRT) without imposing any restrictive assumptions for the underlying stochastic process.

Despite the vast literature on smile consistent models, there are only few studies that test the empirical performance of these models. Dumas \textit{et al.} (1998) examine the predictive and hedging performance of a class of smile-consistent deterministic volatility models using options data of the S&P 500 from June 1988 through December 1993. They find that the smile-consistent models perform no better than an ad-hoc procedure that smoothes the BS implied volatilities across strikes and time to maturity.

Since deterministic volatility functions are used more often for cross-sectional pricing than for out-of-sample pricing and hedging, Brandt and Wu (2002) extends the study of Dumas \textit{et al.} (1998) and tests the cross-sectional performance of option


\textsuperscript{27} There also exist non-parametric methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.
pricing smile-consistent models in which the volatility is a deterministic function of strike and time to maturity. Specifically, they examine the pricing performance of a modified version of the Derman and Kani (DK, 1994) implied tree, the Cox-Ross-Rubinstein (CRR, 1979), and the ad-hoc model by using daily prices of the FTSE 100 index options over the period October 1995 through September 1997. In their application, they employ Legendre polynomials to interpolate the implied volatilities across strikes and maturities. For each date in their sample they calibrate the models using the market prices of European style options, and then they use the calibrated models to price the American options. In line with Dumas et al. (1998), they find that the implied binomial tree model performs no better than the ad-hoc model.

Lim and Zhi (2002) compare the pricing performance and delta hedging performance of the DK model and Jackwerth (1997) generalized binomial model, and the standard binomial CRR model using daily data of the FTSE 100 over the period January through November 1999. They find that the DK model produces least hedging errors and best results for American call options with earlier maturity than the maturity span of the implied trees. On the other hand, the generalized binomial tree performed better for American at-the-money put options for any maturity.

Linaras and Skiadopoulos (2005) adopt the approach taken by Brandt and Wu (2002) and Lim and Zhi (2002) and apply it to a different data set, the liquid S&P 100 option data traded in the Chicago Board of Exchange (CBOE). They investigate the pricing performance of the DK and the Barle and Cakici (BC, 1995) models versus the standard CRR and the ad-hoc model. Two different interpolation methods (linear and cubic) were used to investigate the effect of the interpolation method on the pricing performance of each model. In line with Brandt and Wu (2002) they find that the interpolation method affects the pricing performance, and in fact a more complex interpolation method may not necessarily improve the pricing performance of the model. Also, they conclude that without loss of accuracy in pricing terms, simpler models that lack a theoretical foundation can be used (ad-hoc models). Finally, similar to Lim and Zhi (2002) they find that different methods should be used to different applications and with caution.
In this chapter, we compare the performance of recombining versus non-recombining implied trees, in pricing of American call options. For a recombining model we use a modified version of the \textit{DK} model, which is a very popular model among practitioners. For a non-recombining model we use the \textit{INRT} model proposed in the first chapter. We also include in the comparison an ad-hoc procedure for smoothing BS implied volatilities across strikes, proposed by Dumas \textit{et al.} (1998). The \textit{DK} implied binomial tree captures the volatility smile. However, negative or greater than one probabilities can occur. In this case, the probabilities are overridden causing the tree to lose much of its implied nature. Also, the \textit{DK} tree is sensitive to the interpolation/extrapolation method. On the other hand, the \textit{INRT} model, because of the many degrees of freedom allows for a more flexible underlying asset distribution. This prevents negative and greater than one transition probabilities. Also, in the \textit{INRT} model, there is no need for interpolation and extrapolation. The ad-hoc model is easily and fast implemented but it is internally inconsistent.

We calibrate the models using European call option data of the FTSE 100 index for the year 2003, obtained from LIFFE. For each date in our sample, we calibrate the implied trees using the panel of all European style call options with the same underlying asset and time to maturity and then examine how well these models price the corresponding panel of American call options. In order to investigate the effect of the interpolation method across strikes on the pricing performance of the \textit{DK} model and the ad-hoc model two methods are used, the 2\textsuperscript{nd} order polynomial regression and a cubic spline. Results show that the pricing performance of the \textit{INRT} model is better than that of a modified version of the \textit{DK} model whereas the pricing performance of the ad-hoc procedure is not statistically different from the \textit{INRT} model. Also, similar to Brandt and Wu (2002) and Linaras and Skiadopoulos (2005) we find that simpler interpolation methods give better results.

The remainder of the chapter is as follows: In section 2.2 we describe the data set and the filtering rules applied and section 2.3 describes the different methodologies used. Tests of the models and results are presented in section 2.4. Section 2.5 concludes.
2.2. Data - FTSE 100 index options

We use the daily closing prices of FTSE 100 American and European call options of January 2003 to December 2003 as reported by LIFFE. The strike prices for a given style option (European or American) are spaced at intervals of 50 index points from each other. However, there are no American and European calls with identical strikes. The strike prices for adjacent European and American style calls are spaced at intervals of 25 index points. For example, there are European style calls with strikes 3075, 3125, and 3175 and American style calls with strikes 3100, 3150, and 3200. Also, the longest maturity for the European calls is 2 years while for the American calls is 6 months. Our initial sample (for the 12 months period) consists of 99,051 observations of European calls, and 34,503 observations of American calls. We apply five filtering rules to both American and European calls data. First we exclude calls that violate the no-arbitrage bounds. Second, we eliminate calls with time to maturity less than 6 calendar days, i.e. $T<6$. These options have very small time-premiums and their implied volatilities are inaccurate since they are very sensitive to market microstructure problems and measurement errors (Hentschel, 2001, Brandt, et al., 2002). Third, we eliminate calls if their closing price is less than 0.5 index points. Fourth, we eliminate calls for which the traded volume is zero (since we want highly liquid options for calibration). Finally, we eliminate calls with moneyness greater than 1.1 or moneyness less than 0.8, since deep in the money and deep out of the money call options are expected to be illiquid and not accurately priced. We define as

\[ C_{EU} = \max(0, S e^{-rT} - K e^{-rT}) \leq C_{EU} \leq S \]  

and for American style call options

\[ C_{AM} = \max(C_{EU}, S - K) \leq C_{AM} \leq S. \]  

---

28 FTSE 100 European options are traded with expiries in March, June, September, and December. Additional serial contracts are introduced so that the nearest 4 months are always available for trading. FTSE 100 American options are traded with expiries the nearest of June and December. Additional serial contracts are introduced so that the nearest 3 months are always available for trading. FTSE 100 options expire on the third Friday of the expiry month. FTSE 100 options positions are marked-to-market daily based on the daily settlement price, which is determined by LIFFE and confirmed by the Clearing House. FTSE 100 options are quoted in index points and have an assigned value of £10 per index point.

29 For European style call options $C_{EU} = \max(0, S e^{-rT} - K e^{-rT}) \leq C_{EU} \leq S$ and for American style call options $C_{AM} = \max(C_{EU}, S - K) \leq C_{AM} \leq S$. 

---

Eleni D. Constantinide
moneyness of a call option the ratio (underlying asset price)/(strike price). After these filtering rules, our sample consists of 13,696 observations of European call options and 355 observations of American call options.

For each model calibration, the options used have the same underlying asset and the same time to maturity\textsuperscript{30}. Also, we consider only cases for which the number of options used for calibration, \( N \) is greater than 8 since with fewer options the distribution of the underlying asset obtained will not be reliable. For the pricing of American call options we use only the European call options that correspond to the American calls. Thus, the European sample used for pricing of American calls is reduced to 2,572 observations and the American sample is reduced to 312 observations. In the implementations for time to maturity, \( T \) we use the calendar days to maturity. For the risk-free rate \( r_f \) we use cubic spline interpolation for matching each option contract with a continuous interest rate that corresponds to the option’s maturity, by utilizing the 1-month to 12-month LIBOR offer rates, collected from Datastream. Also, since the underlying asset of the options on FTSE 100 is a futures contract, we make the standard assumption that the dividend yield (\( \delta \)) equals the risk free rate. The models are calibrated every day.

Tables 2.1a and 2.1b describe the cleaned sample. Table 2.1a shows the mean, median, minimum and maximum call option value and also the number of observations of the whole European call option sample, and also of the sub-samples of the out-of-the money (OTM), at-the-money (ATM), in-the-money (ITM), short term (ST), medium term (MT) and long term (LT) options\textsuperscript{31}. Table 2.1b shows the same statistics as Table 2.1a for the American options sample. We observe that the American data are much less than the corresponding European data (312 vs. 2,572 observations). Tables 2.2a and 2.2b show the same statistics for the trading volume of the European and the corresponding American options. Overall we see that the

\textsuperscript{30} In the specific dataset (given by LIFFE), the underlying asset of each call option is a future contract and thus every trading day the options have different underlying asset.

\textsuperscript{31} Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.
European options are highly traded with mean and median volume 343.5 and 60 respectively. The corresponding American options are less liquid with mean and median volume 57 and 10 respectively.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>84.2</td>
<td>33.5</td>
<td>115.1</td>
<td>220.6</td>
<td>81.7</td>
<td>84.7</td>
<td>88.6</td>
</tr>
<tr>
<td>median</td>
<td>47.3</td>
<td>19</td>
<td>112.5</td>
<td>211</td>
<td>40</td>
<td>49</td>
<td>51</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>30.5</td>
<td>72.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>410.5</td>
<td>263.5</td>
<td>230</td>
<td>410.5</td>
<td>398</td>
<td>400.5</td>
<td>410.5</td>
</tr>
<tr>
<td>observations</td>
<td>2,572</td>
<td>1,731</td>
<td>256</td>
<td>585</td>
<td>1,038</td>
<td>1,117</td>
<td>417</td>
</tr>
</tbody>
</table>

Table 2.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003 which have corresponding American call option. Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>72.5</td>
<td>44.0</td>
<td>99.1</td>
<td>212.9</td>
<td>68.2</td>
<td>81.1</td>
<td>61.7</td>
</tr>
<tr>
<td>median</td>
<td>49</td>
<td>35</td>
<td>97.5</td>
<td>208</td>
<td>49</td>
<td>55</td>
<td>42.75</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>33.5</td>
<td>71</td>
<td>0.5</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>354</td>
<td>192</td>
<td>182</td>
<td>354</td>
<td>354</td>
<td>318</td>
<td>262.5</td>
</tr>
<tr>
<td>observations</td>
<td>312</td>
<td>231</td>
<td>42</td>
<td>39</td>
<td>147</td>
<td>118</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 2.1b: Statistics of the FTSE 100 American (AM) style call options for the year 2003. Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>343.5</td>
<td>378.2</td>
<td>578.2</td>
<td>138.1</td>
<td>385.9</td>
<td>327.9</td>
<td>279.7</td>
</tr>
<tr>
<td>median</td>
<td>60</td>
<td>81</td>
<td>178.5</td>
<td>19</td>
<td>98</td>
<td>52</td>
<td>37</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>7,683</td>
<td>7,583</td>
<td>4,837</td>
<td>400</td>
<td>6,930</td>
<td>7,683</td>
<td>3,021</td>
</tr>
<tr>
<td>observations</td>
<td>2,572</td>
<td>1,731</td>
<td>256</td>
<td>585</td>
<td>1,038</td>
<td>1,117</td>
<td>417</td>
</tr>
</tbody>
</table>

Table 2.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call option. Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.
### Table 2.2b: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003.

Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>57</td>
<td>60</td>
<td>38.2</td>
<td>60.1</td>
<td>33.9</td>
<td>32.2</td>
<td>191.7</td>
</tr>
<tr>
<td>median</td>
<td>10</td>
<td>10</td>
<td>5.5</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>3,000</td>
<td>3,000</td>
<td>340</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>300</td>
</tr>
<tr>
<td>observations</td>
<td>312</td>
<td>231</td>
<td>42</td>
<td>39</td>
<td>147</td>
<td>118</td>
<td>47</td>
</tr>
</tbody>
</table>

#### 2.3. Research methodology

**2.3.1 Derman and Kani (DK, 1994) Implied Binomial Tree (IBT)**

Derman and Kani (DK, 1994) made perhaps the most significant contribution in the area of IBT. They show how to construct a binomial tree for the underlying asset price under the assumption that the volatility is a deterministic function of the asset price and time such that the tree correctly reproduces all options used for the construction of the tree. The DK tree is arbitrage free, and thus can be used for the pricing of other less liquid and exotic derivatives. However, in order for the DK algorithm to be correctly specified, additional rules should be imposed. The algorithm that we are going to describe is a modified version of the original DK algorithm that includes the modifications suggested by BC \(^{32}\) and also some other heuristics that help to avoid violations of the no-arbitrage constraints.

---

\(^{32}\) The modifications introduced by BC are the following:

i. Align the centre nodes of the tree with the forward price rather than with the current stock price.

ii. Use the forward price of the previous node to calculate the new option value of the nodes at the next level.

iii. Use BS formula instead of CRR binomial tree method to calculate the interpolated option prices.
Figure 2.1: The $i^{th}$ and the $(i+1)^{th}$ time steps on a binomial tree.

Suppose that we want to build the modified $DK\text{ IBT}$ for $n-1$ steps, on the interval $[1, T]$ with equally spaced steps $\Delta t$, where $\Delta t = \frac{T}{n-1}$. Figure 2.1 shows the $i^{th}$ and the $(i+1)^{th}$ time steps on the (recombining) tree. Let $S_{i,j}$ be the price of the underlying asset of the $i^{th}$ time step at the $j^{th}$ node, and $F_{i,j} = e^{(r-\delta)\Delta t}S_{i,j}$ the forward price at time step $i+1$ and node $j$. Also let $p_{i,j}$ be the risk-neutral transition probability to go from node $(i, j)$ to node $(i+1, j+1)$, $i = 1, ..., n-1$, $j = 1, ..., i$ which is described by formula (2.1):

$$p_{i,j} = \frac{S_{i,j}e^{(r-\delta)\Delta t} - S_{i+1,j}}{S_{i+1,j+1} - S_{i+1,j}}, \quad i = 1, ..., n-1, \quad j = 1, ..., i$$

(2.1)

and $\lambda_{i,j}$ the Arrow-Debreu price at node $(i, j)$\footnote{The Arrow-Debreu price at node $(i, j)$ is the price of an option that pays 1 unit pay-off in one and only one state $j$ at the $i^{th}$ time step, and otherwise pays 0.}. In general, Arrow-Debreu prices can be obtained by the following iterative formula where $\lambda_{i,1} = 1$ by definition.
\[ \lambda_{i,1} = 1 \]

\[ \lambda_{i+1,1} = e^{-\gamma \Delta t} (1 - p_{i,1}) \lambda_{i,1}, \quad i = 1, \ldots, n-1 \]  

\[ \lambda_{i+1,j} = e^{-\gamma \Delta t} \left( \lambda_{i,j} (1 - p_{i,j}) + \sum_{j=1}^{i-1} \lambda_{i,j-1} p_{i,j-1} \right), \quad i = 2, \ldots, n-1, \quad j = 2, \ldots, i \]

\[ \lambda_{i+1,i+1} = e^{-\gamma \Delta t} \lambda_{i,i} p_{i,i}, \quad i = 1, \ldots, n-1 \]

The value of a call option \( C(K, t_{i+1}) \) with strike \( K \) and maturity at the \( i+1 \) time step \( (t_{i+1}) \) is given by the following formula:

\[
e^{\gamma \Delta t} C(K, t_{i+1}) = \left[ \lambda_{i,i} (1 - p_{i,i}) \max(S_{i+1,1} - K, 0) \right.
\]

\[
+ \left( \sum_{j=1}^{i-1} \lambda_{i,j} p_{i,j} + \lambda_{i,i} p_{i,i} \right) \max(S_{i+1,i+1} - K, 0) \right] \]

\[+ \lambda_{i,i} p_{i,i} \max(S_{i+1,i+1} - K, 0) \right] \]

When the strike equals \( F_{i,j} \), the above formula is written in the form:

\[
e^{\gamma \Delta t} C(F_{i,j}, t_{i+1}) = \lambda_{i,j} p_{i,j} (S_{i+1,j+1} - F_{i,j}) + \sum_{z=j+1}^{i} \lambda_{i,z} (F_{i,z} - F_{i,j}) \]  

(2.4)

In similar way, we can derive a formula for the put values, \( P(F_{i,j}, t_{i+1}) \).

\[
e^{\gamma \Delta t} P(F_{i,j}, t_{i+1}) = \lambda_{i,j} (1 - p_{i,j}) (F_{i,j} - S_{i+1,i}) + \sum_{z=1}^{i-1} \lambda_{i,z} (F_{i,z} - F_{i,z}) \]  

(2.5)

Using formulas (2.1) and (2.4), we can derive formula (2.6) that gives the values of the underlying asset at the upper nodes of the tree, if we know the value of the central node of the tree \( S_{i+1,j} \):

\[
S_{i+1,j+1} = \frac{S_{i+1,j} \left( e^{\gamma \Delta t} C(F_{i,j}, t_{i+1}) - \sum_{j=1}^{i} \lambda_{i,j} F_{i,j} (F_{i,j} - S_{i+1,j}) \right)}{e^{\gamma \Delta t} C(F_{i,j}, t_{i+1}) - \sum_{j=1}^{i} \lambda_{i,j} (F_{i,j} - S_{i+1,j})} \]  

(2.6)
where

\[ \text{Sum1} = \sum_{z=j+1}^{i} \lambda_{i,z} \left( F_{i,z} - F_{i,j} \right) \]  \hspace{1cm} (2.7)

\( C(F_{i,j}, t_{i+1}) \) is the interpolated value for a call struck today at strike price \( F_{i,j} \) and time to maturity \( t_{i+1} \).

Using formulas (2.1) and (2.5), we can derive formula (2.8) that gives the values of the underlying asset at the lower nodes of the tree, \( S_{i+1,j} \):

\[
S_{i+1,j} = \frac{\lambda_{i,j} F_{i,j} \left( S_{i+1,j+1} - F_{i,j} \right) - S_{i+1,j+1} \left( e^{r_i \Delta t} P(F_{i,j}, t_{i+1}) - \text{Sum2} \right)}{\lambda_{i,j} \left( S_{i+1,j+1} - F_{i,j} \right) - \left( e^{r_i \Delta t} P(F_{i,j}, t_{i+1}) - \text{Sum2} \right)}
\]  \hspace{1cm} (2.8)

\[ \text{Sum2} = \sum_{z=k-1}^{i} \lambda_{i,z} \left( F_{i,k} - F_{i,z} \right) \]  \hspace{1cm} (2.9)

\( P(F_{i,k}, t_{i+1}) \) is the interpolated value for a put struck today at strike price \( F_{i,k} \) and time to maturity \( t_{i+1} \).

There are \((2i + 1)\) parameters which define the transition from the \( i^{th} \) to the \((i+1)^{th}\) time step, i.e. \((i + 1)\) asset prices of the nodes at the \((i+1)\) time step and \( i \) transition probabilities. Suppose that the \((2i + 1)\) parameters corresponding to the \( i^{th} \) time step are known, the \( S_{i+1,j} \) and \( p_{i,j} \) for \( j = 1, ..., i+1 \) can be calculated using the following principles:
We always start from the centre nodes in one time step:

If $i$ is even, define $S_{i+1,j} = S_{i,j} e^{(r_{j} - \delta)\Delta t}$ for $j = \frac{i}{2} + 1$.

If $i$ is odd, start from the two central nodes $S_{i+1,j}$ and $S_{i+1,j+1}$ for $j = \frac{i+1}{2}$ and suppose

$$S_{i+1,j+1} = \frac{F_{i,j}^{2}}{S_{i+1,j}},$$  \hspace{1cm} (2.10)$$

which adjusts the logarithmic spacing between $S_{i,j}$ and $S_{i+1,j+1}$ to be the same as that between $S_{i,j}$ and $S_{i+1,j}$.

For $i$ odd, substituting formula (2.10) to (2.6) we derive formula (2.11) for the lower central node:

$$S_{i+1,j} = F_{i,j} \frac{\lambda_{i,j} F_{i,j} - \left( e^{r_{i} \Delta t} C(F_{i,j}, i\Delta t) - \text{sum}1 \right)}{\lambda_{i,j} F_{i,j} + \left( e^{r_{i} \Delta t} C(F_{i,j}, i\Delta t) - \text{sum}1 \right)} \text{ for } j = \frac{i+1}{2}$$ \hspace{1cm} (2.11)$$

Once we have the central nodes’ asset prices, we can continue to calculate those at upper nodes using formula (2.6) and at lower nodes using formula (2.8). After computing the asset values at each node of a specific time step, we use formula (2.1) to compute the transition probabilities of that time step.

For identification of the arbitrage opportunities, the following inequalities are checked:

$$F_{i,j} < S_{i+1,j+1} < F_{i,j+1} \text{ or }$$

$$S_{i,j} e^{(r_{j} - \delta)\Delta t} < S_{i+1,j+1} < S_{i,j+1} e^{(r_{j} - \delta)\Delta t}, \text{ for } j = 1, \ldots, i - 1$$ \hspace{1cm} (2.12)$$
In cases of violations of the inequality (2.12) in the BC algorithm they choose the average of $F_{i,j}$ and $F_{i,j+1}$ as a proxy for $S_{i+1,j+1}$.

Additional heuristics on the modified DK algorithm

In order to prevent violations of the no-arbitrage constraints, we proceeded with the implementation of the following heuristics:

i. If the value of the implied volatility obtained from the interpolation was greater than 0.5\(^{34}\) (this happened in cases of extrapolation), we choose as implied volatility the maximum of the implied volatilities of the options used for the model calibration.

ii. If the implied volatility was negative (this happened in cases of extrapolation), then we choose as implied volatility the minimum of the implied volatilities of the options used for the model calibration.

iii. Constraint (2.12) does not cover the cases where the violation is at the node $(i, 1)$, i.e. when $S_{i+1,1} \geq F_{i,1}$. In this case we reduce systematically the value of $S_{i+1,1}$ until the constraint is satisfied.

iv. Constraint (2.12) does not cover the cases where the violation is at the node $(i, i)$, i.e. when $S_{i+1,i+1} \leq F_{i,i}$. In this case we increase the value of $S_{i+1,i+1}$ in a systematic way until the constraint is satisfied.

Binomial trees exhibit by construction a systematic oscillatory behaviour as a function of the number of steps in the tree. A common practise to avoid this problem is to take averages of the call estimations of consecutive number of steps, i.e.

$$C = \frac{C_n + C_{n+1}}{2}$$

---

\(^{34}\) We use 0.5 since it is the maximum value of the BS implied volatility that we have in our sample.
In the implementation, for the interpolation of BS implied volatilities across strikes, we use 2\textsuperscript{nd} order polynomial regression (DK\_Regr model) and cubic spline (DK\_Spline model).

2.3.2 Implied Non-Recombining Tree (INRT)

The INRT, proposed in the first chapter, was created as a response to the need of a non-recombining implied tree. The non-recombining tree due to the large number of degrees of freedom allows for a flexible underlying asset distribution that prevents negative or greater than one transition probabilities. In order to construct the INRT model, we minimize the discrepancy between the observed market prices and the model values with respect to the underlying asset at each node, subject to constraints that maintain risk neutrality and prevent arbitrage opportunities. Thus, in the INRT we have the following optimization problem:

$$\min_{x} \frac{1}{2} \sum_{k=1}^{N}(C_{Mod}(x,k) - C_{Mkt}(k))^{2}$$

(2.13)

Subject to the following constraints:

i) \(g_{1}(i, j) = S_{i,j}e^{(r_{j} - \delta)\Delta t} - S_{i+1,j-1} \geq 0, i = 1, ..., n - 1, j = 1, ..., 2^{i-1}\) (2.14a)

ii) \(g_{2}(i, j) = S_{i+1,2j} - S_{i,j}e^{(r_{j} - \delta)\Delta t} \geq 0, i = 1, ..., n - 1, j = 1, ..., 2^{i-1}\) (2.14b)

iii) \(g_{3}(k) = S_{1,k} - C_{Mod}(k) \geq 0, \quad k = 1, ..., N\) (2.14c)

iv) \(g_{4}(k) = C_{Mod}(k) - \max(S_{1,k}e^{-\delta T} - K(k)e^{-r_{f}T}, 0) \geq 0, \quad k = 1, ..., N\) (2.14d)

v) \(g_{5}(i, j) = S_{i,j} \geq 0, \quad i = 2, ..., n, \quad j = 1, ..., 2^{i-1}\) (2.14e)

Where, \(C_{Mkt}(k), k = 1, ..., N\) denotes the market prices of \(N\) European calls, with strikes \(K(k)\) and single maturities \(T\). \(C_{Mod}(x,k), k = 1, ..., N\) denotes the prices of the \(N\) calls.
obtained using the model. \( x \) denotes a vector containing the variables of the model which are the values of the underlying asset at each node of the tree, excluding its current value. Figure 2.2 shows a typical triplet in a non-recombining (binary) tree.

![Figure 2.2: A typical triplet in a non-recombining (binary) tree.](image)

The subscripted \( (i, j) \) denotes:

- \( i \): the time dimension, \( i = 1, ..., n \).
- \( j \): the asset (time specific) dimension, \( j = 1, ..., 2^{i-1} \).

\( S_{i,j} \) is the value of the underlying asset at node \( (i, j) \), \( r_f \) denotes the annually continuously compounded riskless rate of interest and \( \delta \) denotes the annually continuously compounded dividend yield.

The problem under consideration is a non-convex optimization problem with linear constraints. For the solution of the problem an exterior penalty method (Fiacco and McCormick, 1968) is adopted in order to handle the inequality constraints. The Exterior Penalty Objective function in the INRT is the following:
\[ P(x, \alpha) = \frac{1}{2} \sum_{k=1}^{N} (C_{Mod}(x,k) - C_{Mkt}(k))^2 \]
\[ + \frac{\alpha}{2} \sum_{i=1}^{n} \sum_{j=1}^{2^{s-i}} \left( \min(g_1(i,j),0) \right)^2 + \left( \min(g_2(i,j),0) \right)^2 \] (2.15)
\[ + \frac{\alpha}{2} \sum_{k=1}^{N} \left( \min(g_3(k),0) \right)^2 + \left( \min(g_4(k),0) \right)^2 \]
\[ + \frac{\alpha}{2} \sum_{i=2}^{n} \sum_{j=1}^{2^{s-i}} \left( \min(g_5(i,j),0) \right)^2 \]

For the optimization a Quasi-Newton algorithm (Fletcher, 1987) is used.

For the initialization of the tree the following volatility term structure is used:

\[ \sigma_i = \sigma_i e^{\lambda(i-1)\Delta t}, \quad \lambda \in \mathbb{R}, i = 1, \ldots, n-1 \] (2.16)

where \( \lambda \) is a constant parameter and \( \sigma_i \) is a properly chosen initial value for the volatility\(^{35}\). In order to obtain a feasible initial tree, \( \lambda \) is chosen to belong in the following interval:

\[ \lambda \in \left[ \frac{1}{T} \log \left( \frac{r_f - \delta \sqrt{N\Delta t}}{\sigma_i} \right), +\infty \right) \] (2.17)

The odd nodes of the tree \( S_{i,j} \), are initialized using the following equation:

\[ S_{i,j} = S_{i-1,j+1} d_{i-1,j+1}^{\frac{1}{2}}, \quad i = 2, \ldots, n, \quad j = 1, 3, \ldots, 2^{i-1} - 1 \] (2.18a)

The even nodes of the tree \( S_{i,j} \), are initialized using the following equation:

\[ S_{i,j} = S_{i-1,j} u_{i-1,j}^{\frac{1}{2}}, \quad i = 2, \ldots, n, \quad j = 2, 4, \ldots, 2^{i-1} \] (2.18b)

\(^{35}\) For initial volatility \( (\sigma_1) \) we use the at-the-money implied volatility given by LIFFE.
where,

\[ u_{i,j} = e^{\alpha_i \sqrt{\Delta t}} \]  \hspace{1cm} (2.19a) \\
\[ d_{i,j} = e^{-\alpha_i \sqrt{\Delta t}} = \frac{1}{u_{i,j}} \]  \hspace{1cm} (2.19b)

Equations (2.16) to (2.19) are used only for initialization. Once the optimization process starts, each value of the underlying asset (except from \( S_{i,1} \)) acts as an independent variable in the system.

To achieve the best feasible solution, i.e. the solution that gives a feasible tree with the smallest error function the algorithm draws consecutively values of \( \lambda \) from the specified interval (2.17) until the objective function is smaller than 1.E-4 and also the penalty term equals zero, i.e. we have a feasible solution.

For the upward transition probabilities \( p_{i,j} \) between the various nodes of the tree the risk-neutral probability formula holds:

\[ p_{i,j} = \frac{S_{i,j} e^{(r - \delta) \Delta t} - S_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}}, \quad i = 1, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \]  \hspace{1cm} (2.20)

The respective downward probability is equal to one minus the upward probability. Probability formula (2.20) constrains the model to Markovian processes.

To overcome the problem of non-differentiability of the call option value at the last time step when \( S_{n,j} = K \), the following smoothing approximation to \( C_{n,j} \) is used:
\[
C_a(n, j) = \begin{cases} 
0 & \text{for } S_{n,j} / K \leq 1 - z / 2 \\
\frac{S_{n,j}}{K} - 1 & \text{for } S_{n,j} / K \geq 1 + z / 2 \\
\frac{1}{2z} \left( \frac{S_{n,j}}{K} - 1 \right)^2 + \frac{z}{2} & \text{for } 1 - z / 2 < S_{n,j} / K < 1 + z / 2 
\end{cases} \tag{2.21a}
\]

\[j = 1, ..., 2^{n-1}\]

where \(z\) is a small positive constant, for example 1.E-2 or 1.E-3.

The value of the call at intermediate nodes is given by the following equation:

\[
C_{i,j} = (p_{i,j}C_{i+1,2j} + (1 - p_{i,j})C_{i+1,2j-1})e^{-r_j \Delta t} \tag{2.21b}
\]

\[i = n - 1, ..., 1, \quad j = 1, ..., 2^{i-1}\]

Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem. The main benefit of the model is its analytical structure which enables the use of efficient methods for nonlinear optimization. Although the method uses a large number of variables, due to the fact that efficient methods for optimization are used, the model is not computationally intensive. Also, because of the many degrees of freedom it allows for a flexible underlying asset distribution which prevents the appearance of non-acceptable probability values. Finally, this method does not need any interpolation or extrapolation across strikes and time to find hypothetical options as in the DK model.

2.3.3. Ad-hoc model

For the pricing of American call options we also use an ad-hoc procedure of smoothing BS implied volatilities across strikes that was firstly introduced by Dumas et al. (1998). Specifically, in order to price the American calls, we build a standard
CRR binomial tree. The volatility that we use in the CRR model is obtained by interpolating the implied volatility surface across strikes constructed from the BS implied volatilities of the European style call options. In our implementation, for the interpolation of BS implied volatilities across strikes we use second order polynomial regression \((AH\_Regr)\) which is Dumas et al. (1998) model and cubic spline \((AH\_Spline)\). This procedure is ad-hoc since it is internally inconsistent. The BS model assumes a constant volatility as input, whereas this method uses different implied volatilities depending on the exercise price (and the time to maturity). However, this ad-hoc method is simple and takes into account the volatility smile.

2.4. Testing the models in pricing FTSE 100 American options

In this section we compare the five different models, for pricing American call options. First, we consider the \(INRT\) model for \(n = 7\). Then we examine the two versions of the modified \(DK\) (1994) model \((DK\_Regr, DK\_Spline)\) and the two versions of the ad-hoc benchmark model \((AH\_Regr, AH\_Spline)\). In the \(DK\) and the ad-hoc models we price the American calls using trees of \(n = 50\) and \(n = 51\), and the call option value is computed as the average of the call option values of the two trees\(^{36}\).

Every trading day in the sample, we calibrate the above models using European call options with the same underlying asset and time to maturity. Since we calibrate the models with options that have the same underlying asset and maturity, the implied trees and its distribution that we find, result from a smile the shape of which is independent of expiration time. Then, we use each of the five models to price the corresponding American style call options. For the comparison of the pricing performance of the five models we use two pricing measures.

(i) The root mean square error (RMSE) which is the square-root of the average squared difference between the model and market prices. The RMSE measures how well the model fits in a statistical sense with the usual bias versus variance trade-off.

---

\(^{36}\) This approach improves considerably the accuracy of the \(DK\) model.
(ii) The mean difference error (MDE) which is the average of the difference between the model and the market prices and reveals systematic biases of the model.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INRT</td>
<td>4.3143</td>
<td>-1.5370</td>
</tr>
<tr>
<td></td>
<td>DK_Regr</td>
<td>5.4119</td>
<td>-2.1646</td>
</tr>
<tr>
<td>Full</td>
<td>DK_Spline</td>
<td>6.9588</td>
<td>-3.3461</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.2251</td>
<td>-1.3426</td>
</tr>
<tr>
<td></td>
<td>AH_Spline</td>
<td>4.3714</td>
<td>-1.3451</td>
</tr>
</tbody>
</table>

Table 2.3: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). We measure pricing errors using the root mean square error (RMSE) and the mean difference error (MDE). Results are obtained using the full sample of American call options.

Table 2.3 reports the pricing errors of the five models for the full sample of American call options. Overall, all models seem to systematically underprice the American calls. Both versions of the DK model seem to systematically underprice more than the other models. The DK_Spline gives the highest underpricing with MDE error -3.3461. Looking at the RMSE error, we see that the AH_Regr model has the lowest error 4.2251, followed by the INRT and AH_Spline models with 4.3143 and 4.3714 respectively. The DK_Regr and DK_Spline models come next with RMSE 5.4119 and 6.9588 respectively. Cubic spline interpolation method gives worst results than 2nd order polynomial regression, in both the DK and the ad-hoc model. This is in line with the findings of Brandt and Wu (2002) and Linaras and Skiadopoulos (2005) who find that complex interpolation methods do not necessarily give better results.
Table 2.4: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). We measure pricing errors with (i) the mean absolute difference error (MADE) and (ii) the median absolute difference error (MDADE). The last column contains the p-values from the Wilcoxon sign rank test of whether the distribution of the absolute errors obtained from the INRT model is statistically different from the distribution of the absolute errors of each of the other models.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>MADE</th>
<th>MDADE</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>INRT</td>
<td>2.9970</td>
<td>2.0614</td>
<td>.......</td>
<td></td>
</tr>
<tr>
<td>DK_Regr</td>
<td>3.5267</td>
<td>2.2330</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>DK_Spline</td>
<td>4.4977</td>
<td>3.1109</td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>AH_Regr</td>
<td>2.9632</td>
<td>2.0622</td>
<td>0.8523</td>
<td></td>
</tr>
<tr>
<td>AH_Spline</td>
<td>3.0343</td>
<td>2.1110</td>
<td>0.6974</td>
<td></td>
</tr>
</tbody>
</table>

In order to test whether the distribution of the pricing performance between the different models is statistically different we perform the Wilcoxon sign rank test at 1% level of significance. As a measure of pricing performance we use the absolute difference error of each model. We test whether the distribution of the absolute errors obtained from the INRT model with each of the other models is statistically different. Table 2.4 contains the mean absolute difference errors (MADE) and the median absolute difference errors (MDADE) of the five models. In the last column of Table 2.4 are presented the p-values of the Wilcoxon sign rank test. We observe that the distributions of the absolute errors of the AH_Regr and the AH_Spline models are not statistically different from that of the INRT model (have \textit{p-values} 0.8523 and 0.6974 respectively). However, the distributions of the absolute errors of both DK models are statistically different from that of the INRT model.

\[37\] Wilcoxon sign rank test is a non-parametric alternative to the paired Student’s t-test for the case of two related samples. Specifically, it performs a paired two sided test of hypothesis that the difference between the matched samples comes from a distribution with median zero. In other words, it tests whether the two paired-samples have the same distribution.
Table 2.5: Pricing errors within moneyness categories for the American FTSE 100 index call options using the implied non-recombining tree (INRT) model, two versions of the modified Derman and Kani (DK, 1994) model (DK_Regr, DK_Spline), and two ad-hoc models (AH_Regr, AH_Spline). We measure pricing errors using the root mean square error (RMSE) and the mean difference error (MDE). Results are for the full sample when divided in 3 moneyness categories. OTM denotes the out-of-the-money, ATM denotes the at-the-money and ITM denotes the in-the-money American call options.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTM</td>
<td>INRT</td>
<td>3.6351</td>
<td>-1.7202</td>
</tr>
<tr>
<td></td>
<td>DK_Regr</td>
<td>4.7149</td>
<td>-2.4165</td>
</tr>
<tr>
<td></td>
<td>DK_Spline</td>
<td>6.0783</td>
<td>-3.4369</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>3.5808</td>
<td>-1.6856</td>
</tr>
<tr>
<td></td>
<td>AH_Spline</td>
<td>3.5824</td>
<td>-1.6763</td>
</tr>
<tr>
<td>ATM</td>
<td>INRT</td>
<td>4.3583</td>
<td>-1.2525</td>
</tr>
<tr>
<td></td>
<td>DK_Regr</td>
<td>4.5463</td>
<td>-1.6957</td>
</tr>
<tr>
<td></td>
<td>DK_Spline</td>
<td>4.9478</td>
<td>-2.1372</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.3396</td>
<td>-1.1160</td>
</tr>
<tr>
<td></td>
<td>AH_Spline</td>
<td>4.3809</td>
<td>-1.1241</td>
</tr>
<tr>
<td>ITM</td>
<td>INRT</td>
<td>7.0840</td>
<td>-0.7583</td>
</tr>
<tr>
<td></td>
<td>DK_Regr</td>
<td>8.9655</td>
<td>-1.1778</td>
</tr>
<tr>
<td></td>
<td>DK_Spline</td>
<td>11.9251</td>
<td>-4.1102</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>6.8253</td>
<td>0.4452</td>
</tr>
<tr>
<td></td>
<td>AH_Spline</td>
<td>7.4958</td>
<td>0.3792</td>
</tr>
</tbody>
</table>

In Table 2.5 are presented the RMSE and MDE of the five models when the sample is divided in the three moneyness categories. All the models in the OTM and ATM categories underprice systematically the call values whereas in the ITM category both versions of the ad-hoc model seem to overprice call values. In the OTM group (74% of the full sample) the AH_Regr model gives the lowest RMSE 3.5808 whereas the INRT model is very competitive to it giving RMSE 3.6351. The AH_Spline model has RMSE 3.5824. The DK_Regr model has RMSE 4.7149 whereas the DK_Spline model has RMSE 6.0783. In the ATM group (13.5% of the full sample), the INRT model has the lowest RMSE, 4.3583, followed by the AH_Regr and the AH_Spline models with RMSE 4.3396 and 4.3809. Again, the two versions of the DK model, the DK_Regr and the DK_Spline have the highest RMSE, 4.5463 and 4.9478 respectively. Finally, in the ITM group (12.5% of the full sample), we observe the same pattern as in the OTM group. Specifically, the AH_Regr model gives the lowest RMSE 6.8253 whereas the INRT model and the AH_Spline model come next with
RMSE 7.0840 and 7.4958 respectively. The $DK_{Regr}$ model has RMSE 8.9655 whereas the $DK_{Spline}$ model has RMSE 11.9251.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INRT</td>
<td>4.0187</td>
<td>-1.0044</td>
</tr>
<tr>
<td>ST</td>
<td>$DK_{Regr}$</td>
<td>4.1461</td>
<td>-1.2896</td>
</tr>
<tr>
<td></td>
<td>$DK_{Spline}$</td>
<td>4.3283</td>
<td>-1.5370</td>
</tr>
<tr>
<td></td>
<td>$AH_{Regr}$</td>
<td>4.0873</td>
<td>-0.9561</td>
</tr>
<tr>
<td></td>
<td>$AH_{Spline}$</td>
<td>4.0621</td>
<td>-0.9662</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>4.1982</td>
<td>-2.0377</td>
</tr>
<tr>
<td>MT</td>
<td>$DK_{Regr}$</td>
<td>6.2820</td>
<td>-2.9114</td>
</tr>
<tr>
<td></td>
<td>$DK_{Spline}$</td>
<td>8.8059</td>
<td>-5.1827</td>
</tr>
<tr>
<td></td>
<td>$AH_{Regr}$</td>
<td>3.9093</td>
<td>-1.5785</td>
</tr>
<tr>
<td></td>
<td>$AH_{Spline}$</td>
<td>4.3210</td>
<td>-1.5907</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>5.3665</td>
<td>-1.9459</td>
</tr>
<tr>
<td>LT</td>
<td>$DK_{Regr}$</td>
<td>6.4485</td>
<td>-3.0265</td>
</tr>
<tr>
<td></td>
<td>$DK_{Spline}$</td>
<td>8.2573</td>
<td>-4.3931</td>
</tr>
<tr>
<td></td>
<td>$AH_{Regr}$</td>
<td>5.2804</td>
<td>-1.9588</td>
</tr>
<tr>
<td></td>
<td>$AH_{Spline}$</td>
<td>5.3259</td>
<td>-1.9133</td>
</tr>
</tbody>
</table>

Table 2.6: Pricing errors within maturity categories for the American FTSE 100 index call options using the implied non-recombining tree ($INRT$) model, two versions of the modified Derman and Kani ($DK$, 1994) model ($DK_{Regr}$, $DK_{Spline}$), and two ad-hoc models ($AH_{Regr}$, $AH_{Spline}$). We measure pricing errors using the root mean square error (RMSE) and the mean difference error (MDE). Results are for the full sample when divided in 3 maturity categories. ST denotes the short term, MT denotes the medium term and LT denotes the long term American call options.

In Table 2.6 are presented the RMSE and MDE of the five models when the sample is divided in the three time-to-maturity categories. In all the time-to-maturity categories, all models underprice systematically the call values. In the ST group, (37% of the full sample) the $INRT$ model gives the lowest RMSE 4.0187 whereas the $AH_{Spline}$ and the $AH_{Regr}$ models come next with RMSE 4.0621 and 4.0873 respectively. The $DK_{Regr}$ model has RMSE 4.1461 whereas the $DK_{Spline}$ model has RMSE 4.3283. In the MT group (38% of the full sample), we observe the same pattern as in the ST group. The $INRT$ model has the lowest RMSE, 4.1982, followed by the $AH_{Regr}$ and the $AH_{Spline}$ models with RMSE 3.9093 and 4.3210. Again, the two versions of the $DK$ model, the $DK_{Regr}$ and the $DK_{Spline}$ have the highest RMSE, 6.2820 and 8.8059 respectively. Finally, in the LT group (15% of the full sample), both versions of the ad-hoc model outperform the $INRT$ model. Specifically, the $AH_{Regr}$
and the *AH_Spline* models have RMSE 5.2804 and 5.3259 respectively whereas the *INRT* model has RMSE 5.3665. Again, the *DK_Regr* and the *DK_Spline* have the highest RMSE, 6.4485 and 8.2573 respectively\(^{38}\).

Concluding, results indicate that the *INRT* model outperforms both versions of the *DK* model. Moreover, the distribution of the pricing errors obtained from the *INRT* model is not statistically different from the distribution of the pricing errors obtained using the ad-hoc model. Therefore, we have evidence that the pricing performance of the ad-hoc procedure is not statistically different from the *INRT* model. Finally, overall results indicate that complex interpolation methods do not necessarily give better results.

### 2.5. Conclusions

The appearance and persistence of the volatility smile and term structure in the financial markets have encouraged the development of smile consistent models. One of the most popular deterministic volatility function models to practitioners is that of Derman and Kani (*DK,* 1994). The *DK* model captures the volatility smile and term structure. However, negative or greater than one probabilities can occur. In this case, the probabilities are overridden causing the tree to lose much of its implied nature. The implied non-recombining tree (*INRT*) model proposed in the first chapter, has the benefit of many degrees of freedom which allow for a flexible underlying asset distribution that prevents negative and greater than one transition probabilities. Both models can be calibrated consistently with a set of observed market prices.

In this study we compare the pricing performance of the *INRT* with the recombining binomial tree of *DK*. In the comparison we also include an ad-hoc procedure of smoothing Black-Scholes implied volatilities across strikes. The *DK* and the ad-hoc models are tested using two interpolation methods across strikes, the second order polynomial regression and the cubic spline. We use European and

\(^{38}\) Overall, for the *DK* and the ad-hoc models, the pricing errors are in the same range with that of Brandt and Wu (2002).
American call options of the FTSE 100 index for the year 2003 obtained from LIFFE. For each date in our sample, we calibrate the models using the panel of all European style calls with the same underlying asset and time-to-maturity and then examine how well these models price the corresponding panel of American calls. Results show that the pricing performance of the INRT model is better than that of a modified version of the DK model whereas the pricing performance of the ad-hoc procedure is not statistically different from the INRT model. Also, we find that simpler interpolation methods give better results.
3. Calibration of non-recombining implied trees for the local volatility surface

Abstract

In this chapter we calibrate the non-recombining implied tree in order to extract the local volatility surface. The problem under consideration is a non-convex constrained optimization problem. We elaborate on the initial guess for the volatility term structure and use nonlinear constrained optimization to minimize the least squares error function on market prices, with respect to the local volatility, subject to constraints that maintain risk neutrality and prevent arbitrage opportunities. We test our model using call options data for the FTSE 100 index for the year 2003. Results strongly support our modelling approach. Knowledge of the local volatility surface is especially useful in markets with pronounced smile to measure market sentiment, to compute the evolution of implied volatilities through time, and to value and hedge exotic options.
3.1. Introduction

The Black-Scholes (BS, 1973) model for pricing options is based on the assumption that the underlying asset evolves according to a geometric Brownian motion with a constant volatility at any time and market level. Consequently, options on the same underlying asset must be priced using the same volatility. The success of the BS framework has led traders to quote a call option’s market price in terms of whatever constant volatility makes the BS formula value equal to the market price. This volatility is called implied volatility. However, ever since the 1987 market crash, the market’s implied BS volatilities for index options vary with strike and time to expiration. This is known as the volatility smile. These violations belie the BS theory, which assumes constant volatility (and therefore, constant implied volatility) for all options and suggests that implied volatility should also vary locally with strike and time to expiration. Therefore, apart from an implied volatility surface there is also a local volatility surface.

This problem has motivated the recent literature on “smile consistent” no-arbitrage models. Consistency is achieved by extracting an implied evolution for the underlying asset price from market prices of liquid standard options. There are two classes of methodologies within this approach. Smile consistent deterministic volatility models and stochastic volatility smile consistent models which allow for smile-consistent option pricing under the no-arbitrage evolution of the volatility surface\textsuperscript{39,40}. Also, in the first chapter, we propose a non-recombining implied tree (INRT) without imposing any restrictive assumptions for the underlying stochastic process. Effectively, the INRT is a non-parametric model. Because of the many degrees of freedom, the INRT can allow for flexible underlying asset distribution.


\textsuperscript{40} There also exist non-parametric methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.
The BS implied volatilities, summarize the level of uncertainty over the *entire life* of an option (Rubinstein, 1994). They are a kind of average future volatility of the underlying asset during the option’s lifetime (Derman *et al.* 1996). In this sense, implied volatility is a global measure of volatility, in contrast to the local volatility that corresponds to different asset level and time. Derman and Kani (1994) and Dupire (1994) determined the local volatility surface by first deducing the implied distribution of the underlying asset, using as inputs the option prices of all strikes and time to expiration. Also, Rubinstein (1994) has taken a similar but not identical approach for the determination of the local volatility surface, using only options with the same time to expiration. Derman, Kani and Zou (1996) stated that: “In the same way as fixed income investors analyze the yield curve in terms of forward rates, so index options investors should analyze the volatility smile in terms of local volatilities.” Therefore, they propose a more quantitative and exact way of deducing the future local volatilities. They deduce the local volatility surface, using as inputs instead of the option prices, the BS implied volatilities of several liquid options of various strikes and expirations.

In this chapter, we propose a model for calibrating the non-recombining tree with respect to the local volatility. Here, unlike in the *INRT* model proposed in the first chapter, we calibrate the non-recombining tree assuming that the system variable is the local volatility, instead of the underlying asset. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and use nonlinear constrained optimization to minimize the least-squares error function on market prices. Specifically, in order to handle the inequality constraints, we adopt an exterior penalty method and the optimization is implemented using a Quasi-Newton algorithm. Appropriate constraints allow us to maintain risk neutrality and prevent arbitrage opportunities.

---

41 Rubinstein (1994) argues that the very high (low) local volatilities could imply much lower (higher) BS implied volatilities over longer intervals.
Assuming that the local volatility is the system variable, we reduce the number of variables that the optimization algorithm has to deal with to half, in relation to the INRT model, making the problem less computationally intensive. Moreover, we can easily control the values of the local volatilities at each node of the tree by imposing lower and upper bounds on the values of the local volatility. Knowledge of the local volatility surface is especially useful in markets with pronounced smile to measure market sentiment, to compute the evolution of implied volatilities through time, and to value and hedge exotic options.

We test our model using call options data on the FTSE 100 index, for the year 2003 obtained from LIFFE. Results strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Also, the computational burden is not a major issue. Furthermore, we compare the pricing performance of the proposed model in pricing of American call options relative to the INRT model proposed in the first chapter and to an ad-hoc procedure of smoothing BS implied volatilities across strikes proposed by Dumas et al. (1998). Results indicate that all models are statistically equivalent in pricing of American calls.

The chapter continues as follows: In section 3.2 we describe the proposed model. In section 3.3 we present the analytical formulation of the problem. In section 3.4 we describe the dataset used for testing the model. In section 3.5 we calibrate the model and test the properties of the model, and in section 3.6 we test the performance of the model in pricing of American call options. Section 3.7 concludes. In Appendix 3.1 we provide the derivation of the partial derivatives of the transition probabilities at each node with respect to the local volatility at the same node. In Appendix 3.2 we provide the analytical formulation of an alternative approach for computing the partial derivatives of the call with respect to the local volatility at each node of the tree.
3.2. Description of the model

Assume that the behaviour of the underlying asset is described by a non-recombining tree. Figure 3.1 shows a non-recombining tree with 4 steps \((n = 5)\). The point \((i, j)\) on the tree denotes:

\(i\): the time dimension, \(i = 1, \ldots, n\).

\(j\): the asset (time specific) dimension, \(j = 1, \ldots, 2^{i-1}\).

\(\sigma_{i,j}\) is the value of the local volatility at node \((i, j)\).

![Non-recombining tree with 4 steps](image)

Figure 3.1: Non-recombining tree with 4 steps.

Also, assume that the local volatility \(\sigma_{i,j}\) at each node of the tree is the only variable of the system. The problem we have is to find the values of the local volatility at each point \((i, j)\) for \(i = 1, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1}\) such that the square difference of the option value obtained by the model and the market value is as close to zero as possible, subject to constraints that maintain the risk neutrality of the tree and also prevent arbitrage opportunities.
Figure 3.2: A typical triplet on a non-recombining tree.

Starting from the initial node at which the value of the underlying asset is known, we set the new up and down values of the underlying asset at every node \((i, j)\) as (see Fig. 3.2):

\[
S_{i+1,2j} = S_{i,j} u_{i,j} \quad (3.1a)
\]

\[
i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]

\[
S_{i+1,2j-1} = S_{i,j} d_{i,j} \quad (3.1b)
\]

where

\[
u_{i,j} = e^{\sigma_{i,j} \sqrt{\Delta t}} \quad (3.2a)
\]

\[
i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}
\]

\[
d_{i,j} = \frac{1}{u_{i,j}} = e^{-\sigma_{i,j} \sqrt{\Delta t}} \quad (3.2b)
\]

and
\[ \Delta t = \frac{T}{n-1} \]  

(3.3)

We denote with \( T \) the option’s time to maturity and with \( \Delta t \) the step of the tree.

Because we work in a risk neutral world, the formula for the transition probabilities at each node of the tree is given by:

\[
p_{i,j} = \frac{S_{i,j} e^{(r_f - \delta)\Delta t} - S_{i+1,j-1}}{S_{i+1,2j-1} - S_{i+1,2j-1}} \quad \text{for} \quad i = 1, ..., n-1, \quad j = 1, ..., 2^{i-1} \]  

(3.4)

Where \( r_f \) denotes the annually continuously compounded riskless rate of interest and \( \delta \) denotes the annually continuously compounded dividend yield.

Substituting formulas (3.2a) and (3.2b) to (3.4) we have:

\[
p_{i,j} = \frac{e^{(r_f - \delta)\Delta t} - d_{i,j}}{u_{i,j} - d_{i,j}} \quad \text{(3.5a)}
\]

or

\[
p_{i,j} = \frac{e^{(r_f - \delta)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \quad \text{(3.5b)}
\]

The call option value at the last time step is given by:

\[
C_{n,j} = \max \{ S_{n,j} - K, 0 \}, \quad j = 1, ..., 2^{n-1} \]  

(3.6)

where \( K \) is the strike price of the option.

To overcome the problem of non-differentiability of function (3.6) when \( S_{n,j} = K \), we use the same smoothing approximation to \( C_{n,j} \) as in the INRT model:
\[
\frac{C_a(n,j)}{K} = \begin{cases} 
0 & \text{for } S_{n,j} / K \leq 1 - z/2 \\
\frac{S_{n,j}}{K} - 1 & \text{for } S_{n,j} / K \geq 1 + z/2 \\
\frac{1}{2z} \left( \left( \frac{S_{n,j}}{K} - 1 \right) + \frac{z}{2} \right)^2 & \text{for } 1 - z/2 < S_{n,j} / K < 1 + z/2 
\end{cases}
\]

for \( j = 1, \ldots, 2^{n-1} \)

where \( z \) is a small positive constant, for example 0.01.

The value of the call at intermediate nodes is given by the traditional backward equation:

\[
C_{i,j} = (p_{i,j} C_{i+1,2j} + (1 - p_{i,j}) C_{i+1,2j-1}) e^{-r_j \Delta t}, i = n - 1, \ldots, 1, \quad j = 1, \ldots, 2^{i-1} \quad (3.7b)
\]

**Initialization of the model**

Like in the INRT model, we use the following volatility term structure to initialize the tree:

\[
\sigma_i = \sigma_i e^{\lambda (i-1) \Delta t}, \quad \lambda \in \mathbb{R}, \quad i = 1, \ldots, n-1 \quad (3.8)
\]

where \( \lambda \) is a constant parameter and \( \sigma_i \) is a properly chosen initial value for the volatility. If \( \lambda \) is positive, then volatility increases as we approach maturity and if \( \lambda \) is negative, then volatility decreases as we approach maturity\(^{42}\). In order to keep the probabilities well specified at every time step and hence obtain a feasible initial tree, \( \lambda \) should belong in the following interval:

\[^{42}\text{Other non-monotonic functions could also be used for } \sigma_i \text{ but what we have tried proved adequate for our purposes.}\]
\[
\lambda \in \left[ \frac{1}{T} \log \left( \frac{|r_f - \delta| \sqrt{\Delta t}}{\sigma_i} \right), +\infty \right) \tag{3.9}
\]

By choosing \( \lambda \) from the above interval, we allow the initial volatility to increase or decrease across time. We make several consecutive draws from interval (3.9) until we find the value of \( \lambda \) that gives the "optimal" tree\(^{43}\).

**Risk-neutrality and no-arbitrage constraints**

In order for the transition probabilities \( p_{i,j} \) defined in Eq.(3.5) to be well specified, they should take values between zero and one. This implies the following risk-neutrality constraints:

\begin{align*}
& \quad e^{(r_f - \delta) \Delta t} < e^{\sigma_{i,j} \sqrt{\Delta t}} \\
& \quad \quad \quad \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{j-1} \\
& \quad e^{(r_f - \delta) \Delta t} > e^{-\sigma_{i,j} \sqrt{\Delta t}} \tag{3.10b}
\end{align*}

Risk neutrality constraints in the non-recombining tree prevent nodes \( 2j - 1 \) and \( 2j \) to cross, for \( i=1,\ldots,n \) and \( j=1,\ldots,2^{i-1} \).

Inequalities (3.10a) and (3.10b) can take the following form:

\[
\sigma_{i,j} > |r_f - \delta| \sqrt{\Delta t} \quad \text{for} \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1} \tag{3.11}
\]

- If \( (r_f - \delta) > 0 \) Eq.(3.11) becomes \( \sigma_{i,j} > (r_f - \delta) \sqrt{\Delta t} \) \tag{3.12a}
- If \( (r_f - \delta) < 0 \) Eq.(3.11) becomes \( \sigma_{i,j} > -(r_f - \delta) \sqrt{\Delta t} \) \tag{3.12b}
- If \( (r_f - \delta) = 0 \) Eq.(3.11) becomes \( \sigma_{i,j} > 0 \) \tag{3.12c}

\(^{43}\) Optimal tree is the one that gives the lowest-value objective function subject to the initial constraints.
Like in the INRT model, in order to avoid arbitrage opportunities, we include the no-arbitrage constraints. Specifically, at node (1,1) we impose the following constraint:

\[
\max \left( S_{1.1} e^{-\delta T} - K e^{-r T}, 0 \right) \leq C_{1.1} \leq S_{1.1}
\]  

(3.13)

The problem

Let \( C_{\text{Mkt}}(k), \ k = 1, ..., N \) denote the market prices of \( N \) European calls, with strikes \( K(k) \). Also, let \( C_{\text{Mod}}(x,k), \ k = 1, ..., N \) denote the prices of the \( N \) calls obtained using the model. \( x \) denotes a vector containing the variables of the model which are the values of the local volatility at each node of the tree. The ideal solution is to find the values of the local volatility (the model variables) at each node of the tree such that a perfect match is achieved between the option market prices and those predicted by the tree. However, due to market imperfections and other factors perfect matching may not always be possible.

Therefore, the objective of the problem is to minimize the least squares error function of the discrepancy between the observed market option prices and the model values. Thus, we have the following optimization problem:

\[
\min_x \frac{1}{2} \sum_{k=1}^{N} \left( C_{\text{Mod}}(x,k) - C_{\text{Mkt}}(k) \right)^2
\]  

(3.14)

subject to the constraints:

i) \( g_1(i,j) = \sigma_{i,j} - |r_f - \delta|\sqrt{\Delta T} > 0, \ i = 1, ..., n - 1, \ j = 1, ..., 2^{i-1} \)  

(3.15a)

ii) \( g_2(k) = S_{1.1} - C_{\text{Mod}}(k) \geq 0, \ k = 1, ..., N \)  

(3.15b)

iii) \( g_3(k) = C_{\text{Mod}}(k) - \max(S_{1.1} e^{-\delta T} - K(k) e^{-r T}, 0) \geq 0, \ k = 1, ..., N \)  

(3.15c)
Since the problem under consideration is a nonlinear optimization problem with linear constraints we adopt an exterior penalty method (Fiacco and McCormick, 1968) to convert the nonlinear constrained problem into a nonlinear unconstrained problem. The Exterior Penalty Objective function that we use is the following:

\[
P(x, \alpha) = \frac{1}{2} \sum_{k=1}^{N} (C_{Mod}(x, k) - C_{Mbl}(k))^2 + \\
\quad + \frac{\alpha}{2} \sum_{i=1}^{n-1} \sum_{j=1}^{2^{i-1}} \left[ \min(g_1(i, j), 0) \right]^2 \\
\quad + \frac{\alpha}{2} \sum_{k=1}^{N} \left[ \min(g_2(k), 0) \right]^2 + \left[ \min(g_3(k), 0) \right]^2 \]

The second and third terms in \( P(x, \alpha) \) give a positive contribution if and only if \( x \) is infeasible. Under mild conditions it can be proved that minimizing the above penalty function for strictly increasing sequence \( \alpha \) tending to infinity the optimum point \( x(\alpha) \) of \( P \) tends to \( x^* \), a solution of the constrained problem.

For the optimization we use a Quasi-Newton algorithm. Specifically we use the BFGS formula\(^{44}\) (Fletcher, 1987). For the procedure of Line Search in the algorithm we use the Charalambous (1992) method. To achieve the best feasible solution, i.e. the solution that gives us a feasible tree with the smallest error function we force the algorithm to draw consecutively values of \( \lambda \) from the specified interval (3.9) until the objective function is smaller than 1.E-4 and also the penalty term equals zero, i.e. we have a feasible solution.

An interesting thing to note is that the number of variables required is \( 2^{n-1} - 1 \), which is exactly the half of the variables required in the INRT model. Thus, using the same number of steps, the new model has to solve an optimization problem with the half variables required in the INRT model, making the optimization algorithm less computationally intensive.

\(^{44}\) The BFGS formula was discovered in 1970 independently by Broyden, Fletcher, Goldfarb and Shanno.
3.3. Analytical formulation of the problem

For the implementation of the optimization method, we need to calculate the partial derivatives of $C_{\text{Mod}}(k)$ \footnote{From now on we will use $C_{i,i}$ instead of $C_{\text{Mod}}$.} with respect to the local volatility $\sigma_{i,j}$ at each node, for $k = 1, \ldots, N$ i.e. we want to find $\frac{\partial C(l,k)}{\partial \sigma_{i,j}}$, $i = 1, \ldots, n - 1$, $j = 1, \ldots, 2^{i-1}$ and $k = 1, \ldots, N$.

For notational simplicity in the following, we assume that we have only one call option.

It can be shown that $\frac{\partial C_{i,i}}{\partial \sigma_{i,j}}$, $i = 1, \ldots, n - 1$, $j = 1, \ldots, 2^{i-1}$ is given by the formula:

\[
\frac{\partial C_{i,i}}{\partial \sigma_{i,j}} = \prod \left( \text{probabilities on the path that takes us from node (1,1) to node (i,j)} \right) \frac{\partial C_{i,i}}{\partial \sigma_{i,j}} e^{-(i-1)\gamma N/\Delta}
\]

for $i = 1, \ldots, n - 1$, $j = 1, \ldots, 2^{i-1}$.  \hspace{1cm} (3.17)

Using formula (3.7b) we compute at each intermediate node $(i, j)$ the partial derivatives of the call $C_{i,j}$ with respect to the local volatility $\sigma_{i,j}$, $\frac{\partial C_{i,j}}{\partial \sigma_{i,j}}$, for $i = 1, \ldots, n - 1$, $j = 1, \ldots, 2^{i-1}$.

Hence,

\[
\frac{\partial C_{i,j}}{\partial \sigma_{i,j}} = \left( G_{i,j}^{(1)} + G_{i,j}^{(2)} \right) e^{-\gamma N/\Delta}, \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1}.
\]  \hspace{1cm} (3.18)

where,

\[
G_{i,j}^{(1)} = \left( C_{i+1,2j} - C_{i+1,2j-1} \right) \frac{\partial p_{i,j}(\sigma_{i,j})}{\partial \sigma_{i,j}}
\] \hspace{1cm} (3.19a)
\[ G^{(2)}_{i,j} = p_{i,j} \frac{\partial C_{i+1,2j}}{\partial \sigma_{i,j}} + (1 - p_{i,j}) \frac{\partial C_{i+1,2j-1}}{\partial \sigma_{i,j}} \]  

(3.19b)

Also using formulas (3.5a) and (3.5b) we obtain the partial derivatives \( \frac{\partial p_{i,j}}{\partial \sigma_{i,j}} \) at each node \((i, j)\). Hence,

\[
\frac{\partial p_{i,j}}{\partial \sigma_{i,j}} = \frac{\sqrt{\Delta t}}{u_{i,j} - d_{i,j}} \left( d_{i,j} (1 - p_{i,j}) - u_{i,j} p_{i,j} \right) 
\]

(3.20)

for \(i = 1, \ldots, n-1\), and \(j = 1, \ldots, 2^{i-1}\).

(For proof see Appendix 3.1).

\[
G^{(1)}_{i,j} = -\left( C_{i+1,2j} + C_{i+1,2j-1} \right) \left( \frac{p_{i,j} u_{i,j} - (1 - p_{i,j}) d_{i,j}}{u_{i,j} - d_{i,j}} \right) \sqrt{\Delta t}
\]

\[
= -\left( \frac{C_{i+1,2j} - C_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} \right) e^{-\alpha t} \left( \frac{p_{i,j} S_{i+1,2j} - (1 - p_{i,j}) d_{i,j}}{u_{i,j} - d_{i,j}} \right) e^{\alpha t} \sqrt{\Delta t}
\]

\[
G^{(1)}_{i,j} = -\Delta_{i,j} \left( p_{i,j} S_{i+1,2j} - (1 - p_{i,j}) S_{i+1,2j-1} \right) e^{\alpha t} 
\]

(3.21)

where,

\[
\Delta_{i,j} = \left( \frac{C_{i+1,2j} - C_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} \right) e^{-\alpha t} \text{ Delta ratio at node } (i, j). 
\]

(3.22)

For the computation of \( G^{(2)}_{i,j} \), we consider that we are at time \( i = n - 2 \) (see Fig. 3.3).
Figure 3.3: Last two steps on a non-recombining (binary) tree.

Then,

$$\Gamma_{n-2,j}^{(2)} = p_{n-2,j} \frac{\partial C_{n-1,2j}}{\partial \sigma_{n-2,j}} + (1 - p_{n-2,j}) \frac{\partial C_{n-1,2j-1}}{\partial \sigma_{n-2,j}}$$  \hspace{1cm} (3.23)

But,

$$e^{\gamma_{n-2,j}} \frac{\partial C_{n-1,2j}}{\partial \sigma_{n-2,j}} = p_{n-1,2j} \frac{\partial C_{n,4j}}{\partial \sigma_{n-2,j}} + (1 - p_{n-1,2j}) \frac{\partial C_{n,4j-1}}{\partial \sigma_{n-2,j}}$$

$$= p_{n-1,2j} \left( \frac{\partial C_{n,4j}}{\partial S_{n,4j}} \right) \left( \frac{\partial S_{n,4j}}{\partial \sigma_{n-2,j}} \right) + (1 - p_{n-1,2j}) \left( \frac{\partial C_{n,4j-1}}{\partial S_{n,4j-1}} \right) \left( \frac{\partial S_{n,4j-1}}{\partial \sigma_{n-2,j}} \right)$$
Using the fact that,

\[ S_{n,4,j} = S_{n-2,j} e^{(\sigma_{n-2,j} + \sigma_{n-1,j})\sqrt{\Delta t}} \]

\[ S_{n,4,j-1} = S_{n-2,j} e^{(\sigma_{n-2,j} - \sigma_{n-1,j})\sqrt{\Delta t}} \]

We obtain

\[ \frac{\partial S_{n,4,j}}{\partial \sigma_{n-2,j}} = S_{n,4,j} \sqrt{\Delta t} \]

\[ \frac{\partial S_{n,4,j-1}}{\partial \sigma_{n-2,j}} = S_{n,4,j-1} \sqrt{\Delta t} \]

Hence,

\[ \frac{\partial C_{n-1,2,j}}{\partial \sigma_{n-2,j}} = \left[ p_{n-1,2,j} S_{n,4,j} \left( \frac{\partial C_{n,4,j}}{\partial S_{n,4,j}} \right) + (1 - p_{n-1,2,j}) S_{n,4,j-1} \left( \frac{\partial C_{n,4,j-1}}{\partial S_{n,4,j-1}} \right) \right] e^{-r \Delta t} \]

(3.24a)

Also,

\[ e^{r \Delta t} \frac{\partial C_{n-1,2,j-1}}{\partial \sigma_{n-2,j}} = p_{n-1,2,j-1} \left( \frac{\partial C_{n,4,j-2}}{\partial S_{n,4,j-2}} \right) \left( \frac{\partial S_{n,4,j-2}}{\partial \sigma_{n-2,j}} \right) + (1 - p_{n-1,2,j-1}) \left( \frac{\partial C_{n,4,j-3}}{\partial S_{n,4,j-3}} \right) \left( \frac{\partial S_{n,4,j-3}}{\partial \sigma_{n-2,j}} \right) \]

\[ S_{n,4,j-2} = S_{n-2,j} e^{(-\sigma_{n-2,j} + \sigma_{n-1,j-1})\sqrt{\Delta t}} \]

\[ S_{n,4,j-3} = S_{n-2,j} e^{(-\sigma_{n-2,j} - \sigma_{n-1,j-1})\sqrt{\Delta t}} \]

\[ \frac{\partial S_{n,4,j-2}}{\partial \sigma_{n-2,j}} = -S_{n,4,j-2} \sqrt{\Delta t} \]
\[
\frac{\partial S_{n,4j-3}}{\partial \sigma_{n-2,j}} = -S_{n,4j-3} \sqrt{\Delta t}
\]

Hence,

\[
\frac{\partial C_{n-1,2j-1}}{\partial \sigma_{n-2,j}} = \left[ p_{n-1,2j-1} S_{n,4j-2} \left( \frac{\partial C_{n,4j-2}}{\partial S_{n,4j-2}} \right) + (1 - p_{n-1,2j-1}) S_{n,4j-3} \left( \frac{\partial C_{n,4j-3}}{\partial S_{n,4j-3}} \right) \right] e^{-r_{j,k} \Delta t} \sqrt{\Delta t}
\]

From (3.24a) and (3.24b) it follows that:

\[
G_{n-2,j}^{(2)} = (U_{n-2,j} - D_{n-2,j}) e^{-r_{j,k} \Delta t} \sqrt{\Delta t}
\]

where,

\[
U_{n-2,j} = p_{n-2,j} p_{n-1,2j} S_{n,4j} \left( \frac{\partial C_{n,4j}}{\partial S_{n,4j}} \right) + p_{n-2,j} (1 - p_{n-1,2j}) S_{n,4j-1} \left( \frac{\partial C_{n,4j-1}}{\partial S_{n,4j-1}} \right)
\]

\[
D_{n-2,j} = (1 - p_{n-2,j}) p_{n-1,2j-1} S_{n,4j-2} \left( \frac{\partial C_{n,4j-2}}{\partial S_{n,4j-2}} \right) + (1 - p_{n-2,j}) (1 - p_{n-1,2j-1}) S_{n,4j-3} \left( \frac{\partial C_{n,4j-3}}{\partial S_{n,4j-3}} \right)
\]

Note that,

- \( U_{n-2,j} \) is the expected value of \( \left\{ S_{n,k} \frac{\partial C_{n,k}}{\partial S_{n,k}} \right\} \), given that we are at node \((n-2, j)\) and we follow the up direction, to get to the terminal nodes.

- \( D_{n-2,j} \) is the expected value of \( \left\{ S_{n,k} \frac{\partial C_{n,k}}{\partial S_{n,k}} \right\} \), given that we are at node \((n-2, j)\) and we follow the down direction, to get to the terminal nodes.

Hence,
\[ \frac{\partial C_{n-2,j}}{\partial \sigma_{n-2,j}} = G^{(1)}_{n-2,j} + \left( U_{n-2,j} - D_{n-2,j} \right) e^{-r \Delta t} \sqrt{\Delta t} \]  

(3.27)

For \( i = n - 1 \), \( G^{(2)}_{n-1,j} = p_{n-1,j} \frac{\partial C_{n,2j}}{\partial \sigma_{n-1,j}} + \left( 1 - p_{n-1,j} \right) \frac{\partial C_{n,2j-1}}{\partial \sigma_{n-1,j}} \)

For the computation of \( \frac{\partial C_{n,2j}}{\partial \sigma_{n-1,2j}} \) and \( \frac{\partial C_{n,2j-1}}{\partial \sigma_{n-1,2j-1}} \) for \( j = 1, \ldots, 2^{n-2} \) we use the following formula:

\[
\frac{\partial C_a(n,2j)}{\partial \sigma_{n-1,2j}} = \begin{cases} 
0 & \text{for } S_{n,2j} \leq K(1 - z/2) \\
S_{n,2j} \sqrt{\Delta t} & \text{for } S_{n,2j} \geq K(1 + z/2) \\
\frac{1}{z} \left[ \left( \frac{S_{n,2j}}{K} - 1 \right) + \frac{z}{2} \right] S_{n,2j} \sqrt{\Delta t} & \text{for } K(1 - z/2) < S_{n,2j} < K(1 + z/2) 
\end{cases}
\]

(3.28a)

\[
\frac{\partial C_a(n,2j-1)}{\partial \sigma_{n-1,2j-1}} = \begin{cases} 
0 & \text{for } S_{n,2j-1} \leq K(1 - z/2) \\
-S_{n,2j-1} \sqrt{\Delta t} & \text{for } S_{n,2j-1} \geq K(1 + z/2) \\
-\frac{1}{z} \left[ \left( \frac{S_{n,2j-1}}{K} - 1 \right) + \frac{z}{2} \right] S_{n,2j-1} \sqrt{\Delta t} & \text{for } K(1 - z/2) < S_{n,2j-1} < K(1 + z/2) 
\end{cases}
\]

(3.28b)

The above results can be generalized to the following formula:

\[
\frac{\partial C_{i,j}}{\partial \sigma_{i,j}} = \begin{cases} 
G^{(1)}_{i,j} e^{-r \Delta t} + \left( U_{i,j} - D_{i,j} \right) e^{-r \Delta t} & i = 1, \ldots, n - 2 \\
\left( C^{(l)}_{i,j} + G^{(2)}_{i,j} \right) e^{-r \Delta t} & i = n - 1 
\end{cases}
\]

(3.29)

for \( j = 1, \ldots, 2^{i-1} \)

where,
$U_{i,j}$ is the expected value of $\left\{ S_{n,k} \frac{\partial C_{n,k}}{\partial S_{n,k}} \right\}$, given that we are at node $(i, j)$ and we follow the up direction, to get to the terminal nodes.

$D_{i,j}$ is the expected value of $\left\{ S_{n,k} \frac{\partial C_{n,k}}{\partial S_{n,k}} \right\}$, given that we are at node $(i, j)$ and we follow the down direction, to get to the terminal nodes.

In Appendix 3.2, there is an alternative derivation of the partial derivatives $\frac{\partial C_{i,j}}{\partial \sigma_{i,j}}, i = 1, ..., n-1, j = 1, ..., 2^{i-1}$ using the formula:

$$C_{1,1} = \sum_{j=1}^{2^{n-1}} \lambda_{n,j} C_{n,j}$$

(3.30)

where, $\lambda_{n,j} j = 1, ..., 2^{n-1}$ are the Arrow-Debreu prices$^{46}$.

### 3.4. Data - FTSE 100 index options

We use the daily closing prices of FTSE 100 American and European call options of January 2003 to December 2003 as reported by LIFFE$^{47}$. The strike prices for a given style call (European or American) are spaced at intervals of 50 index points from each other. However, there are no American and European calls with identical strikes. The strike prices for adjacent European and American style calls are

---

$^{46}$ The Arrow-Debreu price at node $(i, j)$ is the price of an option that pays 1 unit pay-off in one and only one state $j$ at the $i^{th}$ time step, and otherwise pays 0.

$^{47}$ FTSE 100 European options are traded with expiries in March, June, September, and December. Additional contracts are introduced serially so that the nearest 4 months are always available for trading. FTSE 100 American options are traded with expiries the nearest of June and December. Additional contracts are introduced serially so that the nearest 3 months are always available for trading. FTSE 100 options expire on the third Friday of the expiry month. FTSE 100 options positions are marked-to-market daily based on the daily settlement price, which is determined by LIFFE and confirmed by the Clearing House. FTSE 100 options are quoted in index points and have an assigned value of £10 per index point.
spaced at intervals of 25 index points. For example, there are European style calls with strikes 3075, 3125, and 3175 and American style calls with strikes 3100, 3150, and 3200. Also, the longest maturity for the European calls is 2 years while for the American calls is 6 months. Our initial sample (for the 12 months period) consists of 99,051 observations of European calls, and 34,503 observations of American calls. We apply five standard filtering rules to both American and European calls data. First we exclude calls that violate the no-arbitrage bounds\(^{48}\). Second, we eliminate calls with time to maturity less than 6 calendar days, i.e. \(T<6\). These calls have very small time-premiums and their implied volatilities are inaccurate since they are very sensitive to market microstructure problems and measurement errors (Hentschel, 2001, Brandt, et al., 2002). Third, we eliminate calls if their closing price is less than 0.5 index points. Fourth, we eliminate calls for which the traded volume is zero (since we want highly liquid options for calibration). Finally, we eliminate calls with moneyness greater than 1.1 or moneyness less than 0.8, since deep in the money and deep out of the money call options are expected to be illiquid and not accurately priced. We define as moneyness of a call option the ratio (underlying asset price)/(strike price). After these filtering rules, our sample consists of 13,696 observations of European call options and 355 observations of American call options.

For the implementation, we consider only cases for which the number of options used for calibration \(N\) is greater than 8 since with fewer options the distribution obtained of the underlying asset will not be reliable. In addition, we eliminate European call options with maturity more than 180 calendar days, since the longest maturity for American calls is 6 months. Thus, the sample of European calls used for the model calibration becomes 8,486. For the pricing of American calls we use only the European calls that correspond to the American calls. Thus, the European sample used for pricing of American calls is reduced to 2,572 observations and the American sample is reduced to 312 observations.

\(^{48}\) For European (EU) style options \(\max(0, S e^{-rT} - K e^{-rT}) \leq C_{EU} \leq S\) and for American (AM) style options \(\max(C_{EU}, S - K) \leq C_{AM} \leq S\).
For time to maturity, \( T \), we use the calendar days to maturity. For the risk-free rate, \( r_f \), we use cubic spline interpolation for matching each option contract with a continuous interest rate that corresponds to the option’s maturity, by utilizing the 1-month to 12-month LIBOR offer rates, collected from Datastream. Also, since the underlying asset of the options on FTSE 100 is a futures contract, we make the standard assumption that the dividend yield (\( \delta \)) equals the risk free rate. The models are calibrated every day. For each implementation, the options used have the same underlying asset and the same time to maturity.

### Table 3.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003, used for the model calibration.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>84.86</td>
<td>33.77</td>
<td>110.80</td>
<td>219.18</td>
<td>82.12</td>
<td>81.08</td>
<td>92.07</td>
</tr>
<tr>
<td>median</td>
<td>48</td>
<td>20</td>
<td>104.5</td>
<td>209.5</td>
<td>39</td>
<td>45.5</td>
<td>56.75</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>17.5</td>
<td>69</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>436.5</td>
<td>283.5</td>
<td>295.5</td>
<td>436.5</td>
<td>398</td>
<td>420</td>
<td>436.5</td>
</tr>
<tr>
<td>observations</td>
<td>8,486</td>
<td>5,614</td>
<td>913</td>
<td>1,959</td>
<td>2,514</td>
<td>3,288</td>
<td>2,684</td>
</tr>
</tbody>
</table>

Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

### Table 3.1b: Statistics of the FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call options.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>84.2</td>
<td>33.5</td>
<td>115.1</td>
<td>220.6</td>
<td>81.7</td>
<td>84.7</td>
<td>88.6</td>
</tr>
<tr>
<td>median</td>
<td>47.3</td>
<td>19</td>
<td>112.5</td>
<td>211</td>
<td>40</td>
<td>49</td>
<td>51</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>30.5</td>
<td>72.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>410.5</td>
<td>263.5</td>
<td>230</td>
<td>410.5</td>
<td>398</td>
<td>400.5</td>
<td>410.5</td>
</tr>
<tr>
<td>observations</td>
<td>2,572</td>
<td>1,731</td>
<td>256</td>
<td>585</td>
<td>1,038</td>
<td>1,117</td>
<td>417</td>
</tr>
</tbody>
</table>

Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

\[49\] In the specific dataset (given by LIFFE), the underlying asset of each call option is a future contract and thus every trading day the options have different underlying asset.
<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mean</strong></td>
<td>72.5</td>
<td>44.0</td>
<td>99.1</td>
<td>212.9</td>
<td>68.2</td>
<td>81.1</td>
<td>61.7</td>
</tr>
<tr>
<td><strong>median</strong></td>
<td>49</td>
<td>35</td>
<td>97.5</td>
<td>208</td>
<td>49</td>
<td>55</td>
<td>42.75</td>
</tr>
<tr>
<td><strong>min</strong></td>
<td>0.5</td>
<td>0.5</td>
<td>33.5</td>
<td>71</td>
<td>0.5</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>354</td>
<td>192</td>
<td>182</td>
<td>354</td>
<td>354</td>
<td>318</td>
<td>262.5</td>
</tr>
<tr>
<td><strong>observations</strong></td>
<td>312</td>
<td>231</td>
<td>42</td>
<td>39</td>
<td>147</td>
<td>118</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 3.1c: Statistics of the FTSE 100 American (AM) style call options for the year 2003. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mean</strong></td>
<td>353.71</td>
<td>394.0</td>
<td>495.76</td>
<td>172.02</td>
<td>367.55</td>
<td>314.11</td>
<td>389.25</td>
</tr>
<tr>
<td><strong>median</strong></td>
<td>52</td>
<td>64</td>
<td>134</td>
<td>21</td>
<td>97.5</td>
<td>42.5</td>
<td>36</td>
</tr>
<tr>
<td><strong>min</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>37,000</td>
<td>13,000</td>
<td>4,837</td>
<td>37,000</td>
<td>6,930</td>
<td>10,100</td>
<td>37,000</td>
</tr>
<tr>
<td><strong>observations</strong></td>
<td>8,486</td>
<td>5,614</td>
<td>913</td>
<td>1,959</td>
<td>2,514</td>
<td>3,288</td>
<td>2,684</td>
</tr>
</tbody>
</table>

Table 3.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, used for the model calibration. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>mean</strong></td>
<td>343.5</td>
<td>378.2</td>
<td>578.2</td>
<td>138.1</td>
<td>385.9</td>
<td>327.9</td>
<td>279.7</td>
</tr>
<tr>
<td><strong>median</strong></td>
<td>60</td>
<td>81</td>
<td>178.5</td>
<td>19</td>
<td>98</td>
<td>52</td>
<td>37</td>
</tr>
<tr>
<td><strong>min</strong></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>max</strong></td>
<td>7,683</td>
<td>7,583</td>
<td>4,837</td>
<td>400</td>
<td>6,930</td>
<td>7,683</td>
<td>3,021</td>
</tr>
<tr>
<td><strong>observations</strong></td>
<td>2,572</td>
<td>1,731</td>
<td>256</td>
<td>585</td>
<td>1,038</td>
<td>3,288</td>
<td>417</td>
</tr>
</tbody>
</table>

Table 3.2b: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003, which have corresponding American call option. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.
<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>57</td>
<td>60</td>
<td>38.2</td>
<td>60.1</td>
<td>33.9</td>
<td>32.2</td>
<td>191.7</td>
</tr>
<tr>
<td>median</td>
<td>10</td>
<td>10</td>
<td>5.5</td>
<td>10</td>
<td>10</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>3,000</td>
<td>3,000</td>
<td>340</td>
<td>350</td>
<td>350</td>
<td>350</td>
<td>300</td>
</tr>
<tr>
<td>observations</td>
<td>312</td>
<td>231</td>
<td>42</td>
<td>39</td>
<td>147</td>
<td>118</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 3.2c: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

Tables 3.1a, 3.1b, 3.1c, 3.2a, 3.2b, and 3.2c describe the cleaned sample. Table 3.1a shows the mean, median, minimum and maximum call option value and also the number of observations for the European call option sample used for the model calibration. The statistics are computed for the full sample of European call options and also for the sub-samples of the out-of-the-money (OTM), at-the-money (ATM), in-the-money (ITM), short term (SM), medium term (MT) and long term (LT) options.

Table 3.1b shows the same statistics as Table 3.1a for the European call options that have corresponding American call option and Table 3.1c shows the statistics for the American options sample. Tables 3.2a 3.2b and 3.2c show the same statistics for the trading volume of the European and American call option samples. Overall we see that the European calls are highly traded with mean and median volume 343.5 and 60 respectively. The corresponding American calls are less liquid with mean and median volume 57 and 10 respectively.

3.5. Calibration and testing of the model

Every trading day, we calibrate the proposed model \((INRT_{LocalVol})\) for \(n = 7\) using European call options with the same underlying asset and time to maturity.

---

50 Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.
Then, we check the model for over-fitting by pricing options with strikes in-between those used for the calibration. Figure 3.4 shows plots of the call prices (market prices and estimated from the model) versus moneyness. For brevity we present only the plots for the first trading day of June, 2003. Similar results are obtained for the other trading days and months. As we see, the estimated call values increase smoothly with increasing moneyness without any evidence of over-fitting.

Figure 3.4: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003.

In order to see how realistic is the distribution obtained from our model for year 2003, we calculate the statistics of the 1-month log-returns obtained from our model and compare them with the historical 1-month log-returns for the year 2003, and the years 2002-2004 and 2001-2005. Specifically, for each calibration for which the options maturity was between 28 and 32 calendar days, we calculate the first four moments (mean, variance, skewness and kurtosis). Then, in order to get a feeling for the representative statistics of 1-month log-returns we provide for each of those moments the mean and the median. For comparison purposes, we also show statistics
for the INRT model proposed in the first chapter for \( n = 6 \) and \( n = 7 \). Statistics are summarized in Table 3.3.

<table>
<thead>
<tr>
<th>( \text{INRT (n=6)} )</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.0021</td>
<td>0.0042</td>
<td>-0.6178</td>
<td>3.7645</td>
<td>58</td>
</tr>
<tr>
<td>Median</td>
<td>-0.0013</td>
<td>0.0027</td>
<td>-0.5318</td>
<td>3.0904</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \text{INRT (n=7)} )</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.0021</td>
<td>0.0042</td>
<td>-0.6270</td>
<td>3.7131</td>
<td>58</td>
</tr>
<tr>
<td>Median</td>
<td>-0.0013</td>
<td>0.0027</td>
<td>-0.5585</td>
<td>3.1819</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \text{INRT}_{\text{LocalVol}} )</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.0022</td>
<td>0.0046</td>
<td>-1.1562</td>
<td>6.3910</td>
<td>58</td>
</tr>
<tr>
<td>Median</td>
<td>-0.0015</td>
<td>0.0031</td>
<td>-0.8857</td>
<td>4.3586</td>
<td>58</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \text{Historical} )</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>0.0106</td>
<td>0.0014</td>
<td>-0.6572</td>
<td>2.7689</td>
<td>12</td>
</tr>
<tr>
<td>2002-2004</td>
<td>-0.0028</td>
<td>0.0020</td>
<td>-1.2699</td>
<td>4.7599</td>
<td>35</td>
</tr>
<tr>
<td>2001-2005</td>
<td>-0.0021</td>
<td>0.0018</td>
<td>-1.1177</td>
<td>4.4749</td>
<td>59</td>
</tr>
</tbody>
</table>

Table 3.3: Implied risk-neutral, and historical statistics for the distribution of the FTSE 100 1-month log-returns, using the proposed model (INRT\(_{\text{LocalVol}}\)) for \( n=7 \), and the non-recombining implied tree model for \( n=6 \) and \( n=7 \) (INRT\((n=6)\), INRT\((n=7)\)).

As we would expect, the mean of the implied risk-neutral and implied real distribution of log-returns differs from that of the historical distribution\(^{51}\). Also, consistently with Jackwerth and Rubinstein (1996) and Liu et al. (2005) in all cases the implied variance is larger than the historical variance. Implied volatility can be thought of as the market’s expected volatility plus some “volatility risk-premium” for other unknown factors such as hedging costs, the inability for perfect hedging, the uncertainty of future volatility etc\(^{52}\). Furthermore, the mean implied skewness of the INRT\(_{\text{LocalVol}}\) model is higher in absolute terms (-1.1562) than the mean implied skewness of the INRT model for both \( n = 6 \) (-0.6178) and \( n = 7 \) (-0.6270). Also the mean implied kurtosis of the INRT\(_{\text{LocalVol}}\) model is higher (6.3910) than the mean implied kurtosis of the INRT for \( n = 6 \) (3.7645) and the INRT for \( n = 7 \) (3.7131). Overall, statistics indicate that the proposed model gives distributions that are more

---

\(^{51}\) Due to rounding errors in the calculations the risk-neutral implied mean is not equal to zero.

\(^{52}\) The annual implied variance estimated from the proposed model is about 23.5\% whereas the annual realized variance is about 15.5\%.
skewed to the left and more leptokurtic than the INRT. Also, the implied skewness and kurtosis of the proposed model are closer to the historical ones\textsuperscript{53}. This is an indication that the implied distribution is realistic.

In order to give further evidence for the implied distributions obtained by our model for \( n = 7 \) versus the INRT model for \( n = 6 \) and \( n = 7 \), representative implied distributions (histograms) for the 1-month log-returns in June 2003 are shown in Figures 3.5a-3.5c. Like in the first chapter, in order to make the histograms of the implied distributions we make use of the Pearson system of distributions\textsuperscript{54} as applied in Matlab\textsuperscript{55}. Using the first four moments of the data it is easy to find in the Pearson system the distribution that matches these moments and to generate a random sample so as to produce a histogram corresponding to the implied distribution. From the figures, it is obvious that the implied distribution obtained from all models has negative skewness and excess kurtosis, with the INRT\textsubscript{LocalVol} model being more leptokurtic. These figures are representative of the vast majority of cases.

\textsuperscript{53} Liu \textit{et al.} (2005) demonstrate that the needed adjustments to get the implied real variance, skewness and kurtosis from the implied risk-neutral ones are minimal. Thus, knowing that our implied risk-neutral moments (beyond the mean) are very close to the implied real ones, we can then compare them with the historical ones (but without expecting the two distributions to be identical).

\textsuperscript{54} In the Pearson system there is a family of distributions that includes a unique distribution corresponding to every valid combination of mean, standard deviation, skewness, and kurtosis.

\textsuperscript{55} Copyright 2005 The MathWorks, Inc.
Figure 3.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) for $n=6$.

Figure 3.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) for $n=7$. 
3.6. Pricing of American options

In this section we use the proposed model for $n = 7$ (63 variables) and the INRT model for $n = 6$ (62 variables) and $n = 7$ (126 variables), for pricing American call options. In terms of computational complexity, the proposed model is comparable to the INRT model for $n = 6$ since they have close number of variables (63 vs. 62). In terms of the step size used, the proposed model is equivalent to the INRT for $n = 7$ since they have the same number of steps. For completeness, we also make a comparison with the ad-hoc model of smoothing BS implied volatilities across strikes proposed by Dumas et al. (1998). Every trading day in the sample, we calibrate the above models using European call options data with the same underlying asset and time to maturity. Then, we use each model to price the corresponding American style calls. For the ad-hoc model, in order to price the American calls, we build standard CRR binomial trees using the implied volatility obtained by interpolating the implied
volatility surface, obtained from European calls, across strikes using a 2nd order polynomial regression (AH_Regr model). We use CRR trees of \( n = 50 \) and \( n = 51 \), and the call option value is computed as the average of the call option values of the two trees reducing thus the impact of oscillations.

For the comparison of the pricing performance of the four models we use two pricing measures.

(i) The root mean square error (RMSE) which is the square-root of the average squared difference between the model and the market prices. The RMSE measures how well the model fits in a statistical sense with the usual bias versus variance trade-off.

(ii) The mean difference error (MDE) which is the average of the difference between the model and the market prices and reveals systematic biases of the model.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>INRT((n=6))</td>
<td>4.7085</td>
<td>-1.6506</td>
</tr>
<tr>
<td></td>
<td>INRT((n=7))</td>
<td>4.3143</td>
<td>-1.5370</td>
</tr>
<tr>
<td></td>
<td>INRT(_{LocalVol}(n=7))</td>
<td>4.4252</td>
<td>-1.5353</td>
</tr>
<tr>
<td></td>
<td>AH(_{Regr})</td>
<td>4.2251</td>
<td>-1.3426</td>
</tr>
</tbody>
</table>

Table 3.4: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree model (INRT) for \( n=6 \) and \( n=7 \), the model proposed in this chapter (INRT\(_{LocalVol}\) for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (AH\(_{Regr}\)). We measure pricing errors by (i) the ad hoc square error (RMSE), and (ii) the mean difference error (MDE). Results are obtained using the full sample of American call options.

Table 3.4 reports model comparisons using the above measures for pricing the American call options using the four models. We observe that on average all models systematically underprice the American calls. The INRT \((n=6)\) model makes the highest underpricing -1.6506, whereas the AH\(_{Regr}\) model makes the least underpricing -1.3426. Looking at the RMSE error, we see that the AH\(_{Regr}\) has the lowest error 4.2251, followed by the INRT\((n=7)\), having RMSE 4.3143. The INRT\(_{LocalVol}\) and the INRT\((n=6)\) come next with RMSE 4.4252 and 4.7085 respectively.
Table 3.5: Pricing errors for the American FTSE 100 index call options using the implied non-recombining tree model (INRT) for \( n = 6 \) and \( n = 7 \), the model proposed in this chapter (INRT_LocalVol) for \( n = 7 \), and the ad-hoc model of Dumas et al. (1998) (AH_Regr). We measure pricing errors with (i) the mean absolute difference error (MADE) and (ii) the median absolute difference error (MDADE). The last two columns contain the p-values from the Wilcoxon sign rank tests. The first column tests whether the distribution of the absolute errors obtained from the INRT_LocalVol is statistically different from the distribution of each of the other models. The last column tests whether the distribution of the absolute errors obtained from the on average best model (AH_Regr) is statistically different from the distribution of each of the other models.

In order to test whether the pricing performance between the different models is statistically different we perform the Wilcoxon sign rank test\(^{56}\). As a measure of pricing performance we use the absolute difference error of each model. We test whether the distribution of the absolute errors obtained from each model is statistically different from the distribution of the absolute errors of the proposed model, INRT_LocalVol and from the distribution of the absolute errors of the on average best model which is the AH_Regr. Table 3.5 contains the mean absolute difference errors (MADE) and the median absolute difference errors (MDADE) for the four models. In the two last columns of Table 3.5 are presented the \( p \)-values of the Wilcoxon sign rank tests. We observe that the distributions of the absolute errors of the INRT model for \( n = 6 \) and \( n = 7 \), and of the AH_Regr model are not statistically different from the corresponding distribution of the INRT_LocalVol model at 1% level of significance. Also the distributions of the absolute errors of the INRT model for \( n = 6 \) and \( n = 7 \), and of the INRT_LocalVol model are not statistically different from the corresponding distribution of the AH_Regr model at 1% level of significance.

In Table 3.6 the RMSE and MDE of the four models are presented when the sample is divided in the three moneyness categories. With respect to underpricing

\(^{56}\) Wilcoxon sign rank test is a non-parametric alternative to the paired Student’s t-test for the case of two related samples. It performs a paired two sided test of the hypothesis that the difference between the matched samples comes from a distribution with median zero. In other words, it tests whether the two paired-samples have the same distribution.
/overpricing results remain the same as in the overall sample except from the ITM category. OTM and ATM options are underpredicted by all models whereas the ITM options are overpredicted by the \textit{AH_Regr} model and underpredicted by the other models. With respect to the RMSE, we observe that in the OTM and the ATM categories, all models give RMSE that are very close. An exception is observed in the ITM category, where there is a variation of the RMSE of each of the models. However, when the Wilcoxon test of Table 3.5 is applied in each moneyness category (untabulated results), we find similar results as in the full sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTM</td>
<td>\textit{INRT}(n=6)</td>
<td>3.6255</td>
<td>-1.7496</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}(n=7)</td>
<td>3.6351</td>
<td>-1.7202</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}_{\text{LocalVol}}(n=7)</td>
<td>3.5884</td>
<td>-1.6987</td>
</tr>
<tr>
<td></td>
<td>\textit{AH_Regr}</td>
<td>3.5808</td>
<td>-1.6856</td>
</tr>
<tr>
<td>ATM</td>
<td>\textit{INRT}(n=6)</td>
<td>4.4578</td>
<td>-1.2448</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}(n=7)</td>
<td>4.3583</td>
<td>-1.2525</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}_{\text{LocalVol}}(n=7)</td>
<td>4.4038</td>
<td>-1.2763</td>
</tr>
<tr>
<td></td>
<td>\textit{AH_Regr}</td>
<td>4.3396</td>
<td>-1.1160</td>
</tr>
<tr>
<td>ITM</td>
<td>\textit{INRT}(n=6)</td>
<td>8.8378</td>
<td>-1.5010</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}(n=7)</td>
<td>7.0840</td>
<td>-0.7583</td>
</tr>
<tr>
<td></td>
<td>\textit{INRT}_{\text{LocalVol}}(n=7)</td>
<td>7.7141</td>
<td>-0.8461</td>
</tr>
<tr>
<td></td>
<td>\textit{AH_Regr}</td>
<td>6.8253</td>
<td>0.4452</td>
</tr>
</tbody>
</table>

Table 3.6: Pricing errors within moneyness categories for the American FTSE 100 index call options using the implied non-recombining tree model (\textit{INRT}) for \(n=6\) and \(n=7\), the model proposed in this chapter (\textit{INRT}_{\text{LocalVol}}) for \(n=7\), and the ad-hoc model of Dumas et al. (1998) (\textit{AH_Regr}). We measure pricing errors by (i) the root mean square error (RMSE), and (ii) the mean difference error (MDE). Results are for the full sample when divided in 3 moneyness categories. OTM denotes the out-of-the-money, ATM denotes the at-the-money and ITM denotes the in-the-money American call options.
Table 3.7: Pricing errors within maturity categories for the American FTSE 100 index call options using the implied non-recombining tree model (INRT) for $n=6$ and $n=7$, the model proposed in this chapter (INRT_LocalVol) for $n=7$, and the ad-hoc model of Dumas et al. (1998) (AH_Regr). We measure pricing errors by (i) the root mean square error (RMSE), and (ii) mean difference error (MDE). Results are for the full sample when divided in 3 maturity categories. ST denotes the short term, MT denotes the medium term and LT denotes the long term American call options.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>INRT($n=6$)</td>
<td>4.0894</td>
<td>-1.0435</td>
</tr>
<tr>
<td></td>
<td>INRT($n=7$)</td>
<td>4.0187</td>
<td>-1.0044</td>
</tr>
<tr>
<td></td>
<td>INRT_LocalVol($n=7$)</td>
<td>4.0317</td>
<td>-0.9935</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.0873</td>
<td>-0.9561</td>
</tr>
<tr>
<td>MT</td>
<td>INRT($n=6$)</td>
<td>5.1448</td>
<td>-2.2709</td>
</tr>
<tr>
<td></td>
<td>INRT($n=7$)</td>
<td>4.1982</td>
<td>-2.0377</td>
</tr>
<tr>
<td></td>
<td>INRT_LocalVol($n=7$)</td>
<td>4.5199</td>
<td>-2.0701</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>3.9093</td>
<td>-1.5785</td>
</tr>
<tr>
<td>LT</td>
<td>INRT($n=6$)</td>
<td>5.3304</td>
<td>-1.9919</td>
</tr>
<tr>
<td></td>
<td>INRT($n=7$)</td>
<td>5.3665</td>
<td>-1.9459</td>
</tr>
<tr>
<td></td>
<td>INRT_LocalVol($n=7$)</td>
<td>5.2789</td>
<td>-1.8870</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>5.2804</td>
<td>-1.9588</td>
</tr>
</tbody>
</table>

In Table 3.7 the RMSE, and the MDE of the four models are presented when the sample is divided in the three time-to-maturity categories. In all categories, the four models on average underprice the American call options. Also the RMSE errors provided by each of the models in the three categories are very close, except in the MT category, where are observed some variations. However, when the Wilcoxon test of Table 3.5 is applied in each maturity category, we find similar results as in the full sample.

Overall, results indicate that the four models in terms of pricing performance of American call options are equally good. Therefore, if traders need to use a non-recombining tree for pricing of American call options, they can chose between the INRT and INRT_LocalVol, secure in the knowledge that the models are equally good as the benchmark model AH_Regr.
3.7. Conclusions

The appearance and persistence of the volatility smile and term structure in the financial markets have encouraged the development of smile consistent models. In this chapter, we propose a model for calibrating the non-recombining tree with respect to the local volatility. In other words, we concentrate on the extraction of the local volatility surface from a non-recombining implied tree. The problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and use nonlinear constrained optimization to minimize the least-squares error function on market prices. Specifically, in order to handle the inequality constraints, we adopt an exterior penalty method and the optimization is implemented using a Quasi-Newton algorithm. Appropriate constraints allow us to maintain risk neutrality and prevent arbitrage opportunities.

Using the local volatility as the system variable has the benefit that, in relation to the implied non-recombining tree (INRT) model proposed in the first chapter, we reduce the number of variables in the optimization algorithm to half, making the problem less computationally intensive. Moreover, we can easily control the values of the local volatilities at each node of the tree by imposing lower and upper bounds on the values of the local volatility. Knowledge of the local volatility surface is especially useful in markets with pronounced smile to measure market sentiment, to compute the evolution of implied volatilities through time, and to value and hedge exotic options.

We test our model using call options data for the FTSE 100 index for the year 2003, obtained from LIFFE. Results strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem, and the derived implied distributions are realistic. Also the computational burden is not a major issue. In pricing of American call options, the proposed model is equally good as the INRT model proposed in the first chapter and the ad-hoc procedure of smoothing Black-Scholes implied volatilities across strikes proposed by Dumas et al. (1998).
Appendix 3.1: Derivation of partial derivative $\frac{\partial p_{i,j}}{\partial \sigma_{i,j}}$, for $i = 1, \ldots, n-1$, and $j = 1, \ldots, 2^{i-1}$.

The risk neutral probability for an up-movement from node $(i, j)$ is given by:

$$p_{i,j} = \frac{e^{(r_j - \delta_j)\Delta t} - d_{i,j}}{u_{i,j} - d_{i,j}} \quad (3.1)$$

Hence,

$$\frac{\partial p_{i,j}}{\partial \sigma_{i,j}} = \left[ \frac{d_{i,j}(u_{i,j} - d_{i,j}) - (u_{i,j} + d_{i,j})(e^{(r_j - \delta_j)\Delta t} - d_{i,j})}{(u_{i,j} - d_{i,j})^2} \right] \sqrt{\Delta t}$$

$$= \left[ \frac{d_{i,j} - (e^{(r_j - \delta_j)\Delta t} - d_{i,j})(u_{i,j} + d_{i,j})}{(u_{i,j} - d_{i,j})^2} \right] \sqrt{\Delta t}$$

$$= -\frac{(p_{i,j}u_{i,j} - (1 - p_{i,j})d_{i,j})}{(u_{i,j} - d_{i,j})} \sqrt{\Delta t} \quad (3.2)$$
Appendix 3.2: Alternative approach for computing \( \frac{\partial C_{i,j}}{\partial \sigma_{i,j}}, i = 1, \ldots, n - 1, \ j = 1, \ldots, 2^{i-1} \)

Here, for the computation of the call option value at point \((1,1)\), we use the following formula:

\[
C_{1,1} = \sum_{j=1}^{2^{i-1}} \lambda_{n,j} C_{n,j}
\]  

where, \( \lambda_{n,j} \ j = 1,\ldots,2^{n-1} \) are the Arrow-Debreu prices.

For the computation of \( \frac{\partial C_{i,j}}{\partial \sigma_{i,j}}, i = 1, \ldots, n - 1, \ j = 1, \ldots, 2^{i-1} \) we implement the following steps:

1st step: Compute the partial derivatives of the risk neutral transition probabilities \( p_{i,j} \) with respect to the local volatility \( \sigma_{i,j} \), \( \frac{\partial p_{i,j}}{\partial \sigma_{i,j}} \), for \( i = 1, \ldots, n - 1, \) and \( j = 1, \ldots, 2^{i-1} \).

\[
\frac{\partial p_{i,j}}{\partial \sigma_{i,j}} = \frac{\sqrt{\Delta t}}{u_{i,j} - d_{i,j}} \left( d_{i,j} (1 - p_{i,j}) - u_{i,j} p_{i,j} \right)
\]  

for \( i = 1, \ldots, n - 1, \) and \( j = 1, \ldots, 2^{i-1} \).

2nd step: Compute the Arrow-Debreu prices at the last time step.

\[
\lambda_{1,1} = 1
\]

\[
\lambda_{n,j} = \begin{cases} \lambda_{i-1,j} \frac{1 - p_{i-1,j}}{2} e^{-r \Delta t} & i = 2, \ldots, n \ j = 1, \ldots, 2^{i-1} - 1 \\ \lambda_{i-1,j} \frac{p_{i-1,j}}{2} e^{-r \Delta t} & i = 2, \ldots, n \ j = 2, 4, \ldots, 2^{i-1} \end{cases}
\]  

Otherwise,
\[ \lambda_{i,j} = \prod \text{(probabilities on the path that takes us from node (1,1) to node (i, j))} e^{r_j (i-1) \Delta} \]

**3rd step:** Compute the partial derivatives \( \frac{\partial \lambda_{n,m}}{\partial \sigma_{i,j}} \) for \( i = 1, ..., n - 1, \ j = 1, ..., 2^{i-1}, \)

\( m = 1, ..., 2^{n-1} \).

\[ \frac{\partial \lambda_{i+1,2j-1}}{\partial \sigma_{i,j}} = -\lambda_{i,j} \frac{\partial p_{i,j}}{\partial \sigma_{i,j}} e^{-r_j \Delta} \quad i = 1, ..., n - 1, \ j = 1, ..., 2^{i-1} \]  

(3.B4a)

or,

\[ \frac{\partial \lambda_{i+1,2j-1}}{\partial \sigma_{i,j}} = -\prod \text{(probabilities on the path that takes us from node (1,1) to node (i, j))} \frac{\partial p_{i,j}}{\partial \sigma_{i,j}} e^{-r_j \Delta} \]

\[ \frac{\partial \lambda_{i+1,2j}}{\partial \sigma_{i,j}} = -\frac{\partial \lambda_{i+1,2j-1}}{\partial \sigma_{i,j}} \quad i = 1, ..., n - 1, \ j = 1, ..., 2^{i-1} \]  

(3.B4b)

\[ \frac{\partial \lambda_{l,m}}{\partial \sigma_{i,j}} = \begin{cases} \frac{\partial \lambda}{\partial \sigma_{i,j}} \left( \frac{l-1}{2} \right) \left( 1 - p_{l-1,m+1/2} \right) e^{-r_l \Delta} & l = i + 2, ..., n, \ \ m = 1, 3, ..., 2^{l-1} - 1 \\ \frac{\partial \lambda}{\partial \sigma_{i,j}} \frac{l-1}{2} p_{l-1,m} e^{-r_l \Delta} & l = i + 2, ..., n, \ \ m = 2, 4, ..., 2^{l-1} \end{cases} \]  

(3.B4c)

or, for \( l = i + 2, ..., n \)

\[ \frac{\partial \lambda_{l,m}}{\partial \sigma_{i,j}} = \begin{cases} \left( \text{probability that takes us from node } (l-1, \frac{m+1}{2}) \text{ to node } (l,m) \right) \frac{\partial \lambda}{\partial \sigma_{i,j}} \left( \frac{l-1}{2} \right) e^{-r_l \Delta} & m = 1, 3, ..., 2^{l-1} - 1 \\ \left( \text{probability that takes us from node } (l-1, \frac{m}{2}) \text{ to node } (l,m) \right) \frac{\partial \lambda}{\partial \sigma_{i,j}} \left( \frac{l-1}{2} \right) e^{-r_l \Delta} & m = 2, 4, ..., 2^{l-1} \end{cases} \]
4th Step: Compute \( \frac{\partial S_{n,m}}{\partial \sigma_{i,j}} \) for \( i = 1, ..., n-1 \), \( j = 1, ..., 2^{i-1} \), \( m = 1, ..., 2^{n-1} \)

\[
\begin{align*}
\frac{\partial S_{n,j}}{\partial \sigma_{n-1,j+1/2}} &= -S_{n,j} \sqrt{\Delta t} \quad j = 1, 3, 5, ..., 2^{n-1} - 1 \\
\frac{\partial S_{n,j}}{\partial \sigma_{n-1,j-1/2}} &= S_{n,j} \sqrt{\Delta t} \quad j = 2, 4, 6, ..., 2^{n-1}
\end{align*}
\] (3.55a) (3.55b)

For \( i \leq n-2 \) and \( j = 1 \)

\[
\frac{\partial S_{n,m}}{\partial \sigma_{i,j}} = \begin{cases} 
-S_{n,m} \sqrt{\Delta t} & m = 1, ..., 2^{n-i-1} \\
S_{n,m} \sqrt{\Delta t} & m = 2^{n-i-1} + 1, ..., 2^{n-i}
\end{cases}
\] (3.55c)

For \( i \leq n-2 \) and \( j \neq 1 \)

\[
\frac{\partial S_{n,m}}{\partial \sigma_{i,j}} = \begin{cases} 
-S_{n,m} \sqrt{\Delta t} & m = 2^{n-i}(j-1) + 1, ..., 2^{n-i}\left( j - \frac{1}{2} \right) \\
S_{n,m} \sqrt{\Delta t} & m = 2^{n-i}\left( j - \frac{1}{2} \right) + 1, ..., 2^{n-i}j
\end{cases}
\] (3.55d)
\[5^{\text{th}} \text{ step: Compute } \frac{\partial C_{n,m}}{\partial \sigma_{i,j}} \text{ for } i = 1, \ldots, n-1, \ j = 1, \ldots, 2^{i-1}, \ m = 1, \ldots, 2^{n-1}\]

\[
\frac{\partial C_\alpha(n,m)}{\partial \sigma_{i,j}} = \begin{cases} 
0 & \text{for } S_{n,m} \leq K(1 - z/2) \\
\frac{\partial S_{n,m}}{\partial \sigma_{i,j}} & \text{for } S_{n,m} \geq K(1 + z/2) \\
\frac{1}{z}\left[\frac{(S_{n,m} - 1)}{K} + \frac{z}{2}\right] \frac{\partial S_{n,m}}{\partial \sigma_{i,j}} & \text{for } K(1 - z/2) < S_{n,m} < K(1 + z/2)
\end{cases}
\]

(3.66)

\[6^{\text{th}} \text{ step: Compute } \frac{\partial C_{1,1}}{\partial \sigma_{i,j}}, \ i = 1, \ldots, n-1, \ j = 1, \ldots, 2^{i-1}\]

\[
\frac{\partial C_{1,1}}{\partial \sigma_{i,j}} = \sum_{m=1}^{2^{i-1}} \frac{\partial \lambda_{n,m}}{\partial \sigma_{i,j}} C_{n,m} + \lambda_{n,m} \frac{\partial C_{n,m}}{\partial \sigma_{i,j}}
\]

(3.67)
4. Option pricing on non-recombining implied trees assuming serial dependence of returns

Abstract

In this chapter we calibrate the non-recombining implied tree taking into account serial dependence of the underlying asset’s returns. Effectively, the model becomes non-Markovian. Unlike typical preference-free option pricing models, a parameter related to the expected return of the underlying asset appears in our model. We calibrate the non-Markovian model using European calls on the FTSE 100 index for the year 2003. Results strongly support our modelling approach. Pricing results are smooth without evidence of an over-fitting problem and the derived implied distributions are realistic. Also, results for the pricing of American call options indicate that the non-Markovian model outperforms the equivalent Markovian model and also an ad-hoc procedure of smoothing Black-Scholes implied volatilities.
4.1. Introduction

In the literature, there are two approaches for developing an option pricing model: the traditional and the smile consistent. In the traditional approach, first we assume the stochastic process of the underlying asset, and then no-arbitrage arguments are used in order to derive the option pricing model. In the smile-consistent approach, the implied distribution of the underlying asset is extracted using liquid European options\textsuperscript{57,58}. Also, in the first chapter, we propose a non-recombining implied tree (INRT) without imposing any restrictive assumptions for the underlying stochastic process. Effectively, the INRT is a non-parametric model. Because of the many degrees of freedom, the INRT can allow for flexible underlying asset distribution. All of the above option pricing models are constrained to Markovian stochastic processes. This means that they do not take account of possible dependence between future and past returns.

However, in the literature there is significant evidence for the predictability of financial assets returns. For example, Fama and French (1988) report negative serial correlation in market returns over observation intervals of three to five years, and Lo and MacKinley (1988) report positive serial correlation in weekly returns. Jegadeesh (1990) finds a highly significant negative first order serial correlation in monthly stock returns and significant positive serial correlation at longer lags, and a particularly strong twelve-month serial correlation. His results suggest that the extent to which security returns can be predicted based on past returns is economically significant and can be attributed to either market inefficiency or to systematic changes in expected stock returns. Also, Jašić and Wood (2006) find evidence of significant autocorrelation in daily returns of FTSE 100 index.


\textsuperscript{58} There also exist non-parametric methods, like Stutzer (1996) who uses the maximum entropy concept to derive the risk neutral distribution from the historical distribution of the asset price and Ait-Sahalia and Lo (1998) who propose a non-parametric estimation procedure for state-price densities using observed option prices.
Lo and Wang (1995) investigate the impact of asset return predictability in the price of an option. They argue that, even though possible predictability of stock returns would be induced by the drift of the underlying process, the drift does not enter the standard option pricing formulas. As a result, even though the market option prices in reality are affected by predictability, standard option pricing formulas do not take account of it. In their study they make an adjustment to the Black-Scholes formula (BS, 1973) that takes into account the predictability of underlying asset’s returns. They find that even small levels of predictability can be important especially for longer maturity options. They also provide several continuous-time linear diffusion processes that can capture different forms of predictability.

In this chapter, we make an extension of Lo and Wang (1995) to smile-consistent option pricing models. To the best of our knowledge, this is the first attempt in the literature of imposing non-Markovian assumptions on the process followed by the underlying asset on an implied tree. We create an implied model built on a non-recombining tree using the methodological framework proposed in the first chapter and impose non-Markovian assumptions on it (NMT model). The non-Markovian assumptions allow us to have a framework with serial dependence. Specifically, we allow the transition probabilities of the non-recombining tree to take into account the serial correlation of the underlying asset’s returns between two consecutive time steps. Therefore, unlike in typical preference-free option pricing models, a parameter related to the expected return of the underlying asset appears in our model. This stands in sharp contrast to the typical preference-free framework where the expected asset return is redundant in the option formula. Our model is built on a non-recombining tree which allows path dependency (and specifically serial dependence) of realized paths\(^{59}\). Thus, like the INRT model, it allows for a very flexible underlying asset distribution, but in addition, because of its non-Markovian nature, it gives a richer and more realistic model than the other implied trees and continuous option pricing models in the literature.

\(^{59}\) We implement a model with serial correlation, but other form of dependence can be implemented.
In the proposed model, we minimize the discrepancy between the observed market option prices and the model values with respect to the underlying asset at each node, subject to constraints that keep the probabilities well specified and also prevent some of the most standard arbitrage opportunities, since in a market with predictable returns arbitrage opportunities cannot be completely excluded. Effectively, the problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure and use nonlinear constrained optimization to minimize the least squares error function on market prices. Specifically, we adopt an exterior penalty method and for the optimization we use a Quasi-Newton algorithm. Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem.

We calibrate different versions (in terms of the values of correlation) of the proposed NMT model using European calls on the FTSE 100 index for the year 2003. The results strongly support our modelling approach. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Furthermore, we compare the pricing performance of the different versions of the NMT model versus the INRT model, which is its corresponding Markovian model, for pricing American call options and also with the ad-hoc method for smoothing BS implied volatilities proposed by Dumas et al. (1998). Again, results support our modelling approach. Overall, the NMT model, calibrated using serial correlation -5% seems to outperform all models. A formal statistical test indicates that the distribution of the absolute errors obtained from this model is statistically the same at 3% level of significance, with the distributions obtained from the other NMT models but statistically different from the distribution obtained by the INRT and the ad-hoc model. Also, consistent with Lo and Wang (1995), we find that predictability adds value to the call options. An interpretation for this can be the fact that when we allow for serial correlation of returns is like adding another source of randomness (in addition to the instantaneous volatility of returns).
The chapter continuous as follows: In section 4.2 we describe the proposed model, in section 4.3 we describe the initialization of the model and in section 4.4 we describe the analytical formulation of the model. In section 4.5 we describe the dataset used and in section 4.6 we calibrate and test the non-Markovian model. In section 4.7 we test the performance of the model in pricing American call options. Section 4.8 concludes. In Appendix 4.1 we prove the feasibility of the tree, and in Appendices 4.2A and 4.2B we give the derivation of some partial derivatives.

4.2. Description of the model

Assume that the behaviour of the underlying asset is described by a non-recombining tree. Figure 4.1 shows a non-recombining tree with 4 steps ($n = 5$).

![Non-recombining tree with 4 steps.](image)

Figure 4.1: Non-recombining tree with 4 steps.

We adopt the methodological framework for implied non-recombining trees proposed in the first chapter and impose non-Markovian assumptions on it. Figures 4.2a and 4.2b show the neighbour-nodes of the node $(i, j)$ for $j$ even and $j$ odd, respectively on a non-recombining tree. The point $(i, j)$ on the tree denotes:

$i$: the time dimension, $i = 1, \ldots, n$.

$j$: the asset (time specific) dimension, $j = 1, \ldots, 2^{i-1}$.
$S_{i,j}$ is the value of the underlying asset at node $(i, j)$.

Figure 4.2a: Neighbour-nodes of the node $(i, j)$ for $j$ even on a non-recombining tree.

Figure 4.2b: Neighbour-nodes of the node $(i, j)$ for $j$ odd on a non-recombining tree.

We denote with $r_{i,j}$ for $i = 2, ..., n$, $j = 1, ..., 2^{i-1}$ the model return at point $(i, j)$ for a period $\Delta t$. For $j$ even, $r_{i,j}$ is the model return from node $(i-1, j/2)$ to node $(i, j)$, and for $j$ odd, $r_{i,j}$ is the model return from node $(i-1, (j+1)/2)$ to node $(i, j)$. Return $r_{i,j}$ is computed by the following formula:
\[ r_{i,j} = \frac{S_{i,j}}{\alpha_{i,j}} - 1, \quad i = 2, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \] (4.1)

where \( \alpha_{i,j} \) is defined as:

\[
\alpha_{i,j} = \begin{cases} 
S_{i-1,\frac{j}{2}} & i = 2, \ldots, n, \quad j = 2, 4, \ldots, 2^{i-1} \\
S_{i-1,\frac{j+1}{2}} & i = 2, \ldots, n, \quad j = 1, 3, \ldots, 2^{i-1} - 1 
\end{cases}
\] (4.2)

Return \( r_{i,j} \) is not observed on the tree. In the implementation we use the actual realized return corresponding to a time period \( \Delta t \) back.

The expected value of the asset at point \((i, j)\) one time step \( \Delta t \) ahead is given by:

\[
S_{i,j}(1 + E(r_{i+1,j})) = S_{i+1,2j}p_{i,j} + S_{i+1,2j-1}(1 - p_{i,j})
\] (4.3)

where, \( E(r_{i+1,j}) \) is the discrete expected return one step ahead, i.e. for the period \( \Delta t \), starting from point \((i, j)\). Hence, transition probabilities \( p_{i,j} \) are given by the following formula:

\[
p_{i,j} = \frac{S_{i,j}(1 + E(r_{i+1,j})) - S_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}}, \quad i = 1, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1}
\] (4.4)

In a risk neutral world, the expected return \( E(r_{i+1,j}) \) equals the risk-free rate. In this case, equation (4.4) is a martingale restriction and the asset process is Markovian. This implies that the correlation of the expected return \( E(r_{i+1,j}) \) with the \( \Delta t \)-period return at point \((i, j)\), \( r_{i,j} \) is zero.
If we assume that there is dependence between \( E(r_{i+1,j}) \) and \( r_{i,j} \), then the stochastic process of the asset is no longer Markovian. In this chapter, we assume that the expected return \( E(r_{i+1,j}) \) is given by the following formula:

\[
E(r_{i+1,j}) = \rho r_{i,j} \quad i = 1, ..., n - 1, \quad j = 1, ..., 2^{i-1}
\]

where, \( \rho \) is the first order serial correlation (or autocorrelation) between the returns \( r_i \) and \( r_{i+1} \). In the implementation, an estimate for the correlation is obtained from real (historical) data\(^6\).

Thus, the general formula for the computation of \( E(r_{i+1,j}) \) is given by:

\[
E(r_{i+1,j}) = \begin{cases} 
\rho r_{i,j} & \rho \neq 0 \\
r_f & \rho = 0 
\end{cases} \quad i = 1, ..., n - 1, \quad j = 1, ..., 2^{i-1}
\]

(4.6)

Therefore, when we impose serial correlation between two consecutive time steps, probability formula (4.4) becomes:

\[
p_{i,j} = \frac{S_{i,j} (1 + \rho r_{i,j}) - S_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} \quad i = 1, ..., n - 1, \quad j = 1, ..., 2^{i-1}
\]

(4.7)

Like in the INRT model the call option values at the last time step \( C_{n,j} \) for \( j = 1, ..., 2^{i-1} \) are approximated by the following smoothing approximation:

\(^6\)For simplicity we assume that returns follow an AR(1) process and \( \rho \) is the coefficient of the first lag variable. \( E(r_{i+1,j}) \) can be approximated by any other autoregressive model.
\[
\frac{C_{a}(n, j)}{K} = \begin{cases} 
0 & \text{for } S_{n,j} / K \leq 1 - z/2 \\
\frac{S_{n,j}}{K} - 1 & \text{for } S_{n,j} / K \geq 1 + z/2 \\
\frac{1}{2z} \left[ \left( \frac{S_{n,j}}{K} - 1 \right) + \frac{z}{2} \right]^2 & \text{for } 1 - z/2 < S_{n,j} / K < 1 + z/2 
\end{cases} 
\]

\(j = 1, \ldots, 2^{n-1}\)

where \(K\) is the option exercise price and \(z\) is a small positive constant, for example 0.01 (see Fig. 4.3).

![Figure 4.3: Smoothing of the option pay-off function at maturity.](image)

The value of the call at intermediate nodes is given by the following equation:

\[
C_{i,j} = \left( p_{i,j}C_{i+1,j} + (1 - p_{i,j})C_{i+1,j-1} \right) e^{-r_j \Delta t} 
\]

\(i = n - 1, \ldots, 1, \quad j = 1, \ldots, 2^{i-1}\)
where \( r_f \) denotes the annually continuously compounded riskless rate of interest. Lo and Wang (1995), in order to take account the predictability of stock returns, use a BS type formula with a volatility input adjusted to account for the serial correlation. For discounting they use the risk free rate. Like in their work, in our model the “local volatility” is affected by the parameter values of serial dependence through the probability formula. Similar to them, we use the risk-free rate for discounting\(^{61}\).

**Constraints**

In order to obtain a feasible probability distribution of the underlying asset, there should be imposed some constraints for the probabilities to be well specified and also for the prevention of some of the most standard arbitrage opportunities. Transition probabilities are well specified when they take values between zero and one. This implies the following constraints:

\[
S_{i,j}(1 + \rho_{i,j}) < S_{i+1,j} \quad (4.9a)
\]

\[
i = 1, \ldots, n-1 \quad , \quad j = 1, \ldots, 2^{j-1}
\]

\[
S_{i,j}(1 + \rho_{i,j}) > S_{i+1,j-1} \quad (4.9b)
\]

Well specification constraints in the non-recombining tree prevent nodes \(2j - 1\) and \(2j\) to cross, for \(i=1,\ldots,n\) and \(j=1,\ldots,2^{i-1}\) (see Fig. 4.2a-4.2b).

To avoid arbitrage opportunities, we include the no-arbitrage constraints. Specifically, a European call with dividends should lie between the following bounds:

\[
\max(S_{1,1}e^{-\delta T} - Ke^{-r_{T}} , 0) < C_{mod} < S_{1,1} \quad (4.10)
\]

\(^{61}\) Since we are not in a risk neutral world, the correct discounting rate to use would be a risk factor determined by the sentiment of investors concerning future price uncertainties and their relationship with a market portfolio (Liu et al., 2005). For that, we make a robustness test using 8\% as a discount factor. Results were in general less accurate than when the risk-free rate (around 4\%) was used for discounting. However, results for autocorrelations less than and including -5\%, were not significantly different.
where $C_{\text{Mod}}$ is the call option value estimated by the model.

Also, every value of the underlying asset on the tree should be greater than zero. Thus, we also impose the following constraint:

$$S_{i,j} > 0, \ i = 2, ..., n, j = 1, ..., 2^{i-1} \quad (4.11)$$

Note that, in a market with serial dependence arbitrage opportunities cannot be completely excluded (see for example Rodgers, 1997). Thus the above no-arbitrage constraints cannot necessarily preclude all possible arbitrage opportunities.

The optimization problem

The objective of the problem is to minimize the least squares error function of the discrepancy between the observed market prices and the values produced by the model. Thus, we have the following optimization problem:

$$\min_{\mathbf{x}} \frac{1}{2} \sum_{k=1}^{N} \left( C_{\text{Mod}}(x,k) - C_{\text{Mkt}}(k) \right)^2 \quad (4.12)$$

where $C_{\text{Mod}}(k)$ and $C_{\text{Mkt}}(k)$ denote the model and market price respectively of the $k^{th}$ call, $k = 1, ..., N$, subject to the constraints:

$$g_1(i,j) = S_{i,j}(1 + \rho r_{i,j}) - S_{i+1,j-1} > 0 \ i = 1, ..., n-1, \ j = 1, ..., 2^{i-1} \quad (4.13a)$$

$$g_2(i,j) = S_{i+1,j} - S_{i,j}(1 + \rho r_{i,j}) > 0 \ i = 1, ..., n-1, \ j = 1, ..., 2^{i-1} \quad (4.13b)$$

$$g_3(i,j) = S_{i,j} > 0 \ i = 2, ..., n, \ j = 1, ..., 2^{i-1} \quad (4.13c)$$

$$g_4(k) = S_{1,1} - C_{1,1}(k) > 0 \ k = 1, ..., N \quad (4.13d)$$

$$g_5(k) = C_{1,1}(k) - \max(S_{1,1}e^{-\delta T} - K(k)e^{-rT},0) > 0 \ k = 1, ..., N \quad (4.13e)$$
Since the problem under consideration is a non-convex optimization problem with linear constraints we adopt an exterior penalty method (Fiacco and McCormick, 1968) to convert the nonlinear constrained problem into a nonlinear unconstrained problem. The Exterior Penalty Objective function that we use is the following:

\[
P(x, \alpha) = \frac{1}{2} \sum_{k=1}^{N} \left( C_{\text{Mod}}(x, k) - C_{\text{Mbt}}(k) \right)^2 + \alpha \sum_{i=1}^{n-1} \sum_{j=1}^{2^{n-1}} \left( \left[ \min(g_1(i, j), 0) \right]^2 + \left[ \min(g_2(i, j), 0) \right]^2 + \left[ \min(g_3(i, j), 0) \right]^2 \right) + \alpha \sum_{k=1}^{N} \left( \left[ \min(g_4(k)) \right]^2 + \left[ \min(g_5(k)) \right]^2 \right)
\]

The second and third terms in \( P(x, \alpha) \) give a positive contribution if and only if \( x \) is infeasible. Under mild conditions it can be proved that minimizing the above penalty function for strictly increasing sequence \( \alpha \) tending to infinity the optimum point \( x(\alpha) \) of \( P \) tends to \( x^* \), a solution of the constrained problem.

For the optimization we use a Quasi-Newton algorithm. Specifically we use the BFGS formula\(^{62}\) (Fletcher, 1987). For the procedure of Line Search in the algorithm we use the Charalambous (1992) method.

4.3. Initialization of the model

We denote with \( u_{i,j} \) and \( d_{i,j} \) the up and down factors by which the underlying asset price can move in the single time step, \( \Delta t \), given that we are at node \((i, j)\). \( \Delta t \), \( u_{i,j} \) and \( d_{i,j} \) factors are given by the following formulas:

\[
\Delta t = \frac{T}{n-1}
\]

\(^{62}\) The BFGS formula was discovered in 1970 independently by Broyden, Fletcher, Goldfarb and Shanno.
\[ u_{i,j} = e^{\sigma_i \sqrt{\Delta t}} \quad (4.16a) \]

\[ d_{i,j} = e^{-\sigma_i \sqrt{\Delta t}} = \frac{1}{u_{i,j}} \quad (4.16b) \]

where \( T \) is the option’s time to maturity and \( \sigma_i \) is the volatility term structure at time step \( i \).

\( S_{1,1} \) is the current value of the underlying asset. The odd nodes of the tree \( S_{i,j} \), are initialized using the following equation:

\[ S_{i,j} = S \frac{j+1}{2} \frac{i+1}{2} \quad i = 2, \ldots, n, \quad j = 1, 3, \ldots, 2^{i-1} - 1 \quad (4.17a) \]

The even nodes of the tree \( S_{i,j} \), are initialized using the following equation:

\[ S_{i,j} = S \frac{j}{2} \frac{i}{2} \quad i = 2, \ldots, n, \quad j = 2, 4, \ldots, 2^{i-1} \quad (4.17b) \]

Like in the INRT model we use the following volatility term structure to initialize the tree:

\[ \sigma_i = \sigma_1 e^{2(i-1)\lambda t}, \quad \lambda \in R, \quad i = 1, \ldots, n-1 \quad (4.18) \]

where \( \lambda \) is a constant parameter and \( \sigma_i \) is a properly chosen initial value for the volatility. If \( \lambda \) is positive, then volatility increases as we approach maturity and if \( \lambda \) is negative, then volatility decreases as we approach maturity\(^63\).

\(^63\) Other non-monotonic functions could also be used for \( \sigma_i \) but what we have tried proved adequate for our purposes.
Equations (4.15) to (4.18) are used only for initialization. Once the optimization process starts, each value of the underlying asset (except from $S_{i,1}$) acts as an independent variable in the system.

In order to keep the probabilities well specified at every time step and hence obtain a feasible initial tree, $\lambda$ should belong in the following interval:

$$
\lambda \in \left[ \frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right), +\infty \right)
$$

(4.19)

where $\xi_m$ is given by:

$$
\xi_m = \frac{1}{\sqrt{\Delta t}} \min_{i,j} \left\{ \log (1 + \rho r_{i,j}) \right\}
$$

(4.20)

$$
i = 1, \ldots, n, \quad j = 1, \ldots, 2^{l-1}
$$

Since $\xi_m$ is a function of $S_{i,j}$ we cannot know a value of $\xi_m$ before building the tree. Thus, for $\xi_m$ we use an arbitrary number close to zero for example 1.E-8. By choosing $\lambda$ from the above interval, we allow the initial volatility to increase or decrease across time. We make several consecutive draws from interval (4.19) until we find the value of $\lambda$ that gives the “optimal” tree\(^{64}\).

4.4. Analytical formulation of the problem

For the implementation of the optimization method, we need to calculate the partial derivatives of $C_{\text{Mod}}(k)$ \(^ {65}\) with respect to the value of the underlying asset at

---

\(^{64}\) Optimal tree is the one that gives the lowest-value objective function subject to the initial constraints.

\(^{65}\) From now on we will use $C_{1,1}$ instead of $C_{\text{Mod}}$. 

Eleni D. Constantinide
each node, for \( k = 1, ..., N \) i.e. we want to find \( \frac{\partial C(1,1,k)}{\partial S_{i,j}}, \quad i = 2, ..., n, \quad j = 1, ..., 2^{i-1} \) and \( k = 1, ..., N \). For notational simplicity in the following, we assume that we have only one call option. For the computation of \( \frac{\partial C_{i,1}}{\partial S_{i,j}}, \forall i, j \) we implement the following steps:

We define the quadruplet vectors (see Fig.4.2a-4.2b):

\[
S_{i,j}^{(1)} = [S_{i+1,j/2}, S_{i,j}, S_{i+1,2j}, S_{i+1,2j-1}^2] \quad i = 2, ..., n, \quad j = 2, 4, ..., 2^{i-1} \tag{4.21a}
\]

\[
S_{i,j}^{(2)} = [S_{i+1,(j+1)/2}, S_{i,j}, S_{i+1,2j}, S_{i+1,2j-1}] \quad i = 2, ..., n, \quad j = 1, 3, ..., 2^{i-1} - 1 \tag{4.21b}
\]

1\textsuperscript{st} step: Compute the partial derivatives of the transition probabilities, \( \frac{\partial p_{i,j}}{\partial a_{i,j}}, \frac{\partial p_{i,j}}{\partial s_{i,j}}, \) for \( i = 2, ..., n-1, \quad j = 1, ..., 2^{i-1} \) and \( \frac{\partial p_{i,j}}{\partial s_{i+1,2j}}, \frac{\partial p_{i,j}}{\partial s_{i+1,2j-1}} \) for \( i = 1, ..., n-1, \quad j = 1, ..., 2^{i-1} \).

We summarize the derivatives in vector forms (4.22a) and (4.22b).

\[
\begin{bmatrix}
\frac{\partial p_{i,j}}{\partial a_{i,j}} \\
\frac{\partial p_{i,j}}{\partial s_{i,j}} \\
\frac{\partial p_{i,j}}{\partial s_{i+1,2j}} \\
\frac{\partial p_{i,j}}{\partial s_{i+1,2j-1}}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{S_{i+1,2j} - S_{i+1,2j-1}^2} \\
\frac{S_{i,j}^2}{1 + 2 \frac{\rho}{a_{i,j}} S_{i,j} - \rho} \\
\frac{1}{S_{i+1,2j} - S_{i+1,2j-1}^2} \\
\frac{1}{S_{i+1,2j} - S_{i+1,2j-1}^2}
\end{bmatrix} \quad i = 2, ..., n-1, \quad j = 1, ..., 2^{i-1} \tag{4.22a}
\]

\[
\begin{bmatrix}
\frac{\partial p_{i,j}}{\partial s_{i+1,2j}} \\
\frac{\partial p_{i,j}}{\partial s_{i+1,2j-1}}
\end{bmatrix} = \begin{bmatrix}
\frac{1}{S_{i+1,2j} - S_{i+1,2j-1}^2} \left[ \frac{-p_{i,j}}{1 - p_{i,j}} \right] \\
\frac{1}{S_{i+1,2j} - S_{i+1,2j-1}^2} \left[ \frac{-p_{i,j}}{1 - p_{i,j}} \right]
\end{bmatrix} \quad i = 1, ..., n-1, \quad j = 1, ..., 2^{i-1} \tag{4.22b}
\]

\[\text{66} \quad \text{We do not calculate} \quad \frac{\partial C(1,1,k)}{\partial S_{i,1}} \text{since} \quad S_{i,1} \text{is a known, fixed parameter, and thus does not take part in the optimization.}\]
2\textsuperscript{nd} step: Compute the partial derivatives \( \frac{\partial C_{i,j}}{\partial \alpha_{i,j}} \), \( \frac{\partial C_{i,j}}{\partial S_{i,j}} \), for \( i = 2, \ldots, n-1 \), \( j = 1, \ldots, 2^{i-1} \), and \( \frac{\partial C_{i,j}}{\partial S_{i+1,2j}} \), \( \frac{\partial C_{i,j}}{\partial S_{i+1,2j-1}} \) for \( i = 1, \ldots, n-1 \), \( j = 1, \ldots, 2^{i-1} \).

Derivatives are given by the following formulas.

\[
\frac{\partial C_{i,j}}{\partial \alpha_{i,j}} = \left( C_{i+1,2j} - C_{i+1,2j-1} \right) \frac{\partial p_{i,j}}{\partial \alpha_{i,j}} e^{-r_j \Delta} \quad i = 2, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \tag{4.23a}
\]

\[
\frac{\partial C_{i,j}}{\partial S_{i,j}} = \left( C_{i+1,2j} - C_{i+1,2j-1} \right) \frac{\partial p_{i,j}}{\partial S_{i,j}} + p_{i,j} \left( \frac{\partial C_{i+1,2j}}{\partial S_{i,j}} - \frac{\partial C_{i+1,2j-1}}{\partial S_{i,j}} \right) + \frac{\partial C_{i+1,2j-1}}{\partial S_{i,j}} \right) e^{-r_j \Delta} = \Delta_{i,j}
\]

\( i = 2, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \) \tag{4.23b}

\[
\begin{bmatrix}
\frac{\partial C_{i,j}}{\partial S_{i+1,2j}} \\
\frac{\partial C_{i,j}}{\partial S_{i+1,2j-1}} \\
\frac{\partial C_{i,j}}{\partial S_{i+1,2j}}
\end{bmatrix}
= \begin{bmatrix}
p_{i,j} \left( \Delta_{i+1,2j} - D_{i,j} \right) e^{-r_j \Delta} \\
\left( 1 - p_{i,j} \right) \left( \Delta_{i+1,2j-1} - D_{i,j} \right) e^{-r_j \Delta}
\end{bmatrix}
\quad i = 1, \ldots, n-1, \quad j = 1, \ldots, 2^{i-1} \tag{4.23c}
\]

where,

\[
D_{i,j} = \frac{C_{i+1,2j} - C_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} \quad \text{and} \quad \Delta_{i,j} = \frac{\partial C_{i,j}}{\partial S_{i,j}}, \text{Delta ratio}
\]
3rd step: Compute the partial derivatives \( \frac{\partial C_{a}(n, j)}{\partial S_{n,j}} \) for \( j = 1, \ldots, 2^{n-1} \). They are given by the following formula:

\[
\frac{\partial C_{a}(n, j)}{\partial S_{n,j}} = \begin{cases} 
0 & \text{for } S_{n,j} \leq K(1 - z/2) \\
1 & \text{for } S_{n,j} \geq K(1 + z/2) \\
\frac{1}{z}\left[\left(\frac{S_{n,j}}{K} - 1\right) + \frac{z}{2}\right] & \text{for } K(1 - z/2) < S_{n,j} < K(1 + z/2)
\end{cases}
\]  

(4.24)

4th step: Compute the partial derivatives \( \frac{\partial C_{1,1}}{\partial S_{i,j}} \) for \( i \geq 3 \).

\[
\frac{\partial C_{1,1}}{\partial S_{i,j}} = \prod \{\text{of the probabilities on the path that take us from node (1,1) to node (i-1,k)}\} \times \frac{\partial C_{i-1,k}}{\partial S_{i,j}} e^{-(i-2)r\Delta t}
\]

(4.25)

\[
k = \begin{cases} 
j/2 & \text{for even } j \\
(j+1)/2 & \text{for odd } j
\end{cases}
\]

4.5. Data - FTSE 100 index options

We use the daily closing prices of FTSE 100 American and European call options of January 2003 to December 2003 as reported by LIFFE.\(^67\) The strike prices

\(^67\) FTSE 100 European options are traded with expiries in March, June, September, and December. Additional contracts are introduced serially so that the nearest 4 months are always available for trading. FTSE 100 American options are traded with expiries the nearest of June and December. Additional contracts are introduced serially so that the nearest 3 months are always available for trading. FTSE 100 options expire on the third Friday of the expiry month. FTSE 100 options positions are marked-to-market daily based on the daily settlement price, which is determined by LIFFE and
for a given style call (European or American) are spaced at intervals of 50 index points from each other. However, there are no American and European calls with identical strikes. The strike prices for adjacent European and American style calls are spaced at intervals of 25 index points. For example, there are European style calls with strikes 3075, 3125, and 3175 and American style calls with strikes 3100, 3150, and 3200. Also, the longest maturity for the European calls is 2 years while for the American calls is 6 months. Our initial sample (for the 12 months period) consists of 99,051 observations of European calls, and 34,503 observations of American calls. We apply five standard filtering rules to both American and European calls data. First we exclude calls that violate the no-arbitrage bounds. Second, we eliminate calls with time to maturity less than 6 calendar days, i.e. $T < 6$. These options have very small time-premiums and their implied volatilities are inaccurate since they are very sensitive to market microstructure problems and measurement errors (Hentschel, 2001, Brandt, et al., 2002). Third, we eliminate calls if their closing price is less than 0.5 index points. Fourth, we eliminate calls for which the traded volume is zero (since we want highly liquid options for calibration). Finally, we eliminate calls with moneyness greater than 1.1 or moneyness less than 0.8, since deep in the money and deep out of the money call options are expected to be illiquid and not accurately priced. We define as moneyness of a call option the ratio (underlying asset price)/(strike price). After these filtering rules, our sample consists of 13,696 observations of European call options and 355 observations of American call options.

For the implementation, we consider only cases for which the number of options used for calibration $N$ is greater than 8 since with fewer options the distribution obtained of the underlying asset will not be reliable. In the literature there have been reported microstructure and thin trading problems in measuring autocorrelations in small intervals (see Lo and MacKinlay, 1988, Jegadeesh, 1989). Thus, in addition we eliminate options with $\Delta t \leq 3$ trading days. Also we remove confirmed by the Clearing House. FTSE 100 options are quoted in index points and have an assigned value of £10 per index point.

$68$ For European (EU) style options $\max(0, S e^{-\delta T} - K e^{-r_f T}) \leq C_{EU} \leq S$ and for American (AM) style options $\max(C_{EU}, S - K) \leq C_{AM} \leq S$. 

confirmed by the Clearing House. FTSE 100 options are quoted in index points and have an assigned value of £10 per index point.
from the sample the calls that trade for the first time because there is no information about the return of the underlying asset one “time-step” back. For the pricing of American options we use only the European options that correspond to the American options (longest maturity 6 months). Thus, the European sample in this case is reduced to 1,579 observations and the American sample to 173.

For time to maturity, \( T \), we use the calendar days to maturity. For the risk-free rate, \( r_f \), we use cubic spline interpolation for matching each option contract with a continuous interest rate that corresponds to the option’s maturity, by utilizing the 1-month to 12-month LIBOR offer rates, collected from Datastream. Also, since the underlying asset of the options on FTSE 100 is a futures contract, we make the standard assumption that the dividend yield \( (\delta) \) equals the risk free rate. The models are calibrated every day. For each implementation, the options used have the same underlying asset and the same time to maturity \(^{69}\).

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>86.00</td>
<td>40.40</td>
<td>134.29</td>
<td>238.55</td>
<td>98.68</td>
<td>84.11</td>
<td>88.63</td>
</tr>
<tr>
<td>median</td>
<td>50</td>
<td>26</td>
<td>133</td>
<td>229.5</td>
<td>55.75</td>
<td>48.75</td>
<td>51</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>58.5</td>
<td>106</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>410.5</td>
<td>263.5</td>
<td>230</td>
<td>410.5</td>
<td>396.5</td>
<td>400.5</td>
<td>410.5</td>
</tr>
<tr>
<td>observations</td>
<td>1,579</td>
<td>1,133</td>
<td>157</td>
<td>289</td>
<td>76</td>
<td>1,086</td>
<td>417</td>
</tr>
</tbody>
</table>

Table 4.1a: Statistics of the FTSE 100 European (EU) style call options for the year 2003. Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

\(^{69}\) In the specific dataset (given by LIFFE), the underlying asset of each call option is a future contract and thus every trading day the options have different underlying asset.
Table 4.1b: Statistics of the FTSE 100 American (AM) style call options for the year 2003. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>76.77</td>
<td>50.17</td>
<td>128.42</td>
<td>229.40</td>
<td>73.95</td>
<td>81.99</td>
<td>64.65</td>
</tr>
<tr>
<td>median</td>
<td>51.5</td>
<td>38</td>
<td>131</td>
<td>230.25</td>
<td>59</td>
<td>55</td>
<td>44</td>
</tr>
<tr>
<td>min</td>
<td>0.5</td>
<td>0.5</td>
<td>78.5</td>
<td>140</td>
<td>5</td>
<td>1.5</td>
<td>0.5</td>
</tr>
<tr>
<td>max</td>
<td>318</td>
<td>191</td>
<td>182</td>
<td>318</td>
<td>229.5</td>
<td>318</td>
<td>262.5</td>
</tr>
<tr>
<td>observations</td>
<td>173</td>
<td>140</td>
<td>13</td>
<td>20</td>
<td>11</td>
<td>115</td>
<td>47</td>
</tr>
</tbody>
</table>

Table 4.2a: Statistics of the trading volumes of FTSE 100 European (EU) style call options for the year 2003. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>EU OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>324.18</td>
<td>357.67</td>
<td>465.80</td>
<td>115.98</td>
<td>447.20</td>
<td>332.66</td>
<td>279.70</td>
</tr>
<tr>
<td>median</td>
<td>50</td>
<td>56</td>
<td>111</td>
<td>14</td>
<td>123.5</td>
<td>52.5</td>
<td>37</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>7,683</td>
<td>7,683</td>
<td>4,837</td>
<td>1,771</td>
<td>5,441</td>
<td>7,683</td>
<td>3,021</td>
</tr>
<tr>
<td>observations</td>
<td>1,579</td>
<td>1,133</td>
<td>157</td>
<td>289</td>
<td>76</td>
<td>1,086</td>
<td>417</td>
</tr>
</tbody>
</table>

Table 4.2b: Statistics of the trading volumes of FTSE 100 American (AM) style call options for the year 2003. Out-of-the-money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.

<table>
<thead>
<tr>
<th>AM OPTIONS</th>
<th>All</th>
<th>OTM</th>
<th>ATM</th>
<th>ITM</th>
<th>ST</th>
<th>MT</th>
<th>LT</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>75.25</td>
<td>81.09</td>
<td>44.31</td>
<td>54.45</td>
<td>25.91</td>
<td>32.38</td>
<td>191.68</td>
</tr>
<tr>
<td>median</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>4</td>
<td>10</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>min</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>max</td>
<td>3,000</td>
<td>3,000</td>
<td>340</td>
<td>350</td>
<td>111</td>
<td>350</td>
<td>3,000</td>
</tr>
<tr>
<td>observations</td>
<td>173</td>
<td>140</td>
<td>13</td>
<td>20</td>
<td>11</td>
<td>115</td>
<td>47</td>
</tr>
</tbody>
</table>

Tables 4.1a, 4.1b, 4.2a, and 4.2b describe the cleaned sample. Table 4.1a shows the mean, median, minimum and maximum call option value and also the number of
observations of the whole European call option sample, and also of the sub-samples of the out-of-the money (OTM), at-the-money (ATM), in-the-money (ITM), short term (SM), medium term (MT) and long term (LT) options\textsuperscript{70}. Table 4.1b shows the same statistics as Table 4.1a for the American calls sample. Tables 4.2a and 4.2b show the same statistics for the trading volume of the European and American calls. Overall we see that the European calls are highly traded with mean and median volume 324.18 and 50 respectively. American calls are less liquid with mean and median volume 75.25 and 7 respectively.

4.6. Calibration and testing of the Non-Markovian (NMT) model

Tests with historical data indicate that there is statistically significant serial autocorrelation of the 4 to 20 trading-day returns of the FTSE 100 index. The correlation is in the range of -5% to -10%. We test the non-Markovian model (NMT) using serial correlations per interval $\Delta t$ -2.5% $(NMT(\rho = -2.5\%))$, -5% $(NMT(\rho = -5\%))$, -7.5% $(NMT(\rho = -7.5\%))$ and -10% $(NMT(\rho = -10\%))$. Every trading day we calibrate the four versions of the NMT and the INRT model (which is the corresponding Markovian model) using European calls data with the same underlying asset and time to maturity. Then, we check the models for over-fitting by pricing options with strikes in-between those used for the optimization (calibration). Figures 4.4a-4.4e show plots of the call prices (market prices and estimated from the model) versus moneyness. For brevity we present only the plots for the first trading day of June for each of the five models. Similar results are obtained for the other trading days and months. The INRT model represents the case with 0% correlation. It is included for comparison purposes. As we see, for all the models the estimated call values increase smoothly with increasing moneyness without any evidence of over-fitting.

\textsuperscript{70} Out-of-the money (OTM) are options with moneyness less than 0.99, at-the-money (ATM) are options with moneyness between and including 0.99 and 1.01, in-the-money (ITM) are options with moneyness greater than 1.01. Short term (ST) are options with maturity less than 30 calendar days, medium term (MT) are options with maturity between and including 30 and 60 calendar days, and long term (LT) are options with maturity greater than 60 calendar days.
Figure 4.4a: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the implied non-recombining tree (INRT) for \( n = 7 \).

Figure 4.4b: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the non-Markovian model (NMT) with correlation -2.5% and for \( n = 7 \).
Figure 4.4c: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -5% and for n=7.

Figure 4.4d: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -7.5% and for n=7.
Figure 4.4e: Plots of the call prices (market and estimated) for the FTSE 100 index, for the 1st trading day of June 2003 using the Non-Markovian model (NMT) with correlation -10% and for $n=7$.

In order to see how realistic is the distribution obtained from the proposed model for year 2003, we calculate the implied statistics\textsuperscript{71} of the 1-month log-returns obtained from the five models with correlations in the interval [-10%- 0%] and as in the first chapter we compare them with the historical 1-month log-returns for the year 2003 and the years 2002-2004 and 2001-2005. Specifically, for each calibration for which the options maturity was between 28 and 32 calendar days, we calculate the first four moments (mean, variance, skewness and kurtosis). Then, in order to get a feeling for the representative statistics of 1-month log-returns we provide for each of those moments the mean and the median. The statistics for the five models are summarized in Table 4.3. As we would expect, the mean of the implied risk-neutral and implied real distribution of log-returns differs from that of the historical

\textsuperscript{71} Liu et al. (2005) discuss the derivations of historical, and implied real and risk-neutral distributions for the FTSE 100 index. They demonstrate that the needed adjustments to get the implied real variance, skewness and kurtosis from the implied risk-neutral ones are minimal. Thus, knowing that our implied risk-neutral moments (beyond the mean) are very close to the implied real ones, we can then compare them with the historical ones (but without expecting the two distributions to be identical).
distribution\textsuperscript{72}. Also, consistently with Jackwerth and Rubinstein (1996) and Liu et al. (2005) in all cases the implied variance is larger than the historical variance. Implied volatility can be thought of as the market’s expected volatility plus some “volatility risk-premium” for other unknown factors such as hedging costs, the inability for perfect hedging, the uncertainty of future volatility etc. Furthermore, the implied NMT skewness in all models is slightly higher in absolute terms (in the range of -0.63 to -0.66) than its corresponding implied risk neutral (-0.57). Also the mean implied NMT kurtosis is slightly higher (in the range of 3.63 to 3.73) than the mean implied risk neutral (3.47) one. Nevertheless, both the implied NMT and risk neutral skewness and kurtosis are consistent with the generally observed historical ones (negative skewness and excess kurtosis). This is an indication that the implied distributions obtained are realistic.

\textsuperscript{72} Because of accumulated rounding errors the estimated implied risk-neutral return is not equal to zero.
<table>
<thead>
<tr>
<th>Correlation</th>
<th>Implied (n=7)</th>
<th>Mean</th>
<th>Variance</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>Mean</td>
<td>-0.0022</td>
<td>0.0044</td>
<td>-0.5736</td>
<td>3.4712</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>-0.0014</td>
<td>0.0027</td>
<td>-0.5259</td>
<td>2.9611</td>
<td>41</td>
</tr>
<tr>
<td>-2.50%</td>
<td>Mean</td>
<td>-0.0025</td>
<td>0.0045</td>
<td>-0.6372</td>
<td>3.7316</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>-0.0017</td>
<td>0.0028</td>
<td>-0.5426</td>
<td>2.9333</td>
<td>41</td>
</tr>
<tr>
<td>-5.00%</td>
<td>Mean</td>
<td>-0.0028</td>
<td>0.0046</td>
<td>-0.6404</td>
<td>3.6778</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>-0.0019</td>
<td>0.0029</td>
<td>-0.5473</td>
<td>2.9157</td>
<td>41</td>
</tr>
<tr>
<td>-7.50%</td>
<td>Mean</td>
<td>-0.0031</td>
<td>0.0047</td>
<td>-0.6562</td>
<td>3.6360</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>-0.0019</td>
<td>0.0031</td>
<td>-0.5888</td>
<td>2.9090</td>
<td>41</td>
</tr>
<tr>
<td>-10.00%</td>
<td>Mean</td>
<td>-0.0034</td>
<td>0.0049</td>
<td>-0.6650</td>
<td>3.7151</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>-0.0020</td>
<td>0.0031</td>
<td>-0.5785</td>
<td>2.8996</td>
<td>41</td>
</tr>
<tr>
<td>Historical</td>
<td>Mean</td>
<td>0.0106</td>
<td>0.0014</td>
<td>-0.6572</td>
<td>2.7689</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Variance</td>
<td>0.0018</td>
<td>-1.1177</td>
<td>4.4749</td>
<td>59</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: Implied risk-neutral, implied Non-Markovian and historical statistics for the distribution of the FTSE 100 1-month log-returns.

In order to give further evidence for the implied distributions obtained by our model versus the INRT model, representative implied distributions (histograms) for the 1-month log-returns in June 2003 are shown in Figures 4.5a-4.5e for the five models for \( n = 7 \). Like in the first chapter, in order to make the histograms of the implied distributions we make use of the Pearson system of distributions\(^{73}\) as applied in Matlab\(^{74}\). Using the first four moments of the data it is easy to find in the Pearson system the distribution that matches these moments and to generate a random

\(^{73}\) In the Pearson system there is a family of distributions that includes a unique distribution corresponding to every valid combination of mean, standard deviation, skewness, and kurtosis.

\(^{74}\) Copyright 2005 The MathWorks, Inc.
sample so as to produce a histogram corresponding to the implied distribution. From the figures, it is obvious that the implied distributions have negative skewness and mostly excess kurtosis which is consistent with historical data. These figures are representative of the vast majority of cases. Another interesting observation is that the shapes of the distributions for all models are very similar.

Figure 4.5a: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the implied non-recombining tree (INRT) model for $n=7$. 
Figure 4.5b: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -2.5% and for $n=7$.

Figure 4.5c: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -5% and for $n=7$. 
Figure 4.5d: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -7.5% and for \( n = 7 \).

Figure 4.5e: Implied probability distribution (histogram) obtained for the 1-month log-return of June 2003 using the Non-Markovian (NMT) model with correlation -10% and for \( n = 7 \).
4.7. Testing the models in pricing FTSE 100 American options

In this section we use the four different versions of the NMT model and the INRT model for \( n = 7 \), for pricing American call options. We also make a comparison with the ad-hoc model of smoothing BS implied volatilities across strikes proposed by Dumas et al. (1998). Every trading day in the sample, we calibrate the above models using European call options data with the same underlying asset and time to maturity. Then, we use each model to price the corresponding American style calls. For the ad-hoc model, in order to price the American call options, we build a standard CRR binomial tree using the implied volatility obtained by interpolating the implied volatility surface, obtained form European call options, across strikes, using 2nd order polynomial regression (AH_Regr model). We use CRR trees of \( n = 50 \) and \( n = 51 \), and the call option value is computed as the average of the call option values of the two trees reducing thus the impact of oscillations.

For the comparison of the pricing performance of the six models we use two pricing measures.

(i) The root mean square error (RMSE) which is the square-root of the average squared difference between the model and the market prices. The RMSE measures how well the model fits in a statistical sense with the usual bias versus variance trade-off.

(ii) The mean difference error (MDE) which is the average of the difference between the model and the market prices and reveals systematic biases of the model.
Table 4.4: Pricing errors for the American FTSE 100 index call options using the Non-Markovian model for correlation -10% (\(NMT(\rho = -10\%))\), for correlation -7.5% (\(NMT(\rho = -7.5\%))\), for correlation -5% (\(NMT(\rho = -5\%))\), and for correlation -2.5% (\(NMT(\rho = -2.5\%))\), the implied non-recombining tree (INRT) model for \(n=7\), and the ad-hoc model of Dumas et al. (1998) (AH_Regr). We measure pricing errors by (i) the root mean square error (RMSE), and (ii) the mean difference error (MDE). Results are obtained using the full sample of American call options.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NMT((\rho = -10%))</td>
<td>5.2863</td>
<td>2.0861</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho = -7.5%))</td>
<td>4.2512</td>
<td>0.8039</td>
</tr>
<tr>
<td>Full</td>
<td>NMT((\rho = -5%))</td>
<td>3.9452</td>
<td>-0.2505</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho = -2.5%))</td>
<td>4.1023</td>
<td>-1.1982</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>4.5113</td>
<td>-1.9984</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.3494</td>
<td>-1.7011</td>
</tr>
</tbody>
</table>

Figure 4.6: Correlation versus the root mean square error (RMSE) smoothed using cubic spline.

Table 4.4 reports model comparisons using the above measures for pricing the American call options using the six models. See also Fig. 4.6 for a graphical depiction of the RMSE versus the correlation of the NMT models. Overall, the models \(NMT(\rho = -5\%), \ NMT(\rho = -2.5\%),\) the INRT model, and the ad-hoc model systematically underprice the American calls whereas the models \(NMT(\rho = -10\%),\)
and \( \text{NMT}(\rho = -7.5\%) \) overprice the American calls. The \text{INRT} model makes the highest underpricing whereas the \( \text{NMT}(\rho = -10\%) \) makes the highest overpricing. The MDE error of the \text{INTR} model and the \( \text{NMT}(\rho = -10\%) \) is -1.9984 and 2.0861 respectively. An interesting observation is that as we go away from 0% correlation, the \text{NMT} model assigns higher value to the options. This suggests that predictability adds value, and it is consistent with the theoretical prediction in Lo and Wang (page 97, 1995). An interpretation for this can be the fact that when we allow for serial correlation of returns is like adding another source of randomness (in addition to the instantaneous volatility of returns). Looking at the RMSE error, we see that the \( \text{NMT}(\rho = -5\%) \) has the lowest error 3.9452, followed by the \( \text{NMT}(\rho = -2.5\%) \), \( \text{NMT}(\rho = -7.5\%) \) having RMSE 4.1023 and 4.2512 respectively. The \text{AH_Regr} and \text{INRT} models come next with RMSE 4.3494 and 4.5113 respectively. The \( \text{NMT}(\rho = -10\%) \) model comes last with RMSE 5.2863.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>MADE</th>
<th>MDADE</th>
<th>p-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>( \text{NMT}(\rho = -10%) )</td>
<td>3.3233</td>
<td>1.8483</td>
<td>0.0707</td>
</tr>
<tr>
<td></td>
<td>( \text{NMT}(\rho = -7.5%) )</td>
<td>2.8063</td>
<td>1.8290</td>
<td>0.6422</td>
</tr>
<tr>
<td></td>
<td>( \text{NMT}(\rho = -5%) )</td>
<td>2.7052</td>
<td>1.7377</td>
<td>.......</td>
</tr>
<tr>
<td></td>
<td>( \text{NMT}(\rho = -2.5%) )</td>
<td>2.7823</td>
<td>1.9154</td>
<td>0.1492</td>
</tr>
<tr>
<td></td>
<td>\text{INRT}</td>
<td>3.0881</td>
<td>2.0027</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>\text{AH_Regr}</td>
<td>2.9970</td>
<td>1.9943</td>
<td>0.0289</td>
</tr>
</tbody>
</table>

Table 4.5: Pricing errors for the American FTSE 100 index call options using the Non-Markovian model for correlation -10\% (\( \text{NMT}(\rho = -10\%) \)), for correlation -7.5\% (\( \text{NMT}(\rho = -7.5\%) \)), for correlation -5\% (\( \text{NMT}(\rho = -5\%) \)), and for correlation -2.5\% (\( \text{NMT}(\rho = -2.5\%) \)), the implied non-recombining tree (\text{INRT}) model for \( n=7 \), and the ad-hoc model of Dumas et al. (1998) (\text{AH_Regr}). We measure pricing errors with (i) the mean absolute difference error (MADE) and (ii) the median absolute difference error (MDADE). The last column contains the p-values from the Wilcoxon sign rank test of whether the distribution of the absolute errors obtained from the on average best model (\( \text{NMT}(\rho = -5\%) \)) is statistically different from the distribution of the absolute errors of each of the other models. Results are obtained using the full sample of American call options.
In order to test whether the pricing performance between the different models is statistically different we perform the Wilcoxon sign rank test\(^{75}\). As a measure of pricing performance we use the absolute difference error of each model. We test whether the distribution of the absolute errors obtained from each model is statistically different from the distribution of the absolute errors of the on average best model, which is the model \(NMT(\rho = -5\%)\). Table 4.5 contains the mean absolute difference errors (MADE) and the median absolute difference errors (MDA\(E\)) for the six models. In the last column of Table 4.5 are presented the \(p\)-values of the Wilcoxon sign rank test. We observe that the distributions of the absolute errors of the \(NMT\) models for all correlations are not statistically different from the corresponding distribution of the \(NMT(\rho = -5\%)\) model at 3\% level of significance. However, the distributions of the absolute errors of both the \(INRT\) and the \(AH\_Regr\) models are statistically different from the corresponding distribution of the \(NMT(\rho = -5\%)\) model. Also, when we apply the Wilcoxon test to check whether the distribution of the absolute errors of the \(INRT\) model is statistically different from the corresponding distribution of the \(AH\_Regr\) model, consistent with the findings in the second chapter we find that they are not statistically different at 1\% level of significance (\(p\)-value=0.2925).

In Table 4.6 the RMSE, and MDE of the six models are presented when the sample is divided in the three moneyness categories. With respect to underpricing /overpricing results remain the same as in the overall sample. However, in the ATM and ITM categories, on average the model \(NMT(\rho = -5\%)\) overprices call values. Overall in the OTM category (81\% of the sample), the \(NMT\) models give lower RMSE with the \(NMT(\rho = -7.5\%)\) giving the lowest RMSE 3.1330. The \(AH\_Regr\) model and the \(INRT\) give RMSE 4.0609 and 4.0998 respectively. In the ATM category (7.5\% of the sample) the \(NMT(\rho = -2.5\%)\) has the lowest RMSE 4.1365 followed by the \(AH\_Regr\) model and the \(INRT\) model with RMSE 4.1210 and 4.5119 respectively. In the ITM

\(^{75}\) Wilcoxon sign rank test is a non-parametric alternative to the paired Student’s t-test for the case of two related samples. It performs a paired two sided test of the hypothesis that the difference between the matched samples comes from a distribution with median zero. In other words, it tests whether the two paired-samples have the same distribution.
category (11.5% of the sample) the AH_Regr model has the lowest RMSE 6.0955, followed by the NMT(\(\rho = -2.5\%\)) and NMT(\(\rho = -5\%\)) with RMSE 6.3516 and 6.6926 respectively and the INRT model with RMSE 6.7196. However, results for the samples ATM and ITM are not representative since they characterize only the 19% of the full sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>OTM</td>
<td>NMT((\rho=-10%))</td>
<td>3.4294</td>
<td>0.9993</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-7.5%))</td>
<td>3.1330</td>
<td>0.1270</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-5%))</td>
<td>3.2760</td>
<td>-0.6153</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-2.5%))</td>
<td>3.6665</td>
<td>-1.3174</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>4.0998</td>
<td>-1.9114</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.0609</td>
<td>-1.8890</td>
</tr>
<tr>
<td>ATM</td>
<td>NMT((\rho=-10%))</td>
<td>6.9232</td>
<td>4.5585</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-7.5%))</td>
<td>5.2461</td>
<td>2.3638</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-5%))</td>
<td>4.7585</td>
<td>0.9128</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-2.5%))</td>
<td>4.1365</td>
<td>-0.8549</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>4.5119</td>
<td>-1.9136</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>4.1210</td>
<td>-1.7158</td>
</tr>
<tr>
<td>ITM</td>
<td>NMT((\rho=-10%))</td>
<td>11.3244</td>
<td>8.0864</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-7.5%))</td>
<td>8.3504</td>
<td>4.5277</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-5%))</td>
<td>6.6926</td>
<td>1.5468</td>
</tr>
<tr>
<td></td>
<td>NMT((\rho=-2.5%))</td>
<td>6.3516</td>
<td>-0.5876</td>
</tr>
<tr>
<td></td>
<td>INRT</td>
<td>6.7196</td>
<td>-2.6624</td>
</tr>
<tr>
<td></td>
<td>AH_Regr</td>
<td>6.0955</td>
<td>-0.3768</td>
</tr>
</tbody>
</table>

Table 4.6: Pricing errors within moneyness categories for the American FTSE 100 index call options using the Non-Markovian model for correlation -10% \(NMT(\rho = -10\%)\), for correlation -7.5% \(NMT(\rho = -7.5\%)\), for correlation -5% \(NMT(\rho = -5\%)\), and for correlation -2.5% \(NMT(\rho = -2.5\%)\), the implied non-recombining tree (INRT) model for \(n=7\), and the ad-hoc model of Dumas et al. (1998) (AH_Regr). We measure pricing errors by (i) the root mean square error (RMSE), and (ii) the mean difference error (MDE). Results are for the full sample when divided in 3 moneyness categories. OTM denotes the out-of-the-money, ATM denotes the at-the-money and ITM denotes the in-the-money American call options.

In Table 4.7 the RMSE, and the MDE of the six models are presented when the sample is divided in the three time to maturity categories. With respect to underpricing / overpricing results remain the same as in the overall sample. In the ST
category (6.5% of the sample) the $NMT(\rho = -2.5\%)$ has the lowest RMSE 2.9480. In the MT (66.5% of the sample) and in the LT (27% of the sample) the $NMT(\rho = -5\%)$ gives the lowest RMSE 3.4946 and 5.0360 respectively.

Overall, results support our modelling approach. The correct choice of the correlation parameter in the $NMT$ model, can give results that outperform the $AH\_Regr$ and the $INTR$ model.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Model</th>
<th>RMSE</th>
<th>MDE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST</td>
<td>$NMT(\rho=-10%)$</td>
<td>3.8531</td>
<td>1.1084</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-7.5%)$</td>
<td>3.6026</td>
<td>0.4593</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-5%)$</td>
<td>2.9591</td>
<td>-0.4282</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-2.5%)$</td>
<td>2.9480</td>
<td>-1.1513</td>
</tr>
<tr>
<td></td>
<td>$INRT$</td>
<td>3.2615</td>
<td>-1.8748</td>
</tr>
<tr>
<td></td>
<td>$AH_Regr$</td>
<td>3.7352</td>
<td>-2.0630</td>
</tr>
<tr>
<td>MT</td>
<td>$NMT(\rho=-10%)$</td>
<td>4.6869</td>
<td>1.9339</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-7.5%)$</td>
<td>3.7690</td>
<td>0.7334</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-5%)$</td>
<td>3.4946</td>
<td>-0.2999</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-2.5%)$</td>
<td>3.7117</td>
<td>-1.2124</td>
</tr>
<tr>
<td></td>
<td>$INRT$</td>
<td>4.2203</td>
<td>-2.0336</td>
</tr>
<tr>
<td></td>
<td>$AH_Regr$</td>
<td>3.9595</td>
<td>-1.5906</td>
</tr>
<tr>
<td>LT</td>
<td>$NMT(\rho=-10%)$</td>
<td>6.7556</td>
<td>2.6874</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-7.5%)$</td>
<td>5.3598</td>
<td>1.0569</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-5%)$</td>
<td>5.0360</td>
<td>-0.0880</td>
</tr>
<tr>
<td></td>
<td>$NMT(\rho=-2.5%)$</td>
<td>5.1186</td>
<td>-1.1746</td>
</tr>
<tr>
<td></td>
<td>$INRT$</td>
<td>5.3706</td>
<td>-1.9410</td>
</tr>
<tr>
<td></td>
<td>$AH_Regr$</td>
<td>5.2920</td>
<td>-1.8870</td>
</tr>
</tbody>
</table>

Table 4.7: Pricing errors within maturity categories for the American FTSE 100 index call options using the Non-Markovian model for correlation -10% ($NMT(\rho = -10\%)$), for correlation -7.5% ($NMT(\rho = -7.5\%)$), for correlation -5% ($NMT(\rho = -5\%)$), and for correlation -2.5% ($NMT(\rho = -2.5\%)$), the implied non-recombining tree ($INRT$) model for $n=7$, and the ad-hoc model of Dumas et al. (1998) ($AH\_Regr$). We measure pricing errors by (i) the root mean square error (RMSE), and (ii) mean difference error (MDE). Results are for the full sample when divided in 3 maturity categories. ST denotes the short term, MT denotes the medium term and LT denotes the long term American call options.
4.8. Conclusions

In the literature there is significant evidence for the predictability of financial assets returns. Even though the market option prices are affected by predictability, standard option pricing formulas do not take account of it. Lo and Wang (1995) investigate the impact of asset return predictability in the prices of an option’s underlying asset and find that, option values are significantly affected by return predictability. In this chapter we propose an implied model for option pricing that takes into account the serial dependence of the underlying asset’s returns, i.e takes into account non-Markovian assumptions. We adapt the methodological framework of a non-recombining implied tree proposed in the first chapter and impose non-Markovian assumptions. The non-Markovian assumptions allow us to have a non-parametric framework that imposes serial dependence. Therefore, unlike typical preference-free option pricing models, a parameter related to the expected return of the underlying asset appears in our model. This is in contrast to the typical preference-free framework where the expected asset return is redundant in the option formula.

We calibrate the non-Markovian model using different values for the correlation parameter estimated from historical data. We use European calls of the FTSE 100 index for year 2003. The results strongly support our modelling approach. Pricing results are smooth without evidence of an over-fitting problem and the derived implied distributions are realistic. Overall, results for the pricing of American call options indicate that the non-Markovian model outperforms the INRT model and also the ad-hoc method of smoothing Black-Scholes implied volatilities proposed by Dumas et al. (1998). Finally, consistent with Lo and Wang (1995), we find that predictability adds value to the call options.

For further research, as an extension of the proposed model it would be interesting to allow serial correlation to be an “implied parameter” determined by the model instead of providing it econometrically.
Appendix 4.1: Feasibility of the initialized non-recombining tree

We initialize the tree using the following volatility term structure:

\[ \sigma_i = \sigma_1 e^{\lambda(t) \Delta t}, \quad \lambda \in R \text{ where } i = 1, \ldots, n \]

The feasibility of the initial tree depends on the right choice of the local volatility term structure; hence to obtain a feasible initial tree we must find an interval with the appropriate values of \( \lambda \). In order to keep the probabilities well specified at every time step, the following constraints must be satisfied:

\[
S_{i,j} (1 + \rho r_{i,j}) \leq S_{i+1,j} \quad (4.1a)
\]
\[
S_{i,j} (1 + \rho r_{i,j}) \geq S_{i+1,j-1} \quad (4.1b)
\]

Also,

\[
S_{i+1,2j} = S_{i,j} u_{i,j} = S_{i,j} e^{\sigma_j \Delta t} \quad (4.2a)
\]
\[
S_{i+1,2j-1} = S_{i,j} d_{i,j} = S_{i,j} e^{-\sigma_j \Delta t} \quad (4.2b)
\]

Substituting (4.2a) and (4.2b) to (4.1a) and (4.1b) respectively we get the following inequalities:

\[
\sigma_i \geq \frac{1}{\sqrt{\Delta t}} \log (1 + r_{i,j}) \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1} \quad (4.3a)
\]
\[
\sigma_i \geq \frac{1}{\sqrt{\Delta t}} \log (1 + r_{i,j}) \quad i = 1, \ldots, n - 1, \quad j = 1, \ldots, 2^{i-1} \quad (4.3b)
\]
Thus we have that

\[ \sigma_i \geq \frac{1}{\sqrt{\Delta t}} \log \left( 1 + r_{i,j} \right) \quad i = 1, \ldots, n-1 \quad j = 1, \ldots, 2^{i-1} \]  

(4.A 4)

For \( \lambda \geq 0 \), since \( \sigma_i = \sigma_1 e^{\lambda (i-1) \Delta t} \) it follows that the values of \( \sigma_{i,j} \) are strictly increasing as \( i \) increases.

Let \( \xi_M = \frac{1}{\sqrt{\Delta t}} \max_{i,j} \left\{ \log \left( 1 + \rho r_{i,j} \right) \right\} \)

Then (4.A 4) holds for every \( i \) if

\[ \min_i \sigma_i \geq \xi_M \quad \text{or} \quad \sigma_1 \geq \xi_M \]  

(4.A 5)

The minimum value of \( \sigma_i \) is for \( i=1 \) \( (\sigma_1) \), thus (4.A 5) is independent of \( \lambda \). Therefore, if \( \lambda \) is positive there is no upper bound for \( \lambda \).

For \( \lambda < 0 \), \( \sigma_i = \sigma_1 e^{\lambda (i-1) \Delta t} \) is strictly decreasing.

Let \( \xi_m = \frac{1}{\sqrt{\Delta t}} \min_{i,j} \left\{ \log \left( 1 + \rho r_{i,j} \right) \right\} \)

Then (4.A 4) holds for every \( i \) if

\[ \min_i \sigma_i \geq \xi_m \]

\[ \sigma_n \geq \xi_m \]

\[ \sigma_1 e^{\lambda (i-1) \Delta t} \geq \xi_m \]

But, \((i-1) \Delta t = T\), thus,

\[ \lambda \geq \frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right) \]  

(4.A 6)
If we allow $\lambda$ to take both negative and positive values, then $\lambda$ should belong in the interval, $\lambda \in \left[ \frac{1}{T} \log \left( \frac{\xi_m}{\sigma_1} \right), +\infty \right)$

\hspace{1cm} (4.A 7)
Appendix 4.2A: Derivation of \( \frac{\partial C_{i,j}}{\partial S_{i+1,j}} \) for \( i = n-1, ..., 1 \) and \( j = 1, ..., 2^{i-1} \)

\[
\frac{\partial C_{i,j}}{\partial S_{i+1,j}} = \left( \frac{\partial p_{i,j} C_{i+1,j}}{\partial S_{i+1,j}} + p_{i,j} \frac{\partial C_{i+1,j}}{\partial S_{i+1,j}} - \frac{\partial p_{i,j} C_{i+1,j-1}}{\partial S_{i+1,j}} \right) e^{-r_j \Delta t} =
\]

\[
\left( \frac{-\partial p_{i,j}}{S_{i+1,j} - S_{i+1,j-1}} C_{i+1,j} + p_{i,j} \Delta_{i+1,j} + \frac{p_{i,j}}{S_{i+1,j} - S_{i+1,j-1}} C_{i+1,j-1} \right) e^{-r_j \Delta t} =
\]

\[
\left( -p_{i,j} \frac{C_{i+1,j} - C_{i+1,j-1}}{S_{i+1,j} - S_{i+1,j-1}} + p_{i,j} \Delta_{i+1,j} \right) e^{-r_j \Delta t} =
\]

\[
p_{i,j} \left( \Delta_{i+1,j} - D_{i,j} \right) e^{-r_j \Delta t} \Rightarrow
\]

\[
\frac{\partial C_{i,j}}{\partial S_{i+1,j}} = p_{i,j} \left( \Delta_{i+1,j} - D_{i,j} \right) e^{-r_j \Delta t}
\]

where,

\[
D_{i,j} = \frac{C_{i+1,j} - C_{i+1,j-1}}{S_{i+1,j} - S_{i+1,j-1}} \quad \text{and} \quad \Delta_{i,j} = \frac{\partial C_{i,j}}{\partial S_{i,j}}, \text{Delta ratio}
\]
Appendix 4.2B: Derivation of $\frac{\partial C_{i,j}}{\partial S_{i+1,2j-1}}$ 

$i = n - 1, \ldots, 1 \quad j = 1, \ldots, 2^{i-1}$

$$\frac{\partial C_{i,j}}{\partial S_{i+1,2j-1}} = \left( \frac{\partial p_{i,j}}{\partial S_{i+1,2j-1}} C_{i+1,2j} - \frac{\partial p_{i,j}}{\partial S_{i+1,2j-1}} C_{i+1,2j-1} + (1-p_{i,j}) \frac{\partial C_{i+1,2j-1}}{\partial S_{i+1,2j-1}} \right) e^{-r_f \Delta}$ =

$$\left( -\frac{1-p_{i,j}}{S_{i+1,2j} - S_{i+1,2j-1}} (C_{i+1,2j} - C_{i+1,2j-1}) + (1-p_{i,j}) \Delta_{i+1,2j-1} \right) e^{-r_f \Delta}$ =

$$\left( - (1-p_{i,j}) \frac{C_{i+1,2j} - C_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} + (1-p_{i,j}) \Delta_{i+1,2j-1} \right) e^{-r_f \Delta}$ =

$$\frac{\partial C_{i,j}}{\partial S_{i+1,2j-1}} = (1-p_{i,j}) (\Delta_{i+1,2j-1} - D_{i,j}) e^{-r_f \Delta}$

where,

$$D_{i,j} = \frac{C_{i+1,2j} - C_{i+1,2j-1}}{S_{i+1,2j} - S_{i+1,2j-1}} \quad \text{and} \quad \Delta_{i,j} = \frac{\partial C_{i,j}}{\partial S_{i,j}}, \text{Delta ratio}$$
Summary and Conclusions

In most options markets, the implied Black-Scholes volatilities vary with both strike and expiration, a relationship commonly known as the volatility smile. The appearance and persistence of the volatility smile and term structure in the financial markets have encouraged the development of smile consistent models or otherwise known as model calibration. Model calibration deals with the inverse problem that option pricing theory deals with. Instead of assuming a stochastic process for the underlying asset, it identifies the (unknown) stochastic process of the underlying asset given information about prices of options. In this thesis we develop and test three smile consistent models for calibrating a non-recombining tree using FTSE 100 index call options data of the year 2003.

In order to capture the implied distribution, in all of the models, we minimize the discrepancy between the observed market prices and the model values, subject to constraints that keep the probabilities well specified and prevent arbitrage opportunities. We built our models on a non-recombining tree so as to allow the local volatility to be a function of the underlying asset and of time and to enable each node of the tree to act as an independent variable. Effectively, the problem under consideration is a non-convex optimization problem with linear constraints. We elaborate on the initial guess for the volatility term structure, and using methods from nonlinear constrained optimization we minimize the least squares error function on market prices. Specifically, we adopt an exterior penalty method and for the optimization we use a Quasi-Newton algorithm. Because of the combinatorial nature of the tree and the large number of constraints, the search for an optimum solution as well as the choice of an algorithm that performs well becomes a very challenging problem.

In the first model (INRT), in order to capture the implied distribution of the underlying asset, we calibrate the non-recombining tree assuming that the system variable is the underlying asset at each node of the tree. This method is flexible since it applies to arbitrary underlying asset distribution, which implies arbitrary local
volatility distribution. In the second model (INRT_LocalVol), instead of the underlying asset, we assume that the variable system is the local volatility at each node of the tree. In this way, we reduce the number of variables that the optimization algorithm has to deal with to half, in relation to the INRT model, making the problem less computationally intensive. Moreover, if required, we can easily control the values of the local volatilities at each node of the tree by imposing lower and upper bounds on the values of the local volatility. Finally, in the third model (NMT), we calibrate the non-recombining tree taking into account the serial dependence of underlying asset’s returns, i.e. we take into account non-Markovian assumptions. To the best of our knowledge, this is the first attempt in the literature of imposing non-Markovian assumptions on the process followed by the underlying asset on an implied tree. This setup gives a richer and more realistic model than the other implied binomial trees and continuous option pricing models in the literature.

The main benefit of the proposed models is their analytical structure which enables us to use efficient methods for nonlinear optimization. Although the models use a large number of variables, due to the constraints imposed and due to the fact that we use efficient methods for optimization the models are not computationally intensive. Also, the extra degrees of freedom allow for more flexibility in the estimation of the underlying asset distribution. This has two implications: It prevents non-acceptable transition probability values and also helps to obtain almost perfect fit between the model and market option values. In addition, because of their binary structure, the models allow for path dependency. This creates a more natural framework for pricing path depended options and also provides a market consistent tree for option replication with transaction costs which requires non-recombining tree (see Edirisinghe et al., 1993). In contrast to Rubinstein (1994), the proposed models can be easily modified to account for European contracts with different maturities. Our models do not need any interpolation or extrapolation across strikes and time to find hypothetical options as opposed to Derman and Kani (1994). The proposed models can be easily modified to capture the observed bid/ask spreads in the market. This is
very useful since the reported closing prices may not always be accurate, or may be inaccurate due to various market frictions (transaction costs, illiquidity, etc.).

All models are calibrated using FTSE 100 European call options for the year 2003. Overall results strongly support our modelling approaches. Pricing results are smooth without the presence of an over-fitting problem and the derived implied distributions are realistic. Also, the computational burden is not a major issue. Models are also tested in pricing of American call options, and also compared with an ad hoc method ($AH_{Regr}$) of smoothing BS implied volatilities proposed by Dumas et al. (1998), and with an improved version of the implied recombining (binomial) tree of Derman and Kani (DK, 1994), a very popular model among practitioners. The $INRT$ and the $INRT_{LocalVol}$ models are equally good in pricing of American call options as the $AH_{Regr}$ model. Moreover, the $NMT$ for specific choices of correlation between the asset’s returns of two consecutive time steps outperforms the equivalent Markovian $INRT$ model, and the $AH_{Regr}$ model. Finally, the $INRT$ model outperforms the $DK$ model.

Possible future research can be concentrated in the following aspects:

i. Check to what extend the implied tree’s local volatility distribution can forecast index (underlying asset) local volatility at future times and index levels.

ii. Create a model of calibration of a non-recombining tree that will allow the extraction of the distribution of implied correlation of returns.
References


