

Markov Tree Options Pricing

Harish S. Bhat* Nitesh Kumar*

March 7, 2010

Abstract

This paper questions one of the fundamental assumptions made in options pricing: that the daily returns of a stock are independent and identically distributed (IID). We apply an estimation procedure to years of daily return data for all stocks in the French CAC-40 index. We find six stocks whose log returns are best modeled by a first-order Markov chain, not an IID sequence. We further propose the Markov tree (MT) model, a modification of the standard binomial options pricing model, that takes into account this first-order Markov behavior. Empirical tests reveal that, for the six stocks found earlier, the MT model's option prices agree very closely with market prices.

1 Introduction

In the Black-Scholes model for the price of a European option, one of the main assumptions is that the price of the underlying asset follows a geometric Brownian motion [8]. If S_t is the underlying asset price at time t , one assumes $dS_t = \mu S_t dt + \sigma S_t dW_t$, where μ and σ are constants and W_t is a Brownian motion. For fixed $t > 0$, define $X_n = \log(S_{(n+1)t}/S_{nt})$. Then

$$X_n = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma (W_{(n+1)t} - W_{nt}).$$

Since W_t is a Brownian motion, $W_{(n+1)t} - W_{nt}$ is normally distributed with mean 0 and variance t . This implies that X_n is normally distributed with mean $(\mu - \sigma^2/2)t$ and variance $\sigma^2 t$, i.e., the distribution of X_n does not depend on n , so each X_n is identically distributed. Moreover, X_{n+1} is independent of X_n , so

$$P(X_n | X_{n-1}) = P(X_n) \tag{1}$$

for all positive integers n , and hence the $\{X_n\}$ sequence is IID (independent and identically distributed).

In fact, (1) follows from assumptions that are much more general than the geometric Brownian motion assumption; for example, if we take $S_t = S_0 \exp(L_t)$ where L_t is *any* Lévy process, (1) still holds. The upshot is that most options pricing models in

*School of Natural Sciences, University of California, Merced. Email: hbhat@ucmerced.edu

Strike	Market	Black-Scholes	Markov Tree
40	34.49	36.57	35.85
48	27.48	29.85	28.09
56	20.90	23.96	20.83
60	17.78	21.36	17.53
64	15.03	18.99	14.53
72	10.00	14.90	9.55
80	6.26	11.60	5.94
88	3.70	8.99	3.53
120	0.32	3.17	0.32
160	0.01	0.87	0.01

Table 1: Market and model prices (in €) for a particular European call option on August 24, 2009.

use—including Black-Scholes, binomial, and most jump-diffusion models—implicitly assume that the daily log returns for any stock are IID. With this in mind, the plan for this paper is as follows:

1. We first check whether (1) is consistent with real data. To do this, we apply order estimators to log return time series data for European stocks in the CAC-40 index. For several stocks, we find that (1) can be rejected in favor of a first-order Markov model for the stock price process.
2. We modify the standard binomial tree model to formulate a method for pricing options that is valid when (1) is not. We introduce first-order Markov behavior of the underlying asset into the tree, by allowing the jumps of the tree to depend on whether the previous jump was an upward or downward jump. We refer to this model as the Markov Tree (MT) model.
3. Finally, we test the MT model against the standard Black-Scholes model. We find that the MT model’s option prices are much closer to market prices than the Black-Scholes model’s prices. As a preview of our results, we present Table 1, which compares model and market prices on August 24, 2009, for a particular European call option.

2 Motivation

Let us discuss Table 1 in greater detail. On August 24, 2009, we obtained from euronext.com the end-of-day market prices for European call options for Air Liquide (symbol: AI) expiring in September 2010. We have tabulated the market prices together with prices calculated using the Black-Scholes [1] model and the MT model introduced in this paper. To calculate prices using the Black-Scholes model, we require two parameters, the risk-free interest rate r and the volatility σ . Using standard estimation procedures from empirical data¹, we obtain $r = 0.00905453$ and $\sigma = 0.41632$. The

¹We estimate the risk-free rate using the no-arbitrage futures pricing formula $F = Se^{rt}$; here F is the futures price, S is the spot price, and t is the time until expiration of the futures contract. On August 24,

MT model uses these two parameters together with σ_+ and σ_- , which are the volatilities on days where the stock's return increased (for σ_+) or decreased (for σ_-) relative to the previous day's return².

Examining Table 1, we find that for a strike of €40, the Black-Scholes model's price is only 6% greater than the market price, but as the strike increases and exceeds the spot price of €75.43, the Black-Scholes model's price diverges considerably. For example, at a strike of €88, the Black-Scholes model's price is 143% greater than the market price. This well-known divergence is usually explained through the dependence of volatility on the strike price. For each strike, one computes the value of the volatility such that the Black-Scholes model price matches the market price. When the resulting implied volatilities are plotted versus strike price, one obtains the classic volatility smile [8, Chap. 16].

We do not dispute that volatility should vary in some way as a function of option strike and time until expiry. However, in the absence of an exact form of this quantitative dependence, we ask: do we know for sure that the discrepancy between Black-Scholes and market prices is due *entirely* to the volatility smile? Our view is that, for certain options, the discrepancy is at least partially due to the market's knowledge that today's returns alter or influence the probability distribution of tomorrow's returns. Unlike commonly used option pricing models, the MT model accounts for this, and as shown in Table 1, it is significantly more accurate than Black-Scholes for out-of-the-money options, with no strike-dependent volatilities used for either model. Though the MT model does not provide an analytical formula for the option price, it is computationally tractable thanks to a large amount of recombination in the price tree for the underlying asset. We revisit these implementation issues later in the paper.

3 Past Work

Before continuing with the plan of the paper given in Section 1, we discuss relevant past work. A k -th order Markov chain is defined as a sequence $\{Y_n\}_{n \geq 1}$ of random variables such that

$$P(Y_n | Y_{n-1}, \dots, Y_1) = P(Y_n | Y_{n-1}, \dots, Y_{n-k}).$$

In a k -th order Markov chain, the current state Y_n is allowed to depend only on the past k states. The *order estimation problem* is to take N observations y_1, \dots, y_N generated by a Markov chain (of unknown order) and return an estimate \hat{k} of the chain's order. The estimator is *consistent* if, as the number of observations N goes to infinity, \hat{k} converges

2009, we found that $S = 75.43$ and $F = 75.658$ for the AI future expiring in December 2009, which also gives $t = 84$ trading days = 0.33 years. This yields an annualized risk-free rate of $r = 0.0090543$. To estimate the volatility, we start with 252 trading days (or one year) of the adjusted closing price for AI, which we represent as $\{S_1, S_2, \dots, S_{252}\}$. We then calculate $\hat{\sigma}$, the standard deviation of the log return sequence $\{\log S_2/S_1, \log S_3/S_2, \dots, \log S_{252}/S_{251}\}$; this yields the annualized volatility $\sigma = \hat{\sigma}\sqrt{252} = 0.41632$. This follows [8, Chap. 13].

²The parameters σ_{\pm} are calculated in precisely the same way as σ , except that for σ_+ we take the standard deviation of log returns on days when the stock's return increased, while for σ_- we take the standard deviation of log returns on days when the stock's return decreased. This is discussed in greater detail in Section 6.2.

to the true order k of the Markov chain. In our work, we make use of the provably consistent BIC order estimator [3].

In the context of jump-diffusion models, the IID assumption has been examined recently by Câmara and Li [2], who discuss several empirical studies that have rejected that stock jumps are IID. The focus of their paper is the development of a jump-diffusion options pricing model that does not assume the jumps are IID. Their work differs from ours in two ways: (1) non-IID behavior is modeled only through the jumps (and not through the diffusion) of the jump-diffusion process that the underlying asset is assumed to follow, and (2) the means and variances of the jumps are allowed to be time-varying. By comparison, because the MT model is discrete in time, every stock price path is a sequence of jumps—non-IID behavior is not confined to one part of the model. We make no claims about the limit of the MT model as the number of steps becomes infinite. However, we do assume that the magnitudes of possible jumps remain constant throughout the price tree.

Markov and semi-Markov processes, including processes with finite state spaces, have been used to price options [9, 4]. Though these works assume that the log return process $\log(S_t/S_{t-1})$ follows some type of discrete-time Markov or semi-Markov process, the tree models that are proposed differ from the MT model in one important regard: starting from any vertex of the tree, the *magnitudes* of the up and down jumps are always the same. The same is true in models where a Markov chain is used to approximate the true underlying process—see [5], for instance. In the MT model, if we start from a vertex such that the jump leading to that vertex *was an upward jump*, then we have different up/down magnitudes as compared with a vertex such that the jump leading to that vertex *was a downward jump*. In other words, the magnitudes of the jumps in the MT model’s tree possess the first-order Markov property. This same property distinguishes the MT model from other tree models that involve trinomial, pentanomial, or more general branching at tree vertices—see [11, 15].

4 Order Estimation: Methodology

Here we describe the methods used to test (1) against real data. We begin with a time series $\{s_0, \dots, s_N\}$ consisting of the adjusted daily closing price of a given stock. We define $x_n = \log(s_n/s_{n-1})$ and obtain the log return time series $\{x_1, \dots, x_N\}$. Note that each element of this time series is real-valued. To apply Markov order estimation, we must first convert the log return time series into a sequence of symbols drawn from a finite set. As in prior work [10, 14], the simplest way to do this is with just two symbols. We therefore define

$$z_n = \begin{cases} u & x_n \geq 0 \\ d & x_n < 0, \end{cases} \quad (2)$$

where the symbols u and d stand for “up” and “down,” respectively. Note that this transformation erases the magnitudes of the upward/downward movements of the stock. Now that we have a sequence $\{z_j\}_{j=1}^N$ of u ’s and d ’s, we can begin to extract maximum likelihood estimates (MLE’s) of Markov chain transition probabilities. Let us describe how this is done.

By the definition given in Section 3, a zeroth-order Markov chain is simply a sequence of IID random variables. Since each z_j in our sequence can be in one of only two possible states, if the sequence was generated by a zeroth-order Markov chain, then each z_j was generated by a Bernoulli random variable with only one parameter: $p = P(u)$, the probability of obtaining u . In this case, $1 - p = P(d)$. In this case, we define the zeroth-order log likelihood

$$L_0(p) = n_u \log p + n_d \log(1 - p),$$

where n_u is the number of u 's observed in the sequence, $n_d = N - n_u$ is the number of d 's observed in the sequence, and N is the total length of the sequence. Solving $dL_0/dp = 0$ for p gives the MLE $\hat{p} = n_u/N$, which is in fact the maximizer of $L_0(p)$. Using this MLE, we can compute the maximum value $L_0(\hat{p})$.

Let us now redo this calculation assuming that the sequence $\{z_j\}_{j=1}^N$ was generated by a first-order Markov chain. Now we require three parameters, $p_1 = P(u)$, $p_2 = P(u|u)$ and $p_3 = P(d|d)$. Note that p_1 is necessary to handle z_1 , the first element of the sequence. Also note that $P(d|u) = 1 - p_2$ and $P(u|d) = 1 - p_3$. Putting it all together, we obtain the first-order log likelihood

$$\begin{aligned} L_1(p_1, p_2, p_3) &= m \log p_1 + (1 - m) \log(1 - p_1) \\ &\quad + n_{uu} \log p_2 + n_{ud} \log(1 - p_2) \\ &\quad + n_{dd} \log p_3 + n_{du} \log(1 - p_3). \end{aligned} \tag{3}$$

Here $m = 1$ if $z_1 = u$ and $m = 0$ if $z_1 = d$. The notation $n_{\sigma\tau}$ denotes, for any choice $\sigma, \tau \in \{u, d\}$, the number of times the string $\sigma\tau$ was observed in the sequence. Now solving $\partial L_1/\partial p_j = 0$ for p_j , $j = 1, 2, 3$ yields the MLE's

$$\begin{aligned} \hat{p}_1 &= m \\ \hat{p}_2 &= \frac{n_{uu}}{n_{uu} + n_{ud}} = \frac{n_{uu}}{\text{\# of } u \text{ in first } N - 1 \text{ slots}} \\ \hat{p}_3 &= \frac{n_{dd}}{n_{dd} + n_{du}} = \frac{n_{dd}}{\text{\# of } d \text{ in first } N - 1 \text{ slots}} \end{aligned}$$

Using these MLE's, we can calculate the maximum value $L_1(\hat{p}_1, \hat{p}_2, \hat{p}_3)$.

Following the same methodology, we can assume that the sequence $\{z_j\}_{j=1}^N$ was generated by a k -th order Markov chain and then write down the k -th order log likelihood function L_k of the unknown Markov transition probabilities $\mathbf{p} = (p_1, \dots, p_M)$. We plug into L_k the frequencies of different strings found in the actual $\{z_j\}_{j=1}^N$ sequence and then maximize L_k over \mathbf{p} . We thereby find both the MLE's $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_M)$ as well as the maximum value of the log likelihood $L_k(\hat{\mathbf{p}})$. The calculations are carried out in the Appendix of this paper, for a sequence that is assumed to be an i -th order Markov chain over Q states. (In the above discussion, we treated only the $Q = 2$ case.)

Armed with this information, we employ the BIC (Bayesian Information Criterion) order estimation method. We calculate $f(j, N) = L_j(\hat{\mathbf{p}}) - 2^{j-1} \log(N)$, looping over values of j from 0 to K . The BIC order estimate \hat{k} equals the value of j that maximizes $f(j, N)$. In the limit where the number of data points is infinite, $N \rightarrow \infty$, it has been proven that the BIC estimate converges to the true order of the Markov chain generating the data [3].

		\hat{k}		
		0	1	2
k	0	966	33	1
	1	180	818	2
	2	28	116	856

Table 2: The (k, \hat{k}) entry equals the number of times the BIC order estimator returned an order of \hat{k} , when applied to a random sequence of length $N = 500$ generated by a randomly generated Markov chain of true order k . For results in this table, transition probabilities were drawn uniformly from $(0, 1) \subset \mathbb{R}$.

Though the theoretical results on BIC order estimation with an infinite amount of data are encouraging, they are obviously not strictly applicable to our situation, where the length of the time series is finite. To remedy this, we study the performance of BIC order estimation on finite, synthetic data sets.

For $k \in \{0, 1, 2\}$, we randomly generate transition probabilities for a k -th order Markov chain. Each probability is drawn uniformly from the interval $(0, 1)$. Using this Markov chain, we generate a sequence of length $N = 500$. We apply BIC order estimation to this sequence and thereby obtain an estimate \hat{k} of the Markov chain's order. The results are summarized in Table 4. When we apply BIC order estimation, we loop over possible orders $j = 0, 1, 2, \dots, 8$. However, in no instance do we find that the estimate \hat{k} exceeds two. This can be seen by noting that for each value of k , we randomly generated exactly 1000 sequences, and each row of the table sums to 1000. Based on the numbers given in Table 4, we make the following estimates:

$$P(k \geq 1 \mid \hat{k} = 1) \approx \frac{818 + 116}{33 + 818 + 116} = 0.9659 \quad (4)$$

$$P(k = 0 \mid \hat{k} = 0) \approx \frac{966}{966 + 180 + 28} = 0.8228 \quad (5)$$

That is to say, if the BIC order estimator equals one for a given sequence, we find there is a greater than 95% chance that the sequence was generated by a Markov chain of *at least* order one, i.e., a 95% chance that the sequence was in fact *not* IID. On the other hand, if the BIC order estimator equals zero for a given sequence, we find that there is an approximately 80% chance that the sequence was in fact IID.

In our observations, after converting real time series for stocks into u/d sequences, the maximum likelihood estimates of the transition probabilities are always between 0.4 and 0.6. This motivates us to rerun the above tests. This time, when we randomly generate transition probabilities for a k -th order Markov chain, we draw each probability uniformly from the interval $(0.4, 0.6)$. Other parameters of the test remain the same. The results are summarized in Table 4. Once again, in no instance do we find that the estimate \hat{k} exceeds two; we ran 1000 tests for each value of k , and each row

		\hat{k}		
		0	1	2
k	0	983	17	0
	1	686	312	2
	2	758	182	60

Table 3: The (k, \hat{k}) entry equals the number of times the BIC order estimator returned an order of \hat{k} , when applied to a random sequence of length $N = 500$ generated by a randomly generated Markov chain of true order k . For results in this table, transition probabilities were drawn uniformly from $(0.4, 0.6) \subset \mathbb{R}$.

sums to 1000. Based on Table 4, we make the following estimates:

$$P(k \geq 1 \mid \hat{k} = 1) \approx \frac{312 + 182}{17 + 312 + 182} = 0.9667 \quad (6)$$

$$P(k = 0 \mid \hat{k} = 0) \approx \frac{983}{983 + 686 + 758} = 0.4050 \quad (7)$$

These results strengthen our conclusion that if the BIC order estimator applied to a sequence yields one, there is a greater than 95% chance that the sequence was in fact *not* IID. Note, however, that drawing the transition probabilities from the interval $(0.4, 0.6)$ —centered at 0.5—has made it very easy for the BIC order estimator to *underestimate* the true order of the Markov chain. This is intuitively clear: if the transition probabilities for either a first- or second-order Markov chain are all close to 0.5, then short sequences generated by the Markov chain will appear to be IID. One will require an extremely long sequence from such a Markov chain in order to distinguish the sequence from an IID sequence; $N = 500$ samples is simply not enough.

The meaning of these results is that we can reliably use the BIC order estimator to *falsify* (1), but never to verify (1). In situations where we apply the BIC order estimator to real financial time series and obtain an estimate of at least one, there is a high probability that (1) is false; if, on the other hand, we obtain an estimate of zero, we should discard it.

5 Order Estimation: Results

We apply the BIC order estimation technique to stocks listed on the French CAC-40 index. Our interest in these stocks stems purely from the fact that European-style options on these stocks are traded on Euronext, and both the classical Black-Scholes model and our MT model are designed to price European-style options. For each stock on the index, we download at least two years of adjusted daily closing prices from Yahoo! Finance. Note that there are 252 trading days in one year, so at least two years' worth of data gives us a time series of length $N \geq 504$. We then apply the methodology of

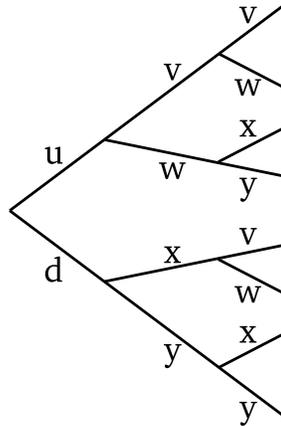


Figure 1: Illustration of the first three steps of the Markov tree. An upward edge always bifurcates into v and w . A downward edge always bifurcates into x and y . In this way, the tree accounts for the first-order Markov nature of the underlying asset's log return time series.

Section 4 and produce BIC order estimates for each time series. We find that there are six French companies for which the BIC order estimate equals one:

- Air Liquide (Euronext: AI), using data from Jan. 1, 2007 to Oct. 2, 2009.
- AXA Group (Euronext: CS, NYSE: AXA), using NYSE data from Feb. 1, 2007 to Oct. 2, 2009.
- L'Oréal Group (Euronext: OR), using data from Jan. 1, 2007 to Oct. 2, 2009.
- Pernod Ricard (Euronext: RI), using data from Jan. 1, 2003 to Oct. 2, 2009.
- Sanofi-Aventis (Euronext: SAN, NYSE: SNY), using either Euronext or NYSE data from June 30, 2007 to June 30, 2009.
- Société Générale (Euronext: GLE), using data from Jan. 1, 2007 to Oct. 2, 2009.

We believe that (1) is false for stock time series for each of these six companies. Later, when we compare the results of the Black-Scholes and MT models against market prices for European call options for these six companies, we will add further evidence to this claim. Note that we could also test (1) by using more traditional time series methods such as ACF and PACF. However, since our focus is obtaining a discrete tree model to price options, it seems natural to convert the original time series into a finite state time series and then test (1). In future work, we shall explore whether there exist time series whose non-IID behavior can be detected correctly by Markov order estimators and *not* by ACF-based methods, and vice versa.

6 Markov Tree Model: Theory

We now describe a tree model that accounts for the first-order Markov dependence in the log return time series. We restrict our model to accommodate only first-order Markov dependence (instead of, say, k -th order Markov dependence) not only to obtain computational tractability but also to maintain parsimony. Like the binomial tree, our tree is generated by working forward from valuation day to expiration of the option. Let S_n be the stock's spot price at time step n . When $n = 0$, we use one step of the standard binomial tree

$$P(S_1 = uS_0) = q \quad (8a)$$

$$P(S_1 = dS_0) = 1 - q. \quad (8b)$$

For $n \geq 1$, let us define two events:

$$S_n^+ = \{S_n \geq S_{n-1}\} \quad (9)$$

$$S_n^- = \{S_n < S_{n-1}\}. \quad (10)$$

In words, the event S_n^+ is the event that the stock price increased from time step $n - 1$ to time step n . The event S_n^- is the complement of S_n^+ , i.e., the event that the stock price decreased from time step $n - 1$ to time step n . We can now write down our model for the evolution of S_n , for $n \geq 1$:

$$P(S_{n+1} = vS_n | S_n^+) = q^+ \quad (11a)$$

$$P(S_{n+1} = wS_n | S_n^+) = 1 - q^+ \quad (11b)$$

$$P(S_{n+1} = xS_n | S_n^-) = q^- \quad (11c)$$

$$P(S_{n+1} = yS_n | S_n^-) = 1 - q^-. \quad (11d)$$

Here we have introduced four symbols, v, w, x and y , which represent different factors by which the stock price at every time step is allowed to change. According to our model, if the stock price increased from step $n - 1$ to step n , then the stock price at step $n + 1$ is vS_n with probability q^+ and wS_n with probability $1 - q^+$. If the stock price decreased from step $n - 1$ to step n , then the stock price at step $n + 1$ is xS_n with probability q^- and yS_n with probability $1 - q^-$.

We remark that we think of $q, q^+,$ and q^- as, respectively, risk-neutral versions of the empirical probabilities $P(u), P(u|u),$ and $P(u|d)$. We shall explain later how, with respect to these risk-neutral probabilities, the stock price process S_n is in fact a martingale.

The first three steps of the tree are illustrated in Figure 1. If we let S_0 denote the initial spot price of the stock, then it is clear that $S_3 \in J_3$ where

$$J_3 = \{S_0uw^2, S_0uvw, S_0uwx, S_0uwy, S_0dxv, S_0dxw, S_0dyx, S_0dy^2\}.$$

In general, let J_n denote the vector of possible states the stock can be in after n steps of the Markov tree. Let $\delta_n : J_n \rightarrow \mathbb{Z}^+$ be the function that counts the number of paths in the tree that lead from S_0 to a given element of J_n . For $\omega \in J_n$, we refer to $\delta_n(\omega)$ as the *duplication number* of state ω . We list without proof these facts:

- J_n contains $n^2 - n + 2$ unique elements.

That is to say, states do recombine. If the stock decreases from S_0uvw , it reaches the same value as if it increases from S_0uvw —in both cases, it reaches S_0uvw . Because there are two possible paths leading from S_0 to S_0uvw , we assign the duplication number $\delta_4(S_0uvw) = 2$. Because states recombine, the number of states does *not* increase like 2^n . In the standard binomial model, the number of states grows linearly in the depth of the tree n . In the MT model, the number of states grows quadratically in the depth of the tree n . This polynomial growth ensures the tractability of the MT model as a computational method.

- $\sum_{\sigma \in J_n} \delta_n(\sigma) = 2^n$.

One can make sense of this intuitively by recalling that if we do not count the duplication of states, then a tree of depth n will contain 2^n states.

- Let p_n denote the polynomial that gives all states in J_n together with their duplication numbers, i.e.,

$$p_n(u, d, v, w, x, y) = \sum_{\omega \in J_n} \delta_n(\omega)\omega.$$

Then p_n may be computed via

$$p_n = u \begin{bmatrix} v & w & 0 & 0 \\ 0 & 0 & x & y \\ v & w & 0 & 0 \\ 0 & 0 & x & y \end{bmatrix}^{n-2} \mathbf{1} + d \begin{bmatrix} 0 & 0 & x & y \\ v & w & 0 & 0 \\ 0 & 0 & x & y \end{bmatrix}^{n-2} \mathbf{1},$$

where $\mathbf{1}$ denotes a column vector with each entry equal to one. The above fact may be derived by writing the adjacency matrix for a directed, weighted graph related to our Markov tree. For now, we merely mention that once we use the above iterative matrix formula to compute p_n for a given n and thereby generate a Markov tree of depth n , we can then reuse this tree many times to price many different options. For different options, S_0 , u , d , v , w , x , and y will be different, but the set of states J_n and the duplication numbers δ_n will always be the same.

For example, carrying out the tree one step further than shown in Figure 1, we find that

$$J_4 = \{S_0uv^3, S_0uv^2w, S_0uvw, S_0uvw, S_0uvw, S_0uvw^2x, S_0uvw^2y, S_0uvw^2z, S_0dvw^2, S_0dvw, S_0dvw^2w, S_0dvw^2x, S_0dvw^2y, S_0dvw^2z, S_0dvw^2x, S_0dvw^2y, S_0dvw^2z\}.$$

We have $\delta_4(S_0uvw) = \delta_4(S_0dvw) = 2$ and $\delta_4(\sigma) = 1$ for all other possible states $\sigma \in J_4$. Note that, as per our formula, there are $4^2 - 4 + 2 = 14$ elements in J_4 , and $\sum_{\sigma \in J_4} \delta_4(\sigma) = 16 = 2^4$.

Next, note that it is simple to calculate the probability that the stock's price path reaches a given state in J_n , starting at S_0 . Let $\sigma = S_0 u^m d^{1-m} v^a w^b x^c y^d$ denote an arbitrary state in J_n . (Clearly either $m = 0$ or $m = 1$, and also the sum of the exponents must equal n , i.e., $1 + a + b + c + d = n$.) Then, by the definitions made in (8) and (11), the probability of reaching σ starting at S_0 is simply equal to

$$P(\sigma) = \delta_n(\sigma) q^m (1-q)^{1-m} \times (q^+)^a (1-q^+)^b (q^-)^c (1-q^-)^d. \quad (12)$$

Let us now explain how we use the tree to price a European call option. Let K denote the strike price and let S_* denote the spot price at the time of expiry. The payoff of the option is denoted by $(S_* - K)_+$, which equals zero unless $S_* - K > 0$, in which case it equals $S_* - K$. Let T denote the time until expiry. We fix the total number of steps N in the tree and set $\Delta t = T/N$. With these definitions, we can see that S_* is a random variable that can take on any of the states $\sigma \in J_N$ with probabilities given by (12). We have enough information to write down the expected value of the option's payoff at the time of expiry:

$$E[(S_* - K)_+] = \sum_{\sigma \in J_N} I_{\sigma > K} (\sigma - K) P(\sigma). \quad (13)$$

Here $I_{\sigma > K}$ is an indicator variable that equals one when $\sigma > K$ and zero when $\sigma \leq K$. Now let r equal the risk-free interest rate. Then we define the MT model's call option price to be expected payoff at the time of expiry, discounted to the present time:

$$C = e^{-rT} E[(S_* - K)_+]. \quad (14)$$

Note that using precisely the same approach, we can price European put options without making use of put-call parity. The payoff of a European put option equals $(K - S_*)_+$. The MT model's put option price is, once again, the discounted expected payoff:

$$U = e^{-rT} E[(K - S_*)_+]. \quad (15)$$

6.1 No Arbitrage.

Let us show that our model does not admit arbitrage. We define

$$q = \frac{\exp(r\Delta t) - d}{u - d}$$

$$q^+ = \frac{\exp(r\Delta t) - w}{v - w}, \quad q^- = \frac{\exp(r\Delta t) - y}{x - y}.$$

One may easily verify that with these three risk-neutral probabilities,

$$E[S_1 | S_0] = uS_0q + dS_0(1-q) = e^{r\Delta t} S_0,$$

and for $n \geq 1$,

$$\begin{aligned}
E[S_{n+1}|S_n, \dots, S_0] &= E[S_{n+1}|S_n, \dots, S_0, S_n^+]P(S_n^+) \\
&\quad + E[S_{n+1}|S_n, \dots, S_0, S_n^-]P(S_n^-) \\
&= [vS_nq^+ + wS_n(1 - q^+)]P(S_n^+) \\
&\quad + [xS_nq^- + yS_n(1 - q^-)]P(S_n^-) \\
&= e^{r\Delta t}S_nP(S_n^+) + e^{r\Delta t}S_nP(S_n^-) \\
&= e^{r\Delta t}S_n.
\end{aligned}$$

This is enough to imply that the discounted stock process $\tilde{S}_n = e^{-rn\Delta t}S_n$ is a martingale under the risk-neutral probabilities given by q , q^+ , and q^- . Then, by the first fundamental theorem of asset pricing (see [13, Chapter 2.4]), there is no arbitrage in the MT model.

6.2 Implementation Notes.

The parameters u , d , v , w , x , and y are estimated as follows. For each date on which we wish to value an option, we start with the time series of one prior year's worth of adjusted closing daily returns for the stock. We scan through this time series and form two disjoint time series: each time a given day's return exceeds or equals the previous day's, we add that return to series 1; each time a given day's return is less than the previous day's, we add that return to series 2. We then take the logarithm of all returns in series 1 and 2 and also in the original time series. Let $\hat{\sigma}_+$ and l_+ denote the standard deviation and length of log return series 1, and let $\hat{\sigma}_-$ and l_- denote the standard deviation and length of log return series 2. Let $\hat{\sigma}$ be the standard deviation of the entire log return series. The standard deviations are then converted to volatilities σ , σ_+ and σ_- using $\sigma = \sqrt{252}\hat{\sigma}$ and $\sigma_{\pm} = \sqrt{l_{\pm}}\hat{\sigma}_{\pm}$. With these volatilities, we set

$$\begin{aligned}
u &= \exp\left(\sigma\sqrt{\Delta t}\right), \\
v &= \exp\left(\sigma_+\sqrt{\Delta t}\right), \quad x = \exp\left(\sigma_-\sqrt{\Delta t}\right),
\end{aligned}$$

where Δt is the duration of each time step in the model. We then set $d = 1/u$, $w = 1/v$, and $y = 1/x$.

7 Tree Model: Results

For 44 trading days from July 17, 2009 to September 17, 2009, we tracked the end-of-day market prices for European-style call options for the six companies listed in Section 5. Data was obtained from euronext.com. We emphasize that all of the tests we are about to describe are out-of-sample tests; at no time did we use past or present market prices of options as inputs to the MT or Black-Scholes models. The fact that the MT model requires no calibration with options price data from real markets is in marked contrast to, say, Rubinstein's implied binomial tree model [12].

On day i of the study, we used stock and futures prices from days before or on day i to estimate parameters that are fed as inputs to the MT and Black-Scholes options pricing models. Specifically, we estimated the risk-free rate r and the volatilities σ , σ_+ , and σ_- , which determine the jumps u , d , v , w , x , and y . With these parameters, we priced all exchange-traded options using both the MT and Black-Scholes models. For the MT model, we used $N = 501$ steps. For each option at hand, we compared the outputs of these options pricing models on day i to the market price of the same option on day i .

7.1 Comparison of Model and Market Prices.

We first consider the day-by-day performance of the MT model versus the Black-Scholes model, averaged across all strikes. Let i be a fixed day. Let \mathbf{b}_i , \mathbf{m}_i , and \mathbf{M}_i be the vectors containing Black-Scholes, MT, and market prices on day i for options of different strikes (but the same expiration date). On each day, we compute

$$\epsilon_i^b = \frac{\|\mathbf{b}_i - \mathbf{M}_i\|_2}{\|\mathbf{M}_i\|_2}, \quad \epsilon_i^m = \frac{\|\mathbf{m}_i - \mathbf{M}_i\|_2}{\|\mathbf{M}_i\|_2}. \quad (16)$$

In each of the six panels of Figure 2, we plot the relative error curves ϵ_i^b (in red) and ϵ_i^m (in blue) versus day i for options from each of the six companies listed in Section 5, respectively.

We then consider the strike-by-strike performance of the MT model versus the Black-Scholes model, averaged across all days. Let j be a fixed strike price. Let \mathbf{b}_j , \mathbf{m}_j , and \mathbf{M}_j be the vectors containing Black-Scholes, MT, and market prices for options with strike j on different days (but the same expiration date). On each day, we compute

$$\gamma_j^b = \frac{\|\mathbf{b}_j - \mathbf{M}_j\|_2}{\|\mathbf{M}_j\|_2}, \quad \gamma_j^m = \frac{\|\mathbf{m}_j - \mathbf{M}_j\|_2}{\|\mathbf{M}_j\|_2}. \quad (17)$$

In each of the six panels of Figure 3, we plot the log relative error curves $\log(\gamma_j^b)$ (in red) and $\log(\gamma_j^m)$ (in blue) versus strike price j for options from each of the six companies listed in Section 5, respectively.

In both Figure 2 and Figure 3, the following symbols are used to denote common expiration dates: “o” means September 2009, “*” means March 2010, “◇” means September 2010, and “+” means March 2011.

Comparing Black-Scholes and MT relative errors for options with the *same* expiration date means comparing blue and red curves with *identical* symbols in Figure 2 and Figure 3. For example, in Figure 2, comparing blue and red curves with “+” symbols shows that the MT model’s prices for options expiring in March 2011 are much closer to market prices than the Black-Scholes model’s prices for options expiring in March 2011. This is true for all six companies.

In fact, comparing blue and red curves in Figure 2 with identical symbols reveals that the only expiration date for which the two models produce comparable results is the September 2009 expiration date, denoted by “o.” In this case, for all six companies, the MT model still produces relative errors ϵ_i^m that are two to ten times smaller than the relative errors ϵ_i^b produced by the Black-Scholes model. For all other expiration

dates, comparing the two models on a day-by-day basis, the MT model’s call option prices are far closer to market prices than the Black-Scholes model’s prices.

Moving to Figure 3, we see that as the strike price increases, the Black-Scholes model’s error increases more rapidly than the MT model’s error. Note that each point on each of the panels in Figure 3 is an aggregate result, averaged (in the sense of the 2-norm) over 44 trading days’ worth of data. For this reason, we believe Figure 3 provides strong evidence that the discrepancy between Black-Scholes and market prices for out-of-the-money options is not entirely due to the dependence of volatility on strike price and time until expiration.

7.2 Comparison of Volatilities.

For each of the six stocks listed in Section 5, we plot in Figure 4 the three volatilities σ , σ_+ , and σ_- on each of the 44 days. These plots show that for three of the six stocks (AI, OR, and GLE before day 40), the difference between σ_+ and σ_- is small, on the order of 1%. For these three stocks, the MT model, in the way we have implemented it with the formulas from Section 6.2, produces option prices close to those produced by a binomial model with volatility given by either σ_+ , σ_- , or perhaps a weighted average of these values. It is noteworthy that a binomial model with *volatility estimated by splitting past historical data based on whether returns were increasing or decreasing relative to the previous day* does far better at tracking market prices than a vanilla Black-Scholes (or, equivalently, binomial) model with volatility σ . The formulas given in Section 6.2 for σ_{\pm} were determined by extensive trial-and-error. In future work, we shall provide a more rigorous theory for estimating the parameters v , w , x , and y that serve as inputs to the MT model.

On the other hand, the plots in Figure 4 also indicate that for three of the six stocks (CS, RI, and SAN), the difference between σ_+ and σ_- is closer to 10%. In this case, one can show that the set of states J_{501} together with (12) yield a probability distribution on the set of stock prices at the time of expiry that is different from the distribution of final stock prices provided by a standard binomial model. For these three stocks, the MT model does not reduce to a binomial model. Depending on the specific values of the parameters, it is possible for the MT model’s final stock price distribution to feature heavier tails and interesting asymmetries relative to the lognormal distribution. We expect that these features, which have been reported elsewhere in the financial time series literature, will appear when we use better methods for estimating σ_{\pm} .

8 Conclusion

Over the past two years, many securities have been subject to large fluctuations in price, and financial modeling assumptions that used to be considered standard should now be called into question. One such assumption is (1). In this paper, we have tested (1) using the BIC order estimation method. The tests have revealed six stocks in the French CAC-40 index whose log return time series is not IID. For these six stocks, and for other stocks whose log return time series is best modeled by a k -th order Markov chain with $k \geq 1$, we propose the MT options pricing model. The number of states in

the Markov tree grows quadratically in the depth of the tree, giving the model computational tractability. Implementing the MT model, we find strong agreement between the MT model's prices and market prices.

In future work, we shall compare the MT model against more sophisticated options pricing models, such as those incorporating stochastic volatility. The first-order Markov dependence of our tree model is a general concept that could be incorporated into discrete-time stochastic volatility models [6], which could further reduce the error between model and market prices. Finally, we shall extend the MT model to price weather derivatives, especially in light of scientific studies that propose Markov chain models for quantities such as rainfall [7].

9 Appendix: MLE's for k -th order Markov chain with Q symbols

Assume that we have a sequence $\{X_1, \dots, X_N\}$ generated by a k -th order Markov chain where every experiment has Q possible outcomes.

Assume that each X_j takes values from a set $\{A_m\}_{m=1}^Q$ of Q distinct symbols. For a k -th order Markov chain, X_j depends on k outcomes prior to the j -th outcome. As in the earlier case with two possible outcomes, here also the Markov property kicks in only if $j > k$. Let S_i be a subsequence of k outcomes prior to the j -th outcome. Since each of these k outcomes in the subsequence is drawn from the set $\{A_m\}_{m=1}^Q$, there are Q^k possible subsequences S_i . Let these sequences be denoted by $\{S_i\}_{i=1}^{Q^k}$. Let us scan the given sequence $\{X_j\}_{j=1}^N$ from left to right and record the following Q^{k+1} numbers:

$$\{\{n_{S_i A_m} = \# \text{ of times we observe } "S_i A_m" \}_{m=1}^Q\}_{i=1}^{Q^k}$$

For $m = 1$ to $m = Q - 1$, let $p_{S_i A_m} = P(A_m | S_i)$. Let $p_{S_i A_Q} = P(A_Q | S_i) = 1 - \sum_{m=1}^{Q-1} p_{S_i A_m}$ and $p_k = P(\{X_j\}_{j=1}^k)$. In words, p_k is the probability of observing the first k terms of the $\{X_j\}_{j=1}^N$ sequence. Putting everything together, the log likelihood for the whole sequence is

$$L = \log p_k + \sum_{i=1}^{Q^k} \left[\sum_{m=1}^{Q-1} n_{S_i A_m} \log(p_{S_i A_m}) \right] + \sum_{i=1}^{Q^k} n_{S_i A_Q} \log \left(1 - \sum_{m=1}^{Q-1} p_{S_i A_m} \right).$$

Let us maximize L over all $p_{S_i A_m}$. Taking partial derivatives of both sides with respect to $p_{S_i A_m}$ for one particular (m, i) , we get

$$\frac{1}{L} \frac{\partial L}{\partial p_{S_i A_m}} = \frac{n_{S_i A_m}}{p_{S_i A_m}} - \frac{n_{S_i A_Q}}{1 - \sum_{m=1}^{Q-1} p_{S_i A_m}}.$$

Setting $\partial L / \partial p_{S_i A_m} = 0$ to maximize L , we get

$$\frac{n_{S_i A_m}}{p_{S_i A_m}} = \frac{n_{S_i A_Q}}{1 - \sum_{m=1}^{Q-1} p_{S_i A_m}}. \quad (18)$$

Note that for a particular value of i , the above equation represents a set of $Q - 1$ linear equations in $Q - 1$ unknowns (the probabilities to be estimated), which give us the following result:

$$\frac{n_{S_i A_1}}{p_{S_i A_1}} = \frac{n_{S_i A_2}}{p_{S_i A_2}} = \dots = \frac{n_{S_i A_{Q-1}}}{p_{S_i A_{Q-1}}} = \frac{n_{S_i A_Q}}{1 - \sum_{m=1}^{Q-1} p_{S_i A_m}}. \quad (19)$$

Solving the linear system given by (18) and (19), we get the MLE for the transition probability:

$$\hat{p}_{S_i A_m} = \frac{n_{S_i A_m}}{\sum_{m=1}^Q n_{S_i A_m}}.$$

We can then use the collection of all \hat{p} 's to find the maximum value of the log likelihood L .

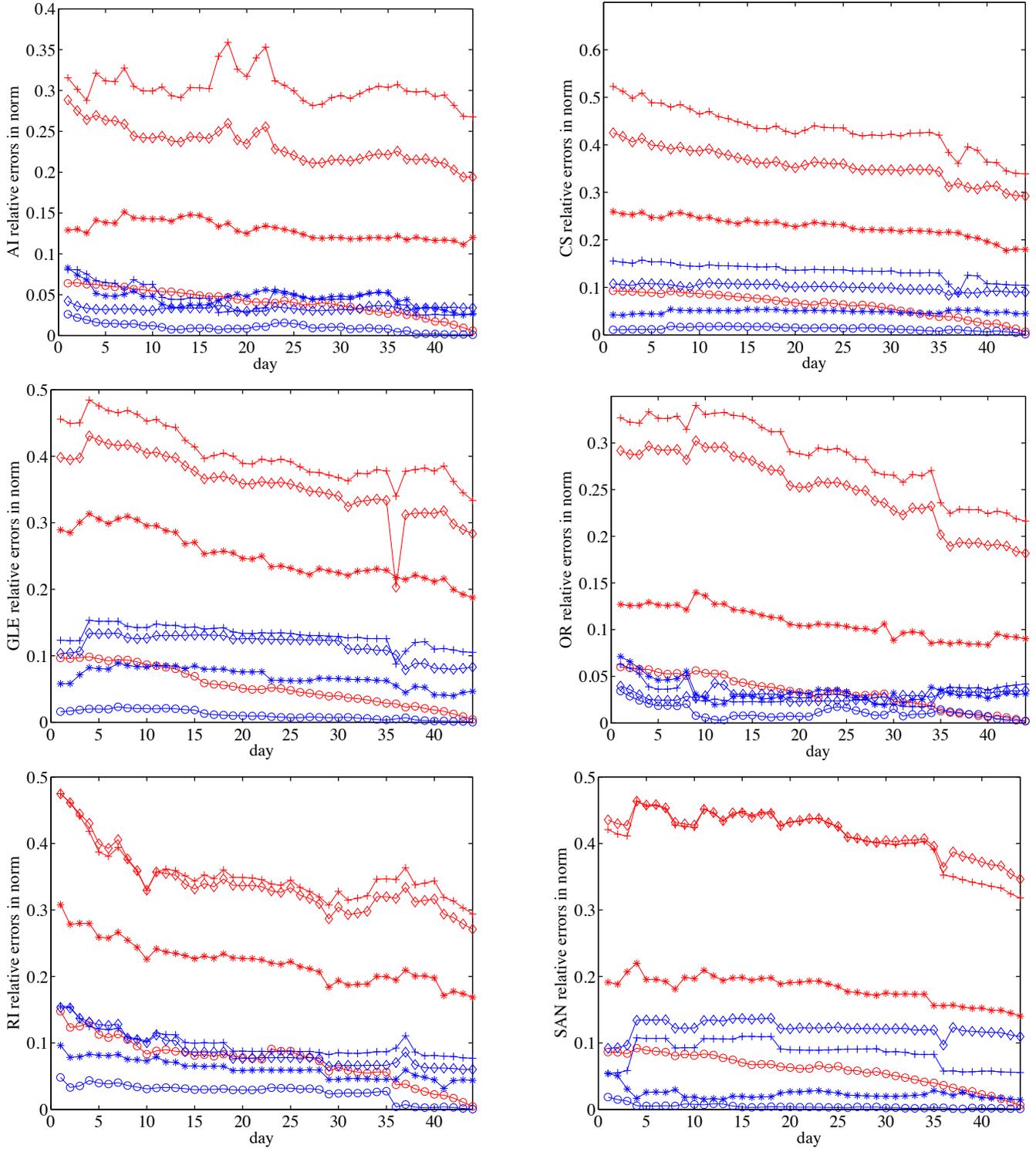


Figure 2: From left to right, top to bottom, we plot model relative errors for the six companies listed in Section 5 in the following order (alphabetical in the Euronext symbols): AI, CS, GLE, OR, RI, and SAN. Each panel displays relative errors ϵ_i^b (Black-Scholes error in red) and ϵ_i^m (MT error in blue) versus day i for options with different expiration dates. The following symbols are used to denote common expiration dates: “o” means September 2009, “*” means March 2010, “◊” means September 2010, and “+” means March 2011. Note that for all expiration dates except September 2009, the MT model’s relative error curves are far below the Black-Scholes relative error curves. For options expiring in September 2009, both models yield nearly identical results.

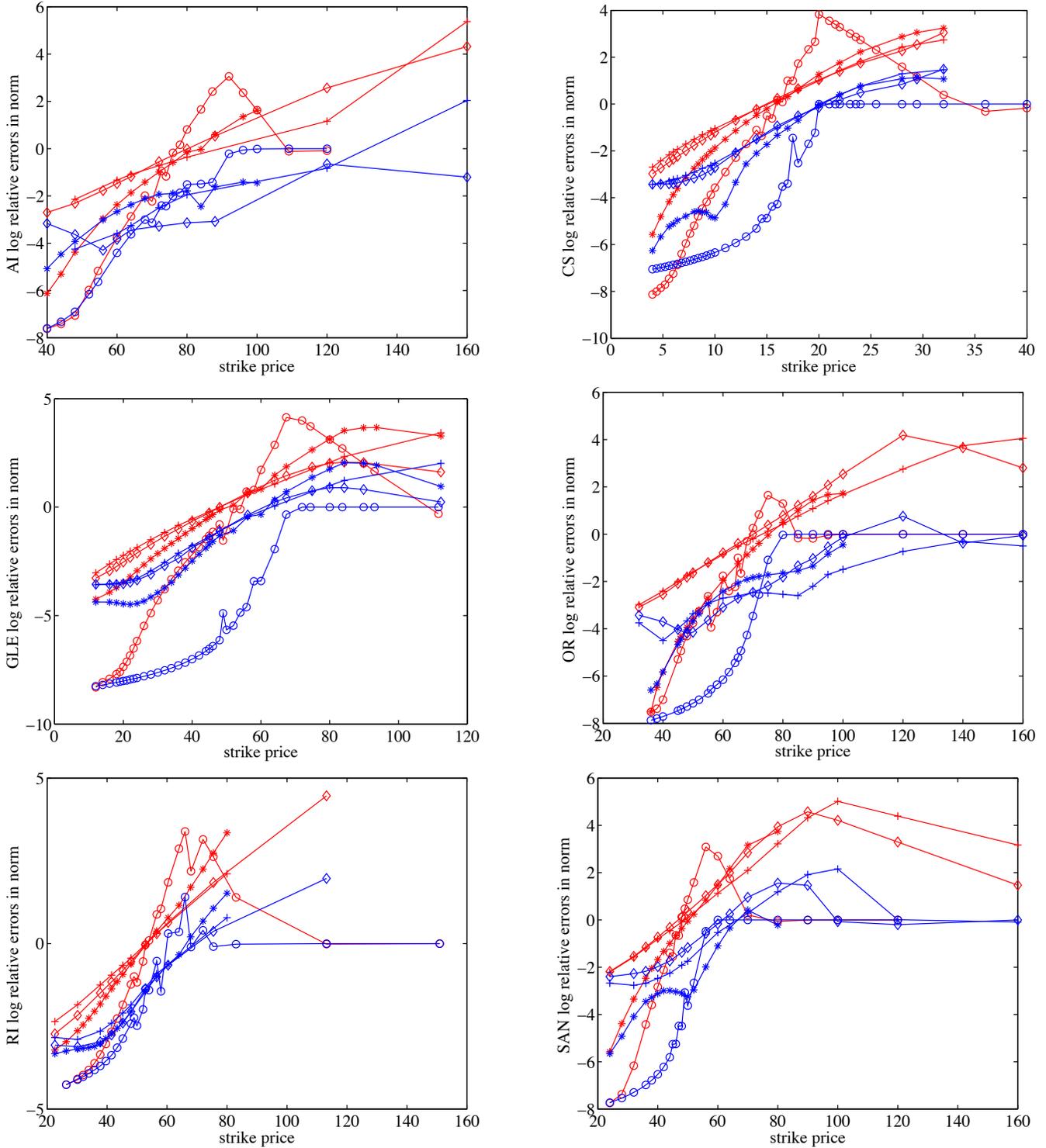


Figure 3: From left to right, top to bottom, we plot model relative errors for the six companies listed in Section 5 in the following order (alphabetical in the Euronext symbols): AI, CS, GLE, OR, RI, and SAN. Each panel displays log relative errors $\log(\gamma_j^b)$ (Black-Scholes error in red) and $\log(\gamma_j^m)$ (MT error in blue) versus strike price j for options with different expiration dates. The following symbols are used to denote common expiration dates: “o” means September 2009, “*” means March 2010, “◇” means September 2010, and “+” means March 2011. Note that for all expiration dates, as the strike price increases, the Black-Scholes model’s relative error curves far exceed the MT model’s relative error curves.

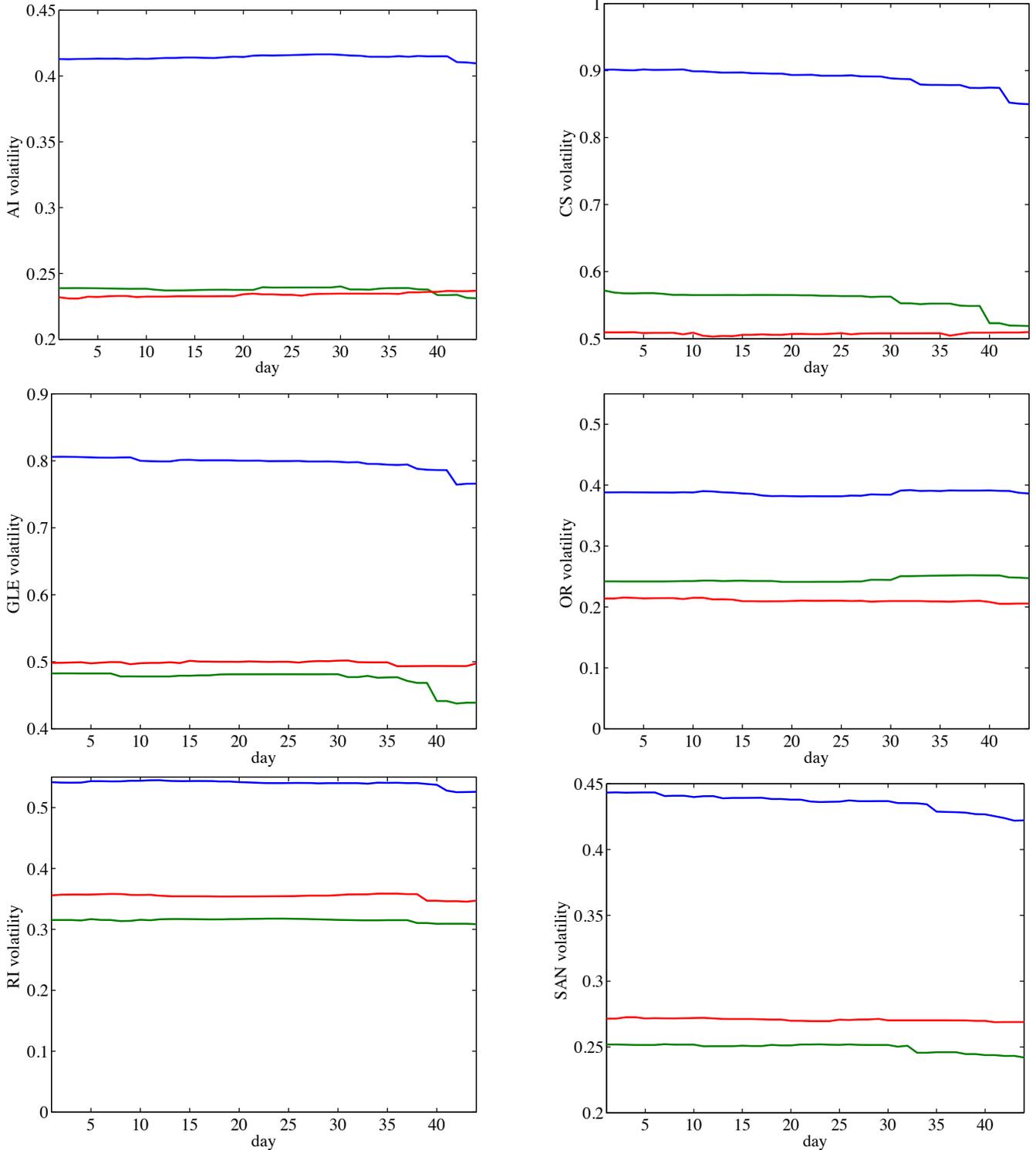


Figure 4: From left to right, top to bottom, we plot the volatilities for six companies listed in Section 5 in the following order (alphabetical in the Euronext symbols): AI, CS, GLE, OR, RI, and SAN. Each panel displays the volatilities σ , σ_+ and σ_- in blue, green and red respectively versus day i . These values σ , σ_+ , and σ_- are used to calculate the jump factors u , v and x , respectively. Recall that the jump factors d , w , and y are the reciprocals of u , v , and x , respectively. Note that in the MT model, volatility is assumed to be independent of the strike price and the expiration date of the option.

References

- [1] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy, 81 (1973), pp. 637–659.
- [2] A. Câmara and W. Li, *Jump-diffusion option pricing without IID jumps*, (2008), Preprint: <http://ssrn.com/abstract=1282882>.
- [3] I. Csiszár and P. C. Shields, *The consistency of the BIC Markov order estimator*, Annals of Statistics, 28 (2000), pp. 1601–1619.
- [4] G. D’Amico, J. Janssen and R. Manca, *European and American options: The semi-Markov case*, Physica A, 388 (2009), pp. 3181–3194.
- [5] J.-C. Duan and J.-G. Simonato, *American option pricing under GARCH by a Markov chain approximation*, Journal of Economic Dynamics and Control, 25 (2001), pp. 1689–1718.
- [6] I. Florescu and F. Viens, *Stochastic volatility: option pricing using a multinomial recombining tree*, Applied Mathematical Finance, 15 (2008), pp. 151–181.
- [7] K. R. Gabriel and J. Neumann, *A Markov chain model for daily rainfall occurrence at Tel-Aviv*, Quarterly Journal of the Royal Meteorological Society, 88 (1962), pp. 90–95.
- [8] J. C. Hull, *Options, Futures, and Other Derivatives*, 6th Ed., Pearson Education, Upper Saddle River, NJ, 2006.
- [9] J. Janssen, R. Manca and G. Di Biase, *Markov and semi-Markov option pricing models with arbitrage possibility*, Applied Stochastic Models and Data Analysis, 13 (1997), pp. 103–113.
- [10] G. McQueen and S. Thorley, *Are stock returns predictable? A test using Markov chains*, Journal of Finance, 46 (1991), pp. 239–263.
- [11] J. Primbs, M. Rathinam and Y. Yamada, *Option pricing with a pentanomial lattice model that incorporates skewness and kurtosis*, Applied Mathematical Finance, 14 (2007), pp. 1–17.
- [12] M. Rubinstein, *Implied binomial trees*, Journal of Finance, 49 (1994), pp. 771–818.
- [13] S. E. Shreve, *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*, Springer-Verlag, New York, 2004.
- [14] B. Tan and K. Yilmaz, *Markov chain test for time dependence and homogeneity: an analytical and empirical evaluation*, European Journal of Operational Research, 137 (2002), pp. 524–543.
- [15] Y. Yamada and J. Primbs, *Distribution based option pricing on lattice asset dynamics models*, International Journal of Theoretical and Applied Finance, 5 (2002), pp. 599–618.