

A Binary Tree, Dynamic Asset Pricing Model to Capture Moving Average and Autoregressive Behavior

Davide Lauria¹, Yuan Hu², W. Brent Lindquist^{3*}, and Svetlozar T. Rachev³

¹Department of Economics, Statistics and Finance, University of Calabria, Arcavacata di Rende, Italy

²Department of Mathematics, University of California San Diego, La Jolla CA 92093-0112, USA

³Department of Mathematics & Statistics, Texas Tech University, Lubbock TX 79409-1042, USA

*Corresponding author, brent.lindquist@ttu.edu

February 29, 2024

Abstract

We introduce a binary tree for pricing contingent claims when the underlying security prices exhibit history dependence, specifically moving average or autoregressive behavior, that is characteristic of price histories induced by market microstructure behavior. Our model is market-complete and arbitrage-free. When passing to the risk-neutral measure, the model preserves all parameters governing the natural world price dynamics, including the instantaneous mean of the asset return and the instantaneous probabilities for the direction of asset price movement. This preservation holds for arbitrarily small, but non-zero, time increments characteristic of market microstructure transactions. In the (unrealistic) limit of continuous trading, the model reduces to continuous diffusion price processes, with the concomitant loss of the microstructure information.

Keywords. binary trees, asset pricing, option pricing, market microstructure, technical analysis

1 Introduction

“At the level of transactions prices, ..., the random walk conjecture is ... a hypothesis that is very easy to reject in most markets even in small data samples. In microstructure, the question is not ‘whether’ transactions prices differ from a random walk, but rather ‘how much’ and ‘why?’ ” (Hasbrouck, 1996)

Dynamic asset pricing theory, as introduced¹ by Black and Scholes (1973) and Merton (1973) (BSM), is based on the concepts of no arbitrage opportunity and replicating portfolios, along with a set of assumptions that can be classified into two groups. The first group of assumptions concerns the microstructure of the market: the rules under which trades are performed; the impact of transaction and timing costs; the role of information and its disclosure; discovery and formation of prices; volatility; liquidity; market maker and investor behavior. Under the assumptions of the BSM model: any trade is executed without taxes, transaction costs, and amount restriction (the market is frictionless); traders are price takers with symmetric information (a perfectly competitive market) and are able to trade any amount (no liquidity constraints) over any infinitesimally small

¹See also the seminal works of Bachelier (1900).

time interval (continuous trading); the market is assumed to be efficient (all relevant information is embedded in the market price), liquid (every order is executed instantaneously at the current equilibrium price), and free of arbitrage opportunity. The second group of assumptions is related to the choice of geometric Brownian motion (GBM) as the stochastic process describing the price dynamics of the security underlying the option contract. The assumption of GBM invokes a strong set of restrictions, including constant volatility, the normal distribution of log returns, and absence of long-memory.²

The BSM model provides analytical solutions and an elegant machinery for computing the price of a European option; however many of the hypotheses upon which it is rooted have been shown to be too restrictive. It is well-known that many of the empirical properties³ of stock price returns are not consistent with the assumption of GBM. Consequently, a range of alternative models have been proposed to include various stylized facts. Another problem with the original BSM model is that it does not provide solutions for more complex contingent claims, such as those with a path dependent pay-off (e.g., American options) or for those whose underlying risk is not fully priced in financial markets, leading to market incompleteness. The most baffling result of the BSM model is that the option price does not depend on the drift of the underlying security. This *puzzle* was clarified in subsequent work of Cox and Ross (1976) and Merton (1976), and then reformulated in terms of the risk-neutral measure by Harrison and Kreps (1979) and Harrison and Pliska (1981). The concept of the risk-neutral price was subsequently accepted and continuous time models have proliferated.

Subsequent developments to the solution of the continuous-time pricing problem can be interpreted as improvements in one of two directions. The first is directed to the stochastic process driving the price dynamics in order to incorporate more statistical features of real price processes. This direction has produced the following strong result, known as the general version of the fundamental theorem of asset pricing (Delbaen and Schachermayer, 1994): “if a stochastic price process \mathcal{S} is a bounded, real-valued semimartingale, there is an equivalent martingale measure for \mathcal{S} if and only if \mathcal{S} satisfies the condition of *no free lunch with vanishing risk* (NFLVR)”, where NFLVR is a generalization of the no arbitrage condition.^{4,5} One consequence of this continuous-time, fundamental theorem is the necessity to work within the confines of stochastic integration theory.

The second direction is to develop methods for solving pricing problems having no known analytical solution, due either to the complexity of the stochastic process or the complexity of the pay-off function associated with a contingent claim. The binomial option pricing model proposed by Cox et al. (1979) (CRR) was the first approach to pricing American options without sacrificing the intellectual machinery developed under the BSM model. CRR utilized a discrete-time, binomial lattice graph to describe the evolution of the price process of the underlying security. The discrete process was designed to converge to GBM as the time interval between two successive trades converged to zero. There was no intention in the CRR model to use the discrete setting to incorporate other stylized facts of asset returns. Other discrete models – utilizing binomial or trinomial lattices, or binary trees – have been developed to numerically price contingent claims

²For the definition of a long-memory process, we refer the interested reader to Mandelbrot (2001) and Beran (2017).

³It is common to define these properties as “stylized facts”. These include: volatility clustering; returns with heavy tailed distributions; tail dependence; leverage effects; and long-term memory. See Mittnik et al. (2007) for a comprehensive exposition on the topic.

⁴It is worth noting that all Lévy processes are semimartingales, and many well-studied models in finance assume that the asset log-returns follow Lévy processes, as seen in Eberlein and Prause (2000); Schoutens (2003); Rachev et al. (2011).

⁵See Samura et al. (2013) for conditions under which real-valued, cadlag processes that satisfy NFLVR must be semimartingales.

under more complex assumptions, such as stochastic volatility or jump processes. (See for instance Boyle (1986); Rubinstein (1994); Derman et al. (1996), and Rubinstein (1998).) Again, these discrete models have been designed to converge to a solution of a continuous-time stochastic process. This is usually ensured by setting moment matching conditions in order to apply Donsker’s invariance principle (Billingsley, 2013). Using discrete models avoids working explicitly with stochastic integration theory.⁶

As noted above, the BSM model is very restrictive with regards to its incorporation of the details of market microstructure.⁷ In seminal work, Roll (1984) showed that, in an efficient market, the effective bid-ask spread can be measured by $spread = 2\sqrt{-cov}$, where cov is the first-order serial covariance of price changes. Crucially, Roll’s reasoning was based upon analysis of a discrete-time model. In his treatise on market microstructure, Hasbrouck (2007) describes several discrete-time empirical market microstructure models (which incorporate Roll’s bid-ask model). The models are designed to capture, in various ways, the price formation process, incorporating the sequence of action and reactions between market makers and traders. Paralleling the basis of Roll’s reasoning and the approach of Hasbrouck, we also adopt a discrete-time approach to handle the complexity of the stochastic processes involved.

The work of Kim et al. (2016, 2019) and Hu et al. (2020a,b) have shown that binomial pricing trees have sufficient flexibility to capture some of the stylized facts of price dynamics for option pricing in complete discrete-time markets. These include the preservation, from the natural world to the risk-neutral valuation, of: the probabilities of the natural-world stock-price directions; the mean and higher moments of returns; and the effects of noisy, informed, and misinformed traders. However *binomial* trees are too simplistic to accommodate either the autoregressive or moving average behavior of asset prices. *Our conclusion is that binary pricing trees are crucial for developing dynamic asset pricing models that incorporate such phenomena.*

To further clarify the need for binary pricing trees, recall the fundamental pricing model in continuous time for a market consisting of a single bond and stock. The continuous time bond price dynamics are given by

$$d\beta_t^{(cts)} = r_t^{(cts)}\beta_t^{(cts)}dt, \quad t \in [0, T], \quad (1)$$

where $\beta_0^{(cts)} = \beta_0 > 0$ and $r_t^{(cts)}$ is a continuous-time riskless rate (Duffie, 2001, p. 102). The stock’s log-price dynamics $L_t^{(cts)} = \ln(S_t^{(cts)})$ with $S_0^{(cts)} = S_0 > 0$, follow a continuous diffusion determined by the Itô process (Aït-Sahalia and Jacod, 2014, Chapter 1)

$$dL_t^{(cts)} = \mu_t dt + \sigma_t dB_t, \quad t \in [0, T], \quad (2)$$

where B_t , $t \in [0, T]$ is a standard Brownian motion whose trajectories generate a canonical filtered probability space (Duffie, 2001, Chapter 5 and Appendix E)

$$(\Omega, \mathbb{F}^{(cts)} = \{\mathcal{F}^{(cts)} = \sigma(B_u, 0 \leq u \leq t)\}, \mathbb{P}).$$

Inclusion of microstructure features modifies the stock log-price dynamics, which can be written in discrete form as

$$L_{t_n}^{(obs)} = L_{t_n}^{(cts)} + \epsilon_{t_n}^{(micro)}, \quad n = 1, \dots, m - 1, \quad (3)$$

⁶However answering questions related to the existence and uniqueness of solutions to the continuous time stochastic PDE that a numerical model may be intending to approximate does require the full machinery of stochastic integration theory.

⁷This term originates with the seminal paper by Garman (1976). See O’Hara (1997) for an extensive overview of market microstructure studies.

where $t_n = t_1, \dots, t_{m-1}$ indicate the times at which the microstructure features associated with a particular market “actor” (such as a trader) are realized.⁸ The microstructure dynamics $\epsilon_{t_n}^{(\text{micro})}$, $n = 1, \dots, m - 1$, are determined by (for example) a moving average process MA(q) of order $q = 0, 1, \dots$ (Mills, 2019, Chapter 3)

$$\epsilon_{t_n}^{(\text{micro})} = \sum_{k=0}^q \phi_k \zeta_{t_n-k}, \quad \phi_0 = 1, \quad \phi_k \in \mathbb{R}, \quad \phi_k \neq 0, \quad n = 1, \dots, m - 1. \quad (4)$$

Here $\zeta_{t_n-k} = 0$ when $k > n$, and ζ_{t_k} , $k = 0, \dots, m - 1$ are independent and identically distributed random variables with zero mean and specified variance. As can be seen, for example as diagrammed by O’Hara (1999, Figure 1), when general microstructure features are included in the observed log-prices, the recombining, binomial pricing tree is no longer an appropriate model for the stock price dynamics and the extension to a binary (i.e., non-recombining) pricing model must be introduced.

The fundamental asset pricing theorem of Delbaen and Schachermayer (1994) requires the ability to trade in continuous time. *Market microstructure phenomena occur at discrete times.* The resultant observed process (see e.g., (3)), being a combination of a semimartingale plus discrete-time microstructure noise, is therefore not a semimartingale and the fundamental asset pricing theorem cannot be applied. Cheridito (2003) showed that fractional Brownian motion can be used as a price process and still maintain NFLVR by restricting the class of trading strategies. Jarrow et al. (2009) extended this work to show that arbitrage-free price processes can be obtained without reliance on semimartingales provided continuous-time trading is not allowed (although the finite time-intervals can be arbitrarily small).

We present here a discrete binary tree approach that is general enough to reproduce the statistical properties of real prices and encompass a class of models that are used in microstructure theory.⁹ To this end, we apply the general approach of Hu et al. (2020a,b) to binary tree option pricing models.¹⁰ In Section 2 we develop a discrete binary tree (the *binary information tree*) supporting non-recombined random walks. In Sections 3 and 4, we describe discrete-time pricing on this tree for a market consisting of a riskless rate, a single bond and a stock. In Section 4.1 we demonstrate that the binary information tree captures moving average and autoregressive features of stock prices. Computation of the risk-neutral measure and pricing of options is discussed in Section 5. An empirical computation of European option prices is presented in Section 5.1.

In general, a random walk on a non-recombined binary tree (which is a particular case of an arbitrary branching process) will converge to a measure-valued diffusion (Skorokhod, 1997; Daley and Vere-Jones, 2003, 2008; Mitov et al., 2009). In Section 6 we show that, under the well-known Donsker-Prokhorov invariance principle (for constant instantaneous mean return and variance) or the Davydov-Rotar invariance principle (for time dependent instantaneous mean return and variance), the restrictions of these invariance principles unfortunately require that a non-recombined random walk approach a classical random walk, which, in the continuum limit, produces price processes such as as GBM (under the Donsker-Prokhorov invariance principle) or continuous diffusion (under the Davydov-Rotar invariance principle), resulting in the concomitant loss of the microstructure information.

⁸As noted in Section 2, we reserve the time points t_0 and t_m for the current time and the terminal time of an option, respectively.

⁹See Easley and O’Hara (1995, 2003); Hasbrouck (2007), Ait-Sahalia and Jacod (2014, Chapter 2), and Fan et al. (2016).

¹⁰We refer the interested reader to Dzhaparidze and van Zuijlen (1996), Shiryaev (1999, Section II), and Cordero et al. (2014). These papers extend the Cox et al. (1979) and Jarrow and Rudd (1982) binomial models to general binary pricing models, without preserving the information on upward stock price probability or mean stock returns. As discussed in Hu et al. (2020a,b), this is a considerable drawback when option trading is performed in discrete times.

As the price-direction probabilities are preserved in the risk-neutral dynamics of our model, in Section 7 we delve into a technical analysis of these probabilities for sequences of price change. We compute sequence probabilities empirically based upon daily closing prices of an ETF and for stocks comprising the Dow Jones Industrial average. Our results indicate the presence of market inefficiencies, even in this daily data.

2 A binary information tree for pricing

We define a discrete-time filtered probability space (discrete time stochastic basis) $\mathbb{ST}^{(d)} = (\Omega, \mathbb{F}^{(d)}, \mathbb{P})$ with the discrete filtration $\mathbb{F}^{(d)} = \{\mathcal{F}^{(n)}, n \in \mathcal{N}_0\}$, $\mathcal{N}_0 \stackrel{\text{def}}{=} \{0, 1, \dots\}$. The sigma-fields $\mathcal{F}^{(n)} = \sigma(\epsilon_n^{(k)}, k = 1, 2, \dots, 2^{n-1})$, $n \in \mathcal{N}$, $\mathcal{F}^{(0)} = \{\emptyset, \Omega\}$, are generated by a sequence of dependent binary random variables $\epsilon_n^{(k)}$, $k = 1, 2, \dots, K_n = 2^{n-1}$, $n \in \mathcal{N}$, $P(\epsilon_n^{(k)} = 1) = 1 - P(\epsilon_n^{(k)} = 0) \in (0, 1)$. The triangular array $\mathfrak{E} = (\epsilon_n^{(k)}, k = 1, 2, \dots, K_n, n \in \mathcal{N})$ of the binary random variables $\epsilon_n^{(k)}$ is defined on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ with predetermined joint distributions,

$$p \begin{pmatrix} m_n^{(1)}, \dots, m_n^{(K_n)} \\ \epsilon_n^{(1)}, \dots, \epsilon_n^{(K_n)} \end{pmatrix} = P(\epsilon_n^{(1)} = m_n^{(1)}, \dots, \epsilon_n^{(K_n)} = m_n^{(K_n)}),$$

$$m_n^{(k)} \in \{0, 1\}, \quad k = 1, \dots, K_n, \quad n \in \mathcal{N},$$

satisfying Kolmogorov's extension theorem (Oksendal, 2013, Theorem 2.1.5, p. 11). The probability space $(\Omega, \mathbb{F}, \mathbb{P})$ is a standard probability space; without loss of generality, we can assume it is the Lebesgue probability space $\Omega = [0, 1]$, $\mathbb{F} = \mathfrak{B}([0, 1])$, $\mathbb{P} = \text{Leb}([0, 1])$.¹¹ The probability distribution $p \begin{pmatrix} m_n^{(1)}, \dots, m_n^{(n)} \\ \epsilon_n^{(1)}, \dots, \epsilon_n^{(n)} \end{pmatrix}$, $m_n^{(k)} \in \{0, 1\}$, $k = 1, 2, \dots, K_n$, will be determined by the market microstructure features embedded in our pricing model.

We define the probability law for $\{\epsilon_n^{(k)}, k = 1, \dots, K_n, n \in \mathcal{N}\}$ sequentially.

$n = 0$. For $n = 0$, set $\epsilon_0^{(0)} = 0$, $\mathcal{E}_0 \stackrel{\text{def}}{=} \{\epsilon_0^{(0)}\}$, and $\mathcal{F}^{(0)} = \sigma(\mathcal{E}_0) = \{\emptyset, \Omega\}$.

$n = 1$. For $n = 1$, set $\mathcal{E}_1 \stackrel{\text{def}}{=} \{\epsilon_0^{(0)}, \epsilon_1^{(1)}\} = \{\mathcal{E}_0, \epsilon_1^{(1)}\}$. Then $\mathcal{F}^{(1)} = \sigma(\mathcal{E}_1)$ with

$$\mathbb{P}(\epsilon_1^{(1)} = 1) = p_1^{((0,1),(0,1))} \in (0, 1),$$

$$\mathbb{P}(\epsilon_1^{(1)} = 0) = p_1^{((0,1),(0,0))} = 1 - p_1^{((0,1),(0,1))}.$$

$n > 1$. The general case is as follows. (For additional clarity, the sequential definitions for $n = 2$ and 3 are provided in full in the appendix.) Set $\mathcal{E}_n \stackrel{\text{def}}{=} \left\{ \mathcal{E}_{n-1}, \left(\epsilon_n^{(1)}, \dots, \epsilon_n^{(2^{n-1})} \right) \right\}$. \mathcal{E}_n is the triangular array of binary random variables $\epsilon_m^{(k_m)}$, $m = 1, \dots, n$, $k_m = 1, \dots, 2^{m-1}$, with $\epsilon_0^{(0)} = 0$. Then

$$\mathcal{F}^{(n)} = \sigma(\mathcal{E}_n), \quad n = 0, 1, \dots, \quad \mathcal{F}^{(0)} = \sigma(\mathcal{E}_0) = \{\emptyset, \Omega\}, \quad (5)$$

¹¹See "Standard probability space", Encyclopedia of Mathematics, EMS Press (2001).

and

$$\begin{aligned}
\mathbb{P}\left(\epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}, 1\right) &= p_n \binom{\left((0, k_1, \dots, k_n), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}, 1\right)\right)}{\left(0, p_{n-1} \binom{\left((0, k_1, \dots, k_{n-1}), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right)}\right)}, \\
\mathbb{P}\left(\epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}, 0\right) &= p_n \binom{\left((0, k_1, \dots, k_n), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}, 0\right)\right)}{\left(0, p_{n-1} \binom{\left((0, k_1, \dots, k_{n-1}), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right)}\right)} \\
&= p_{n-1} \binom{\left((0, k_1, \dots, k_{n-1}), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right)}{\left(0, p_{n-1} \binom{\left((0, k_1, \dots, k_{n-1}), \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}, 1\right)\right)}\right)}.
\end{aligned} \tag{6}$$

The \mathcal{E}_n -conditional probabilities are

$$\begin{aligned}
\mathbb{P}\left(\epsilon_n^{(k_n)} = 1 \mid \mathcal{E}_{n-1} = \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right) &\in (0, 1), \\
\mathbb{P}\left(\epsilon_n^{(k_n)} = 0 \mid \mathcal{E}_{n-1} = \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right) \\
&= 1 - \mathbb{P}\left(\epsilon_n^{(k_n)} = 1 \mid \mathcal{E}_{n-1} = \left(0, \epsilon_1^{(k_1)}, \dots, \epsilon_{n-1}^{(k_{n-1})}\right)\right),
\end{aligned} \tag{7}$$

where $k_1 = 1$ and $k_n = \left(\epsilon_1^{(k_1)} \dots \epsilon_{n-1}^{(k_{n-1})}\right)_{10} + 1$ for $n > 1$.¹²

The $(n + 1)$ -tuple

$$\mathbb{L}_n = (0, l_1, \dots, l_j, \dots, l_n), \quad l_j = 1, \dots, 2^{j-1}, \quad j = 1, \dots, n, \tag{8}$$

together with the discrete filtration $\mathcal{F}^{(n)}$ and the conditional probabilities (7) defines a particular type of binary random tree which we designate as a *binary information tree to level n* , BIT_n .¹³ Specification of the $(n + 1)$ -tuple

$$\mathbb{M}_n = (0, m_1, \dots, m_j, \dots, m_n), \quad m_j = \{0, 1\}, \quad j = 1, \dots, n, \tag{9}$$

defines a unique event (a node)

$$\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} = \mathcal{E}_n^{((0, l_1, \dots, l_j, \dots, l_n), (0, m_1, \dots, m_j, \dots, m_n))} \tag{10}$$

on BIT_n where, in (10), $l_j = (m_0 m_1 \dots m_{j-1})_{10} + 1$, $j = 1, \dots, n$, $m_0 = 0$. Event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ occurs with probability¹⁴

$$\mathbb{P}\left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}\right) = p_n^{(\mathbb{L}_n, \mathbb{M}_n)} \in (0, 1). \tag{11}$$

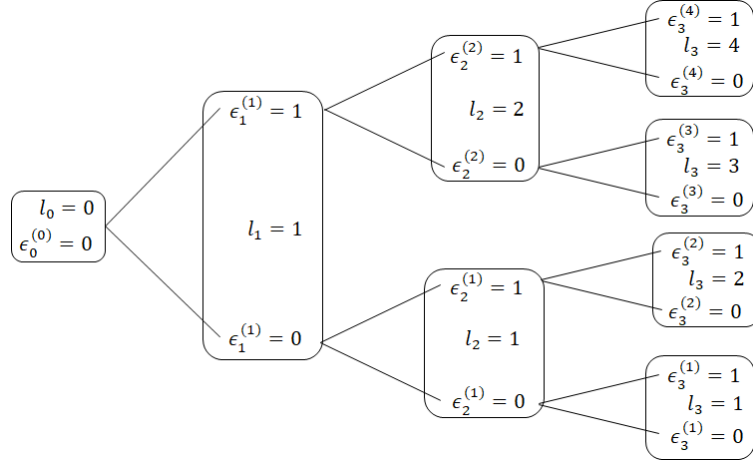
The binary tree indices l_n , $\epsilon_n^{(l_n)}$, and path labels, \mathbb{L}_n , \mathbb{M}_n , for $n = 0, \dots, 3$ on BIT_4 are illustrated in Fig. 1. The figure also provides an illustration of two specific probabilities of the form $p_3^{(\mathbb{L}_3, \mathbb{M}_3)}$.

To develop a random price time series simulating microstructure timing, for any given $m \in \mathcal{N} \stackrel{\text{def}}{=} \{1, 2, \dots\}$ we associate the levels of BIT_m with a sequence of time instances $0 = t_0 < t_1 < \dots < t_{m-1} < t_m = T$ over the finite period $[0, T]$, $T < \infty$. In our application to option pricing, the current time is $t_0 = 0$ (corresponding to the root event of BIT_m), while $t_m = T$ is the terminal time (corresponding to the leaf events of BIT_m). Trades of assets occur only at the times

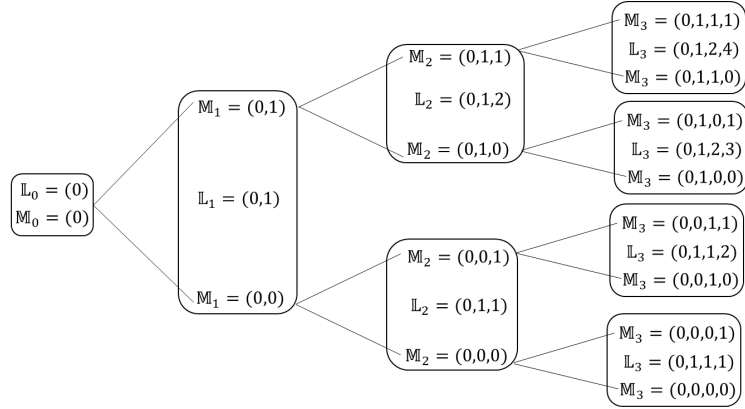
¹² $\left(\epsilon_1^{(k_1)} \dots \epsilon_{n-1}^{(k_{n-1})}\right)_{10}$ denotes the decimal value of the binary string $\epsilon_1^{(k_1)} \dots \epsilon_{n-1}^{(k_{n-1})}$.

¹³ Binary information trees are nested, $\text{BIT}_n \in \text{BIT}_m$ for $n < m$.

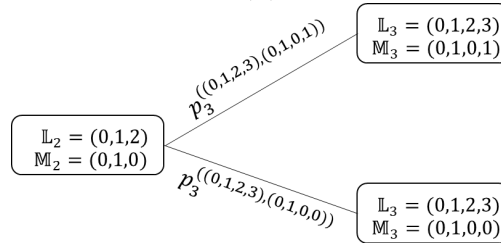
¹⁴ The family of probabilities $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$, $\mathbb{L}_n = (0, l_1, \dots, l_j, \dots, l_n)$, $l_j = 1, \dots, 2^{j-1}$, $\mathbb{M}_n = (0, m_1, \dots, m_j, \dots, m_n)$, $m_j = 0, 1$, $j = 1, \dots, n$, should satisfy Kolmogorov's extension theorem (see Oksendal (2013, Theorem 2.1.5, p. 11)), as illustrated in the cases for $n = 2, 3$ in the appendix.



(a)



(b)



(c)

Figure 1: Nomenclature used to label the binary information tree BIT_4 . (a) Local indices, l_n and $\epsilon_n^{(l_n)}$. (b) Trajectory labels, \mathbb{L}_n and \mathbb{M}_n , $n = 0, \dots, 3$. (c) An illustration of two of the probabilities $p_3^{(\mathbb{L}_3, \mathbb{M}_3)}$.

$t_1 < \dots < t_{m-1}$. These trading instances are fixed and known at time t_0 .¹⁵ The time intervals $[0, t_1)$, (t_n, t_{n+1}) , $n = 1, \dots, m-2$, and $(t_{m-1}, T]$ over which no trades occur are denoted inter-trade periods.¹⁶ We define $\Delta t_n \stackrel{\text{def}}{=} t_n - t_{n-1}$, $n = 1, \dots, m$.

Associated with BIT_m , for all $t \in [0, T]$ we can recursively define random sequences (*paths*) $\mathcal{E}_{n,t}^{(\mathbb{L}_n)}$, $n = 0, \dots, m-1$, as follows:

$$\begin{aligned} \mathcal{E}_{0,t}^{(\mathbb{L}_0)} &= \epsilon_0^{(0)} = 0 && \text{for } t \in [t_0 = 0, t_1), \\ \mathcal{E}_{n,t}^{(\mathbb{L}_n)} &= \mathcal{E}_{n-1,t}^{(\mathbb{L}_{n-1})}, \epsilon_n^{(l_n)} && \text{for } t \in [t_n, t_{n+1}), \quad 1 \leq n \leq m-2, \\ \mathcal{E}_{m-1,t}^{(\mathbb{L}_m)} &= \mathcal{E}_{m-2,t}^{(\mathbb{L}_{m-2})}, \epsilon_{m-1}^{(l_{m-1})} && \text{for } t \in [t_{m-1}, t_m = T]. \end{aligned} \quad (12)$$

We define the market information flow $\mathbb{IF}_{m;[0,T]}$ as

$$\mathbb{IF}_{m;[0,T]} = \left\{ \mathcal{E}_{n,t}^{(\mathbb{L}_n)}; \quad n = 0, \dots, m-1, \quad t \in [0, T], \quad \mathbb{L}_n \in \mathcal{L}_n \right\}, \quad (13)$$

where \mathcal{L}_n is the set of all $(n+1)$ -tuples, \mathbb{L}_n . $\mathbb{IF}_{m;[0,T]}$ generates the BIT_m *stochastic basis* $\mathbb{ST}_{(m;[0,T])}^{(d)} = (\Omega, \mathbb{F}_{m;[0,T]}^{(d)}, \mathbb{P})$ on $[0, T]$, where the filtration is defined by

$$\mathbb{F}_{m;[0,T]}^{(d)} = \left\{ \mathcal{F}_{0;[0,T]}^{(d)} = \{\emptyset, \Omega\}, \quad \mathcal{F}_{n;[0,T]}^{(d)} = \sigma \left\{ \mathcal{E}_{n,t}^{(\mathbb{L}_n)}, \mathbb{L}_n \in \mathcal{L}_n \right\}, \quad n = 0, \dots, m-1 \right\}. \quad (14)$$

Given BIT_m , specification of $\mathbb{M}_m = (0, m_1, \dots, m_n, m_{n+1}, \dots, m_m) \equiv (\mathbb{M}_n, m_{n+1}, \dots, m_m)$, $n = 0, \dots, m-1$, defines a unique, nested set of paths $\mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)}$, $n = 0, \dots, m-1$, $t \in [0, T]$, determined by the specifications $\mathbb{L}_n = (0, l_1, \dots, l_n)$, with $l_j = (m_0, m_1, \dots, m_{j-1})_{10} + 1$, $j = 1, \dots, n$, $m_0 = 0$, and $\epsilon_1^{(l_1)} = m_1, \dots, \epsilon_n^{(l_n)} = m_n$. From (12), this unique set of paths is

$$\begin{aligned} \mathcal{E}_{0,t}^{(\mathbb{L}_0, \mathbb{M}_0)} &= 0 && \text{for } t \in [t_0 = 0, t_1), \\ \mathcal{E}_{1,t}^{(\mathbb{L}_1, \mathbb{M}_1)} &= 0, m_1 && \text{for } t \in [t_1, t_2), \\ \mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)} &= \mathcal{E}_{n-1,t}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, m_n && \text{for } t \in [t_n, t_{n+1}), \quad 2 \leq n \leq m-2, \\ \mathcal{E}_{m-1,t}^{(\mathbb{L}_{m-1}, \mathbb{M}_{m-1})} &= \mathcal{E}_{m-2,t}^{(\mathbb{L}_{m-2}, \mathbb{M}_{m-2})}, m_{m-1} && \text{for } t \in [t_{m-1}, t_m = T]. \end{aligned} \quad (15)$$

The \mathcal{E}_{n-1} -conditional probabilities along path $\mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)}$ are

$$\mathbb{P} \left(\epsilon_j^{(l_j)} = m_j \mid \epsilon_1^{(l_1)} = m_1, \dots, \epsilon_{j-1}^{(l_{j-1})} = m_{j-1} \right), \quad j = 1, \dots, n. \quad (16)$$

The unconditional probability $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ for event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ on path $\mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)}$ is determined by the sequence of conditional probabilities (16).

As there is no trade at $t_m = T$, there are no new events at time t_m on BIT_m ; the last events occur at t_{m-1} . Similarly there is no event at $t_0 = 0$, the first events occur at t_1 . Thus on BIT_m , the events are labeled from $\mathcal{E}_1^{(\mathbb{L}_1, \mathbb{M}_1)}$ to $\mathcal{E}_{m-1}^{(\mathbb{L}_{m-1}, \mathbb{M}_{m-1})}$. From (10) and (15), we have the event-path equivalence

$$\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \stackrel{\text{EP}}{\sim} \mathcal{E}_{n,t_n}^{(\mathbb{L}_n, \mathbb{M}_n)}, \quad n = 1, \dots, m-1,$$

¹⁵While it may be of interest to view the trading times $t_1 < \dots < t_{m-1}$ as stopping times, this is beyond the scope of the current paper. However, such an extension can be done by introducing binary pricing models with dynamics following discrete-time semimartingales that are contaminated by noise occurring at random time instances. See, for example, Jacod and Protter (2012, Chapter 16) and (Aït-Sahalia and Jacod, 2014, Chapter 9).

¹⁶Note the assumption that no trade occurs at $t_0 = 0$ or at the terminal time $t_m = T$ (i.e., there is no new market information at times 0 and T).

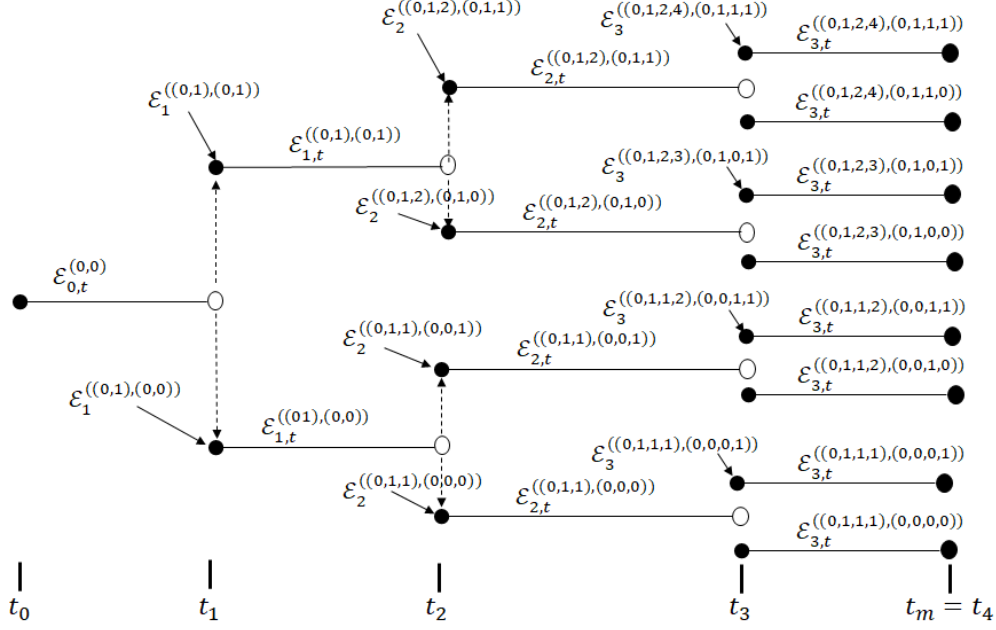


Figure 2: Event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ and path $\mathcal{E}_{n,t_n}^{(\mathbb{L}_n, \mathbb{M}_n)}$ labeling for BIT_4 . To preserve figure clarity, only half of the events are labeled at t_3 .

where $\overset{\text{EP}}{\sim}$ means that event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ occurs at time point t_n on path $\mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)}$. To preserve uniformity of notation on BIT_m , we define the pseudo-event $\mathcal{E}_0^{(\mathbb{L}_0, \mathbb{M}_0)} = \mathcal{E}_0^{(0,0)}$ for the “root” node, leading to the equivalence

$$\mathcal{E}_0^{(0,0)} \overset{\text{EP}}{\sim} \mathcal{E}_{0,t_0}^{(0,0)}.$$

As the “leaf” nodes at $t = t_m$ are path termination points, the terminating pseudo-events will be labeled using path notation $\mathcal{E}_{m-1,t_m}^{(\mathbb{L}_{m-1}, \mathbb{M}_{m-1})}$. We shall refer to all events, pseudo- or otherwise, simply as events. Event and path labels on BIT_4 are illustrated in Fig. 2.

For every fixed $m \in \mathcal{N}$, the set of all $\sum_{n=0}^{m-1} 2^n$ paths $\mathcal{E}_{n,t}^{(\mathbb{L}_n, \mathbb{M}_n)}$, $n = 0, \dots, m-1$, defines an $\mathbb{F}^{(d)}$ -adapted BIT_m , which we denote \mathbb{BT}_m . The probabilities $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ will represent the natural probabilities for the direction of stock movements. \mathbb{BT}_m provides the stochastic dynamics of the market information based on the time instances $0 = t_0 < t_1 < \dots < t_{m-1} < t_m = T$.

We consider a discrete market consisting of a riskless rate $r^{(d,f)}$, a riskless asset (bond)¹⁷ \mathcal{B} , a risky asset (stock) \mathcal{S} , and a European contingent claim¹⁸ \mathcal{C} and develop option pricing on \mathbb{BT}_m . In this discrete market setting, $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ will represent the probability of the direction of price changes at t_n . For example, given price $S_{t_{n-1}}$ corresponding to event $\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}$, then $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ with $\mathbb{L}_n = (\mathbb{L}_{n-1}, l_n)$, $\mathbb{M}_n = (\mathbb{M}_{n-1}, 1)$, and $l_n = (m_1 \dots m_{n-1} 1)_{10} + 1$ represents the probability of a price increase $S_{t_n} - S_{t_{n-1}} > 0$, while $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ with $\mathbb{M}_n = (\mathbb{M}_{n-1}, 0)$ and $l_n = (m_1 \dots m_{n-1} 0)_{10} + 1$ represents the probability of a price decrease $S_{t_n} - S_{t_{n-1}} < 0$.

¹⁷The asset \mathcal{B} may also be interpreted as a riskless bank account.

¹⁸For brevity, we shall refer to the European contingent claim as the option.

2.1 Estimation of probabilities

Assuming that a sufficient history ($t < 0$) of stock prices is available at $t = 0$, we utilize the historical frequency $\hat{p}_1^{((0,1),(0,1);\Delta t_1)}$ of positive price changes over trading periods of size Δt_1 as an estimator for $p_1^{((0,1),(0,1))}$. For example, if $\Delta t_1 = 5$ minutes, then $\hat{p}_1^{((0,1),(0,1);\Delta t_1)}$ is the proportion of positive stock returns observed in a historical sample of 5-minute returns.¹⁹ For $n > 1$, we use the historical frequency $\hat{p}_n^{(\mathbb{L}_n, \mathbb{M}_n; \Delta t_{1,n})}$ as an estimator for $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$, where $\Delta t_{1,n} \stackrel{\text{def}}{=} \Delta t_1, \dots, \Delta t_n$. As an example of the computation of $\hat{p}_n^{(\mathbb{L}_n, \mathbb{M}_n; \Delta t_{1,n})}$, assume equally spaced time intervals, $\Delta t_1 = \dots = \Delta t_n = \Delta t$, where $\Delta t = 1$ day. Consider a data set of \mathcal{T} historical daily returns partitioned into V non-overlapping time periods, each period consisting of n days. In each period, the succession of signs of the n daily returns is compared to the succession of signs implied by the \mathbb{M}_n indexing of the path $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$. If these two sign successions agree, the trial period is marked as a “success” (otherwise a “failure”). Thus the partitioning of the data provides a set of V Bernoulli trials from which to compute

$$\hat{p}_n^{(\mathbb{L}_n, \mathbb{M}_n; \Delta t_{1,n})} = \frac{\text{number of successes in } V \text{ trials}}{V}. \quad (17)$$

For values of $n \gtrsim 6$, a prohibitively extensive history $\mathcal{T} = Vn$ will be required to ensure adequate sampling. In Section 7 we consider the use of bootstrap resampling to provide adequate samples.

3 Riskless rate and bond price dynamics on \mathbb{BT}_m

We consider first the path-dependent dynamics of the riskless rate $r_t^{(d,f)}$ and \mathcal{B}_t on \mathbb{BT}_m . Corresponding to the discrete times t_n , $n = 1, \dots, m$, we consider the $\mathcal{F}^{(n-1)}$ -measurable riskless rates $r_{t_n; \mathcal{E}_{n-1}}^{(d,f)}$ and, without loss of generality, set $r_{t_0}^{(d,f)} = 0$.²⁰ For any time $t \in [0, T]$, we define the $\mathbb{ST}_{m; [0, T]}^{(d)}$ -adapted riskless rate $r_t^{(d,f)}$ as follows.

For $t \in [0, t_1)$, $r_t^{(d,f)} = r_{t_0}^{(d,f)} = 0$.

For $t \in [t_1, t_2)$, $r_t^{(d,f)} = r_{t_1; \mathcal{E}_0^{(0,0)}}^{(d,f)} \equiv r_{t_1; (\epsilon_0^{(0)}=0)}^{(d,f)}$.

For $t \in [t_2, t_3)$,

$$r_t^{(d,f)} = \begin{cases} r_{t_2; \mathcal{E}_1^{((0,1),(0,1))}}^{(d,f)} \equiv r_{t_2; (\epsilon_1^{(1)}=1)}^{(d,f)} & \text{w.p. } p_1^{((0,1),(0,1))}, \\ r_{t_2; \mathcal{E}_1^{((0,1),(0,0))}}^{(d,f)} \equiv r_{t_2; (\epsilon_1^{(1)}=0)}^{(d,f)} & \text{w.p. } p_1^{((0,1),(0,0))}. \end{cases}$$

¹⁹Hung and Swallow (1999) provide a robust test for sample proportions when the Bernoulli trials are dependent. Applying robust estimates for $p_1^{((0,1),(0,1))}$ does not make a significant difference in the numerical examples we consider because our sample size is relatively large and the dependence between the Bernoulli trials is weak.

²⁰The first trading date is t_1 . Thus, the first opportunity for the trader to make a deposit in the riskless bank is t_1 . The value of $r_{t_0; \mathcal{E}_0}^{(d,f)}$ is irrelevant for the trader; without loss of generality, we assume $r_{t_0; \mathcal{E}_0}^{(d,f)} = 0$.

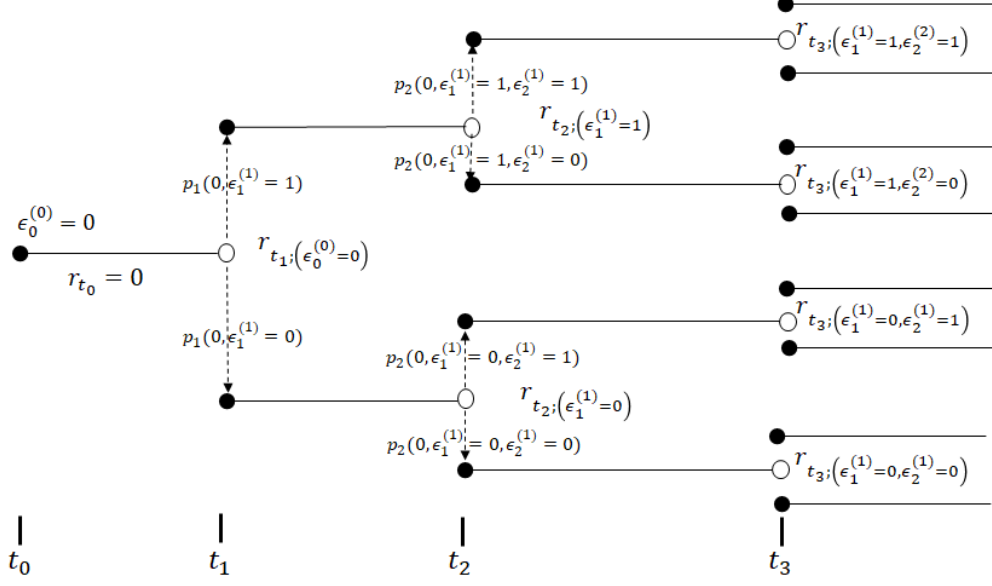


Figure 3: Illustration of the path dependence of the riskless rates $r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d,f)}$. For figure clarity, each riskless rate is indicated by the shortened form $r_{t_n; (\epsilon_1^{l_1} = m_1, \dots, \epsilon_n^{l_n} = m_n)}$.

For $t \in [t_3, t_4)$,

$$r_t^{(d,f)} = \begin{cases} r_{t_3; \mathcal{E}_2^{((0,1,2), (0,1,1))}}^{(d,f)} \equiv r_{t_3; (\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1)}^{(d,f)} & \text{w.p. } p_2^{((0,1,2), (0,1,1))}, \\ r_{t_3; \mathcal{E}_2^{((0,1,2), (0,1,0))}}^{(d,f)} \equiv r_{t_3; (\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0)}^{(d,f)} & \text{w.p. } p_2^{((0,1,2), (0,1,0))}, \\ r_{t_3; \mathcal{E}_2^{((0,1,1), (0,0,1))}}^{(d,f)} \equiv r_{t_3; (\epsilon_1^{(1)} = 0, \epsilon_2^{(2)} = 1)}^{(d,f)} & \text{w.p. } p_2^{((0,1,1), (0,0,1))}, \\ r_{t_3; \mathcal{E}_2^{((0,1,1), (0,0,0))}}^{(d,f)} \equiv r_{t_3; (\epsilon_1^{(1)} = 0, \epsilon_2^{(2)} = 0)}^{(d,f)} & \text{w.p. } p_2^{((0,1,1), (0,0,0))}. \end{cases}$$

In general, for path $\mathcal{E}_{n-1,t}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}$ and $t \in [t_n, t_{n+1})$, $n = 2, \dots, m-1$,

$$r_t^{(d,f)} = r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d,f)} \equiv r_{t_n; (\epsilon_1^{(l_1)} = m_1, \dots, \epsilon_{n-1}^{(l_{n-1})} = m_{n-1})}^{(d,f)} \text{ w.p. } p_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \quad (18)$$

with the understanding that, when $n = m-1$, for any path $\mathcal{E}_{m-2,t}^{(\mathbb{L}_{m-2}, \mathbb{M}_{m-2})}$, (18) holds for the closed time interval $t \in [t_{m-1}, T]$. Fig. 3 illustrates the path dependence of the riskless rates $r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d,f)}$.

For $t \in [t_n, t_{n+1})$, the riskless rate $r_t^{(d,f)}$ is $\mathcal{F}^{(n-1)}$ -measurable. This definition is consistent with the definition of the riskless rate (short rate) dynamics in continuous time (Duffie, 2001, p. 102). Without loss of generality, we can define the path dependent, instantaneous riskless rate $r_t^{(d,f,\text{inst})} > 0$, $t \in [t_n, t_{n+1})$ by the relation

$$r_t^{(d,f)} = r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d,f)} = r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d,f,\text{inst})} \Delta t_n. \quad (19)$$

Having determined the dynamics of the riskless rates $r_t^{(d,f)}$, $t \in [0, T]$, we now turn our attention to the dynamics of the riskless bond \mathcal{B} . The $\mathbb{S}\mathbb{T}_{m;[0,T]}^{(d)}$ -adapted bond price $\beta_t^{(d)}$, $t \in [0, T]$, is defined as follows.

For $t \in [0, t_1)$,

$$\beta_t^{(d)} = \beta_{t_0}^{(d)} = \beta_0. \quad (20)$$

For $t \in [t_1, t_2)$,

$$\beta_t^{(d)} = \beta_{t_1}^{(d)} = \beta_0 \left(1 + r_{t_1}^{(d,f)}\right). \quad (21)$$

For $t \in [t_2, t_3)$,

$$\begin{aligned} \beta_t^{(d)} &= \beta_{t_2}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2}^{(d,f)}\right) \\ &= \begin{cases} \beta_{t_2; (\epsilon_1^{(1)}=1)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=1)}^{(d,f)}\right) \text{w.p. } p_1^{((0,1),(0,1))}, \\ \beta_{t_2; (\epsilon_1^{(1)}=0)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=0)}^{(d,f)}\right) \text{w.p. } p_1^{((0,1),(0,0))}. \end{cases} \end{aligned} \quad (22)$$

For $t \in [t_3, t_4)$,

$$\begin{aligned} \beta_t^{(d)} &= \beta_{t_3}^{(d)} = \beta_{t_2}^{(d)} \left(1 + r_{t_3}^{(d,f)}\right) \\ &= \begin{cases} \beta_{t_3; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=1)}^{(d,f)}\right) \left(1 + r_{t_3; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)}^{(d,f)}\right) \text{w.p. } p_2^{((0,1,2),(0,1,1))}, \\ \beta_{t_3; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=1)}^{(d,f)}\right) \left(1 + r_{t_3; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)}^{(d,f)}\right) \text{w.p. } p_2^{((0,1,2),(0,1,0))}, \\ \beta_{t_3; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=1)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=0)}^{(d,f)}\right) \left(1 + r_{t_3; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=1)}^{(d,f)}\right) \text{w.p. } p_2^{((0,1,1),(0,0,1))}, \\ \beta_{t_3; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=0)}^{(d)} = \beta_{t_1}^{(d)} \left(1 + r_{t_2; (\epsilon_1^{(1)}=0)}^{(d,f)}\right) \left(1 + r_{t_3; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=0)}^{(d,f)}\right) \text{w.p. } p_2^{((0,1,1),(0,0,0))}. \end{cases} \end{aligned} \quad (23)$$

For $t \in [t_n, t_{n+1})$, $n = 2, \dots, m-1$, given path $\mathcal{E}_{n-1,t}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}$, the value of the bond $\beta_t^{(d)} = \beta_{t_n}^{(d)}$, $t \in [t_n, t_{n+1})$ is

$$\begin{aligned} \beta_t^{(d)} &= \beta_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} = \beta_{t_n; \{\epsilon_1^{(l_1)}=m_1, \dots, \epsilon_{n-1}^{(l_{n-1})}=m_{n-1}\}}^{(d)} \\ &= \beta_{t_{n-1}; \{\epsilon_1^{(l_1)}=m_1, \dots, \epsilon_{n-2}^{(l_{n-2})}=m_{n-2}\}}^{(d)} \left[1 + r_{t_n; \{\epsilon_1^{(l_1)}=m_1, \dots, \epsilon_{n-1}^{(l_{n-1})}=m_{n-1}\}}^{(d,f)}\right] \\ &= \beta_{t_1}^{(d)} \prod_{k=2}^n \left[1 + r_{t_k; \{\epsilon_1^{(l_1)}=m_1, \dots, \epsilon_{k-1}^{(l_{k-1})}=m_{k-1}\}}^{(d,f)}\right] \text{w.p. } p_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \end{aligned} \quad (24)$$

with the understanding that, when $n = m-1$, for any path $\mathcal{E}_{m-2,t}^{(\mathbb{L}_{m-2}, \mathbb{M}_{m-2})}$, (24) holds for the closed time interval $t \in [t_{m-1}, T]$.

For $[t_n, t_{n+1})$, the bond price $\beta_t^{(d)}$ is $F^{(n-1)}$ -measurable. This definition is consistent with the definition of the riskless bond price dynamics $\beta_t, t \in [0, T]$ in continuous time. More precisely, as shown in Section 6, in continuous time the bond price $\beta_t, t \in [0, T]$ is defined on the filtered probability space $(\Omega, \mathbb{F}_{[0, T]} = \{\mathcal{F}_t = \sigma(B_u, 0 \leq u \leq t)\}_{t \in [0, T]}, \mathbb{P})$, where $B_t, t \in [0, T]$, is a standard Brownian motion on $[0, T]$, and its price dynamics are determined by $d\beta_t = r_t \beta_t dt$, $\beta_0 > 0$, where the short rate $r_t \geq 0$ is \mathcal{F}_t measurable with $\mathbb{P}\left(\sup_{t \in [0, T]} \left\{r_t + \frac{1}{r_t}\right\} < \infty\right) = 1$. Thus, β_t is *instantaneously riskless*; in other words, $\beta_u = \beta_t$ for $u \in [t, t + dt)$. This definition of the riskless bond pricing in continuous time was the motivation to define riskless bond pricing in discrete time by equations (20) through (24).

4 Stock price dynamics on BIT_m

The price dynamics $S_t^{(d)}, t \in [0, T]$ of \mathcal{S} on the stochastic basis $\mathbb{S}\mathbb{T}_{m; t \in [0, T]}^{(d)}$ is an $\mathbb{F}_{m; t \in [0, T]}^{(d)}$ -adapted process. We define $S_t^{(d)}, t \in [0, T]$ on $\mathbb{S}\mathbb{T}_{m; t \in [0, T]}^{(d)}$ as follows.

For $t \in [0, t_1)$,

$$S_t^{(d)} = S_{t_0}^{(d)} = S_0 > 0 \quad (25)$$

For $t \in [t_1, t_2)$,

$$S_t^{(d)} = S_{t_1}^{(d)} = \begin{cases} S_{t_1}^{(d)} = S_0 s_{t_1; (\epsilon_1^{(1)}=1)} \text{ w.p. } p_1^{((0,1),(0,1))} \\ S_{t_1}^{(d)} = S_0 s_{t_1; (\epsilon_1^{(1)}=0)} \text{ w.p. } p_1^{((0,1),(0,0))} \end{cases} \quad (26)$$

for values $s_{t_1; (\epsilon_1^{(1)})} > 0$. Let

$$r_{t_1}^{(d)} = \frac{S_{t_1}^{(d)} - S_0}{S_0} = \begin{cases} s_{t_1; (\epsilon_1^{(1)}=1)} - 1 \text{ w.p. } \mathbb{P}(\epsilon_1^{(1)} = 1), \\ s_{t_1; (\epsilon_1^{(1)}=0)} - 1 \text{ w.p. } \mathbb{P}(\epsilon_1^{(1)} = 0), \end{cases} \quad (27)$$

be the discrete (arithmetic) return of the stock at t_1 . We assume that the mean $\mathbb{E}[r_{t_1}^{(d)}]$ and variance $\text{Var}[r_{t_1}^{(d)}]$ of the return

$$\begin{aligned} \mathbb{E}[r_{t_1}^{(d)}] &= \mu_{t_1}^{(r)} \Delta t_1, & \text{for some } \mu_{t_1}^{(r)} > r_{t_1}^{(d, f, \text{inst})}, \\ \text{Var}[r_{t_1}^{(d)}] &= (\sigma_{t_1}^{(r)})^2 \Delta t_1, & \text{for some } \sigma_{t_1}^{(r)} > 0, \end{aligned} \quad (28)$$

are known (estimated from observed stock prices). In (28), we interpret $\mu_{t_1}^{(r)}$ as the instantaneous mean return and $(\sigma_{t_1}^{(r)})^2$ as the instantaneous variance at t_1 . The moment conditions (28) imply that the constants $s_{t_1; (\epsilon_1^{(1)})}$ in (27) are determined by

$$\begin{aligned} s_{t_1; (\epsilon_1^{(1)}=1)} &= 1 + \mu_{t_1}^{(r)} \Delta t_1 + \sigma_{t_1}^{(r)} \sqrt{\frac{\mathbb{P}(\epsilon_1^{(1)}=0)}{\mathbb{P}(\epsilon_1^{(1)}=1)}} \Delta t_1, \\ s_{t_1; (\epsilon_1^{(1)}=0)} &= 1 + \mu_{t_1}^{(r)} \Delta t_1 - \sigma_{t_1}^{(r)} \sqrt{\frac{\mathbb{P}(\epsilon_1^{(1)}=1)}{\mathbb{P}(\epsilon_1^{(1)}=0)}} \Delta t_1. \end{aligned} \quad (29)$$

For $t \in [t_2, t_3)$,

$$S_t^{(d)} = S_{t_2}^{(d)} = \begin{cases} S_{t_2; \mathcal{E}_2^{((0,1,2),(0,1,1))}}^{(d)} = S_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)}^{(d)} & \text{w.p. } p_1^{((0,1,2),(0,1,1))}, \\ S_{t_2; \mathcal{E}_2^{((0,1,2),(0,1,0))}}^{(d)} = S_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)}^{(d)} & \text{w.p. } p_1^{((0,1,2),(0,1,0))}, \\ S_{t_2; \mathcal{E}_2^{((0,1,1),(0,0,1))}}^{(d)} = S_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=1)}^{(d)} & \text{w.p. } p_1^{((0,1,1),(0,0,1))}, \\ S_{t_2; \mathcal{E}_2^{((0,1,1),(0,0,0))}}^{(d)} = S_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=0)}^{(d)} & \text{w.p. } p_1^{((0,1,1),(0,0,0))}. \end{cases} \quad (30)$$

From (26) and (29),

$$S_{t_2}^{(d)} = \begin{cases} S_{t_1; (\epsilon_1^{(1)}=1)} S_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)} & \text{when } \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1, \\ S_{t_1; (\epsilon_1^{(1)}=1)} S_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} & \text{when } \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0, \\ S_{t_1; (\epsilon_1^{(1)}=0)} S_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=1)} & \text{when } \epsilon_1^{(1)} = 0, \epsilon_2^{(2)} = 1, \\ S_{t_1; (\epsilon_1^{(1)}=0)} S_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=0)} & \text{when } \epsilon_1^{(1)} = 0, \epsilon_2^{(2)} = 0, \end{cases} \quad (31)$$

for values $s_{t_2; (\epsilon_1^{(1)}, \epsilon_2^{(2)})} > 0$. Let

$$\begin{aligned} r_{t_2}^{(d)} \Big|_{(\epsilon_1^{(1)}=1)} &= \frac{S_{t_2}^{(d)} \Big|_{(\epsilon_1^{(1)}=1)} - S_{t_1}^{(d)} \Big|_{(\epsilon_1^{(1)}=1)}}{S_{t_1}^{(d)} \Big|_{(\epsilon_1^{(1)}=1)}} = \begin{cases} \frac{S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)} - S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)}}{S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)}} \\ S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} - S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} \\ \frac{S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} - S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)}}{S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)}} \\ S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} - 1 \text{ w.p. } \mathbb{P}(\epsilon_2^{(2)} = 1 | \epsilon_1^{(1)} = 1), \\ S_{t_1; (\epsilon_1^{(1)}=1)}^{(d)} s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} - 1 \text{ w.p. } \mathbb{P}(\epsilon_2^{(2)} = 0 | \epsilon_1^{(1)} = 1), \end{cases} \end{aligned}$$

be the conditional arithmetic return at t_2 given that $\epsilon_1^{(1)} = 1$. We assume that the conditional mean and conditional variance

$$\begin{aligned} \mathbb{E} \left[r_{t_2}^{(d)} \Big|_{(\epsilon_1^{(1)}=k)} \right] &= \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)} \Delta t_2, & \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)} &> r_{t_2; (\epsilon_1^{(1)}=k)}^{(d, f, \text{inst})}, \\ \text{Var} \left[r_{t_2}^{(d)} \Big|_{(\epsilon_1^{(1)}=k)} \right] &= \left(\sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)} \right)^2 \Delta t_2, & \sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)} &> 0, \end{aligned} \quad (32)$$

are known. In (32), $\mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)}$ is the instantaneous conditional mean return and $\left(\sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=k)} \right)^2$ is the instantaneous conditional variance of the return at t_2 given $\epsilon_1^{(1)} = k$. The moment conditions

(32) imply that, in (31),

$$\begin{aligned}
s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=1)} &= 1 + \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=1)} \Delta t_2 + \sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=1)} \sqrt{\frac{\mathbb{P}(\epsilon_2^{(2)}=0 | \epsilon_1^{(1)}=1)}{\mathbb{P}(\epsilon_2^{(2)}=1 | \epsilon_1^{(1)}=1)}} \Delta t_2, \\
s_{t_2; (\epsilon_1^{(1)}=1, \epsilon_2^{(2)}=0)} &= 1 + \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=1)} \Delta t_2 - \sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=1)} \sqrt{\frac{\mathbb{P}(\epsilon_2^{(2)}=1 | \epsilon_1^{(1)}=1)}{\mathbb{P}(\epsilon_2^{(2)}=0 | \epsilon_1^{(1)}=1)}} \Delta t_2, \\
s_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=1)} &= 1 + \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=0)} \Delta t_2 + \sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=0)} \sqrt{\frac{\mathbb{P}(\epsilon_2^{(2)}=0 | \epsilon_1^{(1)}=0)}{\mathbb{P}(\epsilon_2^{(2)}=1 | \epsilon_1^{(1)}=0)}} \Delta t_2, \\
s_{t_2; (\epsilon_1^{(1)}=0, \epsilon_2^{(2)}=0)} &= 1 + \mu_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=0)} \Delta t_2 - \sigma_{t_2}^{(r)} \Big|_{(\epsilon_1^{(1)}=0)} \sqrt{\frac{\mathbb{P}(\epsilon_2^{(2)}=1 | \epsilon_1^{(1)}=0)}{\mathbb{P}(\epsilon_2^{(2)}=0 | \epsilon_1^{(1)}=0)}} \Delta t_2.
\end{aligned} \tag{33}$$

For $t \in [t_n, t_{n+1})$, $n = 2, \dots, m-1$, given the path $\mathcal{E}_{n-1, t}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}$, the stock price is

$$\begin{aligned}
S_t^{(d)} &= S_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} \\
&= \begin{cases} S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=1)}^{(d)} = S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=1)}, \\ S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=0)}^{(d)} = S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=0)}, \end{cases} \tag{34}
\end{aligned}$$

for values $s_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=k)}$, $k = 0, 1$. Again, it is understood that, when $n = m-1$, for any path $\mathcal{E}_{m-2, t}^{(\mathbb{L}_{m-2}, \mathbb{M}_{m-2})}$, (34) holds for the closed time interval $t \in [t_{m-1}, T]$. Let

$$\begin{aligned}
r_{t_n}^{(d)} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}} &= \frac{S_{t_n}^{(d)} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}} - S_{t_{n-1}}^{(d)} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}}{S_{t_{n-1}}^{(d)} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}} \\
&= \begin{cases} \frac{S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=1)} - S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)}}{S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)}} \stackrel{\text{def}}{=} r_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n=(\mathbb{M}_{n-1}, 1))}}^{(d)} \\ S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=0)} - S_{t_{n-1}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} \stackrel{\text{def}}{=} r_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n=(\mathbb{M}_{n-1}, 0))}}^{(d)} \end{cases} \\
&= \begin{cases} s_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=1)} - 1 & \text{w.p. } \mathbb{P}(\epsilon_n^{(l_n)} = 1 \mid \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}), \\ s_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=0)} - 1 & \text{w.p. } \mathbb{P}(\epsilon_n^{(l_n)} = 0 \mid \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}), \end{cases} \tag{35}
\end{aligned}$$

be the conditional arithmetic return at t_n given the path $\mathcal{E}_{n-1, t}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}$.²¹ We assume the conditional

²¹We note the equivalent notations $S_{t_n; (\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)}=1)}^{(d)} = S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n=(\mathbb{M}_{n-1}, 1))}}^{(d)}$,

mean and conditional variance of $r^{(d)}$
 $t_n \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right.$,

$$\begin{aligned} \mathbb{E} \left[r_{t_n}^{(d)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right] &= \mu_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \Delta t_n, \\ \text{Var} \left[r_{t_n}^{(d)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right] &= \left(\sigma_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)^2 \Delta t_n, \end{aligned} \quad (36)$$

are known for some

$$\mu_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. > r_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d, f, \text{inst})}, \quad \sigma_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. > 0.$$

Conditions (36) imply that, in (34),

$$\begin{aligned} s_{t_n; \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 1 \right)} &= 1 + \mu_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \Delta t_n \\ &\quad + \sigma_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \sqrt{\frac{\mathbb{P} \left(\epsilon_n^{(l_n)} = 0 \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)}{\mathbb{P} \left(\epsilon_n^{(l_n)} = 1 \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)}} \Delta t_n, \\ s_{t_n; \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 0 \right)} &= 1 + \mu_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \Delta t_n \\ &\quad - \sigma_{t_n}^{(r)} \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \sqrt{\frac{\mathbb{P} \left(\epsilon_n^{(l_n)} = 1 \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)}{\mathbb{P} \left(\epsilon_n^{(l_n)} = 0 \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)}} \Delta t_n. \end{aligned} \quad (37)$$

The riskless rate $r_{t_n}^{(d, f)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right)$, bond price $\beta_{t_n}^{(d)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right)$, stock price $S_{t_n}^{(d)} \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right)$, constant $s_{t_n; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right)}$, return $r_{t_n}^{(d)} \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right)$, and the moments, $\mu_{t_n}^{(r)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right)$ and $\sigma_{t_n}^{(r)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right)$, corresponding to each of two events, $\mathcal{E}_n^{(\mathbb{L}_n, (\mathbb{M}_{n-1}, m_n=1))}$ and $\mathcal{E}_n^{(\mathbb{L}_n, (\mathbb{M}_{n-1}, m_n=0))}$, are illustrated in Fig. 4. The bond price, riskless rate and the two moments are the same for each of the two events.

Stock price dynamics in the natural world are determined by (34) and (36) which contain the following sets of model parameters: $\mathcal{I}^{(\mathbb{P})}$ - the probabilities for stock upward direction, $\mathbb{P} \left(\epsilon_n^{(l_n)} = 1 \left| \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right. \right)$; $\mathcal{I}^{(\mu)}$ - the conditional means $\mu_{t_n}^{(r)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 1 \right)$; and $\mathcal{I}^{(\sigma)}$ - the conditional variances $\sigma_{t_n}^{(r)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})} \right)$ on each node of the pricing tree. In Section 5, we show that the risk-neutral tree dynamics of the stock preserves $\mathcal{I}^{(\mathbb{P})}, \mathcal{I}^{(\mu)}, \mathcal{I}^{(\sigma)}$. This is important given that the information on $\mathcal{I}^{(\mathbb{P})}$ and $\mathcal{I}^{(\mu)}$ is lost in passing to the continuous time limit in the natural world and then using Black-Scholes-Merton risk-neutral valuation. Using a discrete pricing tree rather than a continuous time pricing model allows us to introduce richer, more flexible models for the price dynamics to accommodate market-microstructure features in option pricing.

$$s_{t_n; \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 1 \right)} = s_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n = (\mathbb{M}_{n-1}, 1))}}, \quad r_{t_n}^{(d)} \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 1 \right) = r_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n = (\mathbb{M}_{n-1}, 1))}}^{(d)}, \text{ etc.}$$

$\mathbb{L}_{n-1} = (0, l_1, l_2, \dots, l_{n-1})$	$\mathbb{M}_{n-1} = (0, m_1, m_2, \dots, m_{n-1})$	$\mathbb{M}_n^{(0)} = (\mathbb{M}_{n-1}, m_n = 0)$
$\mathbb{L}_n = (\mathbb{L}_{n-1}, l_n)$	$\mathbb{M}_n^{(1)} = (\mathbb{M}_{n-1}, m_n = 1)$	

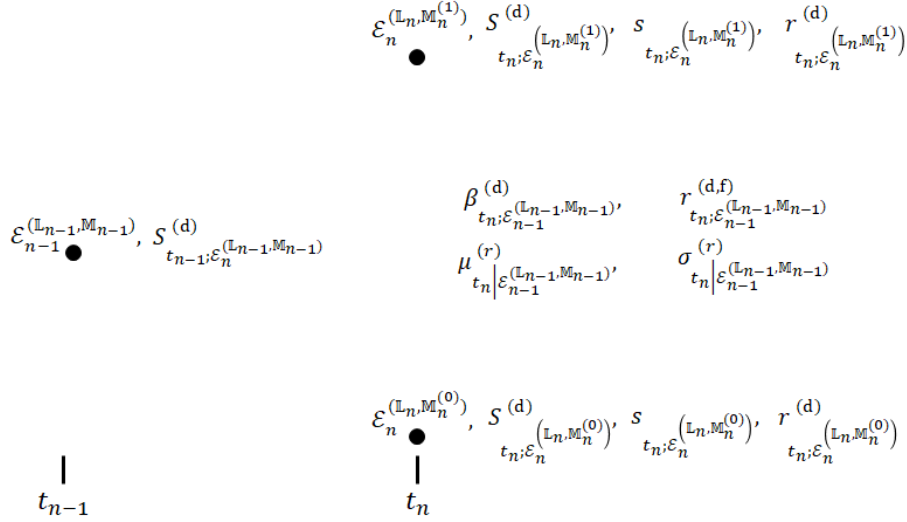


Figure 4: Illustration of the riskless rate $r^{(d,f)}$, bond price $\beta^{(d)}$, stock price $S^{(d)}$, constant s , return $r^{(d)}$, and moments $\mu^{(r)}$ and $\sigma^{(r)}$, at time t_n corresponding to each of two events, $\mathcal{E}_n^{(\mathbb{L}_n, (\mathbb{M}_{n-1}, m_n=1))}$ and $\mathcal{E}_n^{(\mathbb{L}_n, (\mathbb{M}_{n-1}, m_n=0))}$.

4.1 Stock pricing - special cases

The binary tree pricing model presented here encompasses several time series processes that have been proposed to model stock prices. We consider three such processes, *all of which assume constant time spacing*, $\Delta t_n = \Delta t$, $n = 1, \dots, m$. These examples include the case in which the first difference of the price process is assumed to display either moving average or autoregressive behavior; time-discrete models that are of particular importance in microstructure theory. Examples in the literature include the seminal Roll model (Roll, 1984) used for empirical estimation of the bid-ask spread, and the subsequent models of Hasbrouck (1988); see also Hasbrouck (2007, Chapter 8) for an overview and discussion of other closely related models.

Binary white noise. Consider the case in which the conditional probabilities at each node of BIT_m are the same:

$$\left. \begin{aligned} \mathbb{P} \left(\epsilon_j^{(l_j)} = 1 \mid \mathcal{E}_{j-1} = (0, \epsilon_j^{(l_1)} = m_1, \dots, \epsilon_{j-1}^{(l_{j-1})} = m_{j-1}) \right) &= p_1 \\ \mathbb{P} \left(\epsilon_j^{(l_j)} = 0 \mid \mathcal{E}_{j-1} = (0, \epsilon_j^{(l_1)} = m_1, \dots, \epsilon_{j-1}^{(l_{j-1})} = m_{j-1}) \right) &= p_0 = 1 - p_1 \end{aligned} \right\}, \quad j = 1, \dots, K_j. \quad (38)$$

Under this assumption, we define *binary white noise* (BWN) on BIT_m as the $\mathbb{F}^{(d)}$ -adapted process

$$\begin{aligned} z_0^{(d)} &= z_{t_0}^{(d)} = 0, \\ z_t^{(d)} &= z_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} \stackrel{\text{def}}{=} \begin{cases} z_{t_n; \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 1 \right)}^{(d)} & := z_u, \\ z_{t_n; \left(\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}, \epsilon_n^{(l_n)} = 0 \right)}^{(d)} & := z_d, \end{cases} \quad n = 1, \dots, m-1. \end{aligned} \quad (39)$$

As the process (39) on BIT_m is path independent, the notation for the BWN process can be simplified to $z_{t_n; \mathcal{E}_{n-1}}^{(d), (L_{n-1}, M_{n-1})} \equiv z_{t_n}^{(l_n)}$, where l_n takes on the values $k = 1, \dots, K_n$. Imposing (38) and the moment matching conditions $\mathbb{E} [z_{t_n}^{(l_n)}] = 0$ and $\text{Var} [z_{t_n}^{(l_n)}] = \sigma_z^2$, $\sigma_z^2 \in (0, \infty)$, (39) becomes

$$\begin{aligned} z_t^{(d)} &= z_0^{(0)} = 0, & t \in [0, t_1 = \Delta t), \\ z_t^{(d)} &= z_{t_n}^{(l_n)} = \begin{cases} z_u := \sqrt{\frac{1-p_1}{p_0}} \sigma_z^2 \text{ w.p. } p_1, \\ z_d := \sqrt{\frac{1-p_0}{p_1}} \sigma_z^2 \text{ w.p. } p_0 = 1 - p_1, \end{cases} & t \in [t_n, t_{n+1} = t_n + \Delta t), \\ & & l_n = 1, \dots, K_n, \quad n = 1, \dots, m - 1. \end{aligned} \quad (40)$$

It is straightforward to show that the BWN process $z_t^{(d)}$ has the mean $\mathbb{E} [z_{t_n}^{(l_n)}] = 0$, variance $\text{Var} [z_{t_n}^{(l_n)}] = \sigma_z^2$, and first-order autocorrelation $\mathbb{E} [z_{t_n}^{(l_n)} z_{t_{n-1}}^{(l_{n-1})}] = 0$ properties of a general white noise process.²²

Moving Average of order 1. We define the random process $\Delta S_t^{(d)} = S_t^{(d)} - S_{t-\Delta t}^{(d)}$ to be a *binary moving average process of order one* by setting

$$\Delta S_t^{(d)} = \Delta S_{t_n; \mathcal{E}_n^{((L_{n-1}, l_n), (M_{n-1}, m_n))}}^{(d)} \stackrel{\text{def}}{=} \begin{cases} c + \theta z_{t_{n-1}}^{(l_{n-1})} + z_u, & \text{if } m_n = 1, \\ c + \theta z_{t_{n-1}}^{(l_{n-1})} + z_d, & \text{if } m_n = 0, \end{cases} \quad (41)$$

for $t \in [t_n, t_{n+1})$, $n = 1, \dots, m - 1$, and values $c, \theta \in \mathbb{R}^+$. In (41), $z_{t_{n-1}}^{(l_{n-1})}$ is the BWN process (40). The expected value, variance, and first-order autocorrelation of $\Delta S_t^{(d)}$ follows the usual MA(1) process; for $s < n$, $s \in \mathcal{N}$,

$$\mathbb{E} [\Delta S_{t_n}^{(d)}] = c, \quad \text{Var} [\Delta S_{t_n}^{(d)}] = (1 + \theta^2) \sigma_z^2, \quad \mathbb{E} [S_{t_s}^{(d)} S_{t_n}^{(d)}] = \begin{cases} \theta \sigma_z^2, & \text{if } s = n - 1, \\ 0, & \text{otherwise.} \end{cases}$$

In practice, the coefficients θ and σ_z of the MA(1) process can be estimated from observed price differences. Let $\hat{\gamma}_0$ and $\hat{\gamma}_1$ denote, respectively, the empirical variance and first-order autocorrelation, of the risky asset price differences. Setting

$$\hat{\gamma}_0 = (1 + \theta^2) \sigma_z^2, \quad \hat{\gamma}_1 = \theta \sigma_z^2,$$

and solving for θ and σ_z^2 gives the solution pairs θ_+, σ_{z+}^2 and θ_-, σ_{z-}^2 ,

$$\begin{aligned} \theta_+ &= \frac{\hat{\gamma}_0 + \sqrt{\hat{\gamma}_0^2 - 4\hat{\gamma}_1^2}}{2\hat{\gamma}_1}, & \sigma_{z+}^2 &= \frac{2\hat{\gamma}_1^2}{\hat{\gamma}_0 + \sqrt{\hat{\gamma}_0^2 - 4\hat{\gamma}_1^2}}, \\ \theta_- &= \frac{\hat{\gamma}_0 - \sqrt{\hat{\gamma}_0^2 - 4\hat{\gamma}_1^2}}{2\hat{\gamma}_1}, & \sigma_{z-}^2 &= \frac{2\hat{\gamma}_1^2}{\hat{\gamma}_0 - \sqrt{\hat{\gamma}_0^2 - 4\hat{\gamma}_1^2}}, \end{aligned} \quad (42)$$

having the properties $\theta_+ \theta_- = 1$ and $\sigma_{z+}^2 \sigma_{z-}^2 = \hat{\gamma}_1^2$. Equations (42) require the condition $|\hat{\gamma}_0 / (2\hat{\gamma}_1)| \geq 1$, which also guarantees $\sigma_{z\pm}^2 > 0$. Addition of the constraint $|\theta| < 1$ to guarantee invertibility of the MA(1) process restricts the solution pair to θ_-, σ_{z-}^2 .

²²See, for instance, Hamilton (2020, Chapter 3).

Autoregressive of order 1. The BNW process can also be used to model the first difference of the price process as autoregressive of the first order, AR(1). Let

$$\Delta S_t^{(d)} = \begin{cases} \Delta S_{t_1; \mathcal{E}_1^{(l_1, m_1)}}^{(d)} & \stackrel{\text{def}}{=} \begin{cases} c + z_u, & \text{if } m_1 = 1, \\ c + z_d, & \text{if } m_1 = 0, \end{cases} & \text{for } t \in [t_1, t_2), \\ \Delta S_{t_n; \mathcal{E}_n^{(L_{n-1}, l_n), (M_{n-1}, m_n)}}^{(d)} & \stackrel{\text{def}}{=} \begin{cases} c + \phi \Delta S_{t_{n-1}; \mathcal{E}_{n-1}^{(L_{n-1}, M_{n-1})}}^{(d)} + z_u, & \text{if } m_n = 1, \\ c + \phi \Delta S_{t_{n-1}; \mathcal{E}_{n-1}^{(L_{n-1}, M_{n-1})}}^{(d)} + z_d, & \text{if } m_n = 0, \end{cases} & \text{for } t \in [t_n, t_{n+1}), \quad n = 2, \dots, m-1, \end{cases} \quad (43)$$

where $c, \phi \in \mathbb{R}$. Requiring $\phi \in (-1, 1)$ ensures that the process has MA(∞) representation. The expected value, variance, and autocorrelation functions of ΔS_{t_n} are

$$\begin{aligned} \mathbb{E}[\Delta S_{t_n}] &= c(1 + \phi + \phi^2 + \dots + \phi^n), \\ \text{Var}[\Delta S_{t_n}] &= \sigma_z^2 \sum_{i=0}^n \phi^{2i} = \sigma_z^2 \frac{1 - \phi^{2(n+1)}}{1 - \phi^2}, \\ \mathbb{E}[\Delta S_{t_k} \Delta S_{t_n}] &= \sigma_z^2 \phi^{n-k} \sum_{i=0}^{k-1} \phi^{2i} = \sigma_z^2 \phi^{n-k} \frac{1 - \phi^{2k}}{1 - \phi^2}, \end{aligned}$$

for $k < n$, $k \in \mathcal{N}$.

5 Risk-neutral dynamics on BIT_m : Option pricing

The option \mathcal{C} has discrete price dynamics $f_{t_n}^{(d)} = f(S_{t_n}^{(d)}, t_n)$, $n = 0, \dots, m-1$, on BIT_m for some $f(x, t) \in \mathbb{R}$, $x > 0$, $t \in [0, T]$, with terminal time $T = t_m$ and terminal value $f_T = g(S_T)$ for some $g(x) \in \mathbb{R}$, $x > 0$.²³ As the price S_T is fixed by the conditions of the last event at t_{m-1} (that is, S_t has constant value S_T over the time interval $[t_{m-1}, t_m = T]$), the terminal value $g(S_T)$ determines the option price $f_{t_{m-1}}^{(d)}$ which remains constant over $[t_{m-1}, T]$. Recall that the path dependent development of stock prices is given by (34). Following the general methodology on locally risk-neutral option valuation in Kao (2012),²⁴ consider the price

$$f_{t_n; \mathcal{E}_n^{(L_n, M_n)}}^{(d)} = D_{t_n; \mathcal{E}_n^{(L_n, M_n)}}^{(d)} S_{t_n; \mathcal{E}_n^{(L_n, M_n)}}^{(d)} + \beta_{t_n; \mathcal{E}_{n-1}^{(L_{n-1}, M_{n-1})}}^{(d)} \quad (44)$$

of the replicating portfolio at event $\mathcal{E}_n^{(L_n, M_n)}$ with

$$f_{t_n; \mathcal{E}_n^{(L_n, M_n)}}^{(d)} = f\left(S_{t_n; \mathcal{E}_n^{(L_n, M_n)}}^{(d)}, t_n\right) \quad (45)$$

²³The functions f on $(0, \infty) \times [0, T]$ and g on $(0, \infty)$ satisfy the usual regularity conditions; see Duffie (2001, Chapter 5). These conditions will only be needed when we consider the limiting option price process as $\max(\Delta t_n) \rightarrow 0$.

²⁴See also Duan (1995); Duan et al. (2006); and Chorro et al. (2012).

and delta-position $D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)}$. By the risk-neutrality condition²⁵ we have

$$\begin{aligned}
D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} & S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} + \beta_{t_n; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)} - f_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} \\
&= D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} + \beta_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} - f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} \\
&= D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)} + \beta_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} - f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)}.
\end{aligned} \tag{46}$$

From (46), the delta-position $D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)}$ at event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ is given by

$$\begin{aligned}
D_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} &= \frac{f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} - f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)}}{S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} - S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)}} \\
&= \left[\frac{f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} - f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)}}{S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} \sigma_{t_{n+1}}^{(r)} \Big|_{\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}} \sqrt{\Delta t_{n+1}}} \right] \\
&\quad \times \sqrt{\mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 1 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) \mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 0 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right)}, \tag{47}
\end{aligned}$$

where the final equality in (47) is obtained using (34) and (37). From (46), (47), (24), and the representation (19), we obtain the recurrence relation for the risk-neutral option value at event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$,

$$\begin{aligned}
f_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)} &= \frac{1}{\left(1 + r_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d, f, \text{inst})} \Delta t_{n+1} \right)} \left(q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} \right. \\
&\quad \left. + q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)} f_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)} \right), \tag{48}
\end{aligned}$$

where the conditional risk-neutral probabilities at node $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ at t_n are given by

$$\begin{aligned}
q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} &= \mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 1 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) - \theta_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}} \sqrt{\mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 0 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) \mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 1 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) \Delta t_{n+1}}, \\
q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)} &= 1 - q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} \\
&= \mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 0 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) + \theta_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}} \sqrt{\mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 0 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) \mathbb{P} \left(\epsilon_{n+1}^{(l_{n+1})} = 1 \mid \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)} \right) \Delta t_{n+1}}, \tag{49}
\end{aligned}$$

²⁵Although we should refer to (46) as a local risk-neutrality condition, we take this to be understood and omit the reference to “local” when referring to our risk-neutral option price valuation.

where

$$\theta_{t_{n+1}} \Big|_{\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}} = \frac{\mu_{t_{n+1}}^{(r)} \Big|_{\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}} - r_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d, f, \text{inst})}}{\sigma_{t_{n+1}}^{(r)} \Big|_{\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}}$$

is the market price of risk. Equations (44) through (49) hold on $\mathbb{B}\mathbb{T}_m$ for $n = 0, \dots, m-2$. We note that the natural world conditional probabilities, and the mean and variance of the return dynamics, of \mathcal{S} are retained in the risk-neutral price dynamics of \mathcal{C} . This represents the tremendous advantage of binary option pricing over its limiting continuous-time model; under the latter information about the probabilities and the mean of the return process is lost (Section 6).

The extension to the pricing of an American option follows from the classical approach for valuation of an American option on a binomial tree.²⁶ The market value of the American option at event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ is given by

$$f_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{American})} = \max \left(f_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)}, S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{exercise})} \right), \quad (50)$$

where $S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{exercise})}$ is the exercise value of the stock, which is known at event $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$. We emphasize that $S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{exercise})}$ is the market value of the stock and not its fair-holding value $S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{fair-holding})}$. The risk-neutral tree stock dynamics $S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{fair-holding})}$, $n = 1, \dots, m$, are determined via the recursion

$$S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{fair-holding})} = \frac{1}{\left(1 + r_{t_{n+1}; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d, f, \text{inst})} \Delta t_{n+1} \right)} \left(q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d)} S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 1 \right)}^{(d; \text{fair-holding})} + q_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d)} S_{t_{n+1}; \left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}, \epsilon_{n+1}^{(l_{n+1})} = 0 \right)}^{(d; \text{fair-holding})} \right), \quad (51)$$

$n = 0, \dots, m-2$, based on the terminal value²⁷

$$S_T^{(d; \text{fair-holding})} = S_{t_{m-1}; \mathcal{E}_{m-1}^{(\mathbb{L}_{m-1}, \mathbb{M}_{m-1})}}^{(d; \text{fair-holding})} = S_{T; \mathcal{E}_{m-1}^{(\mathbb{L}_{m-1}, \mathbb{M}_{m-1})}}^{(d)}.$$

In contrast to (50), Breen (1991) uses $S_{t_n; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{fair-holding})}$ to define the fair-value of the American option at the node $\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}$,

$$f_{t_n, N; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{American; fair})} = \max \left(f_{t_n, N; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d)}, S_{t_n, N; \mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}}^{(d; \text{fair-holding})} \right). \quad (52)$$

A trader in search of statistical arbitrage opportunities relative to an American option could compare (50) with (52) when seeking potential mispricing in the market value of the option.

²⁶See, for example, Hull (2006, Section 12.5).

²⁷Recall discussion immediately following (34).

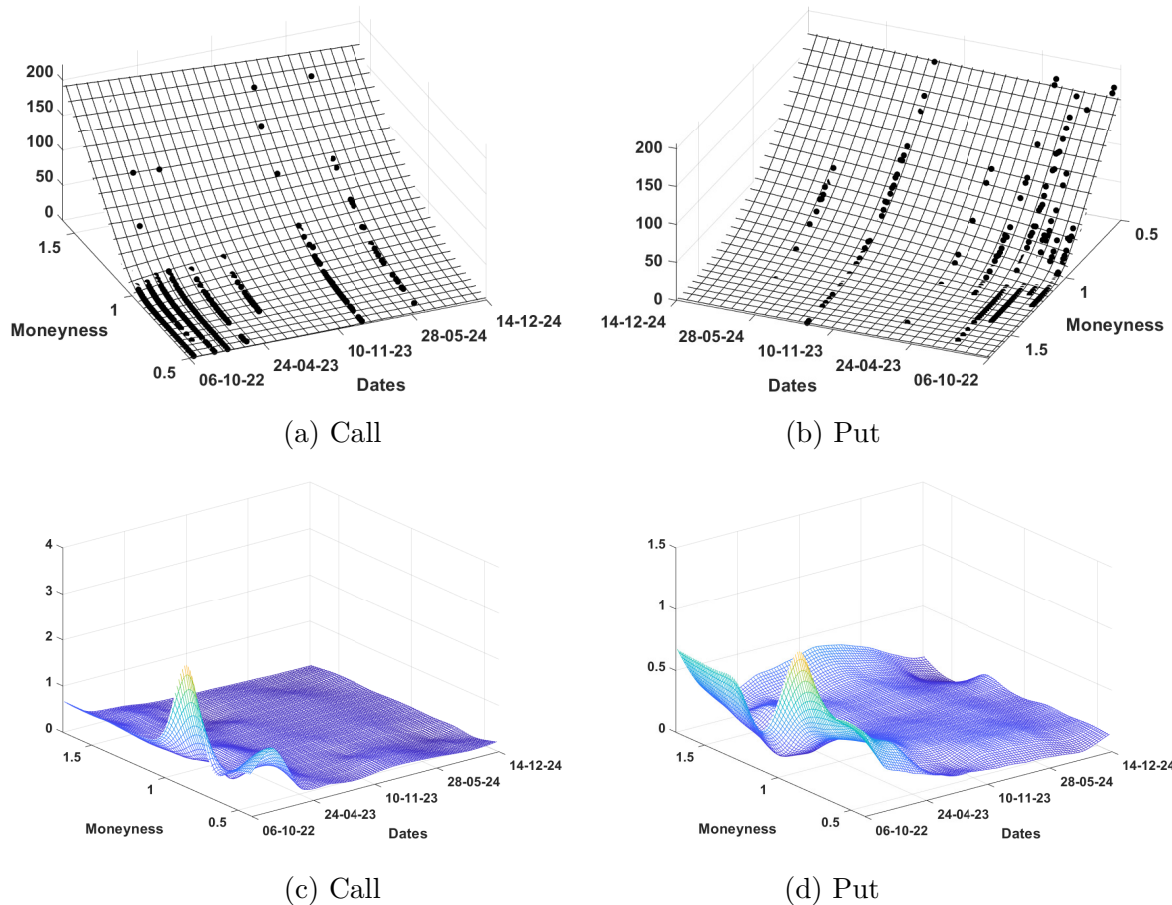


Figure 5: (a,b) Model MA(1) computed option price surfaces and historical market option prices (black dots) for DIA on 08/26/2022. (c,d) Implied volatility surfaces obtained by fitting the the MA(1) model to DIA market option prices on 08/26/2022.

5.1 Empirical example

We demonstrate that the MA(1) binary tree described in (41) can be used efficiently to price options. In spite of its simplicity, (it does not incorporate the full features of the general binary tree model proposed here), we demonstrate that it is capable of producing better results than the Black-Scholes formula. The choice of the MA(1) binary tree is motivated by the fact that it represents a generalization²⁸ of the first microstructure model introduced by Roll (1984). We apply the MA(1) model to price European options on the SPDR Dow Jones Industrial Average ETF Trust (DIA) on 26-August-2022. We estimate the probability p_1 in (38) using (17) and the parameters θ and σ_z^2 by solving (42)²⁹ using a series of price first-differences, $\Delta S_1^{(d)}$, over the preceding 510 trading days.

Prices for European call and put options obtained from the MA(1) binary tree are shown in Fig. 5 along with corresponding historical market option prices (black dots). We also computed

²⁸See, for instance, Hasbrouck (2007, Chapter 4, paragraph 2 and Chapter 8).

²⁹The constraint $|\theta| < 1$ was imposed to guarantee invertibility of the MA(1) process and produce a unique $\theta - \sigma_z$ pair.

option prices (not shown) using the Black-Scholes equation.³⁰ We computed the mean absolute difference (MAD) between (a) the MA(1) model and the historical option prices and (b) between the Black-Scholes model and the historical prices. For the MA(1) model, the MAD values for call and put options were 8.8 and 3.6, respectively; for the Black-Scholes model, the MAD values for call and put options were 10 and 4.3, respectively. The highest observed differences between either of these model predictions and realized prices occurred for put prices at low values of moneyness.

Using the value of θ obtained from the solution of (42), we then computed, as a function of time to maturity and moneyness, the implied volatility σ_z that minimized the distance between the model and historical option prices. The implied volatility surfaces for call and put options are provided in Fig. 5.

6 Limiting dynamics of binary pricing trees

In this section, we investigate the limiting dynamics of the binary tree pricing tree by assuming that the lengths of trading intervals uniformly vanish at a rate $\frac{1}{N}$ as $N \rightarrow \infty$. To support this, we generalize our notation for the trading times as follows. For any given $N \in \mathcal{N} = \{1, 2, \dots\}$, we consider the fixed time instances $0 = t_{0,N} < t_{1,N} < \dots < t_{n_N-1,N} < t_{n_N,N} = T < \infty$. As previously, the current time is $t_{0,N} = 0$ and the terminal time is $t_{n_N,N} = T$; trades of \mathcal{S} and \mathcal{B} occur only at the times $t_{1,N} < \dots < t_{n_N-1,N}$. The time intervals are denoted $\Delta t_{n,N} = t_{n,N} - t_{n-1,N}$, $n = 1, \dots, n_N$. It is straightforward to adapt the results of Sections 2 through 5 to this notational change for t . The resultant $\mathbb{F}^{(d)}$ -adapted binary information tree is now denoted \mathbb{BT}_{n_N} . We impose the restriction,

$$\Delta_N = \max\{\Delta t_{n,N}, \quad n = 1, \dots, n_N\} = O\left(\frac{1}{N}\right). \quad (53)$$

To determine the continuum limit behavior, we apply the Donsker-Prokhorov invariance principle (DPIP) for continuous diffusions.³¹ To apply DPIP, we assume that for each $n = 1, \dots, n_N$, the random variables $\epsilon_n^{(1)}, \dots, \epsilon_n^{(K_n)}$, determining the probabilities that (11) and (16), are independent. Therefore the probabilistic structure of the triangular array $\mathfrak{E} = \left(\epsilon_n^{(k)}, k = 1, \dots, K_n, n \in \mathcal{N}\right)$ is determined by the joint distributions,

$$P\left(\begin{matrix} m_n^{(1)}, \dots, m_n^{(K_n)} \\ \epsilon_n^{(1)}, \dots, \epsilon_n^{(K_n)} \end{matrix}\right) = P(\epsilon_n^{(1)} = m_n^{(1)}, \dots, \epsilon_n^{(K_n)} = m_n^{(K_n)}) = \prod_{k=1}^{K_n} \mathbb{P}\left(\epsilon_n^{(k)} = m_n^{(k)}\right),$$

for $m_n^{(k)} \in \{0, 1\}$, $k = 1, \dots, K_n$, $n \in \mathcal{N}$. This assumption of independence results in simplified

³⁰For the Black-Scholes option price computation we estimated σ using historical volatility.

³¹The DPIP is also known as the Functional Limit Theorem. We will apply the DPIP for continuous diffusions only, see Davydov and Rotar (2008). Extensions to more general DPIP, where the limiting price process is a semimartingale, are known; see Cherny et al. (2003); Duan et al. (2006) and Hu et al. (2020a). It will be of interest to study DPIP when the limiting pricing process is a semimartingale plus noise. These types of DPIP could be obtained by applying limiting results as studied in Jacod and Protter (2012), but that line of research is beyond the scope of this paper. Unfortunately, as pointed out in Hu et al. (2020a,b), the limiting stock price dynamics erases important information contained in the discrete pricing model; specifically the probabilities for the direction of stock price moment and, in the case of option pricing, the mean return of the stock. For this critical reason we view this section on the continuum limit of the discrete dynamics mainly as an extension to the classical CRR and Jarrow and Rudd (1982) option pricing models. As these limiting results reveal, incorporation of market microstructure features requires full use of the discrete binary tree pricing model.

expressions for (11) and (16):

$$\left. \begin{aligned} \mathbb{P}\left(\epsilon_n^{(k)} = 1\right) &\stackrel{\text{def}}{=} p_n^{(k;1)}, \\ \mathbb{P}\left(\epsilon_n^{(k)} = 0\right) &\stackrel{\text{def}}{=} p_n^{(k;0)} = 1 - p_n^{(k;1)}, \end{aligned} \right\} \quad k = 1, \dots, K_n, \quad (54)$$

$$\mathbb{P}\left(\mathcal{E}_n^{(\mathbb{L}_n, \mathbb{M}_n)}\right) = p_n^{(\mathbb{L}_n, \mathbb{M}_n)} = \prod_{k=1}^n \mathbb{P}\left(\epsilon_k^{(l_k)} = m_k\right) = \prod_{k=1}^n p_k^{(l_k; m_k)}, \quad (55)$$

$$\mathbb{P}\left(\epsilon_j^{(l_j)} = m_j \mid \left(\epsilon_1^{(l_1)} = m_1, \dots, \epsilon_{j-1}^{(l_{j-1})} = m_{j-1}\right)\right) = \mathbb{P}\left(\epsilon_j^{(l_j)} = m_j\right) = p_j^{(l_j; m_j)}. \quad (56)$$

As $\epsilon_n^{(k)} \stackrel{d}{=} \text{Bernoulli}\left(p_n^{(k;1)}\right)$, then

$$\delta_n^{(k)} \stackrel{\text{def}}{=} \frac{\epsilon_n^{(k)} - p_n^{(k;1)}}{\sqrt{p_n^{(k;0)}}}, \quad k = 1, \dots, K_n, \quad n = 1, \dots, n_N,$$

has $\mathbb{E}\left[\delta_n^{(k)}\right] = 0$ and $\text{Var}\left[\delta_n^{(k)}\right] = 1$. We now view the filtration (5) as generated by the triangular series of $\delta_n^{(1)}, \dots, \delta_n^{(n)}$, $n \in \mathcal{N}$; that is,

$$\mathcal{F}^{(n)} = \sigma\left(\delta_1^{(1)}, \left(\delta_2^{(1)}, \delta_2^{(2)}\right), \dots, \left(\delta_n^{(1)}, \dots, \delta_n^{(n)}\right)\right), \quad n = 1, \dots, n_N.$$

Next, for a given sequence $\mathbb{L} = (l_1, \dots, l_n, \dots)$, $l_n = 1, \dots, n$, $n = 1, \dots, n_N$, we consider the random walk $\delta_n^{(\mathbb{L})} = \sum_{k=1}^n \delta_k^{(k)}$. By the DPIP, the sequence of $D[0, \infty)$ -processes

$$\mathbb{B}_n^{(\mathbb{L}; [0, \infty))} = \left\{ B_{t;n}^{\mathbb{L}} = \frac{\delta_{\lfloor nt \rfloor}^{(\mathbb{L})}}{\sqrt{n}}, \quad t \geq 0 \right\}$$

converges in law to a standard Brownian motion $\mathbb{B}_{[0, \infty)} = \{B_t, t \geq 0\}$ in the Skorokhod J1-topology.³² By denoting the canonical filtration $\mathfrak{F}^{\mathbb{B}_{[0, \infty)}} = \{\sigma(B_u, 0 \leq u \leq t), t \geq 0\}$, we can assume, by the Skorokhod embedding theorem,³³ that $\mathbb{F}^{(d)} \subset \mathfrak{F}^{\mathbb{B}_{[0, \infty)}}$ and the triangular series $\epsilon_1^{(1)}, \left(\epsilon_2^{(1)}, \epsilon_2^{(2)}\right), \dots, \left(\epsilon_n^{(1)}, \dots, \epsilon_n^{(n)}\right), \dots$ are in the same stochastic basis space $(\Omega, \mathcal{F}^{\mathbb{B}_{[0, \infty)}}, \mathbb{P})$.

For the time interval $[0, T]$, we now consider the limiting behavior of the discrete riskless rates $r_{t_n, N}^{(d, f)} = r_{t_n, N}^{(d, f, \text{inst})} \Delta t_{n, N}$ (19). We assume that the discrete instantaneous riskless rate process

$$r_{[0, T]; N}^{(d, f, \text{inst})} = \left\{ r_{t, N}^{(d, f, \text{inst})} = r_{t_{n-1, N}}^{(d, f, \text{inst})}, \quad t \in [t_{n-1, N}, t_{n, N}), \quad n = 1, \dots, n_N, \quad r_{T, N}^{(d, f, \text{inst})} = r_{t_{n_N, N}}^{(d, f, \text{inst})} \right\}$$

converges uniformly to a continuous time instantaneous riskless rate $r_{[0, T]}^{(f)} = \left\{ r_t^{(f)}, t \in [0, T] \right\}$, where³⁴

1. $r_{[0, T]}^{(f)}$ has strictly positive continuous trajectories on $[0, T]$, and
2. $\sup \left\{ \left| r_t^{(f)} - r_{t, N}^{(d, f, \text{inst})} \right|, t \in [0, T] \right\} = O\left(\frac{1}{N}\right)$.

³²See, for example, Jacod, and Protter (2012), p. 49, and Cherny, Shiryayev and Yor (2003) and the references therein.

³³See for example, Kallenberg (2002).

³⁴The limiting riskless rate is also assumed to satisfy $\mathbb{P}\left(\sup_{t \in [0, T]} \left\{ r_t + \frac{1}{r_t} \right\} < \infty\right) = 1$. See Duffie (2001, p. 102) for the extension to a stochastic short rate $r_{[0, T]}^{(f)}$ under additional regularity conditions.

Then, the discrete bond price $\beta_{t,N}^{(d)}$, $t \in [0, T]$ (24) converges uniformly to the continuous time bond dynamics

$$\beta_t = \beta_0 e^{\int_0^t r_u^{(f)} du}, \quad t \in [0, T],$$

where the deterministic instantaneous riskless rate (short rate) process $r_{[0,T]}^{(f)}$ is $\mathfrak{F}^{\mathbb{B}[0,T]}$ -adapted.

Consider the discrete mean and volatility processes for \mathcal{S} ,

$$\mu_{[0,T];N} = \begin{cases} \mu_{t,N} = \mu_{t_{n,N}} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}, & t \in [t_{n,N}, t_{n+1,N}), \quad n = 1, \dots, n_N - 1, \\ \mu_{T,N} = \mu_{t_{n_N-1,N}} \Big|_{\mathcal{E}_{n_N-2}^{(\mathbb{L}_{n_N-2}, \mathbb{M}_{n_N-2})}}, & \end{cases}$$

$$\sigma_{[0,T];N} = \begin{cases} \sigma_{t,N} = \sigma_{t_{n,N}} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}, & t \in [t_{n,N}, t_{n+1,N}), \quad n = 1, \dots, n_N - 1, \\ \sigma_{T,N} = \sigma_{t_{n_N-1,N}} \Big|_{\mathcal{E}_{n_N-2}^{(\mathbb{L}_{n_N-2}, \mathbb{M}_{n_N-2})}}. & \end{cases}$$

Assume that $\mu_{[0,T];N}$ and $\sigma_{[0,T];N}$ converge uniformly on $[0, T]$ to μ_t and σ_t , respectively, such that

$$\sup \{ |\mu_{t,N} - \mu_t| + |\sigma_{t,N} - \sigma_t|, \quad t \in [0, T] \} = O\left(\frac{1}{N}\right),$$

Further, assume that μ_t and σ_t are $\mathfrak{F}^{\mathbb{B}[0,T]}$ adapted, with $\mu_{[0,T]} = \{\mu_t, t \in [0, T]\}$ and $\sigma_{[0,T]} = \{\sigma_t, t \in [0, T]\}$ having continuous trajectories on $[0, T]$.³⁵

We define the $\mathfrak{D}[0, T]$ -price process

$$\mathbb{S}_{[0,T];N} = \begin{cases} S_{t,N} = S_{t_{n,N}}^{(d)} = S_{t_{n,N}; \mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}^{(d)}, & t \in [t_{n,N}, t_{n+1,N}), \quad n = 1, \dots, n_N - 2, \\ S_{t,N} = S_{t_{n_N-1,N}}^{(d)} = S_{t_{n_N-1,N}; \mathcal{E}_{n_N-1}^{(\mathbb{L}_{n_N-1}, \mathbb{M}_{n_N-1})}}^{(d)}, & t \in [t_{n_N-1,N}, t_{n_N,N} = T]. \end{cases} \quad (57)$$

As in Hu et al. (2020a), a non-standard invariance principle (Davydov and Rotar, 2008) can be used to show that (57) converges weakly in $\mathfrak{D}[0, T]$ topology (Skorokhod, 2005) to a process $S_{[0,T]} = \{S_t, t \in [0, T]\}$ governed by a cumulative return process $R_{[0,T]} = \{R_t, t \in [0, T]\}$ satisfying

$$dR_t = \mu_t dt + \sigma_t dB_t, \quad R_0 = 0, \quad (58)$$

where B_t is a standard Brownian motion and $dS_t = S_t dR_t$. (See Duffie (2001, Appendix 6D).) By (58), $S_{[0,T]}$ is a continuous diffusion and, if μ_t and σ_t are constant, then $S_{[0,T]}$ is a GBM. In the risk-neutral world, the limiting cumulative return process obeys (58) with μ_t replaced by $r_t^{(f)}$. We note that the discrete model is much more informative than the continuous time model as it preserves the path dependent probabilities $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$, with no assumption on their (in)dependence.

³⁵Relaxing the assumptions on $\mu_{[0,T]}$ and $\sigma_{[0,T]}$ requires an extension of the non-standard DPIIP by Davydov and Rotar (2008) for general continuous diffusions, which is beyond the scope of the current work. The reason we do not pay significant attention to the limiting behavior of the binary asset pricing dynamics is that the continuous dynamics of the return process $R_{[0,T]} = R_t, t \in [0, T]$ loses the important information regarding the probabilities for the direction of stock price movements $p_n^{(\mathbb{L}_n, \mathbb{M}_n)} = \prod_{j=1}^n p_j^{(l_j, m_j)}$, $n = 0, \dots, n_N, N \in \mathcal{N}$. Even worse, when passing to risk-neutral continuous dynamics, the extremely valuable information about the mean stock returns $\mu_{t_{n,N}} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}$ will also be lost. When discussing market microstructure option pricing models, losing information on $p_n^{(\mathbb{L}_n, \mathbb{M}_n)}$ and $\mu_{t_{n,N}} \Big|_{\mathcal{E}_{n-1}^{(\mathbb{L}_{n-1}, \mathbb{M}_{n-1})}}$ hardly seems justifiable. Thus, in this work we concentrate our attention on (discrete) binary asset pricing, and pass to the limit as $\Delta_N = O\left(\frac{1}{N}\right), N \uparrow \infty$, only to provide a comparison with the classical Black-Scholes-Merton asset pricing continuous time dynamics.

7 Technical analyses of the probability estimates

In Section 2.1 we noted the desirability of estimating the direction-of-price-change probabilities $p_n^{(\mathbb{L}_n, \mathbb{M}_n; \Delta t_{1,n})}$ from historical data using (17) on a set of V non-overlapping binomial sequences, each of length n . However, we also noted that, even for relatively small values of n , a prohibitively extensive history of returns would be required to ensure an adequate sample to determine each of the 2^n probabilities for a given value of n . Table 1 codifies this problem using daily return data for DIA covering the period of prices from 01/20/1998 through 05/05/2023. This data set provides 6,365 return values, that is a sequence of 6,365 values of 0 (down) or 1 (up) price changes. From the table it is evident that, by $n = 6$, V is too small to ensure an adequate sample size for determining each of the 2^6 possible sequences. It is not clear the the sample size is adequate even for $n = 5$.

There are models that can be applied to a historical return time series to generate “mimicking” time series. Of course, all such models add additional “model error” to the process. One approach is to fit the historical return time series to an ARMA(l, m)-GARCH(p, q) model combined with a distribution model for the ARMA residuals and then use the ARMA-GARCH-distribution model with fitted parameters to generate adequate numbers, V , of mimicking return series from which to compute the required probabilities. Such an approach involves fitting a significant number of required parameters. It generates return time series, which is a step removed from the 0, 1 sequences required by (17). We prefer to utilize bootstrap resampling directly on the 0, 1 sequence determined by the historical data set. Bootstrapping has the advantage of constantly resampling the historical sequence.³⁶ For a given value of n , we required $V \geq 10,000 \times 2^n$ bootstrapped samples (i.e. each of the 2^n probabilities is determined based upon a expectation of 10,000 occurrences for each of the 2^n possible sequences).

We compared our bootstrap resampled results against those computed from the historical time series with no resampling (i.e. with expected number of sequence occurrences given in Table 1). For plotting convenience, we have developed the following labeling system for each of the 2^n possible 0, 1 sequences of length n . We illustrate the general notation using specific examples. Consider the $n = 5$ sequences. They can be mapped to the $2^5 = 32$ values $x = -15.5, -14.5, \dots, -0.5, 0.5, \dots, 14.5, 15.5$. For example: 01101 is mapped to $x = (01101)_{10} + 1 - (2^5 + 1)/2 = -2.5$; 00000 is mapped to $x = (00000)_{10} + 1 - (2^5 + 1)/2 = -15.5$; and 11111 to $x = (11111)_{10} + 1 - (2^5 + 1)/2 = 15.5$. This labeling has the property that $-x$ and x label binary complement sequences (i.e., 01101 corresponds to $x = -2.5$ and 10010 to $x = 2.5$). A positive value of x indicates a binary string beginning with 1 (i.e., 1...), while a negative value of x indicates a binary string beginning with 0 (i.e., 0...).

³⁶Specifically we employed the R program `ts_boot()` using block resampling with block lengths having a geometric distribution with mean length n .

Table 1: Expected number of occurrences of each sequence in patterns of length n in the DIA return data set covering the period of prices from 01/20/1998 through 05/05/2023.

pattern length n	number sequences in pattern (2^n)	V	expected number of DIA occurrences per pattern sequence
4	16	1,591	99.4
5	32	1,273	39.8
6	64	1,060	16.6
8	256	795	3.1
10	1024	636	0.6

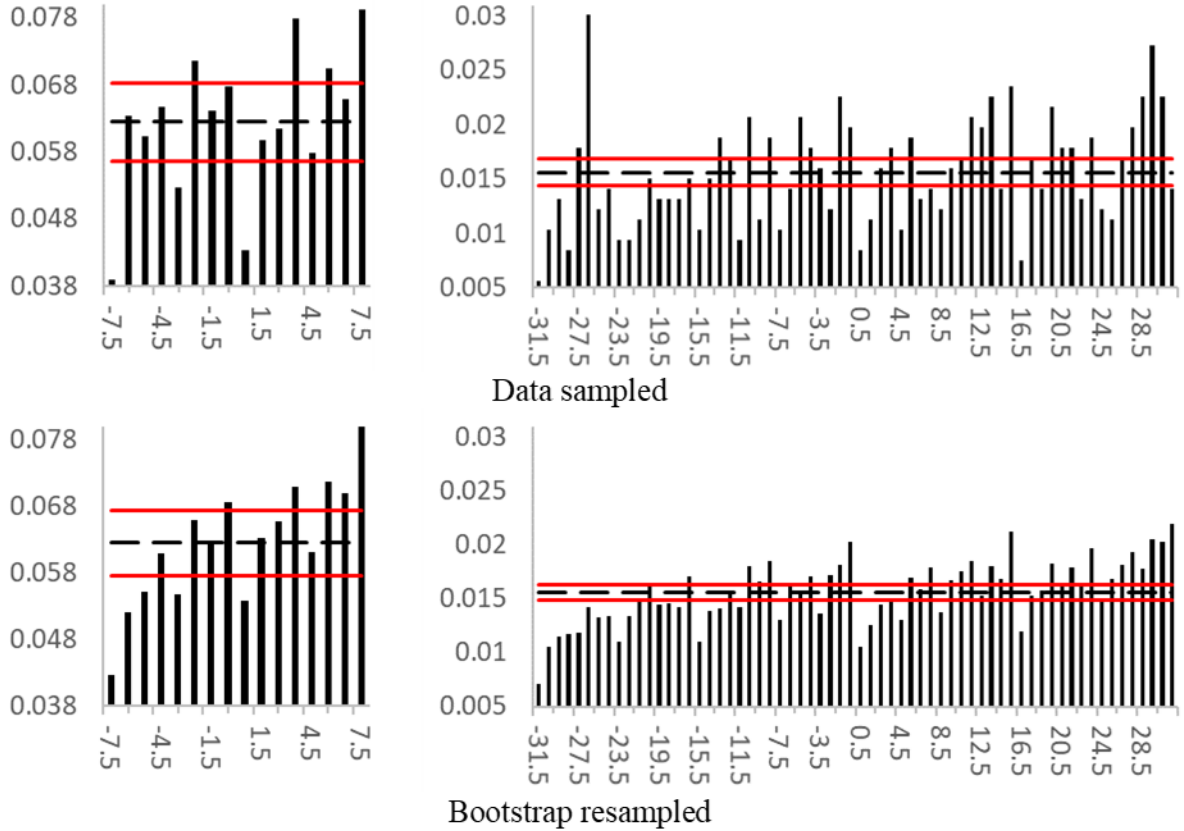


Figure 6: The probabilities $p_n^{(x)}$ obtained from the DIA data set without and with bootstrap resampling for $n = 4$ and 6 . The dashed horizontal line denotes 2^{-n} . The red horizontal lines denote the 95% confidence interval.

Using the values x to represent the sequences corresponding to the various events $\mathcal{E}_n^{(L_n, M_n)}$ with $\Delta t_{1,n} = \Delta t = 1$, Figs. 6 and 7 compare the results obtained for the probabilities $p_n^{(x)}$ computed via (17) for the DIA data set without and with bootstrap resampling for $n = 4, 6$ and 8 . There is reasonable agreement between the results without and with bootstrap resampling for $n = 4$; significant differences develop for $n = 6$, which are then clearly revealed for $n = 8$. In particular, without bootstrap resampling, when $n = 8$ the paucity of data results in the probabilities $p_8^{(x)}$ taking on only nine possible values. The 95% confidence intervals in Figs. 6 and 7 are based upon the results that the probabilities $p_n^{(x)}$ are well-described by a normal distribution (Fig. 8).

Analysis of the probabilities of individual sequences fall within the area of technical pattern analysis (Lo et al., 2000). We continue this analysis by examining the highest and lowest probability paths. Specifically, for fixed n , we consider whether the highest probability sequences specify paths that are “closely grouped” (with a similar statement for the lowest probability paths). If the highest probability paths occur randomly, this would provide further confirmation of the efficient market hypothesis. However clustered paths suggest the presence of pronounced patterns as argued by Lo et al. (2000). A related consideration is whether the highest probability path for $n = n_1$ is a projection of the highest probability path for $n = n_2 > n_1$.

Fig. 9 displays the observed results for the grouping of paths. For all 2^n sequences of length n , we plot the n highest probability path sequences (colored blue with the highest probability

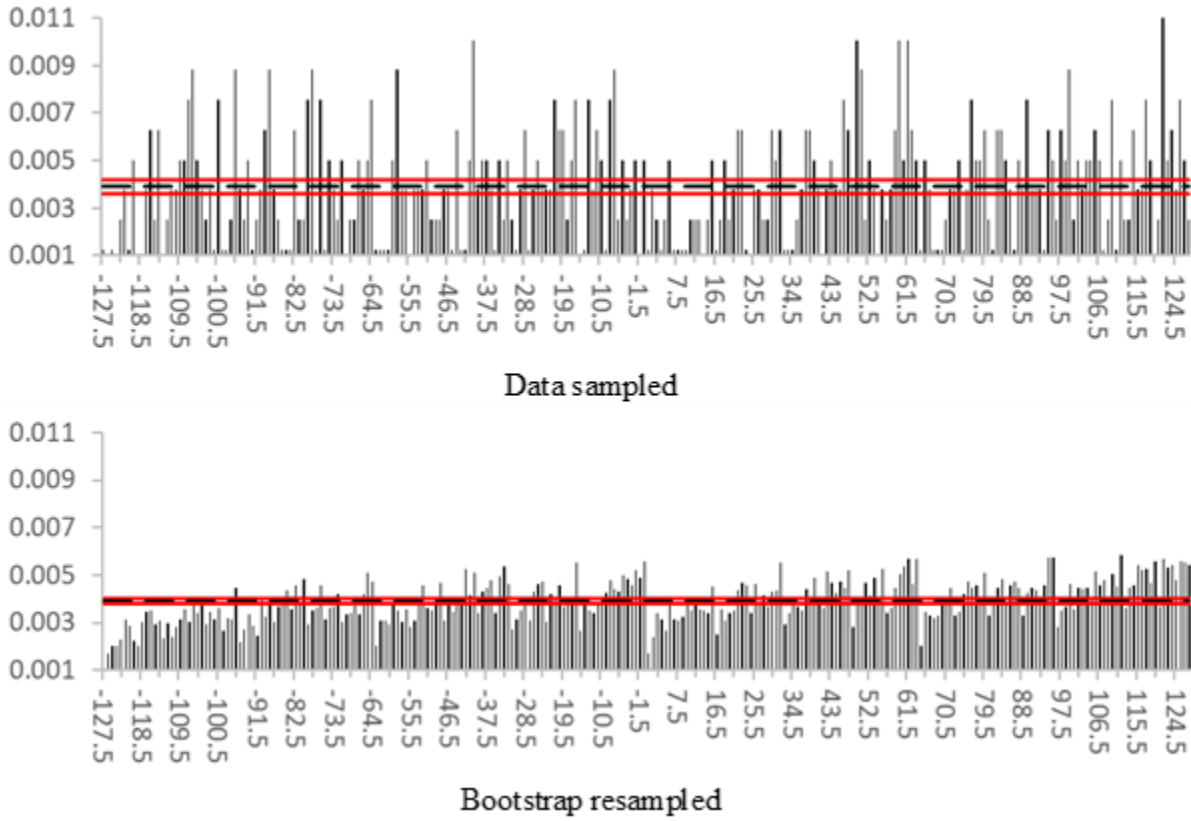


Figure 7: The probabilities $p_n^{(x)}$ obtained from the DIA data set without and with bootstrap resampling for $n = 8$. The dashed horizontal line denotes 2^{-n} . The red horizontal lines denote the 95% confidence interval.

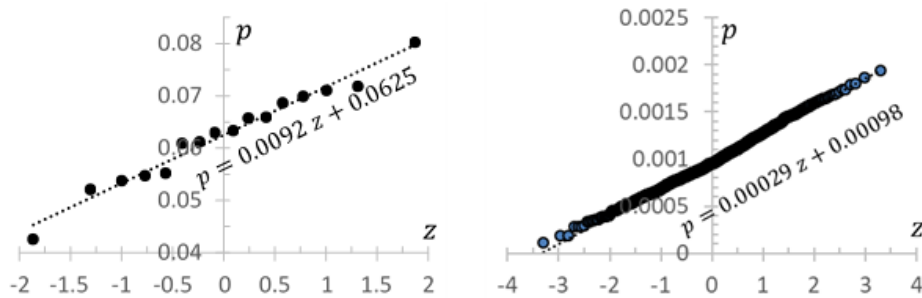


Figure 8: Normal probability plots for the DIA probabilities computed using bootstrap resampling for (left) $n = 4$ and (right) $n = 10$.

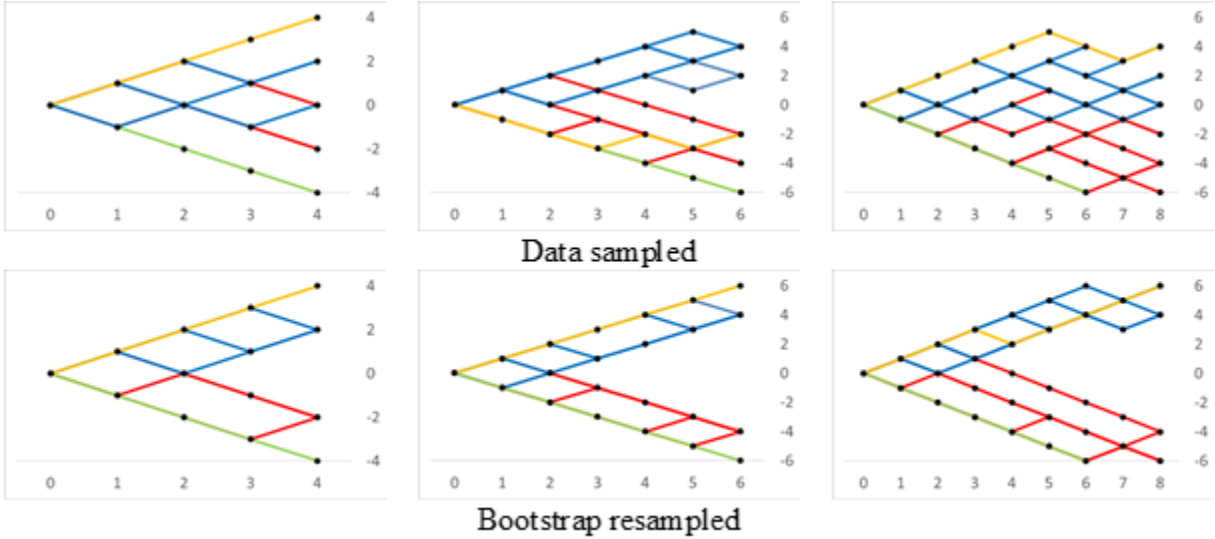


Figure 9: Plotted for sequences of length n are the n highest (blue, gold) and n lowest (red, green) probability paths for the data (top) without and (bottom) with bootstrap resampling. The highest probability path is colored gold, the lowest probability path is green.

path colored gold) as well as the n lowest probability path sequences (colored red with the lowest probability path colored green). We consider $n = \{4, 6, 8\}$ and plot results for the data without and with bootstrap resampling. Due to the data limitations with no resampling, there is a more “random” distribution of high and low probability paths. (Note in particular the highest and lowest probability paths for $n = 6$ in the case with no resampling.) With bootstrap resampling improving sample sizes, there is a more distinct grouping of the high and low probability paths, with the high probability paths characterized by more consistent price “upturns” and the low probability paths characterized by more consistent “downturns”.

Fig. 10 displays the observed results for the projections of the highest and lowest probability paths. The highest probability path for $n = \{4, 5, 6, 7, 8, 10\}$ are each plotted on the same graph. Similarly for the respective lowest probability path for each value of n . It is clear that, over this

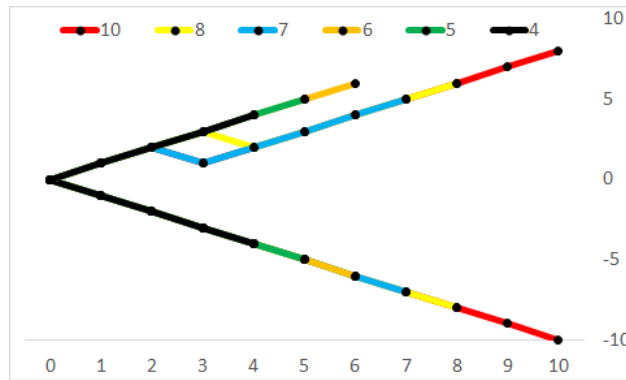


Figure 10: The highest and lowest probability path plotted for $n = \{4, 5, 6, 7, 8, 10\}$. Note that the color used for a smaller- n path obscures the color used for a larger- n path if both have a segment occurring on the same branch of the tree.

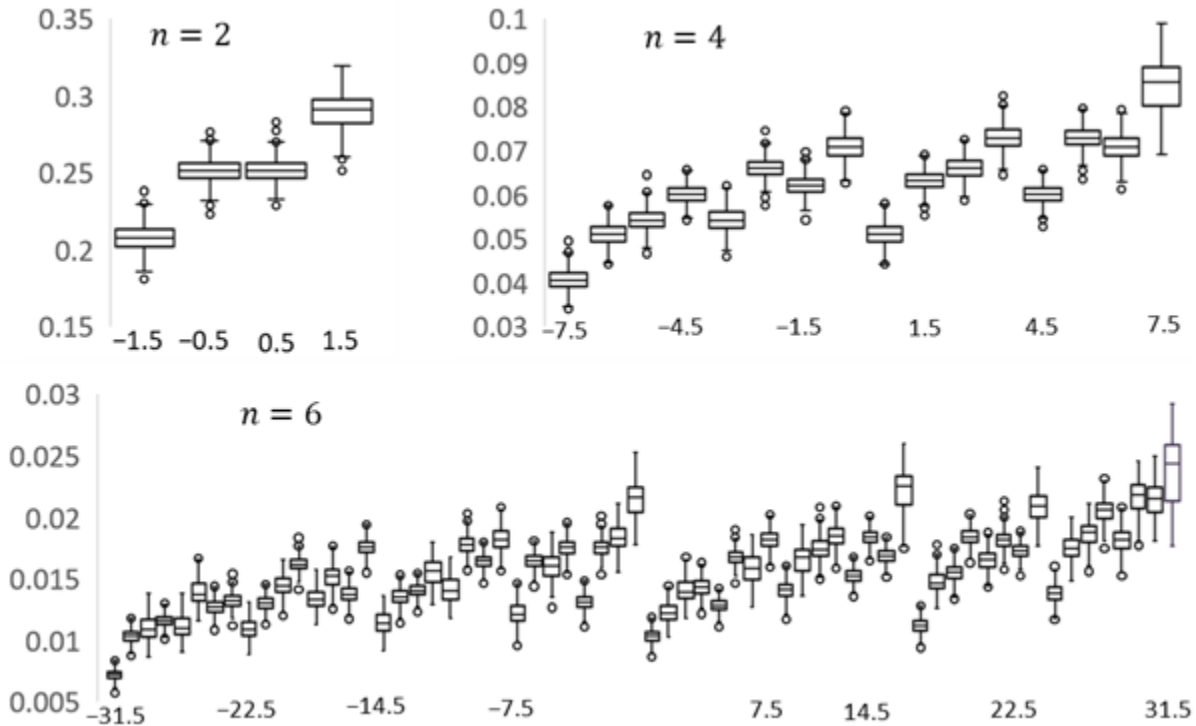


Figure 11: Box-whisker summaries of the computed distribution of probability estimates for each of the 2^n sequences for $n = \{2, 4, 6\}$. The empirical distributions were obtained from the DIA data set using a rolling window of 15 years.

range of values of n , the lowest probability path for $n = n_1$ is simply the projection (truncation) of the highest probability path for $n = n_2 > n_1$. In case of the highest probability path, there is a “discontinuity” in the projection. For $n < 6$ the highest probability path is a truncation of that for $n = 6$. For $n = 7, 8$, the highest probability path is almost a truncation of that for $n = 10$ (with a slight difference occurring for $n = 8$ at t_3).

To test the dynamic stability of such estimates, we redid the probability estimation procedure for the DIA data set using a rolling window of length 15 years (3,780 trading days). This generated 2,586 windows. For each window, sequence probabilities were computed for $n = \{2, 4, 6\}$ using bootstrap resampling to ensure adequate sample sizes. (To speed the computation, we employed $1,000 \times 2^n$ bootstrapped samples in each window.) For each choice of n , the rolling windows produced an empirical distribution of probability estimates for each of the 2^n sequences. These distributions are summarized as box-whisker plots in Fig. 11. Figs. 6 and 11 show very similar structure, indicating relative stability between the rolling window and global estimates of the sequence probabilities.

We now address the substructure that is apparent in Fig. 11 (and in Figs. 6 and 7). Fig. 12 replots the $n = \{4, 6\}$ box-whisker plots with the sequences placed in categories according to the number of zeros (price downturns; equivalently the number of ones (price upturns)) each contains. Within each category, the sequences are still labeled from smallest to largest numerical label, x , as indicated in the top plot of Fig. 12. The substructure seen in Fig. 11 has largely vanished from Fig. 12 indicating that the number of price downturns (equivalently upturns) is the major driver of

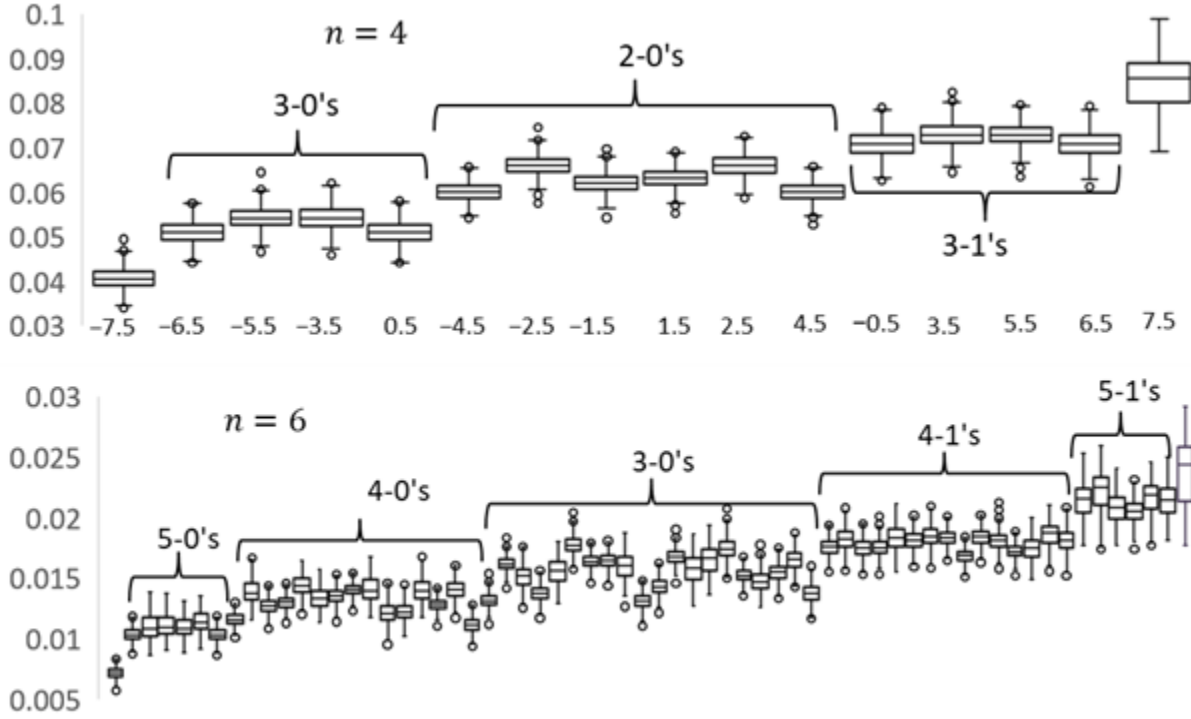


Figure 12: The box-whisker summaries of Fig. 11 reordered into categories based on the number of price up- and downturns occurring in the sequence.

the sequence probabilities. *The uniformity of the ranges of the probability distributions within a category is indicative of an efficient market hypothesis operating within each category. It is in the difference in the ranges of the probability distributions between categories that market inefficiencies are seen.*³⁷ For $n = 4$, there is no overlap between the range of the empirical probability distribution for the sequence 0000 and the range of any distribution for sequences containing two or more upturns. For $n = 6$ there is no overlap between the range of the empirical distribution for the sequence 000000 and the range of any other distribution. Furthermore, the range of any distribution for a sequence containing five or six downturns has no overlap with the range of any distribution for a sequence containing four or more upturns.

We compare sequence probability estimates among different assets using the 30 components comprising (as of August 31, 2020) the Dow Jones Industrial Average (DJIA) index. Price data was used for the period 01/03/2000 through 08/26/2022 with the exception of Visa (price data beginning 03/18/2008) and Dow (price data beginning 02/20/2019). This provided 5,699 return values for 28 of the assets (3,637 returns for Visa and 868 returns for Dow). Sequence probabilities were computed for these assets for $n = \{1, 2, 3\}$. For these small values of n , the probabilities were computed from the data without bootstrap resampling. The estimated probabilities for sequences of length $n = 1, 2$ are presented in Table 2. The probabilities for sequences of length $n = 3$ are presented in Table 3. Significant p -values obtained from the one-sided z-test are also indicated. For comparison, sequence probabilities for the SPDR DJIA ETF data are also provided in these tables.

³⁷We note that these results are based upon daily closing prices. We make no inferences for returns based upon other price intervals.

For $n = 1$, for all 31 assets, the probability of a negative return is smaller than the probability of a positive return. For 24 of these assets, this relationship is significant at a level $\leq 5\%$. For assets where the probability of the sequences 00 ($n = 2$) or 000 ($n = 3$) are significant at the level $\leq 5\%$, these sequences have the smallest probability. This holds for 26 of 31 assets ($n = 2$) and 21 of 31 assets ($n = 3$). For the sequences 11 and 111, the results are not as strong. These sequences are significant at the level $\leq 5\%$ in 11 of 31 assets ($n = 2$) and 9 of 31 assets ($n = 3$). However, these sequences represent, respectively, the highest probability for only 10 of the 11 ($n = 2$) and 7 of the 9 ($n = 3$) of these assets. If, instead, we consider sequences that contain at most a single negative return (a single “0”-value), then the highest probability sequence that has significance at the $\leq 5\%$ level occurs for 15 of 31 ($n = 2$) and 20 of 31 ($n = 3$) assets. These observations are consistent with the DIA results in Fig. 10 for $n = \{4, 5, 6, 7, 8, 10\}$.

Table 2: Path probabilities for sequences of length (columns 2 and 3) $n = 1$ and (columns 4 to 7) $n = 2$ for assets in the DJIA index and for the SPDR DJIA ETF. Also indicated are the probabilities having significant p -values from the one-sided z -test.

Symbol	Path Label		Path Label			
	-0.5	0.5	-1.5	-0.5	0.5	1.5
	Path Probability		Path Probability			
AAPL	0.477***	0.523***	0.230**	0.248	0.246	0.276***
AMGN	0.499	0.501	0.245	0.256	0.253	0.246
AXP	0.490	0.510	0.232*	0.260	0.257	0.251
BA	0.486*	0.514*	0.229**	0.260	0.254	0.257
CAT	0.486*	0.514*	0.239	0.247	0.248	0.266*
CRM	0.487*	0.513*	0.222***	0.271**	0.257	0.249
CSCO	0.484*	0.516*	0.229**	0.262	0.247	0.261
CVX	0.475***	0.525***	0.221***	0.250	0.257	272**
DIS	0.487*	0.513*	0.233*	0.258	0.251	0.259
DOW	0.486	0.514	0.239	0.268	0.225	0.268
GS	0.488*	0.512*	0.229**	0.256	0.262	0.253
HD	0.480**	0.520**	0.229**	0.248	0.255	0.268*
HON	0.477***	0.523***	0.219***	0.265*	0.250	0.266*
IBM	0.488*	0.512*	0.232*	0.265*	0.248	0.255
INTC	0.488*	0.512*	0.232*	0.258	0.256	0.255
JNJ	0.486*	0.514*	0.225**	0.265*	0.257	0.253
JPM	0.493	0.507	0.233*	0.254	0.267*	0.246
KO	0.482**	0.518**	0.230**	0.244	0.260	0.266*
MCD	0.464***	0.536***	0.207***	0.262	0.252	0.279***
MMM	0.475***	0.525***	0.215***	0.259	0.261	0.265*
MRK	0.494	0.506	0.246	0.242	0.253	0.258
MSFT	0.486*	0.514*	0.225**	0.263	0.258	0.253
NKE	0.483**	0.517**	0.224***	0.253	0.264	0.258
PG	0.481**	0.519**	0.225**	0.262	0.250	0.263
TRV	0.477***	0.523***	0.216***	0.254	0.268*	0.262
UNH	0.477***	0.523***	0.221***	0.260	0.251	0.267*
V	0.462***	0.538***	0.195***	0.275**	0.259	0.270*

Table 2 – Continued on next page

Table 2 – Continued from previous page

Symbol	Path Label		Path Label			
	–0.5	0.5	–1.5	–0.5	0.5	1.5
	Path Probability		Path Probability			
VZ	0.493	0.517	0.234*	0.264	0.255	0.248
WBA	0.499	0.501	0.244	0.258	0.251	0.246
WMT	0.487*	0.513*	0.224***	0.263	0.264	0.250
SPDR DJIA	0.455***	0.545***	0.201***	0.262	0.246	0.291***

Table 3: Path probabilities for sequences of length $n = 3$ for assets in the DJIA index and for the SPDR DJIA ETF. Also indicated are the probabilities having significant p -values from the one-sided z -test.

Symbol	Path Label							
	–3.5	–2.5	–1.5	–0.5	0.5	1.5	2.5	3.5
	Path Probability							
AAPL	0.102**	0.125	0.122	0.122	0.128	0.123	0.132	0.146**
AMGN	0.112	0.119	0.148*	0.122	0.131	0.121	0.121	0.125
AXP	0.118	0.121	0.132	0.131	0.114	0.132	0.121	0.131
BA	0.111*	0.134	0.113	0.142*	0.118	0.130	0.121	0.131
CAT	0.119	0.114	0.124	0.131	0.121	0.126	0.128	0.137
CRM	0.110*	0.116	0.127	0.127	0.122	0.145*	0.130	0.124
CSCO	0.095***	0.139*	0.136	0.114	0.128	0.114	0.132	0.142*
CVX	0.101***	0.114	0.128	0.132	0.118	0.144**	0.128	0.136
DIS	0.112	0.102**	0.134	0.139*	0.138	0.115	0.121	0.139*
DOW	0.115	0.137	0.077*	0.132	0.150	0.124	0.128	0.137
GS	0.111*	0.126	0.129	0.132	0.116	0.133	0.125	0.129
HD	0.104**	0.131	0.120	0.129	0.124	0.129	0.122	0.141*
HON	0.102**	0.107*	0.129	0.140*	0.122	0.138	0.129	0.132
IBM	0.112*	0.116	0.123	0.131	0.133	0.125	0.130	0.131
INTC	0.115	0.123	0.126	0.127	0.118	0.126	0.134	0.131
JNJ	0.112	0.135	0.125	0.122	0.117	0.132	0.114	0.144**
JPM	0.116	0.114	0.131	0.150***	0.121	0.127	0.122	0.119
KO	0.116	0.115	0.123	0.130	0.115	0.131	0.132	0.138*
MCD	0.095***	0.108*	0.107*	0.132	0.129	0.147**	0.139*	0.142*
MMM	0.102**	0.112*	0.118	0.138*	0.131	0.130	0.129	0.140*
MRK	0.118	0.132	0.118	0.126	0.130	0.130	0.111*	0.134
MSFT	0.103**	0.113	0.136	0.136	0.126	0.134	0.129	0.123
NKE	0.107*	0.122	0.118	0.132	0.123	0.135	0.136	0.127
PG	0.096***	0.122	0.126	0.121	0.135	0.140*	0.128	0.132
TRV	0.099***	0.122	0.124	0.135	0.119	0.148**	0.120	0.132
UNH	0.101***	0.124	0.126	0.121	0.115	0.142*	0.137	0.135
V	0.079***	0.111	0.126	0.144*	0.119	0.156***	0.139	0.125
VZ	0.106**	0.122	0.136	0.141*	0.126	0.128	0.122	0.118
WBA	0.111*	0.133	0.132	0.119	0.132	0.123	0.128	0.122
WMT	0.103**	0.121	0.127	0.118	0.126	0.146**	0.139*	0.119
SPDR DJIA	0.090***	0.109*	0.118	0.138	0.125	0.129	0.124	0.167***

Appendix: Sequential definition of the probability law on the BIT for $n = 2$ and 3

$n = 2$. Set $\mathcal{E} \stackrel{\text{def}}{=} \{\epsilon_0^{(0)}, \epsilon_1^{(1)}, (\epsilon_2^{(1)}, \epsilon_2^{(2)})\} = \{\mathcal{E}_1, (\epsilon_2^{(1)}, \epsilon_2^{(2)})\}$. Then $\mathcal{F}^{(1)} = \sigma(\mathcal{E}_2)$ with

$$\begin{aligned} \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1) &= p_2^{((0,1,2),(0,1,1))} \in (0, p_1^{((0,1),(0,1))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0) &= p_1^{((0,1,2),(0,1,0))} = p_1^{((0,1),(0,1))} - p_2^{((0,1,2),(0,1,1))}, \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1) &= p_2^{((0,1,1),(0,0,1))} \in (0, p_1^{((0,1),(0,0))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0) &= p_2^{((0,1,1),(0,0,0))} = p_1^{((0,1),(0,0))} - p_2^{((0,1,1),(0,0,1))}. \end{aligned}$$

To estimate $p_2^{((0,1,2),(0,1,1))}$, we use the historical frequency $\hat{p}_1^{((0,1),(0,1);\Delta t_1, \Delta t_2)}$ of observing “a positive price change over a trading period of size Δt_1 followed by a positive price change over a trading period of size Δt_2 ”. Estimates for the remaining three probabilities of two-step stock movements are computed analogously. The \mathcal{E}_1 -conditional probabilities are

$$\begin{aligned} \mathbb{P}(\epsilon_2^{(2)} = 1 \mid \epsilon_1^{(1)} = 1) &\in (0, 1), \\ \mathbb{P}(\epsilon_2^{(2)} = 0 \mid \epsilon_1^{(1)} = 1) &= 1 - \mathbb{P}(\epsilon_2^{(2)} = 1 \mid \epsilon_1^{(1)} = 1), \\ \mathbb{P}(\epsilon_2^{(1)} = 1 \mid \epsilon_1^{(1)} = 0) &\in (0, 1), \\ \mathbb{P}(\epsilon_2^{(1)} = 0 \mid \epsilon_1^{(1)} = 0) &= 1 - \mathbb{P}(\epsilon_2^{(1)} = 1 \mid \epsilon_1^{(1)} = 0). \end{aligned}$$

$n = 3$. Set

$$\mathcal{E}_3 \stackrel{\text{def}}{=} \{\epsilon_0^{(0)}, \epsilon_1^{(1)}, (\epsilon_2^{(1)}, \epsilon_2^{(2)}), (\epsilon_3^{(1)}, \epsilon_3^{(2)}, \epsilon_3^{(3)}, \epsilon_3^{(4)})\} = \{\mathcal{E}_2, (\epsilon_3^{(1)}, \epsilon_3^{(2)}, \epsilon_3^{(3)}, \epsilon_3^{(4)})\}.$$

Then $\mathcal{F}^{(3)} = \sigma(\mathcal{E}_3)$, with

$$\begin{aligned} \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1, \epsilon_3^{(4)} = 1) &= p_3^{((0,1,2,4),(0,1,1,1))} \in (0, p_2^{((0,1,2),(0,1,1))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1, \epsilon_3^{(4)} = 0) &= p_3^{((0,1,2,4),(0,1,1,0))} \\ &= p_2^{((0,1,2),(0,1,1))} - p_3^{((0,1,2,4),(0,1,1,1))}, \\ \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0, \epsilon_3^{(3)} = 1) &= p_3^{((0,1,2,3),(0,1,1,1))} \in (0, p_2^{((0,1,2),(0,1,0))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0, \epsilon_3^{(3)} = 0) &= p_3^{((0,1,2,3),(0,1,0,0))} \\ &= p_2^{((0,1,2),(0,1,0))} - p_3^{((0,1,2,3),(0,1,0,1))}, \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1, \epsilon_3^{(2)} = 1) &= p_3^{((0,1,1,2),(0,0,1,1))} \in (0, p_2^{((0,1,1),(0,0,1))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0, \epsilon_3^{(2)} = 0) &= p_3^{((0,1,1,2),(0,0,1,0))} \\ &= p_2^{((0,1,1),(0,0,1))} - p_3^{((0,1,1,2),(0,0,1,1))}, \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1, \epsilon_3^{(1)} = 1) &= p_3^{((0,1,1,1),(0,0,0,1))} \in (0, p_1^{((0,1,1),(0,0,0))}), \\ \mathbb{P}(\epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0, \epsilon_3^{(1)} = 0) &= p_3^{((0,1,1,1),(0,0,0,0))} \\ &= p_2^{((0,1,1),(0,0,0))} - p_3^{((0,1,1,1),(0,0,0,1))}. \end{aligned} \tag{A1}$$

The \mathcal{E}_2 -conditional probabilities are

$$\begin{aligned}
& \mathbb{P}\left(\epsilon_3^{(4)} = 1 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1\right) \in (0, 1), \\
& \mathbb{P}\left(\epsilon_3^{(4)} = 0 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1\right) = 1 - \mathbb{P}\left(\epsilon_3^{(4)} = 1 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 1\right), \\
& \mathbb{P}\left(\epsilon_3^{(3)} = 1 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0\right) \in (0, 1), \\
& \mathbb{P}\left(\epsilon_3^{(3)} = 0 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0\right) = 1 - \mathbb{P}\left(\epsilon_3^{(3)} = 1 \mid \epsilon_1^{(1)} = 1, \epsilon_2^{(2)} = 0\right), \\
& \mathbb{P}\left(\epsilon_3^{(2)} = 1 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1\right) \in (0, 1), \\
& \mathbb{P}\left(\epsilon_3^{(2)} = 0 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1\right) = 1 - \mathbb{P}\left(\epsilon_3^{(2)} = 1 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 1\right), \\
& \mathbb{P}\left(\epsilon_3^{(1)} = 1 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0\right) \in (0, 1), \\
& \mathbb{P}\left(\epsilon_3^{(1)} = 0 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0\right) = 1 - \mathbb{P}\left(\epsilon_3^{(1)} = 1 \mid \epsilon_1^{(1)} = 0, \epsilon_2^{(1)} = 0\right).
\end{aligned} \tag{A2}$$

Note that the unconditional probabilities (A1) can be written concisely as

$$\begin{aligned}
\mathbb{P}\left(\epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}, 1\right) &= p_3^{\left(\left(0, k_1, k_2, k_3\right), \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}, 1\right)\right)} \in \left(0, p_2^{\left(\left(0, k_1, k_2\right), \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}\right)\right)}\right), \\
\mathbb{P}\left(\epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}, 0\right) &= p_3^{\left(\left(0, k_1, k_2, k_3\right), \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}, 0\right)\right)} \\
&= p_2^{\left(\left(0, k_1, k_2\right), \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}\right)\right)} - p_3^{\left(\left(0, k_1, k_2, k_3\right), \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}, 1\right)\right)},
\end{aligned}$$

where $k_1 = 1$, $k_2 = (\epsilon_1^{(k_1)})_{10} + 1$, $k_3 = (\epsilon_1^{(k_1)} \epsilon_2^{(k_2)})_{10} + 1$, with $(\epsilon_1^{(k_1)})_{10}$ and $(\epsilon_1^{(k_1)} \epsilon_2^{(k_2)})_{10}$ denoting the decimal values of the binary strings $\epsilon_1^{(k_1)}$ and $\epsilon_1^{(k_1)} \epsilon_2^{(k_2)}$. Similarly, the conditional probabilities (A2) can be written concisely as

$$\begin{aligned}
& \mathbb{P}\left(\epsilon_3^{(k_3)} = 1 \mid \mathcal{E}_2 = \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}\right)\right) \in (0, 1), \\
& \mathbb{P}\left(\epsilon_3^{(k_3)} = 0 \mid \mathcal{E}_2 = \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}\right)\right) = 1 - \mathbb{P}\left(\epsilon_3^{(k_3)} = 1 \mid \mathcal{E}_2 = \left(0, \epsilon_1^{(k_1)}, \epsilon_2^{(k_2)}\right)\right).
\end{aligned}$$

References

- Aït-Sahalia, Y. and Jacod, J. (2014). *High-Frequency Financial Econometrics*. Princeton University Press, Princeton, NJ.
- Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17):21–86.
- Beran, J. (2017). *Statistics for Long-Memory Processes*. Chapman & Hall/CRC, Boca Raton, FL.
- Billingsley, P. (2013). *Convergence of Probability Measures*. John Wiley & Sons, New York.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.
- Boyle, P. P. (1986). Option valuation using a three jump process. *International Options Journal*, 3:7–12.

- Breen, R. (1991). The accelerated binomial option pricing model. *Journal of Financial and Quantitative Analysis*, 26(2):153–164.
- Cheridito, P. (2003). Arbitrage in fractional Brownian motion models. *Finance and Stochastics*, 7:533–553.
- Cherny, A. S., Shiryaev, A. N., and Yor, M. (2003). Limit behavior of the "horizontal-vertical" random walk and some extensions of the Donsker-Prokhorov invariance principle. *Theory of Probability & Its Applications*, 47(3):377–394.
- Chorro, C., Guégan, D., and Ielpo, F. (2012). Option pricing for GARCH-type models with generalized hyperbolic innovations. *Quantitative Finance*, 12(7):1079–1094.
- Cordero, F., Klein, I., and Perez-Ostafe, L. (2014). Binary markets under transaction costs. *International Journal of Theoretical and Applied Finance*, 17(5):1450030.
- Cox, J. C. and Ross, S. A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1):145–166.
- Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263.
- Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes, Volume I: Elementary Theory and Methods*. Springer, Berlin.
- Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes, Volume II: General Theory and Structure*. Springer, Berlin.
- Davydov, Y. and Rotar, V. (2008). On a non-classical invariance principle. *Statistics & Probability Letters*, 78(14):2031–2038.
- Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520.
- Derman, E., Kani, I., and Chriss, N. (1996). Implied trinomial trees of the volatility smile. Quantitative Strategies Research Notes, Goldman Sachs, New York.
- Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance*, 5(1):13–32.
- Duan, J.-C., Ritchken, P., and Sun, Z. (2006). Approximating GARCH-jump models, jump-diffusion processes, and option pricing. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 16(1):21–52.
- Duffie, D. (2001). *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, NJ.
- Dzhaparidze, K. O. and van Zuijlen, M. C. A. (1996). Introduction to option pricing in a securities market I: Binary models. *CWI Quarterly*, 9(4):319–355.
- Easley, D. and O’Hara, M. (1995). Market microstructure. *Handbooks in Operations Research and Management Science*, 9:357–383.
- Easley, D. and O’Hara, M. (2003). Microstructure and asset pricing. *Handbook of the Economics of Finance*, 1:1021–1051.

- Eberlein, E. and Prause, K. (2000). The generalized hyperbolic model: Financial derivatives and risk measures. In Geman, H., Madan, D., Pliska, S. R., and Vorst, T., editors, *Mathematical Finance — Bachelier Congress 2000*, Springer Finance Series, pages 245–267, Berlin. Springer.
- Fan, J., Imerman, M. B., and Dai, W. (2016). What does the volatility risk premium say about liquidity provision and demand for hedging tail risk? *Journal of Business & Economic Statistics*, 34(4):519–535.
- Garman, M. B. (1976). Market microstructure. *Journal of Financial Economics*, 3(3):257–275.
- Hamilton, J. D. (2020). *Time Series Analysis*. Princeton University Press, Princeton, NJ.
- Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408.
- Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11(3):215–260.
- Hasbrouck, J. (1988). Trades, quotes, inventories, and information. *Journal of Financial Economics*, 22(2):229–252.
- Hasbrouck, J. (1996). Modeling market microstructure time series. In Maddala, G. S. and Rao, C. R., editors, *Statistical Methods in Finance*, volume 14 of *Handbook of Statistics*, chapter 22, pages 647–692. Elsevier, Amsterdam.
- Hasbrouck, J. (2007). *Empirical Market Microstructure: The Institutions, Economics, and Economics of Securities Trading*. Oxford University Press, Oxford, UK.
- Hu, Y., Shirvani, A., Lindquist, W. B., Fabozzi, F. J., and Rachev, S. T. (2020a). Option pricing incorporating factor dynamics in complete markets. *Journal of Risk and Financial Management*, 13(12):321.
- Hu, Y., Shirvani, A., Stoyanov, S., Kim, Y. S., Fabozzi, F. J., and Rachev, S. T. (2020b). Option pricing in markets with informed traders. *International Journal of Theoretical and Applied Finance*, 23(6):2050037.
- Hull, J. (2006). *Options, Futures, & Other Derivatives: Solutions Manual*. Prentice Hall International, Hoboken, NJ.
- Hung, M. and Swallow, W. H. (1999). Robustness of group testing in the estimation of proportions. *Biometrics*, 55(1):231–237.
- Jacod, J. and Protter, P. (2012). *Discretization of Processes*, volume 67 of *Stochastic Modeling and Applied Probability*. Springer, Berlin.
- Jarrow, R. and Rudd, A. (1982). Approximate option valuation for arbitrary stochastic processes. *Journal of Financial Economics*, 10(3):347–369.
- Jarrow, R. A., Protter, P., and Sayit, H. (2009). No arbitrage without semimartingales. *The Journal of Applied Probability*, 19(2):596–616.
- Kao, L.-J. (2012). Locally risk-neutral valuation of options in GARCH models based on variance-gamma process. *International Journal of Theoretical and Applied Finance*, 15(2):1250015.

- Kim, Y. S., Stoyanov, S., Rachev, S., and Fabozzi, F. (2016). Multi-purpose binomial model: Fitting all moments to the underlying geometric Brownian motion. *Economics Letters*, 145:225–229.
- Kim, Y. S., Stoyanov, S., Rachev, S., and Fabozzi, F. J. (2019). Enhancing binomial and trinomial equity option pricing models. *Finance Research Letters*, 28:185–190.
- Lo, A. W., Mamaysky, H., and Wang, J. (2000). Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. *The Journal of Finance*, 55(4):1705–1765.
- Mandelbrot, B. B. (2001). Stochastic volatility, power laws and long memory. *Quantitative Finance*, 1(6):558–559.
- Merton, R. C. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183.
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.
- Mills, T. C. (2019). *Applied Time Series Analysis: A Practical Guide to Modeling and Forecasting*. Academic Press, London.
- Mitov, G. K., Rachev, S. T., Kim, Y. S., and Fabozzi, F. J. (2009). Barrier option pricing by branching processes. *International Journal of Theoretical and Applied Finance*, 12(7):1055–1073.
- Mittnik, S., Fabozzi, F. J., Focardi, S. M., Rachev, S. T., and Jašić, T. (2007). *Financial Econometrics: From Basics to Advanced Modeling Techniques*. John Wiley & Sons, New York.
- O’Hara, M. (1997). *Market Microstructure Theory*. John Wiley & Sons, New York.
- O’Hara, M. (1999). Making market microstructure matter. *Financial Management*, 28(2):83–90.
- Oksendal, B. (2013). *Stochastic Differential Equations: An Introduction with Applications*. Springer, Heidelberg.
- Rachev, S. T., Kim, Y. S., Bianchi, M. L., and Fabozzi, F. J. (2011). *Financial Models with Lévy Processes and Volatility Clustering*. John Wiley & Sons, New York.
- Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127–1139.
- Rubinstein, M. (1994). Implied binomial trees. *The Journal of Finance*, 49(3):771–818.
- Rubinstein, M. (1998). Edgeworth binomial trees. *Journal of Derivatives*, 5(3):20–27.
- Samura, S. K., Mao, J., and Yao, D. (2013). Semimartingale property and its connection to arbitrage. *Journal of Mathematical Finance*, 3:237–241.
- Schoutens, W. (2003). *Lévy Processes in Finance: Pricing Financial Derivatives*. John Wiley & Sons, New York.
- Shiryayev, A. N. (1999). *Essentials of Stochastic Finance: Facts, Models, Theory*, volume 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific, Singapore.
- Skorokhod, A. V. (1997). Measure-valued diffusion. *Ukrainian Mathematical Journal*, 49:506–513.

Skorokhod, A. V. (2005). *Basic Principles and Applications of Probability Theory*. Springer, Heidelberg.