Applied Mathematics and Computation 471 (2024) 128611

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Full Length Article

Dividend based risk measures: A Markov chain approach ☆

Guglielmo D'Amico [a], Riccardo De Blasis [b,*]

[a] Department of Economics, University of Chieti-Pescara, Viale Pindaro, 42, Pescara, 65127, Italy
[b] Department of Management, Marche Polytechnic University, Piazzale Martelli, 8, Ancona, 60121, Italy

ARTICLE INFO

Keywords:
Stock valuation
Fundamental price
Maximum entropy
Price moments
Risk measures

ABSTRACT

Computations of risk measures in the context of the dividend valuation model is a crucial aspect to deal with when investors decide to buy a share of common stock. This is achieved by using a Markov chain model of growth-dividend evolution, imposing an assumption that controls the growth of the dividend process and in turn allows for the computation of the moments of the price process and the fulfillment of a set of transversality conditions which allows avoiding the presence of speculative bubbles in the market. The probability distribution of the fundamental value of the stock is recovered by solving a moment problem, based on the solution of a maximum-entropy approach from which it is possible to compute classical risk measures based on these fundamental variables. The methodology is applied to real dividend data from the S&P 500 index. Results show that our model provides complete information about the fundamental price not limited to its expectation.

1. Introduction

Even though the Market Efficiency Hypothesis (EMH) states that the fundamental analysis should 'not work' in predicting future values [1], the past literature has discovered a lot of anomalies that led to successful trading strategies linked to earning or returns momentum, earning surprises, stock issuance, and others which allowed investors to make profits due to the differences between the market price and the true value of the stock [2,3].

The first evidence of a misalignment between market prices and fundamental values can be dated back to Williams [4], who was the first to acknowledge the existence of these differences. Since then, the literature has attempted to address two main questions. The first one is related to the estimation of the fundamental or fair value of a stock. To this extent, many scholars addressed their interest in the intrinsic long-term value by proposing different valuation approaches, starting from the classical Gordon growth model [5] and its extensions, which are all based on a deterministic behavior of the dividend growth [see, e.g., 6–9]. However, because the expected dividend growth, which is positively correlated with expected returns, cannot be considered constant [10], other authors introduced a stochastic process to model this growth. In particular, Hurley and Johnson [11,12] modeled the growth as a Markov dividend stream, and Yao [13] extended their works using a trinomial dividend valuation. On this path, Ghezzi and Piccardi [14] generalize the previous approaches by modeling the dividend growth with an n-state Markov chain, while D'Amico [15] moves a step forward with the introduction of a semi-Markov process. For a comprehensive review of the dividend discount models, we refer the reader to [16]. All Markov chain based models are flexible and can be designed to reproduce some stylized facts of the stocks, e.g.,

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author.
E-mail address: r.deblasis@univpm.it (R. De Blasis).

https://doi.org/10.1016/j.amc.2024.128611
Received 20 March 2023; Received in revised form 1 February 2024; Accepted 15 February 2024

Available online 26 February 2024

momentum, mean reversion, and cyclic behaviors [see, e.g., 17]. Moreover, a more recent work introduces a stochastic multi-period model adopting a compound non-homogenous Poisson process for the dividend growth [18].

The second question attains to the possibility of using the signals from the fundamental analysis to build investment portfolio selection strategies. Indeed, this last question is more interesting to practitioners; however, it has recently received little attention from academia. For example, Zhang and Yan [3] connect the theory of the fundamental analysis with the portfolio selection strategies. The authors derive a closed-form information ratio of the investors who can gain information about the fundamentals of a security, allowing them to solve a portfolio choice problem in a continuous-time framework. To this extent, the authors point out that the fundamental analysis theory should not be limited to the assessment of fair prices (or returns) and whether there are misalignments with the market prices, but it can also help in the valuation of the fundamental risk as compared to its market counterpart. On a similar perspective, Barbu et al. [19], extending the work of [14], concentrate their attention on the second moment of the price process including the computation of the risk advanced by [20]. Moreover, D'Amico and De Blasis [21] propose a similar approach to a multivariate setting departing from the explicit formula for the covariance between random stock price by [22]. Finally, D'Amico [23] performs the same computation but employs a semi-Markov model with general phase space.

This paper extends previous work on the Markov chain dividend model from a mathematical point of view, departing from the results concerning the first two moments and moving to general results on the moments of any order. This extension requires the advancement of a general condition that allows us to control the expansion of the moments of the dividend growth process. This condition is sufficient for the fulfillment of the set of transversality conditions, thus ruling out bubbles from the financial markets. The moments of the price process are finally expressed in terms of a recurrent type system of linear equations. From a methodological point of view, we recover the density function of the price process and the consequential computation of some risk indicators solving a maximum entropy problem. Finally, from a financial point of view, our results allow the investors to compare the market risk perception and the true risk motivated by the fundamental variable; the latter may be used for accurate investment decisions. To our knowledge, our paper is the first one contributing to the literature with the analysis of higher moments, thus it might help fill this gap, which is also highlighted by [3].

The paper unfolds as follows. Section 2 describes how the Markov chain is implemented in the stock valuation, Section 3 introduces the computation of the moments of the price process, while Section 4 explains how these moments are employed to retrieve the probability density function and thus computing some risk measures. Section 5 shows the results of the empirical application of the proposed model. Finally, Section 6 closes the manuscript.

2. The Markov chain dividend valuation model

Consider a stock that pays dividends on schedule. Since the dividend value is unknown in advance, it is usually assumed that it is generated by the discrete-time stochastic process $D = \{D(t)\}_{t\in\mathbb{N}}$. The latter is defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ expressing the natural filtration produced by the process D and a physical (or real–world) probability measure $\mathbb{P}$.

The random variable $P(t)$ that describes the fundamental value of a stock at time t, is related to the dividend process by means of the fundamental financial relation:

$$P(t) = \frac{D(t+1) + P(t+1)}{r}, \tag{1}$$

being r a certain constant discounting factor, i.e., one plus the required rate of return for the stock. The price at time t is expressed according to $p(t) := \mathbb{E}_t[P(t)]$ which translates to

$$p(t) = \frac{\mathbb{E}_t[D(t+1) + P(t+1)]}{r}, \tag{2}$$

with $\mathbb{E}_t$ being the conditional expectation operator based on the information known at time t.

It should be highlighted that the value $P(t)$ cannot be quantified in terms of the information set $\mathcal{F}_t$ that the economy had access to up until time t. To demonstrate this point, we observe that $P(t)$ depends on the upcoming dividend $D(t+1)$ as well as the upcoming value $P(t+1)$, where the latter is dependent on upcoming payouts and values as it can be seen through an iteration of equation (2).

To avoid the presence of bubbles in the market [see, 24], Samuelson [25] assumes the validity of the transversality condition

$$\lim_{i\to\infty} \frac{\mathbb{E}_t[P(t+i)]}{r^i} = 0. \tag{3}$$

Thus, the solution of the price function in (1) can be expressed as

$$p(t) = \sum_{i=1}^{\infty} \frac{\mathbb{E}_t[D(t+i)]}{r^i}. \tag{4}$$

Relation (4) relates the fundamental value of a stock to the discounted present value of the future dividend it generates.

In the traditional stock valuation models, several assumptions about the dividend process have been made. For example, Gordon [5] assumes a constant growth rate while other authors introduce 2-stage and 3-stage models [see, e.g., 6–9].

Contrary to the traditional literature, the Markov chain stock model considers the dividend growth rate as a sequence of discrete random variables $\{G(t)\}_{t\in\mathbb{N}}$ described by a Markov process with a finite state-space E [see, e.g., 11–14,19]. The set E contains the possible values assumed by the growth-dividend process at any time.

Specifically, assuming that the dividend series obeys the difference equation

$$D(t+1) = G(t+1)D(t), \tag{5}$$

the price can be described as follows [14]

$$p(t) = d(t) \sum_{i=1}^{\infty} \frac{\mathbb{E}_{(t)}\left[\prod_{j=1}^{i} G(t+j)\right]}{r^i} =: d(t)\psi(g(t)), \tag{6}$$

where $d(t)$ and $g(t)$ are the values assumed by the dividend and the dividend growth process at time t, and $\psi(g(t))$ is the price-dividend ratio. It is important to note that in all analyzed models the discounting factor r is considered constant.

Ghezzi and Piccardi [14] and Barbu et al. [19] started from equation (5) to develop the Markov stock valuation model. Departing from the same formulation, we extend the previous work by computing the nth moment of the price process, which requires a more complex demonstration approach.

3. The computation of the moments of the price process

The objective of this section is to demonstrate how to calculate the higher-order moments of the price process. We advance a sufficient condition for the finiteness of the moments and the fulfillment of the transversality conditions. The price-dividend ratio of any order is successively determined using a recurrent set of linear equations, which provides the moments of the price process.

Let us consider again the fundamental relationship between prices and dividends in an efficient market:

$$P(t) = \frac{D(t+1) + P(t+1)}{r}. \tag{7}$$

Raising to the nth power both on the right and left sides, and substituting into $P(t+1)$ the corresponding representation given by relation (7), gives:

$$\begin{aligned} P^{(n)}(t) &:= (P(t))^n = \left(\frac{D(t+1)+P(t+1)}{r}\right)^n \\ &= \left(\frac{D(t+1)}{r} + \frac{1}{r}\cdot\frac{D(t+2)+P(t+2)}{r}\right)^n \end{aligned}$$

A simple iterative argument gives:

$$P^{(n)}(t) = \left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i} + \frac{P(t+N)}{r^N}\right)^n,$$

which can be further expanded using Newton's binomial formula:

$$\begin{aligned} P^{(n)}(t) &= \sum_{k=0}^{n}\binom{n}{k}\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^k \cdot \left(\frac{P(t+N)}{r^N}\right)^{n-k} \\ &= \left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^n + \sum_{k=0}^{n-1}\binom{n}{k}\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^k \cdot \left(\frac{P(t+N)}{r^N}\right)^{n-k}. \end{aligned}$$

The previous relationship is particularly informative from a financial point of view. It affirms that the nth order moment of the price process does depend on the nth power of the sum of the discounted dividends plus an additional term that involves lower order powers of the price process itself at time $t+N$. This last part tells us that the risk, measured by the nth power of the price process, depends on its own temporal paths, which are represented by the power of any order up to the nth. This implies that the risk may increase purely as a result of an anticipated rise in the future risk represented by the values $(P(k+N), P^{(2)}(k+N), \dots, P^{(n)}(k+N))$, with no connection to the inherent risk of the dividend process that is detailed in the addend $\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^n$.

Previous relations are expressed in terms of random variables. Our next step is to move to the corresponding expected values. Accordingly, we obtain

$$\begin{aligned} p^{(n)}(t) &:= \mathbb{E}_{(t)}\left[P^{(n)}(t)\right] \\ &= \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^n\right] + \sum_{k=0}^{n-1}\binom{n}{k}\mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^k \cdot \left(\frac{P(t+N)}{r^N}\right)^{n-k}\right]. \end{aligned} \tag{8}$$

Now, if $\forall k = 0, 1, \dots, n-1$ it results that

$$\lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N}\frac{D(t+i)}{r^i}\right)^k \cdot \left(\frac{P(t+N)}{r^N}\right)^{n-k}\right] = 0, \tag{9}$$

then we obtain that

$$p^{(n)}(t) = \lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N} \frac{D(t+i)}{r^i}\right)^n\right] = \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{\infty} \frac{D(t+i)}{r^i}\right)^n\right], \tag{10}$$

where the final equality results from the monotone convergence theorem being applied to the series of random variables $X_N := \left(\sum_{i=1}^{N} \frac{D(t+1)}{r^i}\right)$ which are non-negative and increasingly converging to $\sum_{i=1}^{\infty} \frac{D(t+1)}{r^i}$.

Now, if we consider the multiplicative nature of our dividend model (see equation (5)) we have the final representation

$$p^{(n)}(t) = \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{\infty} \frac{\prod_{j=1}^{i} G(t+j)d_t}{r^i}\right)^n\right]. \tag{11}$$

The relations (10) and (11) express the moments of the price process as dependent on the expectation of the dividend process, making it the desired arrival point in an efficient financial market. As a result, we have a representation of the price and risk measures that simply takes into account the dividends. Hence, the dividend dynamic is the fundamental variable that explains the behavior of the price process.

We observe that in Equation (11), d_t is known, as well as $G(t) = g(t)$, therefore

$$\frac{p^{(n)}(t)}{d_t^n} = \psi_n(g(t)) = \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{\infty} \frac{\prod_{j=1}^{i} G(t+j)}{r^i}\right)^n\right]. \tag{12}$$

The quantity $\psi_n(g(t))$ is the price-dividend ratio of order n which will be a key variable in the computation of the price moments. In particular, the price-dividend ratio is dependent on the dividend growth level at time t, which represents a known initial condition. As the dividend growth varies, it provides different price-dividend levels, thus different price moments.

Our objective is to advance sufficient conditions on the parameters of the model such that the moments of the price process are finite and the transversality conditions (9) are fulfilled. To this end, given the Markov chain model with state space $E = \{g_1, g_2, \ldots, g_s\}$ and transition probabilities matrix $\boldsymbol{P} = (p_{ij})$, $i, j \in E$, we assume that

Assumption $\mathbf{A_n}$. $\bar{g}^{(k)} := \max_{i\in E}\left\{\sum_{j\in E} p_{ij} g_j^k\right\} < r^k, \quad \forall k = 1, 2, \ldots, n.$

It is worth noting that this assumption is the natural extension of those considered in [14] and [19] that were set to compute the fundamental price (expectation of the price process) and a preliminary measure of risk such as the second order moment of the price process within a Markov chain dividend valuation model.

Theorem 3.1. *Let $D(t) = d_t > 0$ and $G(t) = g_i \in E$. Under the hypothesis*

$$\mathbf{A_n}: \ \bar{g}^{(k)} := \max_i \left\{\sum_{j\in E} p_{ij} g_j^k\right\} < r^k, \quad \forall k = 1, 2, \ldots, n,$$

we have that the series

$$p^{(n)}(t) = \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{\infty} \frac{\prod_{j=1}^{i} G(t+j)d_t}{r^i}\right)^n\right] < \infty,$$

and the following asymptotic condition is satisfied $\forall k = 0, \ldots, n-1$

$$\lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N} \frac{D(t+i)}{r^i}\right)^k \left(\frac{P(t+N)}{r^N}\right)^{n-k}\right] = 0. \tag{13}$$

Proof. See Appendix A.1. □

Remark 1. Theorem 3.1 presents a sufficient condition that satisfies the transversality conditions (13). Consequently, the presence of speculative bubbles in the model is avoided and therefore the representation of moments of the price process as convergent series that depend only on the dividend process is permitted. Furthermore, assumption $\mathbf{A_n}$ controls each moment of the dividend growth process in order to be able to provide a corresponding result on the moments of the price process.

Remark 2. Theorem 3.1 provides a broad generalization of the results given in [14] and [19] where only the first- and second-order moments were considered, respectively. Additionally, it should be noted that the proof uses a different approach because the Cauchy-Schwarz inequality does not apply in our context, making it impossible to directly extend the techniques employed in [19]. Indeed, when considering the moments of order $k > n/2$ the inequality requires information on the moment of order $2k$ which is greater than the maximum order n considered.

According to the results obtained in Theorem 3.1, we can evaluate the moments of the price process directly by computing the corresponding series involving the expectation of the dividend process. However, below we show an alternative method of calculation that can be conveniently represented by introducing a set of auxiliary functions called price-dividend ratios of higher orders. This method was previously used for Markov and semi-Markov chain models for the computation of the price [14,15,23] for the risk measured with the second-order moment [19,23] and in a multivariate setting [21].

Before showing the representation of the price-dividend ratio of order n, we introduce some matrix notation. Let $\mathbf{I}$ be the identity matrix of dimension $s \times s$. For any $r \neq 0$, we define $\mathbf{I}_r := r\mathbf{I}$ and more in general $\mathbf{I}_r^n = \mathbf{I}_{r^n}$.

Moreover, for any column vector $\mathbf{g} = (g_1, \dots, g_s)^\top$, $\mathbf{g}^n = (g_1^n, \dots, g_s^n)^\top$ with $()^\top$ as the transpose of a vector, we define by

$$\mathbf{I}_\mathbf{g} = (I_g(i,j))_{i,j\in E},\ \ I_\mathbf{g}(i,j) = \begin{cases} g_i, & \text{if } i = j, \\ 0, & \text{if } \ i \neq j, \end{cases} \tag{14}$$

and more in general it results that $\mathbf{I}_\mathbf{g}^n = \mathbf{I}_{\mathbf{g}^n}$ and $\mathbf{I}_\mathbf{g}^{-1} = \mathbf{I}_{\mathbf{g}^{-1}}$.

According to Equation (12), we define the vector $\boldsymbol{\Psi}_n$ of price-dividend ratios of order n.

Definition 3.1. The generic element of the column vector $\boldsymbol{\Psi}_n = (\psi_n(g_1), \dots, \psi_n(g_s))^\top$, given $G(t) = g_k$ and $D(t) = d_t$, is

$$\psi_n(g_k) = \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{\infty} \frac{\prod_{h=1}^{i} G(t+h)}{r^i}\right)^n\right].$$

Observe that $\boldsymbol{\Psi}_0 = (1, \dots, 1)^\top$.

Proposition 3.2. *Suppose that hypothesis* $\mathbf{A_n}$ *holds. Then, the vector* $\boldsymbol{\Psi}_n$ *is the unique and non-negative solutions of the linear system of equations*

$$\left(\boldsymbol{I}_r^n - \boldsymbol{P}\cdot \boldsymbol{I}_g^n\right)\cdot \boldsymbol{\Psi}_n = \boldsymbol{P}\cdot\left(\sum_{m=1}^{n-1}\binom{n}{m}\boldsymbol{g}^n \circ \boldsymbol{\Psi}_{n-m} + \boldsymbol{g}^n\right), \tag{15}$$

where $\cdot$ *denotes the usual row-by-column matrix product and* $\circ$ *is the Hadamard element-by-element product.*

Proof. See Appendix A.2. $\square$

Finally, we observe that the price-dividend ratio for each fixed order n satisfies the system (15), which is of recurrent nature in the sense that when solving it for the nth order we need to know the solutions of the corresponding systems at lower order moments. For example, the price-dividend ratio of order three can be obtained according to

$$\boldsymbol{\Psi}_3 = \left(\boldsymbol{I}_r^3 - \boldsymbol{P}\cdot \boldsymbol{I}_g^3\right)^{-1}\cdot \boldsymbol{P}\cdot\left(3\boldsymbol{g}^3\circ\boldsymbol{\Psi}_2 + 3\boldsymbol{g}^3\circ\boldsymbol{\Psi}_1 + \boldsymbol{g}^3\right),$$

where inverse matrix of $\left(\boldsymbol{I}_r^3 - \boldsymbol{P}\cdot \boldsymbol{I}_g^3\right)$ exists according to Proposition 3.2. This can be done conditionally to the knowledge of the price-dividend ratios of first and second order that can be obtained from the same proposition setting $n = 1$ and $n = 2$ obtaining:

$$\boldsymbol{\Psi}_1 = \left(\boldsymbol{I}_r - \boldsymbol{P}\cdot \boldsymbol{I}_g\right)^{-1}\cdot \boldsymbol{P}\cdot \boldsymbol{g},$$
$$\boldsymbol{\Psi}_2 = \left(\boldsymbol{I}_r^2 - \boldsymbol{P}\cdot \boldsymbol{I}_g^2\right)^{-1}\cdot \boldsymbol{P}\cdot\left(2\boldsymbol{g}^2\circ\boldsymbol{\Psi}_1 + \boldsymbol{g}^2\right).$$

4. Risk measures based on the moment problem

The knowledge of the moments $p^{(k)}(g_i, d_t)$, $k = 1, \dots, n$, conditionally on the observed dividends $D(t) = d_t$ and their growth values $G(t) = g_i$, enables us to look for the associated probability density function (PDF) $f(\cdot; g_i, d_t)$ of the price process at time t. For the rest of this section, we will suppress the use of the dividend and growth values at time t to facilitate notation and denote the PDF of the price according to $f(\cdot)$.

To this extent, we can consider the classical moment problem in which the density is derived from the knowledge of its moments $p^{(n)}(t)$,

$$\int_{\Omega} x^n f(x)dx = p^{(n)}(t), \quad n = 0, 1, 2, \dots, \tag{16}$$

where Ω is the support of the distribution.

The PDF $f(x)$ allows us to calculate any risk measures related to the price process. The most common risk measures employed by the financial literature are the value-at-risk (VaR) and the expected shortfall (ES), also known as conditional value-at-risk (cVaR) [see, e.g. 26,27].

In our context, the VaR at level α represents the smallest price that will not be exceeded with probability $1-\alpha$. Therefore, given the PDF $f(x)$ with its cumulative distribution function (CDF) $F(x)$, we can write

$$VaR_\alpha(P) = sup\{x \in \mathbb{R} : F(x) \leq 1-\alpha\} = F^{-1}(1-\alpha), \tag{17}$$

where F^{-1} is the inverse of the CDF.

The VaR measure presents some limitations because it does not take into account the shape of the tail of the price distribution because it only considers the worst price at a certain confidence level. On the contrary, the ES measures the expected price given that the VaR has been reached [see, 28]. It can be computed as the average VaR as

$$ES_\alpha(P) = \frac{1}{\alpha}\int_0^\alpha VaR_\gamma(P)d\gamma. \tag{18}$$

Therefore, to compute the previous risk measures, we have to solve the general inverse problem in (16) which poses some difficulties in finding a unique density, $f(x)$, as in practical application we only know $N+1$ moments, which lead to infinite functions with the same first $N+1$ moments.

To overcome this problem, the common approach in the literature is to find an approximation procedure to construct specific sequences of functions $f_N(x)$, which will converge to the true distribution $f(x)$ as $N \to \infty$. Following this scheme, Mead and Papanicolau [29] employed the maximum-entropy approach to identify a definite procedure for the construction of a sequence of approximations. They consider the positive density $f(x)$ as a probability density and then maximize its entropy under the condition that the first $N+1$ moments be equal to the true moments $p^{(n)}(t), n = 0, 1, ..., N$.

The procedure consists in maximizing the information entropy of the distribution $f(x)$, defined as

$$S = -\int_\Omega f(x)\ln f(x)dx, \tag{19}$$

subject to the known $N+1$ moments,

$$\int_\Omega x^n f(x)dx = p^{(n)}(t), \quad n = 0, 1, 2, ..., N, \tag{20}$$

with $p^{(0)}(t) = 1$.

Introducing the Lagrange multipliers, we can maximize the following entropy functional

$$H = S + \sum_{n=0}^{N}\lambda_n\left(\int_\Omega x_n f(x)dx - p^{(n)}(t)\right), \tag{21}$$

by taking its derivatives with respect to λ_n and $f(x)$ and setting them to zero. $\frac{\partial H}{\partial f(x)} = 0$ gives us the constraints defined in (20), however, the partial derivative $\frac{\partial H}{\partial \lambda_n} = 0$ allows us to obtain the density as a function of the Lagrange multipliers, as

$$f = f_N(x) = e^{-1+\sum_{n=1}^N \lambda_n x^n} = e^{\sum_{n=0}^N \lambda_n x^n}, \tag{22}$$

where we can set the value of λ_0 as $(\lambda_0 - 1)$.

The maximum entropy is obtained by solving the following non-linear system of equations

$$\int_\Omega x^n e^{\sum_{n=0}^N \lambda_n x^n} dx = p^{(n)}(t), \quad n = 0, 1, 2, ..., N, \tag{23}$$

which can be solved numerically using a globally convergent Newton solver, as proposed by Saad and Ruai [30]. The authors implement the procedure in their software *PyMaxEnt* developed in Python language. Moreover, they suggest setting the initial guesses as follows

$$\lambda_i = \begin{cases} -\ln\sqrt{2\pi} & i = 0 \\ 0 & \text{otherwise,} \end{cases}$$

based on empirical observations and a Gaussian distribution with zero mean and unit variance.

Finally, the obtained PDF allows us to compute the risk measures defined in (17) and (18).

Table 1
Summary statistics of the S&P 500 index price and dividend series.

	Price	Dividend	P_t/P_{t-1}	D_t/D_{t-1}
Obs	1824	1824	1823	1823
mean	357.941	7.282	1.005	1.003
std	763.238	13.216	0.041	0.011
min	2.730	0.180	0.735	0.911
25%	7.920	0.423	0.985	1.000
50%	17.770	0.912	1.007	1.004
75%	168.675	7.541	1.028	1.008
max	4674.773	66.920	1.503	1.060

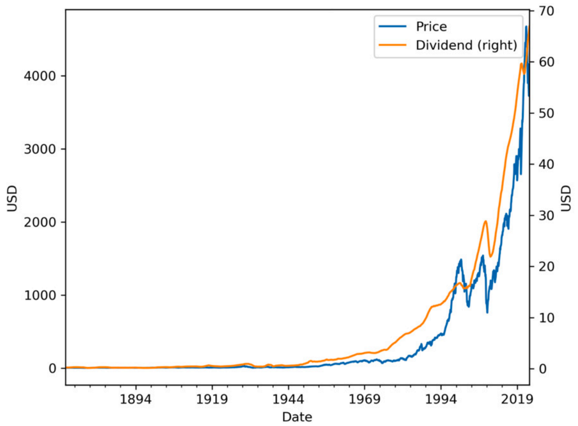

Fig. 1. S&P 500 index price and dividend series.

Table 2
5-states dividend discretization and median value for each state.

	State 1	State 2	State 3	State 4	State 5
Intervals	[0.911, 0.996)	[0.996, 1.001)	[1.001, 1.007)	[1.007, 1.012)	[1.012, 1.06]
Median	0.991	1.000	1.004	1.009	1.018

5. Application

We apply the proposed model to the series of dividends from the S&P 500 index from Shiller [31]. The data, ranging from January 1871 to December 2022, are regularly updated and available at http://www.econ.yale.edu/~shiller/data.htm. The dividends are generally paid quarterly, however, for the estimation of the Markov chain transition probabilities, we use the monthly dividends from the dataset which are computed as a linear interpolation from quarter data since 1926, while the dividends before 1926 are interpolated from annual data. The index prices are monthly averages of daily closing prices. The summary statistics of the monthly dividends and prices up to December 2022, along with their growth series, are reported in Table 1 while their trends are shown in Fig. 1. To allow for a direct comparison, the figure shows two different scales for prices and dividends, with values reported on the left and right axes, respectively, and both expressed in US dollars. In particular, we notice that both trends are comparable to each other showing a direct relationship between the two quantities.

According to the model described in Section 2, we assume that the dividend growth series follows a Markov chain. Thus, to estimate the transition probabilities, we have to discretize the growth series. To this extent, we follow the procedure in [19] which consider a central state containing all continuous values falling within half standard deviation radius from the zero growth. Then, the adjacent states are defined at steps of one standard deviation up to the extremes of the distribution. As in [19], we follow this approach because the highest frequencies of the distribution are concentrated in the middle of the distribution; therefore, it is natural to include them within the central state of the discretized series. On the contrary, we leave the tails of the distribution to the external states. A visual representation of the discretization, along with the histogram of the dividend growth process, is shown in Fig. 2 in which the continuous distribution is divided into 5 states (right chart). Furthermore, Table 2 reports the intervals of the 5-state dividend growth discretization along with the median value for each state. The latter can be used as the representative value for each state.

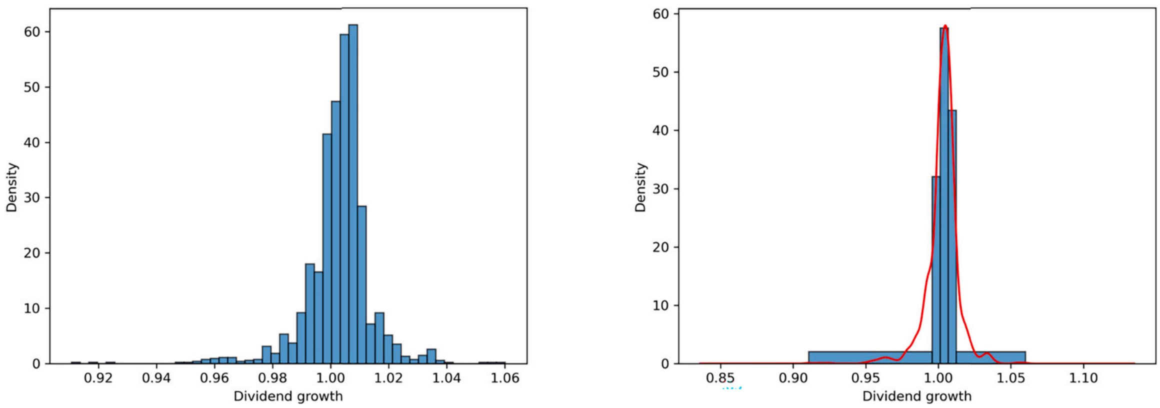

Fig. 2. Dividend growth histogram (left) and 5-state discretization (right).

Once we obtain the discretized growth values, we can estimate the transition probabilities from the monthly data using the non-parametric estimator $p_{ij} = N(i,j)/N(i)$ where $N(i,j)$ is the number of transitions from state g_i to state g_j and $N(i)$ is the total visits to state g_i:

$$\boldsymbol{P} = \begin{pmatrix} 0.917 & 0.042 & 0.025 & 0.013 & 0.003 \\ 0.037 & 0.842 & 0.090 & 0.025 & 0.006 \\ 0.012 & 0.048 & 0.862 & 0.061 & 0.017 \\ 0.009 & 0.019 & 0.080 & 0.857 & 0.035 \\ 0.017 & 0.006 & 0.046 & 0.091 & 0.840 \end{pmatrix}. \tag{24}$$

We note that the matrix in (24) is diagonally dominant, that is

$$\forall i \in E,\ p_{ii} \geq 1 - p_{ii},$$

highlighting that the dividend growth process shows low mobility, which means that the probability of being in the same state is never lower than 84%. Moreover, the probability distributions appearing on the different rows of the matrix $\boldsymbol{P}$ are stochastically ordered in the sense that

$$\mathbf{p}_{i\cdot} \leq_{st} \mathbf{p}_{j\cdot},\ \forall i < j.$$

The symbol $\leq_{st}$ denotes the usual stochastic ordering relation between two probability distributions. This means that $\forall i < j$ and $\forall k \in \{1, \ldots, 5\}$ it results that

$$\sum_{h=1}^{k} p_{ih} \geq \sum_{h=1}^{k} p_{jh}.$$

There is only a slight violation of this empirical evidence that appears for $i = 4, j = 5$ and $k = 1$ having that $p_{41} = 0.009 < p_{51} = 0.017$. Hence, from a practical point of view, we can say that being in a lower state of the dividend growth process implies lower chances of arrival in higher states of the process as compared to starting from a higher state.

The next step is solving the system of equations in (15) to find the unique solutions, i.e., the price-dividend ratios $\psi_n(g_i)$. Now, because the dividends are paid quarterly, to obtain more realistic values we modify the required parameters to a quarterly scale. We compute the third power of the transition probability matrix to have quarterly transitions, then we estimate the required discounting factor from the quarterly observed price returns as the average of the median returns of each state. Then, using the transition probability matrix, the state values reported in Table 2, and the obtained value for the discounting factor $r = 1.019$, we controlled for assumption $\mathbf{A_n}$ and computed the price-dividend values up to the 10^{th} moment for each state of the Markov chain. Results are reported in Table 3 and scaled for readability. According to Definition 3.1 and for each moment's order, the table presents a price-dividend ratio for each state of the dividend growth process. Specifically, $\psi_1 = 62.731$ gives us the price-dividend ratio of the first order when the dividend growth process $G(t) = g_1$. We highlight that the price-dividend ratios increase monotonically at each order with respect to the states g_i, similar to what is reported in [19]. Moreover, we note that this behavior directly translates from the transition probabilities, which are monotonic among the states of the Markov chain.

The price dividend ratios allow computing the price moments $p^{(n)}(g_i, d_t)$, which depend on the state of the dividend growth and the dividend at time t. As an example, Table 4 reports the first ten price moments for each state as of December 2022, when the dividend value was \$66.92. The values of the first moment are reported in thousands of US dollars, while higher moments are scaled by the nth power of thousands of US dollars, e.g., the second moment is expressed in millions of US dollars, and so on. This scaling is necessary to keep the values within a reasonable range for the numerical computation of the maximum entropy. To read the table, we need to compute the state of the dividend growth process as of December 2022. Specifically, the value of the dividend in November

Table 3
Price-dividend ratios of the S&P 500 index.

	ψ_1	$\psi_2 \cdot 10^2$	$\psi_3 \cdot 10^4$	$\psi_4 \cdot 10^6$	$\psi_5 \cdot 10^8$	$\psi_6 \cdot 10^9$	$\psi_7 \cdot 10^{11}$	$\psi_8 \cdot 10^{13}$	$\psi_9 \cdot 10^{15}$	$\psi_{10} \cdot 10^{17}$
State 1	62.731	39.809	25.554	16.592	10.895	72.355	48.591	32.996	22.654	15.726
State 2	64.670	42.276	27.933	18.653	12.588	85.834	59.138	41.165	28.949	20.566
State 3	65.583	43.467	29.112	19.701	13.470	93.044	64.924	45.761	32.579	23.427
State 4	66.244	44.348	30.002	20.509	14.164	98.826	69.655	49.591	35.661	25.901
State 5	67.035	45.437	31.137	21.570	15.104	106.899	76.459	55.263	40.361	29.784

The quarterly discounting factor $r = 1.019$ is computed from the observed price returns.

Table 4
Prices moments of the S&P 500 index as of December 2022.

	$p^{(1)}$	$p^{(2)}$	$p^{(3)}$	$p^{(4)}$	$p^{(5)}$	$p^{(6)}$	$p^{(7)}$	$p^{(8)}$	$p^{(9)}$	$p^{(10)}$
State 1	4.198	17.827	76.582	332.748	1462.243	6498.404	29204.166	132710.063	609756.274	2832549.326
State 2	4.328	18.932	83.712	374.088	1689.357	7708.913	35543.151	165569.553	779186.729	3704402.086
State 3	4.389	19.466	87.246	395.109	1807.818	8356.496	39020.679	184051.939	876877.295	4219608.927
State 4	4.433	19.860	89.914	411.306	1900.951	8875.787	41864.054	199455.789	959841.374	4665309.738
State 5	4.486	20.348	93.314	432.595	2027.135	9600.805	45953.461	222269.851	1086337.847	5364704.636

$p^{(1)}$ prices in thousands of US dollars. Last paid dividend in December 2022: \$ 66.92.

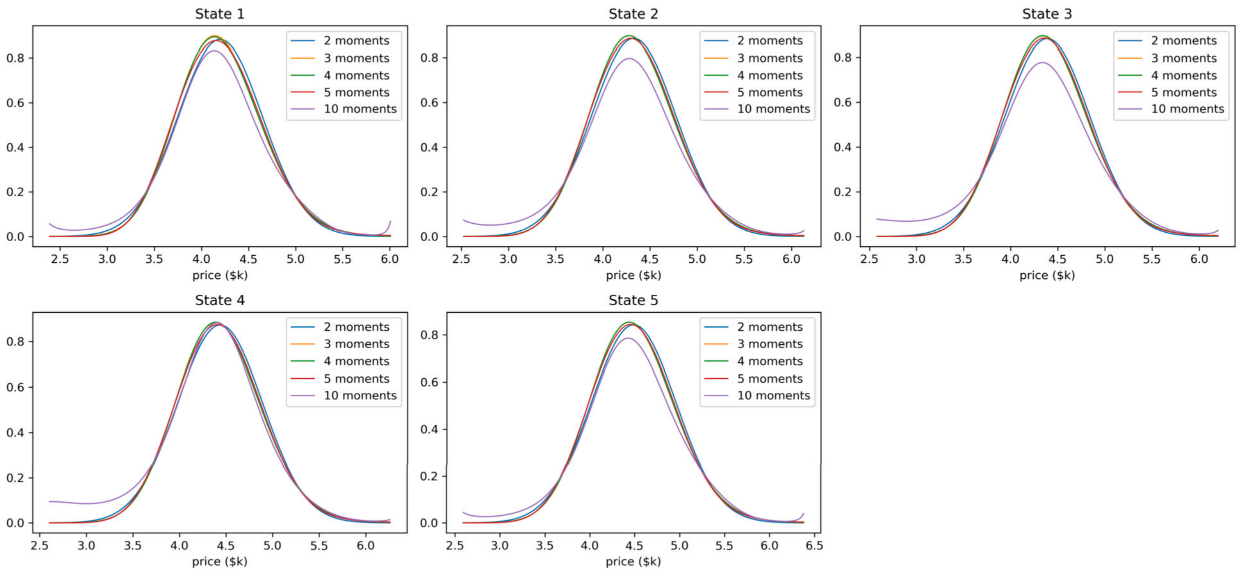

Fig. 3. Probability density functions of the S&P 500 index price as of December 2022.

Table 5
Statistics of the price process distribution.

	Mean	Std	Skew	Kurt
State 1	4197.94	452.51	0.26	3.09
State 2	4327.70	450.77	0.22	3.10
State 3	4388.78	452.11	0.22	3.11
State 4	4433.02	456.97	0.21	3.11
State 5	4485.95	473.44	0.21	3.10

Prices in US dollars.
Last paid dividend in December 2022: \$ 66.92.

2022 was \$66.39, therefore the dividend growth $G(t) = D(t)/D(t-1) = 66.92/66.39 = 1.008$ corresponds to state 4 of the Markov chain. Finally, we consider the price $p^{(1)}(g_4, 66.92) = \$4,433.00$ as the corresponding price of the index for that specific time.

The results of the price moments allow us to obtain information on the shape of the distribution. Specifically, we can compute the mean, standard deviation, skewness and kurtosis, that we report in Table 5. Moreover, from the computed price moments, we can infer the PDFs, as well as the CDFs, of the price process corresponding to each state for a specific date using the maximum entropy approach described in Section 4. We perform several computations with different $N+1$ moments, starting from the first two

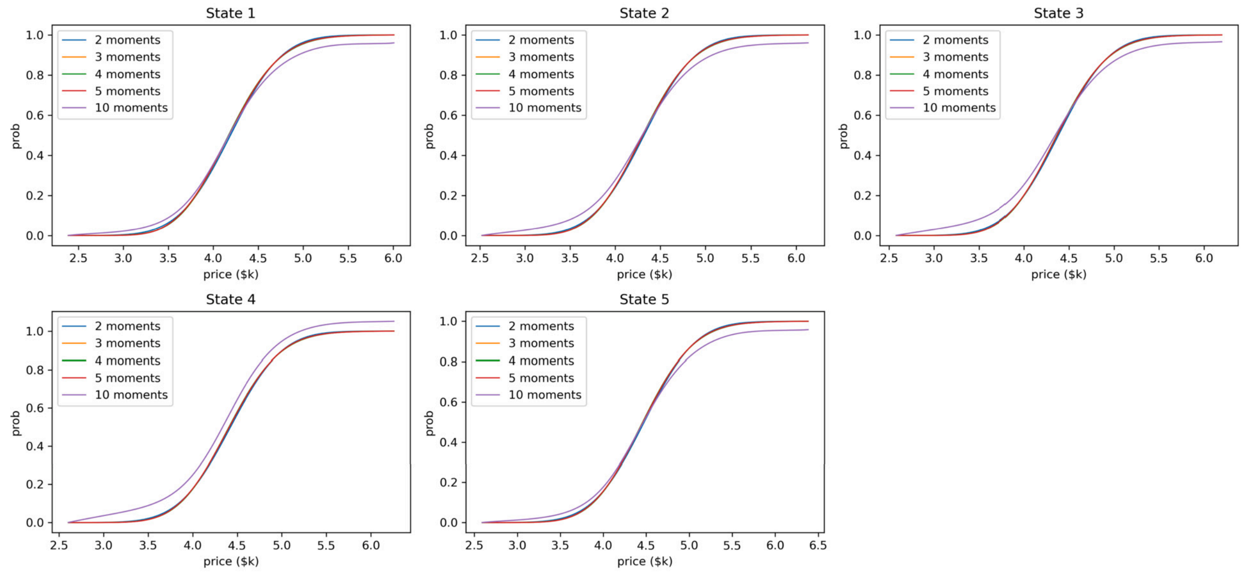

Fig. 4. Cumulative distribution functions of the S&P 500 index price as of December 2022.

Table 6
95% Value-at-risk and Expected Shortfall of the S&P 500 index price as of December 2022.

	2 moments	3 moments	4 moments	5 moments	10 moments
Panel A: 95% Value-at-Risk					
State 1	3440	3476	3476	3476	3302
State 2	3572	3609	3609	3601	3291
State 3	3631	3660	3660	3660	3240
State 4	3667	3697	3697	3697	3155
State 5	3693	3723	3723	3723	3556
Panel B: 95% Expected Shortfall	2 moments	3 moments	4 moments	5 moments	10 moments
State 1	3157	3215	3200	3288	2892
State 2	3286	3438	3438	3399	2885
State 3	3343	3370	3372	3465	2845
State 4	3376	3413	3415	3507	2801
State 5	3394	3432	3434	3523	3165

moments up to the first ten moments. Results are shown in Figs. 3 and 4 in which prices are expressed in thousands of US dollars. The Lagrange multipliers derived from the maximization of the entropy functional are reported in the appendix. It is worth noting that the PDFs up to the first five moments are quite comparable, with a small difference occurring when using the first ten moments. In particular, we observe a fatter left tail showing a higher mass concentrated in the lower range of prices, especially for the three central states. The same behavior is noticeable from the CDFs charts.

The knowledge of the PDFs and CDFs of the price process allows us to compute the risk measures introduced in Section 4. In particular, we compute the 95% Value-at-risk and Expected Shortfall as of December 2022 for each state of the Markov chain. Results are reported in Table 6. Unsurprisingly, the ES values are lower than the VaR ones, as they are more sensitive to the shape of the tail of the distribution.

These results could be useful in different ways. For example, we can directly compare these risk measures with their market counterparts to understand whether a stock or index is under- or over-valued. More specifically, if the VaR (or ES) computed from the market prices is higher than the fundamental one, then we could state that the market is attributing a higher risk to a stock with lower fundamentals and vice versa. Also, we could implement portfolio selection strategies based on the previously cited risk measures. In particular, this approach could be relevant for long-term investment strategies in which the role of the fundamentals has a greater impact.

6. Conclusion

In this paper, we extended the Markov stock model analyzed in [19] with the computation of the nth-order moment of the price process after advancing sufficient conditions for the finiteness of the moments and controlling for the transversality conditions. From the knowledge of the first $N + 1$ moments, we were able to construct the probability density function and the related cumulative distribution function of the price process using the maximum entropy approach described in [29]. Therefore, we computed some measures of risk, such as the Value-at-Risk and the Expected Shortfall. The Application to the price and dividend series of the S&P 500 index showed that the use of higher moments helps in the construction of the PDFs and CDFs of the price process.

From a practitioner point of view, an investor can compare the true risk motivated by the dividend process, as computed by our model, with the market risk perception. The former may be used for accurate investment decision, e.g., portfolio selection strategies.

Further development of this research might consider a different specification of the assumption $\mathbf{A_n}$ to allow more flexibility or the relaxation of the fixed discount rate assumption with the inclusion of a stochastic behavior and its relations with the dividend process.

Data availability

The authors have shared the link to the data.

Acknowledgements

The authors would like to thank the anonymous reviewers for their useful comments and suggestions.

Appendix A

A.1. Proof of Theorem 3.1

First, we are going to prove that $p^{(n)}(t) < \infty$. To this extent, we start with the assumption that $G(t) = g_i$ and $D(t) = d_t$ and we denote by

$$p^{(n)}(g_i, d_t) = p^{(n)}(t) = \lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^{N} \frac{D(t+i)}{r^i}\right)^n\right],$$

so that the dependence on the value of the dividend process and of the growth dividend process at the current time t is highlighted. The nth-order moment can be represented according to

$$\begin{aligned} p^{(n)}(g_i, d_t) &= \lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i_1=1}^{N} \frac{D(t+i_1)}{r^{i_1}}\right)\cdot\left(\sum_{i=1}^{N} \frac{D(t+i)}{r^i}\right)^{n-1}\right] \\ &= \lim_{N\to\infty} \sum_{i_1=1}^{N} \mathbb{E}_{(t)}\left[\frac{D(t+i_1)}{r^{i_1}}\cdot\left(\sum_{i=1}^{N} \frac{D(t+i)}{r^i}\right)^{n-1}\right]. \end{aligned}$$

An iteration of the previous argument gives

$$p^{(n)}(g_i, d_t) = \lim_{N\to\infty} \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\cdots\sum_{i_n=1}^{N} \mathbb{E}_{(t)}\left[\frac{D(t+i_1)D(t+i_2)\dots D(t+i_n)}{r^{\sum_{j=1}^n i_j}}\right].$$

Now, we consider the addend $\mathbb{E}_{(t)}\left[\frac{D(t+i_1)D(t+i_2)\dots D(t+i_n)}{r^{\sum_{j=i}^n i_j}}\right]$ where $i_j \in \{1,2,\dots,N\}, \forall j = 1,\dots,n$. In general, given a n-tuple $(i_1, i_2, \dots, i_n)$, we denote as $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$ one of its increasing permutations such that $\theta_i \le \theta_j, \forall i \le j$.

Note that one permutation $(\theta_1, \theta_2, \dots, \theta_n)$ can correspond to a different n-tuple $(i_1, i_2, \dots, i_n)$. For example, $(\theta_1 = 1, \theta_2 = 2, \theta_3 = 2, \theta_4 = 3)$ corresponds to $(i_1 = 1, i_2 = 3, i_3 = 2, i_4 = 2)$ but also to $(i_1 = 3, i_2 = 2, i_3 = 1, i_4 = 2)$.

In general, the number of n-tuples that have the same increasing permutation depends on the number of elements n and the number of different elements k which are part of the n-tuple $(i_1, i_2, \dots, i_n)$; we denote this number by $B(n; k, \boldsymbol{\theta})$.

Therefore, we have that

$$\begin{aligned} &\lim_{N\to\infty} \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\cdots\sum_{i_n=1}^{N} \mathbb{E}_{(t)}\left[\frac{D(t+i_1)D(t+i_2)\dots D(t+i_n)}{r^{\sum_{j=1}^n i_j}}\right] \\ &= \lim_{N\to\infty} \sum_{\theta_1=1}^{N}\sum_{\theta_2\ge\theta_1}^{N}\cdots\sum_{\theta_n\ge\theta_{n-1}}^{N} \frac{B(n;k,\boldsymbol{\theta})\cdot\mathbb{E}_{(t)}\left[D(t+\theta_1)D(t+\theta_2)\dots D(t+\theta_n)\right]}{r^{\sum_{j=1}^n i_j}}. \end{aligned}$$

We observe that $B(n; k, \boldsymbol{\theta})$ does not depend on N and can be bounded according to $B(n; k, \boldsymbol{\theta}) \le (n!)^n$. See the appendix for proof and additional details.

Also, we note that $\sum_{j=1}^n i_j = \sum_{j=1}^n \theta_j$, therefore,

$$\frac{\mathbb{E}_{(t)}\left[D(t+\theta_1)D(t+\theta_2)\dots D(t+\theta_n)\right]}{r^{\sum_{j=1}^n \theta_j}}$$

$$= \frac{\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right] d_t^n}{r^{\sum_{j=1}^n \theta_j}}, \tag{A.1}$$

with $\prod_{h_i=\theta_{i-1}+1}^{\theta_i} \bullet = 1$ if $\theta_{i-1}+1 > \theta_i$.

The expected value in (A.1) can now be computed using the tower property of the conditional expectation:

$$\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n-1} G(t+h_n)\mathbb{E}_{(t+\theta_n-1)}[G(t+\theta_n)]\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}},$$

but $\mathbb{E}_{(t+\theta_n-1)}[G(t+\theta_n)] \le \bar{g}^{(1)}$ from hypothesis $\mathbf{A_n}$ with $k=1$. Therefore, we have that

$$\begin{aligned}&\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}}\\ &\le \mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n-1} G(t+h_n)\right]\frac{\bar{g}^{(1)} d_t^n}{r^{\sum_{j=1}^n \theta_j}}.\end{aligned}$$

Iterating this strategy with the tower property of conditional expected value at each time $t+m$, with $m \in \{\theta_{n-1}+1,\dots,\theta_{n-1}\}$, we obtain

$$\begin{aligned}&\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}}\\ &\le \mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_{n-1}=\theta_{n-2}+1}^{\theta_{n-1}} G^2(t+h_{n-1})\right]\\ &\cdot \frac{\left(\bar{g}^{(1)}\right)^{\theta_n-\theta_{n-1}}}{r^{\theta_n-\theta_{n-1}}}\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j-(\theta_n-\theta_{n-1})}}.\end{aligned}$$

The iteration can continue for each time $t+m$, with $m \in \{\theta_{n-2}+1,\dots,\theta_{n-1}\}$ as follows

$$\begin{aligned}&\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}}\\ &\le \mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_{n-1}=\theta_{n-2}+1}^{\theta_{n-1}-1} G^2(t+h_{n-1})\mathbb{E}_{(t+\theta_{n-1})}[G^2(t+\theta_{n-1})]\right]\\ &\cdot \frac{\left(\bar{g}^{(1)}\right)^{\theta_n-\theta_{n-1}}}{r^{\theta_n-\theta_{n-1}}}\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j-(\theta_n-\theta_{n-1})}}.\end{aligned}$$

We observe that $\mathbb{E}_{(t+\theta_{n-1})}[G^2(t+\theta_{n-1})] \le \bar{g}^{(2)}$ from hypothesis $\mathbf{A_n}$ with $k=2$ and by iteration we have

$$\begin{aligned}&\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}}\\ &\le \mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_{n-2}=\theta_{n-3}+1}^{\theta_{n-2}} G^3(t+h_{n-2})\right]\\ &\cdot \left(\frac{\bar{g}^{(2)}}{r^2}\right)^{\theta_{n-1}-\theta_{n-2}}\left(\frac{\bar{g}^{(1)}}{r}\right)^{\theta_n-\theta_{n-1}}\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j-(\theta_n-\theta_{n-1})-2(\theta_{n-1}-\theta_{n-2})}}.\end{aligned}$$

This strategy can be still applied again on all times, this gives

$$\begin{aligned}&\mathbb{E}_{(t)}\left[\prod_{h_1=1}^{\theta_1} G^n(t+h_1)\prod_{h_2=\theta_1+1}^{\theta_2} G^{n-1}(t+h_2)\dots\prod_{h_n=\theta_{n-1}+1}^{\theta_n} G(t+h_n)\right]\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j}}\\ &\le \left(\frac{\bar{g}^{(n)}}{r^n}\right)^{\theta_1}\cdot\left(\frac{\bar{g}^{(n-1)}}{r^{n-1}}\right)^{\theta_2-\theta_1}\cdot\ldots\cdot\left(\frac{\bar{g}^{(2)}}{r^2}\right)^{\theta_{n-1}-\theta_{n-2}}\cdot\left(\frac{\bar{g}^{(1)}}{r}\right)^{\theta_n-\theta_{n-1}}\\ &\cdot\frac{d_t^n}{r^{\sum_{j=1}^n \theta_j-\sum_{j=1}^n j(\theta_{n+1-j}-\theta_{n-j})}},\end{aligned}$$

where $\theta_0 := 0$.

We note that the last factor is equal to one considering $\sum_{j=1}^{n} \theta_j - \sum_{j=1}^{n} j(\theta_{n+1-j} - \theta_{n-j}) = 0$, see the appendix for a proof of this last relation.

Therefore, we established that

$$p^{(n)}(g_i, d_t) = \lim_{N\to\infty} \mathbb{E}_{(t)} \left[\left(\sum_{i=1}^{N} \frac{D(t+1)}{r^i} \right)^n \right]$$
$$\leq \lim_{N\to\infty} \sum_{\theta_1=1}^{N} \sum_{\theta_2 \geq \theta_1}^{N} \dots \sum_{\theta_n \geq \theta_{n-1}}^{N} B(n;k,\boldsymbol{\theta}) \left(\frac{\bar{g}^{(n)}}{r^n} \right)^{\theta_1} \cdot \dots \cdot \left(\frac{\left(\bar{g}^{(1)}\right)}{r} \right)^{\theta_n - \theta_{n-1}} d_t^n$$
$$= \lim_{N\to\infty} \sum_{S_1\geq 1}^{N} \sum_{S_2 \geq 0}^{N} \dots \sum_{S_n \geq 0}^{N} B(n;k,\boldsymbol{\theta}) \left(\frac{\bar{g}^{(n)}}{r^n} \right)^{s_1} \cdot \dots \cdot \left(\frac{\left(\bar{g}^{(1)}\right)}{r} \right)^{s_n} d_t^n < \infty,$$

with $s_i = \theta_i - \theta_{i-1}$, considering that $\sum_{s_i\geq 0} \left(\frac{\bar{g}^{(n-1+1)}}{r^{n-i+1}} \right)^{s_i} < \infty$ being $\bar{g}^{(a)} < r^a, \forall a = 1,2,\dots,n$ from condition $\mathbf{A_n}$ and the fact that $B(n;k,\boldsymbol{\theta}) \leq (n!)^n$ independent of N.

Now, it remains to verify the validity of equation (9). First, let us consider the case $k = 0$. The equation becomes

$$\lim_{N\to\infty} \mathbb{E}_{(t)} \left[\left(\frac{P(t+N)}{r^N} \right)^n \right] = 0.$$

To verify its validity, we note that

$$p^{(n)}(g_i, d_t) = \mathbb{E}_{(t)} \left[\left(\sum_{i=1}^{\infty} \frac{D(t+i)}{r^i} \right)^n \right] < \infty,$$

but $D(t+i) = \prod_{h=1}^{i} G(t+h) d_t$, therefore

$$p^{(n)}(g_i, d_t) = \mathbb{E}_{(t)} \left[\left(\sum_{i=1}^{\infty} \frac{\prod_{h=1}^{i} G(t+h)}{r^i} \right)^n d_t^n \right]$$
$$= \mathbb{E}_{(t)} \left[\left(\sum_{i=1}^{\infty} \frac{\prod_{h=1}^{i} G(t+h)}{r^i} \right)^n \right] d_t^n =: \psi_n\left(g_i\right) d_t^n,$$

having set in general

$$\psi_n\left(G(t)\right) = \frac{p^{(n)}(G(t), d_t)}{d_t^n}.$$

Now, set $\bar{\psi}_n(t) = \max_{i\in E} \{\psi_n(g_i)\}$, then

$$0 \leq \mathbb{E}_{(t)} \left[\frac{P^n(t+j)}{r^{jn}} \right] \leq \bar{\psi}_n(t) \mathbb{E}_{(t)} \left[\frac{D^n(t+j)}{r^{jn}} \right]. \tag{A.2}$$

However, because we noted that $\mathbb{E}_{(t)} \left[\left(\sum_{j=1}^{\infty} \frac{D(t+i)}{r^i} \right)^n \right] < \infty$ and considering that "surely"

$$\sum_{j=1}^{\infty} \frac{D^n(t+j)}{r^{jn}} \leq \left(\sum_{j=1}^{\infty} \frac{D(t+j)}{r^j} \right)^n, \tag{A.3}$$

as the mixed products, which are non-negative random variables, are missing; from the relation in (A.3), it follows that

$$\sum_{j=1}^{\infty} \mathbb{E}_{(t)} \left[\frac{D^n(t+j)}{r^{jn}} \right] \leq \mathbb{E}_{(t)} \left[\left(\sum_{j=1}^{\infty} \frac{D(t+j)}{r^j} \right)^n \right] < \infty.$$

Therefore, in (A.2) we have that

$$0 \leq \lim_{j\to\infty} \mathbb{E}_{(t)} \left[\frac{P^n(t+j)}{r^{jn}} \right] \leq \lim_{j\to\infty} \mathbb{E}_{(t)} \left[\frac{D^n(t+j)}{r^{jn}} \right] = 0$$
$$\Rightarrow \lim_{j\to\infty} \mathbb{E}_{(t)} \left[\frac{P^n(t+j)}{r^{jn}} \right] = 0.$$

This proves the validity of the transversality condition for $k = 0$. Now, we prove the validity of (9) $\forall k = 1,2,\dots,n-1$.

We observe that because $P(t) = \frac{D(t+1)+P(t+1)}{r}$, with $r > 1$, then $P(t+1) = rP(t) - D(t+1) \leq rP(t)$, and given the non-negativity of $D(t+1)$

$$\frac{P(t+1)}{r} \leq P(t) \Rightarrow \frac{P(t+2)}{r} \leq P(t+1)$$
$$\Rightarrow \frac{P(t+2)}{r^2} \leq \frac{P(t+1)}{r} \leq P(t),$$

and in general we can say that $\left\{ \frac{P(t+N)}{r^N} \right\}_{N\in\mathbb{N}}$ is a sequence of decreasing non-negative random variables in N.

Also, because we proved that $\lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\frac{P(t+N)}{r^N}\right)^n\right] = 0$, then

$$Y_N := \frac{P(t+N)}{r^N} \xrightarrow[N\to\infty]{L^n} 0$$

where the symbol $\xrightarrow[N\to\infty]{L^n}$ denotes the convergence in L^n as N goes to infinity. Since the convergence in L^n implies the convergence in L^k for $k \leq n$, we have

$$\frac{P(t+N)}{r^N} \xrightarrow[N\to\infty]{L^k} 0, \quad \forall k \leq n.$$

Now, we consider the random variable $X_N := \left(\sum_{i=1}^N \frac{D(t+i)}{r^i}\right)^k$, and because

$$\lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^N \frac{D(t+i)}{r^i}\right)^k\right] \leq \lim_{N\to\infty} \mathbb{E}_{(t)}\left[\left(\sum_{i=1}^\infty \frac{D(t+i)}{r^i}\right)^k\right] = p^{(k)}(g_i, d_t) < \infty,$$

we have that

$$\left(\sum_{i=1}^N \frac{D(t+i)}{r^i}\right)^k \xrightarrow[N\to\infty]{L^k} p^{(k)}(g_i, d_t) \in \mathbb{R}.$$

It follows the convergence in law of the random vector

$$\begin{pmatrix} Y_N \\ X_N \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} 0 \\ p^{(k)}(g_i, d_t) \end{pmatrix},$$

and from the continuous mapping theorem, taken the function $h(x, y) = x \cdot y$, we have

$$\mathbb{E}_{(t)}\left[h(Y_N, X_N)\right] = \mathbb{E}_{(t)}\left[Y_N \cdot X_N\right] \xrightarrow[N\to\infty]{} \mathbb{E}_{(t)}\left[0 \cdot p^{(k)}(g_i, d_t)\right] = 0.$$

A.2. Proof of Proposition 3.2

Let $D(t) = d_t$ and $G(t) = g_i, \forall n \in \mathbb{N}$, then according to equations (7) and (12) we have

$$\psi_n(g_i) = \frac{p^{(n)}(g_i, d_t)}{d_t^n} \Rightarrow \psi_n(g_i) d_t^n = p^{(n)}(g_i, d_t) = \mathbb{E}_{(t)}\left[\left(\frac{D(t+1) + P(t+1)}{r}\right)^n\right]$$
$$= \sum_{m=0}^n \binom{n}{m} \mathbb{E}_{(t)}\left[D^m(t+1) P^{n-m}(t+1)\right] \frac{1}{r^n}.$$

Considering the generic expected value

$$\mathbb{E}_{(t)}\left[D^m(t+1) P^{n-m}(t+1)\right] = \mathbb{E}_{(t)}\left[G^m(t+1) d_t^m \Psi_{n-m}(G(t+1)) G^{n-m}(t+1) d_t^{n-m}\right]$$
$$= d_t^n \sum_{j\in E} p_{ij} g_j^n \Psi_{n-m}(g_j),$$

and substituting, we obtain

$$\psi_n(g_i) = \frac{1}{r^n} \sum_{m=0}^n \binom{n}{m} \sum_{j\in E} p_{ij} g_j^n \Psi_{n-m}(g_j)$$
$$\Rightarrow r^n \psi_n(g_i) - \sum_{j\in E} p_{ij} g_j^n \Psi_n(g_j) = \sum_{j\in E} p_{ij} g_j^n \left(\sum_{m=0}^n \binom{n}{m} \Psi_{n-m}(g_j)\right).$$

The previous equation can be simply arranged in matrix form to give (15).

The uniqueness of the solution of (15) can be established by observing that $d_t \in \mathbb{R}$, $p^{(n)}(g_i, d_t) < \infty$ then $\psi_n(g_i) = \frac{p^{(n)}(g_i,d_t)}{d_t^n}$ exists and is unique in virtue of the fact that the series expressing $p^{(n)}(g_i, d_t)$ converges to a unique value. The non-negativity is a consequence of the fact that both $p^{(n)}(g_i, d_t)$ and d_t are non-negative.

A.3. Extra proofs

This appendix contains the proof of two properties that have been used inside the proof of Theorem 3.1. The first is related to the number of n-tuples that have the same increasing permutation, which here we explain in detail.

Let $\boldsymbol{i} = (i_1, i_2, \dots, i_n) \in \mathbb{N}^n(N)$ be a n-tuple of natural numbers such that $i_j \le N, \forall j = 1, \dots, n$. Let $\mathbb{O}^n(N) = \{(\theta_1, \dots, \theta_n) : \theta_i \in \mathbb{N}, \forall i = 1, \dots, n, \text{ and } \theta_i \le \theta_{i+1}, \forall i = 1, \dots, n-1\}$ be the set of n-tuples of non-decreasing natural numbers not greater than N.

Let us define a function that for each n-tuple $\boldsymbol{i} \in \mathbb{N}^n(N)$ assigns one element $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n) \in \mathbb{O}^n(N)$, which is its non-decreasing permutation,

$$\Pi_N : \mathbb{N}^n(N) \to \mathbb{O}^n(N), \quad \text{such that} \quad \Pi_N((i_1, \dots, i_n)) = (\theta_1, \dots, \theta_n),$$

with $\theta_1 := min\{i_1, \dots, i_n\}$, $\theta_j := min\{i_1^{(j)}, i_2^{(j)}, \dots, i_{n-(j+1)}^{(j)}\}$, $\forall j = 2, 3, \dots, n$, and $i^{(j)}$ being the vector obtained from $\boldsymbol{i}$ removing the elements $\theta_1, \theta_2, \dots, \theta_{j-1}$. Obviously, $i^{(j)} \in \mathbb{N}^{n-(j+1)}(N)$.

For example, if $\boldsymbol{i} = (1, 2, 8, 2, 5) \in \mathbb{N}^5(8)$ then $\Pi_8(\boldsymbol{i}) = (1, 2, 2, 5, 8) = \boldsymbol{\theta}$. The relation Π_N is a non-injective function, indeed if we consider the vectors $\boldsymbol{i} = (1, 2, 8, 2, 5) \in \mathbb{N}^5(8)$ and $\boldsymbol{i} = (2, 8, 2, 5, 1) \in \mathbb{N}^5(8)$, we have that $\Pi_8((1, 2, 8, 2, 5)) = (1, 2, 2, 5, 8)$ and $\Pi_8((2, 8, 2, 5, 1)) = (1, 2, 2, 5, 8)$.

Let $\Pi_N^{-1}(\boldsymbol{\theta})$ be the complete inverse image of the ordered tuple $\boldsymbol{\theta}$. We are interested in establishing the cardinality of $\Pi_N^{-1}(\boldsymbol{\theta})$. To this extent the following result holds.

Proposition A.1. *Let $\boldsymbol{\theta} \in \mathbb{O}^n(N)$ and $k \le n$ be the number of the different elements composing $\boldsymbol{\theta}$. Let $m_1, m_2, \dots, m_k$ be the respective number of elements such that $\sum_{l=1}^k m_l = n$.*

Let us denote with $B(n; k, \boldsymbol{m})$ the cardinality of $\Pi_N^{-1}(\boldsymbol{\theta})$, then

$$B(n; k, \boldsymbol{m}) = \prod_{j=i}^{k} \binom{n - \sum_{i=1}^{j-1} m_i}{m_j},$$

where $\sum_{j=1}^{0} \bullet = 0$ by convention.

Proof. The proof is based on a simple application of the fundamental counting principle. Because the vector $\boldsymbol{\theta}$ has m_1 elements which are equal to the smallest element of $\boldsymbol{\theta}$, they can assume $\binom{n}{m_1}$ possible positions. For each of them, the remaining m_2 elements equal to the second smallest element of $\boldsymbol{\theta}$ can be put in the $n - m_1$ remaining positions following $\binom{n-m_1}{m_2}$ different combinations. We can continue with the other elements of $\boldsymbol{\theta}$ up to the m_k elements which assume the maximum value and can be put in $\binom{n-\sum_{i=1}^{k-1} m_i}{m_k} = \binom{m_k}{m_k} = 1$ different positions.

We apply the counting principle and obtain

$$B(n; k, \boldsymbol{m}) = \binom{n}{m_1} \cdot \binom{n - m_1}{m_2} \cdot \dots \cdot \binom{n - \sum_{i=1}^{k-1} m_i}{m_k} = \prod_{j=i}^{k} \binom{n - \sum_{i=1}^{j-1} m_i}{m_j}.$$

We note that such value is independent of N. □

Now, we prove a second result we used in the proof of Theorem 3.1.

Proposition A.2. *For each $\boldsymbol{\theta} \in \mathbb{O}^n(N)$ the following identity holds*

$$\sum_{j=1}^{n} \theta_j - \sum_{j=1}^{n} j(\theta_{n+1-j} - \theta_{n-j}) = 0, \tag{A.4}$$

where we set $\theta_0 = 0$ by convention.

Proof. For $n = 1$ the equation holds as we have $\theta_1 - 1(\theta_1 - \theta_0) = 0$ which is evidently true because $\theta_0 = 0$.
Let us assume that the relation (A.4) holds true for $n - 1$ such that

$$\sum_{j=1}^{n-1} \theta_j - \sum_{j=1}^{n-1} j(\theta_{n-1+1-j} - \theta_{n-1-j}) = 0. \tag{A.5}$$

Now, we prove the equation for n. We first observe that

Table A.7
Lagrange multiplier for the computation of the maximum entropy.

		State 1	State 2	State 3	State 4	State 5
2 moments	λ_0	−43.112	−46.159	−47.190	−47.140	−45.012
	λ_1	20.480	21.275	21.448	21.206	19.992
	λ_2	−2.439	−2.458	−2.443	−2.392	−2.228
3 moments	λ_0	−78.784	−79.561	−81.254	−80.514	−76.080
	λ_1	46.239	44.691	44.990	44.046	41.019
	λ_2	−8.568	−7.870	−7.809	−7.547	−6.919
	λ_3	0.480	0.413	0.403	0.384	0.345
4 moments	λ_0	−100.465	−80.116	−75.872	−73.935	−72.590
	λ_1	66.951	45.207	40.055	38.071	37.884
	λ_2	−15.905	−8.048	−6.130	−5.533	−5.875
	λ_3	1.623	0.440	0.152	0.085	0.192
	λ_4	−0.066	−0.002	0.014	0.016	0.008
5 moments	λ_0	−445.956	−406.558	−396.959	−380.883	−336.378
	λ_1	479.663	425.331	408.923	387.430	334.937
	λ_2	−211.071	−183.321	−173.951	−162.999	−138.283
	λ_3	47.296	40.447	37.954	35.223	29.398
	λ_4	−5.356	−4.523	−4.202	−3.866	−3.180
	λ_5	0.243	0.202	0.186	0.170	0.138
10 moments	λ_0	1263.731	860.586	−40.089	−83.965	3745.537
	λ_1	−1870.641	−1520.880	44.719	115.608	−7096.398
	λ_2	957.738	1097.986	−101.707	−164.609	5782.093
	λ_3	−122.501	−403.449	154.590	184.792	−2641.560
	λ_4	−47.767	75.685	−113.984	−115.568	740.867
	λ_5	3.301	−8.254	45.210	39.316	−133.310
	λ_6	10.856	2.544	−10.033	−6.902	16.536
	λ_7	−4.546	−1.070	1.168	0.395	−1.734
	λ_8	0.828	0.221	−0.046	0.058	0.184
	λ_9	−0.074	−0.022	−0.003	−0.010	−0.015
	λ_{10}	0.003	0.001	0.000	0.000	0.001

$$\sum_{i=1}^{n} \theta_i = \sum_{i=1}^{n-1} \theta_i + \theta_n,$$

and using the inductive hypothesis (A.5) we have that

$$\begin{aligned}
\sum_{i=1}^{n} \theta_i &= \sum_{i=1}^{n-1} i(\theta_{n-i} - \theta_{n-1-i}) + \theta_n \\
&= \sum_{j=2}^{n} (j-1)(\theta_{n+1-j} - \theta_{n-j}) + \theta_n \\
&= \sum_{j=2}^{n} j(\theta_{n+1-j} - \theta_{n-j}) - \sum_{j=2}^{n} \theta_{n+1-j} + \sum_{j=2}^{n} \theta_{n-j} + \theta_n \\
&= \sum_{j=2}^{n} j(\theta_{n+1-j} - \theta_{n-j}) + \theta_n - \theta_{n-1} - \sum_{j=3}^{n} \theta_{n+1-j} + \sum_{j=2}^{n} \theta_{n-j} \\
&= \sum_{j=1}^{n} j(\theta_{n+1-j} - \theta_{n-j}) - \sum_{h=2}^{n-1} \theta_{n-h} + \sum_{j=2}^{n-1} \theta_{n-j} + \theta_0 \\
&= \sum_{j=1}^{n} j(\theta_{n+1-j} - \theta_{n-j}),
\end{aligned}$$

being $\theta_0 = 0$. □

A.4. Lagrange multipliers

Table A.7 reports the Lagrange multipliers derived from the computation of the maximum entropy for different first $N+1$ moments and for each state of the Markov chain.

References

[1] S.M. Bartram, M. Grinblatt, Agnostic fundamental analysis works, J. Financ. Econ. 128 (1) (2018) 125–147, https://doi.org/10.1016/j.jfineco.2016.11.008.

[2] S. Richardson, İ. Tuna, P. Wysocki, Accounting anomalies and fundamental analysis: a review of recent research advances, J. Account. Econ. 50 (2) (2010) 410–454, https://doi.org/10.1016/j.jacceco.2010.09.008.
[3] H. Zhang, C. Yan, Modelling fundamental analysis in portfolio selection, Quant. Finance 18 (8) (2018) 1315–1326, https://doi.org/10.1080/14697688.2017.1418520.
[4] J.B. Williams, The Theory of Investment Value, Harvard University Press, Cambridge, MA, 1938.
[5] M.J. Gordon, The Investment, Financing, and Valuation of the Corporation, RD Irwin, Homewood, IL, 1962.
[6] B.G. Malkiel, Equity yields, growth, and the structure of share prices, Am. Econ. Rev. 53 (5) (1963) 1004–1031, www.jstor.org/stable/1812046.
[7] R.J. Fuller, C.-C. Hsia, A simplified common stock valuation model, Financ. Anal. J. 40 (5) (1984) 49–56, https://doi.org/10.2469/faj.v40.n5.49.
[8] N. Molodovsky, C. May, S. Chottiner, Common stock valuation: principles, tables and application, Financ. Anal. J. 21 (2) (1965) 104–123, https://doi.org/10.2469/faj.v21.n2.104.
[9] R. Brooks, B. Helms, An n-stage, fractional period, quarterly dividend discount model, Financ. Rev. 25 (4) (1990) 651–657, https://doi.org/10.1111/j.1540-6288.1990.tb01303.x.
[10] M. Lettau, S.C. Ludvigson, Expected returns and expected dividend growth, J. Financ. Econ. 76 (3) (2005) 583–626, https://doi.org/10.1016/j.jfineco.2004.05.008.
[11] W.J. Hurley, L.D. Johnson, A realistic dividend valuation model, Financ. Anal. J. 50 (4) (1994) 50–54, https://doi.org/10.2469/faj.v50.n4.50.
[12] W.J. Hurley, L.D. Johnson, Generalized Markov dividend discount models, J. Portf. Manag. 25 (1) (1998) 27–31, https://doi.org/10.3905/jpm.1998.409658.
[13] Y.F. Yao, A trinomial dividend valuation model, J. Portf. Manag. 23 (4) (1997) 99–103, https://doi.org/10.3905/jpm.1997.409618.
[14] L.L. Ghezzi, C. Piccardi, Stock valuation along a Markov chain, Appl. Math. Comput. 141 (2) (2003) 385–393, https://doi.org/10.1016/S0096-3003(02)00263-1.
[15] G. D'Amico, A semi-Markov approach to the stock valuation problem, Ann. Finance 9 (4) (2013) 589–610.
[16] G. D'Amico, R. De Blasis, A review of the dividend discount model: from deterministic to stochastic models, in: Statistical Topics and Stochastic Models for Dependent Data with Applications, John Wiley & Sons, Ltd., 2020, pp. 47–67, Ch. 3.
[17] J. Giner, V. Zakamulin, A regime-switching model of stock returns with momentum and mean reversion, Econ. Model. 122 (2023) 106237, https://doi.org/10.1016/j.econmod.2023.106237.
[18] B. Gankhuu, J. Kleinow, A. Lkhamsuren, A. Horsch, Dividends and compound Poisson processes: a new stochastic stock price model, Int. J. Theor. Appl. Finance 25 (03) (2022) 2250014, https://doi.org/10.1142/S0219024922500145.
[19] V.S. Barbu, G. D'Amico, R. De Blasis, Novel advancements in the Markov chain stock model: analysis and inference, Ann. Finance 13 (2) (2017) 125–152, https://doi.org/10.1007/s10436-017-0297-9.
[20] A. Agosto, E. Moretto, Variance matters (in stochastic dividend discount models), Ann. Finance 11 (2) (2015) 283–295, https://doi.org/10.1007/s10436-014-0257-6.
[21] G. D'Amico, R. De Blasis, A multivariate Markov chain stock model, Scand. Actuar. J. 2020 (4) (2020) 272–291, https://doi.org/10.1080/03461238.2019.1661280.
[22] A. Agosto, A. Mainini, E. Moretto, Stochastic dividend discount model: covariance of random stock prices, J. Econ. Finance 43 (3) (2019) 552–568, https://doi.org/10.1007/s12197-018-9455-9.
[23] G. D'Amico, Stochastic dividend discount model: risk and return, Markov Process. Relat. Fields 23 (3) (2017) 349–376.
[24] O.J. Blanchard, M.W. Watson, Bubbles, rational expectations and financial markets, in: P. Wachtel (Ed.), Crises in the Economic and Financial Structure, D.C. Heathand Company, Lexington, MA, 1982.
[25] P.A. Samuelson, Proof that properly discounted present values of assets vibrate randomly, Bell J. Econ. Manag. Sci. 4 (2) (1973) 369–374, https://doi.org/10.2307/3003046.
[26] M. Bertocchi, R. Giacometti, S. Ortobelli, S. Rachev, The impact of different distributional hypothesis on returns in asset allocation, Finance Lett. 3 (1) (2005) 17–27.
[27] S. Ortobelli, N. Kouaissah, T. Tichý, On the impact of conditional expectation estimators in portfolio theory, Comput. Manag. Sci. 14 (4) (2017) 535–557, https://doi.org/10.1007/s10287-017-0282-9.
[28] D. Tasche, Expected shortfall and beyond, J. Bank. Finance 26 (7) (2002) 1519–1533, https://doi.org/10.1016/S0378-4266(02)00272-8.
[29] L.R. Mead, N. Papanicolaou, Maximum entropy in the problem of moments, J. Math. Phys. 25 (8) (1984) 2404–2417, https://doi.org/10.1063/1.526446.
[30] T. Saad, G. Ruai, PyMaxEnt: a python software for maximum entropy moment reconstruction, SoftwareX 10 (2019) 100353, https://doi.org/10.1016/j.softx.2019.100353.
[31] R.J. Shiller, Irrational Exuberance, 3rd edition, Princeton University Press, 2016, https://press.princeton.edu/books/paperback/9780691173122/irrational-exuberance.