# Math 5760/6890: Introduction to Mathematical Finance

## The Cox-Ross-Rubinstein model

See Petters and Dong 2016, Section 5.2

Akil Narayan[1]

[1] Department of Mathematics, and Scientific Computing and Imaging (SCI) Institute
University of Utah

Fall 2024

We have modeled a security's price $S_j = S(t_j)$ via,

$$S_{j+1} = G_{j+1} S_j, \qquad G_j = \begin{cases} u, & \text{with probability } p \\ d, & \text{with probability } 1-p \end{cases}$$

From this model, we've concluded:

- $L := \log(S_n/S_0)$ is a scaled/shifted $\text{Binomial}(n,p)$ random variable.
- $S_n = S_0 e^L$ is the exponential of a scaled/shifted Binomial random variable
- The triple $(p, u, d)$ determines the distribution entirely.

We have modeled a security's price $S_j = S(t_j)$ via,

$$S_{j+1} = G_{j+1} S_j, \qquad G_j = \begin{cases} u, & \text{with probability } p \\ d, & \text{with probability } 1-p \end{cases}$$

From this model, we've concluded:

- $L := \log(S_n/S_0)$ is a scaled/shifted $\text{Binomial}(n,p)$ random variable.
- $S_n = S_0 e^L$ is the exponential of a scaled/shifted Binomial random variable
- The triple $(p,u,d)$ determines the distribution entirely.

The CRR model places the following additional constraints on our standard Binomial tree model:

- Geometric symmetry of tree prices: $u = 1/d$
- The continuous-time limit of the expected log-return matches the real-world drift:

$$\mu = \lim_{n\to\infty} \frac{1}{h_n} \mathbb{E} L_j$$

- The continuous-time limit of the variance of the log-return matches the real-world (squared) volatility:

$$\sigma^2 = \lim_{n\to\infty} \frac{1}{h_n} \text{Var} L_j$$

Hence, for finite $n$, $(p,u,d)$ should depend on the time discretization parameter $n$. I.e.:

$$(p,u,d) = (p_n, u_n, d_n).$$

Goal: use CRR constraints to choose $(p_n, u_n, d_n)$.

## The problem setup



We seek to construct a fixed, finite-$n$ Binomial tree model over the time period $[0, T]$. I.e., we seek to compute $(p_n, u_n, d_n)$ for a fixed $n$ and $h_n = T/n$.

We assume that the (continuous-time) drift and volatility parameters $(\mu, \sigma)$ are available to us.
(E.g., we've computed approximations to them from historical data.)

# The problem setup



We seek to construct a fixed, finite-$n$ Binomial tree model over the time period $[0, T]$. I.e., we seek to compute $(p_n, u_n, d_n)$ for a fixed $n$ and $h_n = T/n$.

We assume that the (continuous-time) drift and volatility parameters $(\mu, \sigma)$ are available to us.
(E.g., we've computed approximations to them from historical data.)

Our starting point should be to understand how the CRR assumptions/constraints affect $(p_n, u_n, d_n)$:

- $d_n = 1/u_n \implies$ need only determine $(p_n, u_n)$.

## The problem setup



We seek to construct a fixed, finite-$n$ Binomial tree model over the time period $[0, T]$. I.e., we seek to compute $(p_n, u_n, d_n)$ for a fixed $n$ and $h_n = T/n$.

We assume that the (continuous-time) drift and volatility parameters $(\mu, \sigma)$ are available to us.
(E.g., we've computed approximations to them from historical data.)

Our starting point should be to understand how the CRR assumptions/constraints affect $(p_n, u_n, d_n)$:

- $d_n = 1/u_n \implies$ need only determine $(p_n, u_n)$.
- Drift matching: Ideally, we have,

$$\frac{\mathbb{E}L_j}{h_n} = \mu_n \xrightarrow{n \uparrow \infty} \mu.$$

  We can't really do this directly since we only want to construct $(p_n, u_n)$ for finite $n$. Hence, we will instead impose:

$$\frac{\mathbb{E}L_j}{h_n} = \mu_n \approx \mu.$$

We seek to construct a fixed, finite-$n$ Binomial tree model over the time period $[0, T]$. I.e., we seek to compute $(p_n, u_n, d_n)$ for a fixed $n$ and $h_n = T/n$.

We assume that the (continuous-time) drift and volatility parameters $(\mu, \sigma)$ are available to us.
(E.g., we've computed approximations to them from historical data.)

Our starting point should be to understand how the CRR assumptions/constraints affect $(p_n, u_n, d_n)$:

- $d_n = 1/u_n \Longrightarrow$ need only determine $(p_n, u_n)$.
- Drift matching: Ideally, we have,

$$\frac{\mathbb{E}L_j}{h_n} = \mu_n \xrightarrow{n\uparrow\infty} \mu.$$

  We can't really do this directly since we only want to construct $(p_n, u_n)$ for finite $n$. Hence, we will instead impose:

$$\frac{\mathbb{E}L_j}{h_n} = \mu_n \approx \mu.$$

- We have a similar condition for matching the volatility:

$$\frac{\mathrm{Var} L_j}{h_n} = \sigma_n^2 \approx \sigma^2$$

# The CRR setup, I

Our CRR conditions require statistics of the inter-period log-returns, e.g.,

$$\mathbb{E}L_j, \qquad \mathrm{Var}L_j.$$

$$L_j = \begin{cases} \log u_n, & \text{w/prob } p_n \\ -\log u_n, & \text{w/prob } 1-p_n \end{cases} \qquad (d_n = 1/u_n)$$

Our CRR conditions require statistics of the inter-period log-returns, e.g.,

$$\mathbb{E}L_j, \qquad \mathrm{Var}L_j.$$

Of course, we know how to compute these: We have,

$$L_j = \log d_n + X_j \log \frac{u_n}{d_n} = -\log u_n + 2X_j \log u_n,$$

where $X_j \sim \mathrm{Bernoulli}(p_n)$.

# The CRR setup, I

Our CRR conditions require statistics of the inter-period log-returns, e.g.,

$$\mathbb{E}L_j, \qquad \mathrm{Var}L_j.$$

Of course, we know how to compute these: We have,

$$L_j = \log d_n + X_j \log \frac{u_n}{d_n} = -\log u_n + 2X_j \log u_n, = (2X_j - 1) \log u_n$$

where $X_j \sim \mathrm{Bernoulli}(p_n)$. (Recall: $\mathbb{E}X_j = p_n$ , $\mathrm{Var}\, X_j = p_n(1-p_n)$ )

Therefore:

$$\mathbb{E}L_j = (2p_n - 1)\log u_n, \qquad \mathrm{Var}L_j = 4p_n(1-p_n)(\log u_n)^2$$

Therefore, our CRR constraints are the following:

$$d_n = \frac{1}{u_n}$$
$$\mu = \frac{2p_n - 1}{h_n} \log u_n \qquad (\mu = \mathbb{E}\, L/h)$$
$$\sigma^2 = \frac{4p_n(1-p_n)}{h_n} (\log u_n)^2. \qquad (\sigma^2 = \tfrac{1}{h} \mathrm{Var}\, L)$$

where we have replaced some instances of "≈" with "=".

We can solve this (nonlinear) system of equations for $(p_n, u_n)$, then $d_n$.

Instead: we'll come up w/ approximate solutions.

Idea: for large $n$, $p_n \to 1/2 \implies 4p_n(1-p_n) \to 1$

So we approximate $4p_n(1-p_n) \approx 1$

Then the conditions above reduce to

$$\mu = \frac{2p_n - 1}{h_n} \log u_n \quad (☆)$$

$$\sigma^2 = \frac{(\log u_n)^2}{h_n}$$

$\Rightarrow$ $u_n =$ ~~$\exp(\sigma / \sqrt{h_n})$~~ $\exp(\sigma \sqrt{h_n})$

$d_n =$ ~~$\exp(-\sigma / \sqrt{h_n})$~~ $\exp(-\sigma \sqrt{h_n})$

(☆): $$p_n = \left( \frac{h_n \mu}{\log u_n} + 1 \right) \frac{1}{2}$$

$$= \left( \frac{h_n \mu}{\sigma \sqrt{h_n}} + 1 \right) \frac{1}{2}$$

$$= \frac{1}{2} \left( 1 + \sqrt{h_n} \frac{\mu}{\sigma} \right)$$

# The CRR setup, II

Therefore, our CRR constraints are the following:

$$\begin{aligned} d_n &= \frac{1}{u_n} \\ \mu &= \frac{2p-1}{h_n} \log u_n \\ \sigma^2 &= \frac{4p_n(1-p_n)}{h_n} (\log u_n)^2. \end{aligned}$$

where we have replaced some instances of "$\approx$" with "$=$".

After some computations and approximations, we arrive at the following *real-world CRR equations*:

$$p_n \approx \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h_n}\right), \qquad u_n \approx \exp(\sigma\sqrt{h_n}), \qquad d_n \approx \exp(-\sigma\sqrt{h_n}).$$

# The real-world CRR tree

A "real-world" CRR tree/model is therefore constructed in the following way:

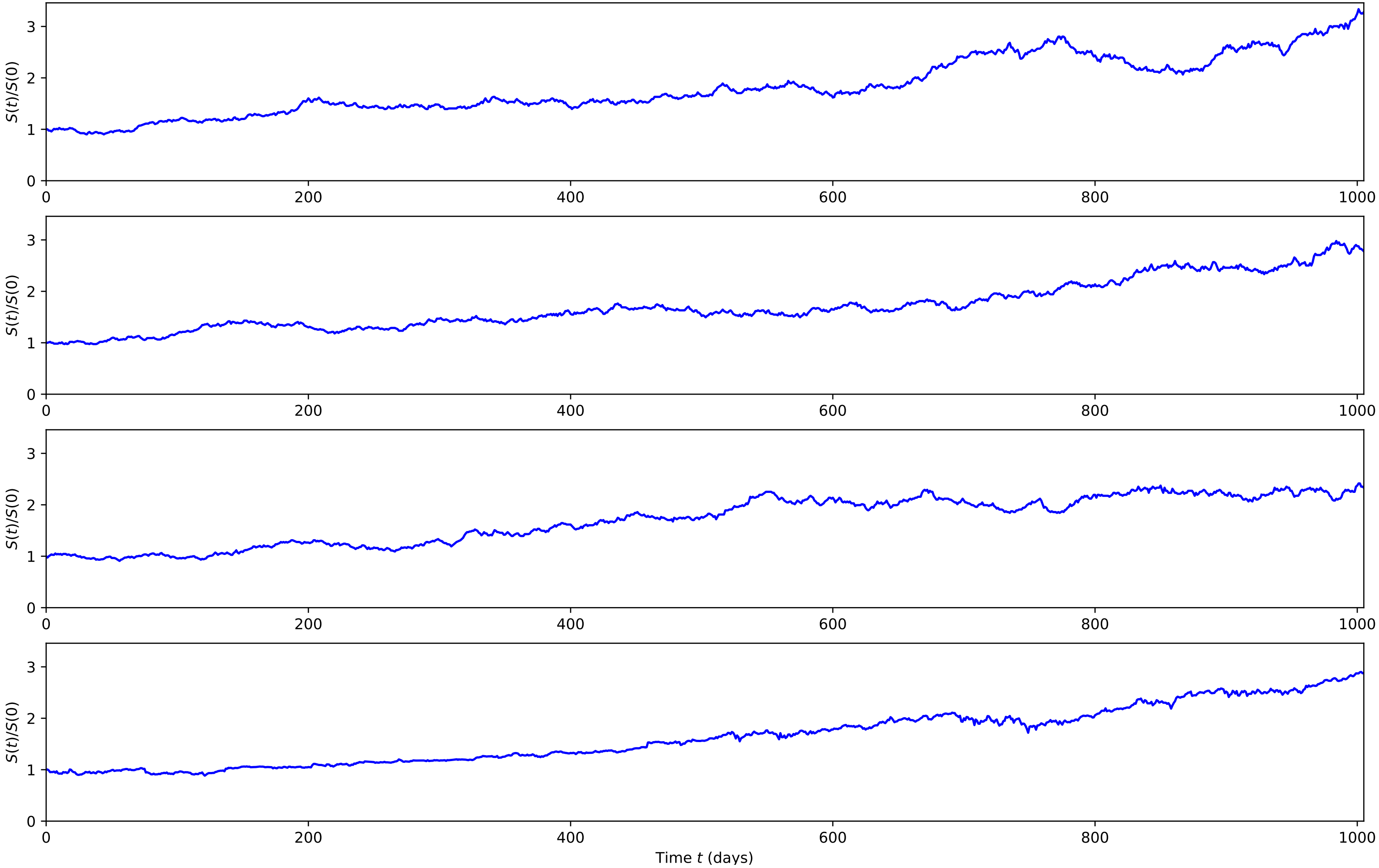
- Historical data is used to compute an asset's continuous-time drift and volatility $(\mu, \sigma)$
- The terminal time $T$ and number of periods $n$ is determined. $h_n = T/n$.
- The real-world CRR equations are used to set $(p_n, u_n, d_n)$:

$$p_n = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h_n}\right), \qquad u_n = \exp(\sigma\sqrt{h_n}), \qquad d_n = \exp(-\sigma\sqrt{h_n}).$$

# Let's play a game



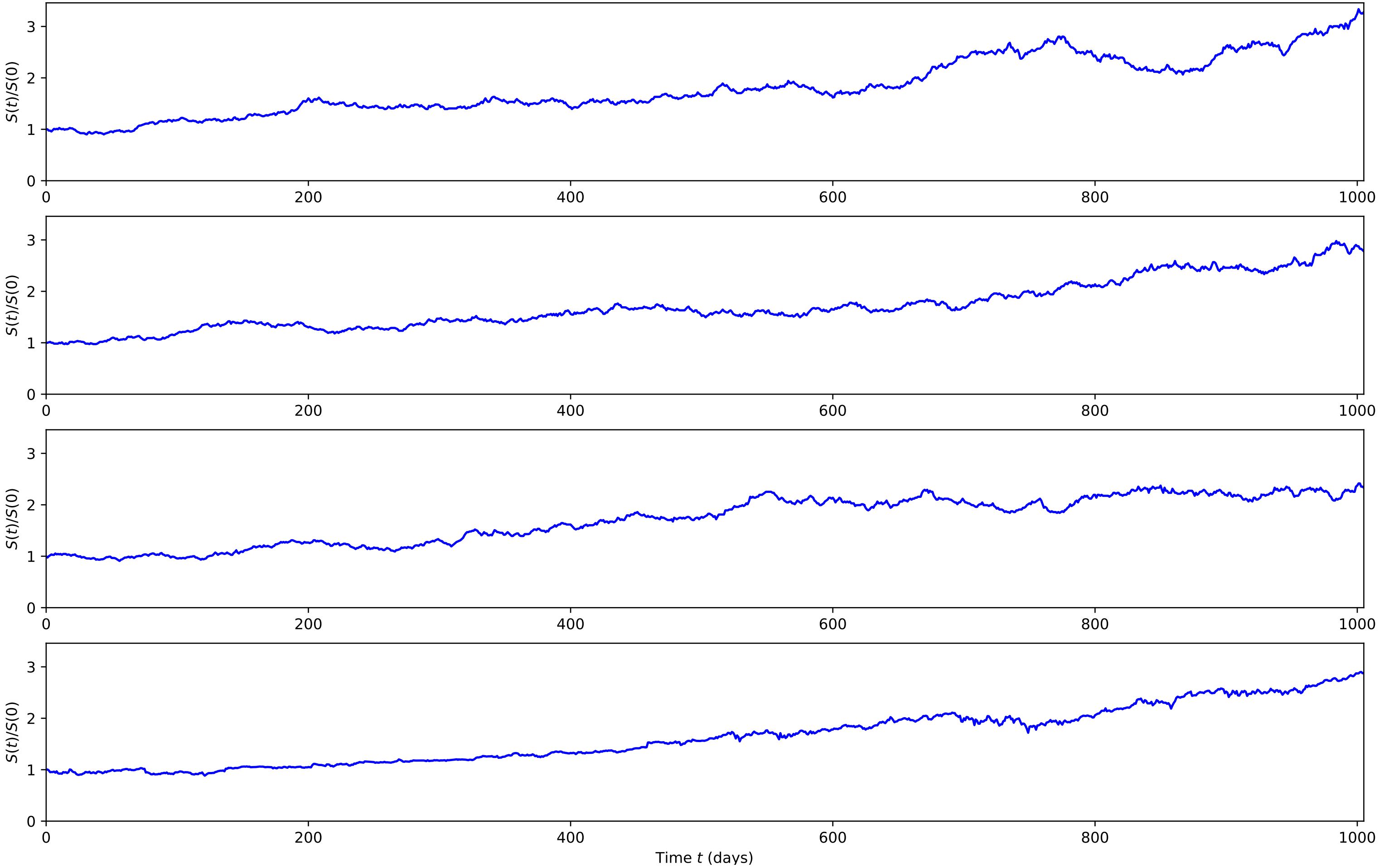
Which is the simulated price?



# Let's play a game



Which is the simulated price?

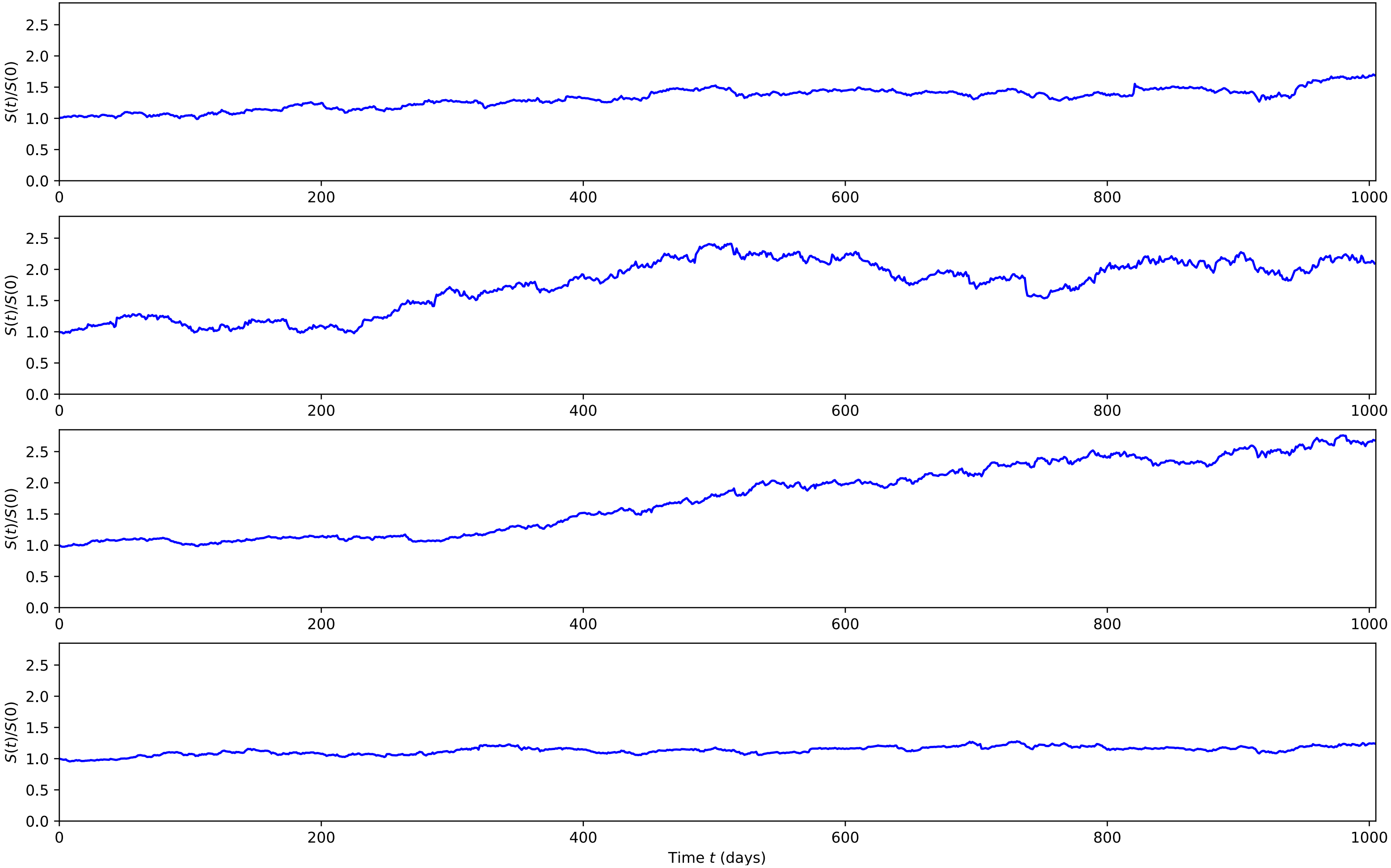



MSFT

# Let's play a game, II

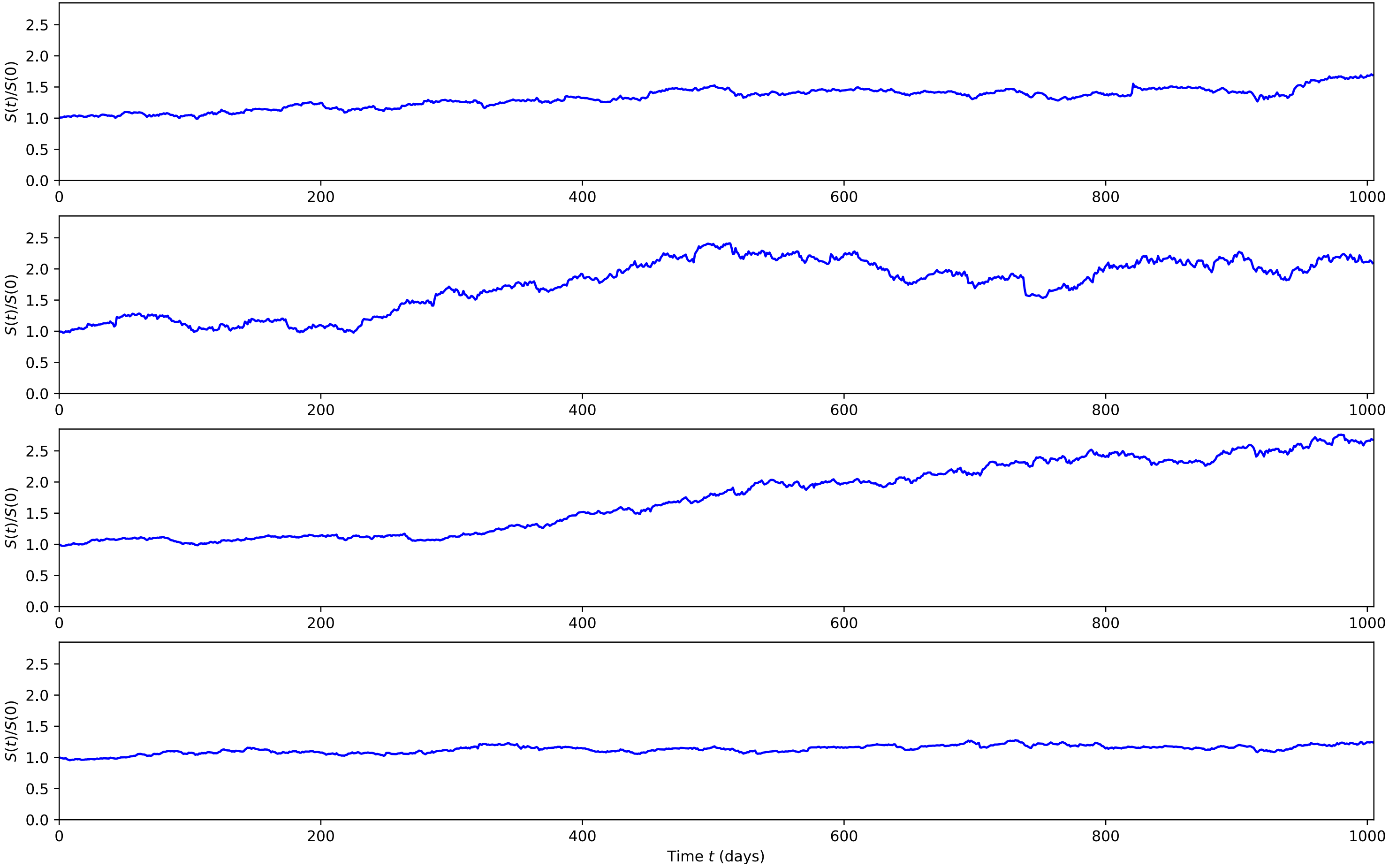


Which prices are simulated?



# Let's play a game, II

Which prices are simulated?




GE

Airbus

Lockheed Martin

Coca cola

# Some CRR properties, I

Some initial observations about the tuple $(p_n, u_n, d_n)$ of the real-world CRR model:

- Because $u_n = \exp(\sigma\sqrt{h_n})$, and $h_n = T/n$, then

$$\lim_{n\uparrow\infty} u_n = 1,$$

  and similarly for $d_n = 1/u_n$.
  I.e., the uptick and downtick geometric rates become very close to unity for large $n$.

# Some CRR properties, I

Some initial observations about the tuple $(p_n, u_n, d_n)$ of the real-world CRR model:

- Because $u_n = \exp(\sigma\sqrt{h_n})$, and $h_n = T/n$, then

$$\lim_{n\uparrow\infty} u_n = 1,$$

  and similarly for $d_n = 1/u_n$.
  I.e., the uptick and downtick geometric rates become very close to unity for large $n$.

- Because $p_n = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h_n}\right)$, then

$$\lim_{n\uparrow\infty} p_n = \frac{1}{2},$$

  so that for large $n$ the CRR tree tends toward fair coin flips.

# Some CRR properties, II

What kind of statistics does $S_n$ have under this model? We have,

$$S_n = S_0 e^L = S_0 \exp(\sum_{j=1}^{n} L_j).$$

We've seen that

$$\mathbb{E} S_n = S_0 \left(p_n u_n + (1 - p_n) d_n\right)^n ,$$

and we have the real-world CRR equations:

$$p_n = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h_n}\right), \qquad u_n = \exp(\sigma\sqrt{h_n}), \qquad d_n = \exp(-\sigma\sqrt{h_n}).$$

$d_n = 1/u_n$

$$\mathbb{E} S_n = S_0 \left(p_n u_n + (1-p_n) d_n\right)^n = S_0 \left[ u_n \left(\frac{1}{2} + \frac{\mu}{2\sigma}\sqrt{h_n}\right) + \frac{1}{u_n}\left(\frac{1}{2} - \frac{\mu}{2\sigma}\sqrt{h_n}\right)\right]^n$$

$d_n$: $1/u_n$

$1 - p_n$: $\frac{1}{2} - \frac{\mu}{2\sigma}\sqrt{h_n}$

$$= S_0 \left[ \frac{1}{2}\left(u_n + \frac{1}{u_n}\right) + \frac{\mu}{2\sigma}\sqrt{h_n}\left(u_n - \frac{1}{u_n}\right) \right]^n$$

$$u_n = \exp(\sigma\sqrt{h_n})$$
$$\approx 1 + \sigma\sqrt{h_n} + \frac{(\sigma\sqrt{h_n})^2}{2!} + \cdots$$

$$\frac{1}{u_n} = \exp(-\sigma\sqrt{h_n}) \approx 1 - \sigma\sqrt{h_n} + \frac{(\sigma\sqrt{h_n})^2}{2!} + \cdots$$

$$\Longrightarrow u_n + \frac{1}{u_n} \approx \cancel{1}\,2 + 0 + \sigma^2 h_n$$

$$u_n - \frac{1}{u_n} \approx 2\sigma\sqrt{h_n}$$

$$\Longrightarrow \mathbb{E}S_n \xrightarrow{n\uparrow\infty} S_0 \left[ \frac{1}{2}\left(2 + \sigma^2 h_n\right) + \frac{\mu}{2\sigma}\sqrt{h_n}\left(2\sigma\sqrt{h_n}\right) \right]^n$$

$$= S_0 \left[ 1 + \frac{\sigma^2}{2} h_n + \mu h_n \right]^n$$

$$= S_0 \left[ 1 + \left(\mu + \frac{\sigma^2}{2}\right)\frac{T}{n} \right]^n \qquad \left(\text{recall: } \lim_{n\to\infty}\left(1 + \frac{k}{n}\right)^n = e^k\right)$$

$$\overset{n\uparrow\infty}{=} S_0 \exp\left(\left(\mu + \frac{\sigma^2}{2}\right)T\right)$$

# Some CRR properties, II

What kind of statistics does $S_n$ have under this model? We have,

$$S_n = S_0 e^L = S_0 \exp(\sum_{j=1}^{n} L_j).$$

We've seen that

$$\mathbb{E} S_n = S_0 \left(p_n u_n + (1 - p_n) d_n\right)^n,$$

and we have the real-world CRR equations:

$$p_n = \frac{1}{2}\left(1 + \frac{\mu}{\sigma}\sqrt{h_n}\right), \qquad u_n = \exp(\sigma\sqrt{h_n}), \qquad d_n = \exp(-\sigma\sqrt{h_n}).$$

These allow us to conclude:

$$\lim_{n\uparrow\infty} \mathbb{E} S_n = S_0 \exp\left[\left(\mu + \frac{\sigma^2}{2}\right) T\right],$$

i.e., there is a well-defined limit independent of the discretization parameters $n$ and $h_n$.

# Some CRR properties, III



What kind of distribution does $S_n$ have? It will be useful to write $S_n$ in terms of *standardizations* of $L_j$.

# Some CRR properties, III

What kind of distribution does $S_n$ have? It will be useful to write $S_n$ in terms of *standardizations* of $L_j$.

A standardization of $L_j$ (or of any random variable) is

$$\widetilde{L_j} = \frac{L_j - \mathbb{E}L_j}{\sqrt{\mathrm{Var}L_j}},$$

i.e., it is a centered version of $L_j$, inversely scaled by its standard deviation: standardizations of random variables are mean-0 and variance-1.

With the standardization of the $L_j$ variables, we have,

$$S_n = S_0 \exp\left(\sum_{j=1}^{n} \left(\mathbb{E}L_j + \sqrt{\mathrm{Var}L_j}\widetilde{L_j}\right)\right).$$

# References I

Petters, Arlie O. and Xiaoying Dong (2016). *An Introduction to Mathematical Finance with Applications: Understanding and Building Financial Intuition*. Springer. ISBN: 978-1-4939-3783-7.