Binomial models for the term structure of interest

Bjarne Astrup Jensen*

6th October 2005


Noten er beregnet på dels at give baggrunden for disse “no arbitrage” baserede modeller, dels at præsentere den algoritme, som i praksis benyttes til at løse disse modeller numerisk. Det er dog ikke muligt at få en ordentlig fornemmelse af modellernes virkemåde uden aktiv medvirken. Det fulde udbytte er betinget af, at læseren for sig selv prøver at implementere og eksperimentere med modellerne, enten i det regneark eller i det programmeringssprog, som man føler sig mest fortrolig med, herunder at få modellernes rentestrukturer vist grafisk.

*Department of Finance, Copenhagen Business School. I am grateful to former students and to Ken Bechmann for useful comments and corrections to an earlier version of this teaching note.
1 The conventional binomial model for shares and currency:

One period

In the usual binomial model we have the following result for any derivative of \( S \), including trivially \( S \) itself:

\[
C = (1 + r)^{-1} \left[ qC^u + (1 - q)C^d \right]
\]  

Precisely because \( S \) itself must fulfill (1), the “probability” \( q \) is uniquely determined as

\[
q = \frac{(1 + r)S - S^d}{S^u - S^d}
\]  

In the currency case a modification is necessary due to the interest bearing character of the underlying asset. Since one unit of currency next period is not equivalent to one unit of currency today, but rather equivalent to the present value of one unit of currency — i.e. \( S(1 + r^*)^{-1} \), where \( r^* \) is the interest rate related to the currency in question — the analogous equation for \( q \) in the currency case is

\[
q = \frac{(1 + r)S(1 + r^*)^{-1} - S^d}{S^u - S^d}
\]  

The pricing relies on a well-known arbitrage argument via the usual technique of duplicating portfolios. The arbitrage argument furnishes the issuer with a recipe on how to hedge the position — pricing and hedging are mirror images of each other.

Assume that one somehow knows the prices of the two derivatives, whose patterns of payment are given by

\[
C^u = 1, \quad C^d = 0 \quad \text{resp.} \quad C^u = 0, \quad C^d = 1
\]

Denoting these prices as \( L^u \) and \( L^d \), respectively, we have from the rhs of (1) that they are equal to \( q(1 + r)^{-1} \) and \( (1 - q)(1 + r)^{-1} \), respectively. These two derivatives carry the name Arrow-Debreu securities in the literature, and their prices are termed Arrow-Debreu prices. Since all derivatives are portfolios of these two Arrow-Debreu securities, the knowledge of these two prices makes it possible to price everything that falls within the framework of the model: The price of a portfolio is the weighted sum of the prices of the individual components of the portfolio.

How could one find these fundamental building blocks? Try the underlying asset in combination with the riskless asset:

\[
S = L^u S^u + L^d S^d
\]

and solve for the Arrow-Debreu prices. The solution is the one given in (1)!

Arrow-Debreu prices will prove to be computationally important prices later in this note.

To encircle why this procedure does not immediately generalize to a model for the pricing of bonds and interest rate derivatives it is useful to consider the required input to and the derived output from the usual binomial model. This is done in figure 1.

The conventional model is primarily a framework to characterize how the arbitrage based pricing looks like, i.e. the pricing relation (1). These characteristics make up a checklist that does not change because one partly interchanges input and output in order to ask other questions different from the usual ones asked in figure 1. One way of doing this is shown in figure 2.
It is not possible to complete this variant of the model without further information. There are infinitely many models with this structure where the requirement “no arbitrage” is fulfilled. The two events \((S^u, S^d)\) need to be related by one restriction in order to get one specific model with one specific price process.

The additional piece of information needed could be information on volatility, either as the ratio \(S^u/S^d\) or as the difference \(S^u - S^d\). As a matter of fact there are infinitely many ways to supply this information. The important thing is that the two degrees of freedom must be reduced to just one degree of freedom.

If we define \(S^u \equiv hS^d\), where \(h \in (0, \infty)\) is a volatility measure, the other inputs to the model determine the only price process \((S^u, S^d)\) in accordance with an arbitrage free price formation:

\[
S^d = (1 + r) \frac{1}{1 - q + qh} S \\
S^u = (1 + r) \frac{h}{1 - q + qh} S
\]

(6)

It is easy to see that the expected rate of return, relative to the probabilities \(q\) for \(S^u\) and \(1 - q\) for \(S^d\), is the short rate \(r\).
2 The standard binomial model in two and several periods

In a similar multi-period context prices of all derivatives are determined by a recursive backwards calculation. For a two-period model this is spelled out in detail in (7)-(10).

\[ C = (1+r)^{-1} \left[ qC^u + (1-q)C^d \right] \]  \hspace{1cm} (7)

\[ C^u = (1+r^u)^{-1} \left[ q^u C^{uu} + (1-q^u)C^{ud} \right] \]  \hspace{1cm} (8)

\[ C^d = (1+r^d)^{-1} \left[ q^d C^{du} + (1-q^d)C^{dd} \right] \]  \hspace{1cm} (9)

\[ C = \frac{C^{uu}}{(1+r)(1+r^u)} + q(1-q^u)\frac{C^{ud}}{(1+r)(1+r^u)} + (1-q)\frac{C^{du}}{(1+r)(1+r^d)} + (1-q^d)\frac{C^{dd}}{(1+r)(1+r^d)} \]  \hspace{1cm} (10)

where the risk neutral probabilities at each node (not necessarily equal!) are again determined from the property that the underlying asset itself is a derivative.

The standard model has \( C^{ud} = C^{du} \) and constant probabilities \( q \). The first amounts to an assumption of path-independent derivatives; the latter is a consequence of the assumption of constant return distributions and constant interest rates and not an integral part of the binomial model.

The general structure introduces securities and prices on alltogether four paths: \( \{ uu, ud, du, dd \} \). The derivative \( C^{ud} = 1, C^{uu} = C^{du} = C^{dd} = 0 \), is a special case of (10) and its price is

\[ \left[ q(1+r)^{-1} \right] \left[ (1-q^u)(1+r^u)^{-1} \right] \]  \hspace{1cm} (11)

which can be obtained as the product of two consecutive one-period Arrow-Debreu prices.

There is a portfolio composition counterpart to this multiplicative property of path prices. If the path-specific security in question is not directly marketed, it can be duplicated or hedged perfectly through the following dynamic portfolio:

At time 0 buy \( (1-q^u)(1+r^u)^{-1} \) units of the Arrow-Debreu security \( C^u = 1, C^d = 0 \). If \( d \) is realized, the money invested is lost. But the value of the path-specific security to be hedged is also 0 – it expires worthless for sure.

If \( u \) is realized, the portfolio position is realized with value \( (1-q^u)(1+r^u)^{-1} \), which exactly matches the price – given \( u \) – for one unit of that claim, which pays 1 provided \( u \) is followed by \( d \) and zero otherwise. This is exactly the payments from the path-specific security in question.

Consequently this dynamic portfolio policy produces the desired result – precisely as if the security had been issued directly from the beginning. The dynamic portfolio policy provides a recipe for a selffinancing portfolio, so we are able to price the path-specific security in a meaningful way.

Figure 3 illustrates the calculation of one of the path prices, namely for the path-specific security with the final payment \( C^{ud} = 1, C^{uu} = C^{du} = C^{dd} = 0 \):

\[ \left[ q(1+r)^{-1} \right] \left[ (1-q^u)(1+r^u)^{-1} \right] \]

\[ (1-q^u)(1+r^u)^{-1} \to 0 \]

\[ \to 1 \]

\[ 0 \]

\[ 0 \]

\[ 0 \]

Figure 3
This is a basic property of the binomial model in several periods and it can be exploited intensively in numerical algorithms, e.g. in the so-called \textit{forward algorithm}, which is described in detail in sections 6 and 7 and which plays a dominant role for the construction of discrete time term structure models.

This model works well for an arbitrary number of periods. It is possible to back out the relevant \( q \)-probabilities by plugging in the underlying asset \textbf{and} the development of the interest rates. As long as the interest rate is "no more stochastic than the underlying" this works fine. The weakness is that a possible stochastic variation in the interest rate can only be allowed to rely on the stochastic properties of a single traded asset. Hence, the methodology is not robust in the sense that it should be able to work for a more general variety of underlying assets.

In the standard model it is required that (i) the short rate of interest as well as the multiplicative parameters for the underlying asset are constants and (ii) that the price of any derivative asset is \textit{path-independent}, i.e. \( C^{ud} = C^{du} \). Under these circumstances the expression for the price is normally viewed as the result of \textit{three} different scenarios instead of four:

\[
C = (1 + r)^{-2} \left[ q^2 C^{uu} + 2q(1 - q)C^{ud} + (1 - q)^2 C^{dd} \right] 
\]  

(12)

As demonstrated above the binomial model, based upon an underlying asset with a path-independent price development, can actually price path-dependent securities as well. However, the reason for the factor 2 in the “ud=du” situation is \textbf{not} that the final payoff is the same; rather it is due to the fact that the path prices for the two possible paths are identical.

\section{What is different about interest rates in the binomial model?}

Interest rates come in a large variety. Usually, the \textbf{zero-coupon} rates, denoted here by \( Y(T; 0) \),\footnote{The letter \( Y \) for "yield".} are taken as the primitives. They are related to \textbf{discount factors} \( D(T; 0) \), which convert one unit of money paid at the future point in time \( T \) to its present value at time 0. The discount factor also serves as the price of a so-called \textbf{zero-coupon bond}; the special type of bonds that have no coupon payments and only one payment, the principal, which falls due at maturity \( T \). Discount factors and zero-coupon interest rates at time 0 are related by\footnote{Or \( D(T; 0) = e^{-Y(T;0)T} \) in a continuously compounded interpretation.}

\[
D(T; 0) = (1 + Y(T; 0))^{-T} \]  

(13)

Assume that we have zero-coupon bonds and zero-coupon rates for a certain range of maturities 0, 1, ..., \( n \). If we try to do the same thing in the interest rate market as with stocks and currencies we should start at \( T = 2 \) and see whether we can find the relevant probability \( q \). We denote the one period interest rates by \( r_... \), where \( ... \) is a list of the movements up and down since time 0:

\[
D(2; 0) = (1 + Y(2; 0))^{-2} = (1 + Y(1; 0))^{-1} \left[ q(1 + r_u)^{-1} + (1 - q)(1 + r_d)^{-1} \right] 
\]  

(14)

Just like in the stock model this will determine \( q \) once we fix the development of the interest rate, i.e. determine the values of \( r_u \) and \( r_d \), resp.

Continuing we would like to do the same thing in the up-node and the down-node, i.e. determine the development of the interest rate and calculate the values of \( q^u \) and \( q^d \), resp. Unfortunately, this is not possible, since we would need to determine the two two-period interest rates at time 1. The only information we have is the initial term structure. Trying this leads to the equation

\[
D(3; 0) = (1 + Y(3; 0))^{-3} = (1 + Y(1; 0))^{-1} \cdot 
\]  

\[
q(1 + r_u)^{-1} \left\{ q_u (1 + r_{uu})^{-1} + (1 - q_u)(1 + r_{ud})^{-1} \right\} + 
\]  

\[
(1 - q)(1 + r_d)^{-1} \left\{ q_d (1 + r_{du})^{-1} + (1 - q_d)(1 + r_{dd})^{-1} \right\} 
\]  

(15)
Even if we determined completely the process for the interest rates and made the development path-independent by forcing \( r_{ud} = r_{du} \), we would still be in the situation that there are two \( q \)-probabilities, \( q_u \) and \( q_d \), to be determined and only one equation. And this gets even worse if we progress further out in the lattice; the number of \( q \)-variables to be determined is equal to the number of nodes at any given point in time. Of course, one could try to close the gap by, e.g., forcing the probabilities to be identical, i.e. to require \( q_u = q_d \) for \( t = 2 \). However, only accidentally will this common value be identical to the value of \( q \) for the first period, and such a procedure appears quite arbitrary.

4 Modelling the term structure

Modelling the dynamics for the term structure has followed another path instead. The trick is to fix the \( q \) values and let the process for possible future realizations of the short term rate of interest be determined endogenously in such a manner that the initial term structure (the values of \( Y(T; 0) \)) is matched exactly. The procedure in the “no arbitrage” based term structure models is illustrated in figure 4.

![Diagram](image)

In contrast to the traditional binomial models - with price processes one at a time for individual assets - the movement through the binomial lattice involves a *simultaneous change of the entire spectrum of zero-coupon bond prices*. As a by-product the prices of all term structure derivatives are obtained.

![Diagram](image)
Hence, it is not appropriate to identify nodes in the lattice with the price development for a single zero-coupon bond. Instead we need a coordinate system and refer to nodes as \((t, s)\), where \(t = 0, 1, 2, \ldots\) denotes time and \(s\) is a “counting variable”, keeping track of the number of “ups” which have occurred.

The lattice structure is illustrated in figure 5. The starting point is \((0, 0)\), and from a given node \((t, s)\) it is possible to move either to \((t + 1, s)\) or to \((t + 1, s + 1)\).

Figure 6 shows the zero-coupon term structure \(T \rightarrow Y(T; 0)\) (equivalent to the discount function \(T \rightarrow D(T; 0)\)) as the initial market input:

![Figure 6](image)

The first step of this procedure is shown in figure 7. The initial term structure in figure 4 becomes one of the two term structures shown — which one depends on whether an “up move” or a “down move” takes place.

![Figure 7](image)
In order to see how this model develops when moving forward in the binomial lattice, consider first \( h \) with identical “volatility parameter” \( q \). \( h \) was the first model to appear in the literature.\(^4\) It fixes the volatility structure by requiring that one period bond prices follow a geometric series:

\[
D(T; t, s) = D(t + 1; t, s) [qD(T; t + 1, s + 1) + (1 - q)D(T; t + 1, s)] \\
\forall(T, t, s) \ 0 \leq s \leq t \leq T
\]

(16)

\[
D(T; T, s) = 1 \ \forall(T, s) \ 0 \leq s \leq T
\]

(17)

\[
T \rightarrow D(T; 0) \text{ given}
\]

(18)

The Ho-Lee model was the first model to appear in the literature.\(^4\) It fixes the volatility structure by requiring that one period zero-coupon bond prices follow a geometric series:

\[
D(t + 1; t, s + 1) = hD(t + 1; t, s) \ \forall(t, s)
\]

(19)

with identical “volatility parameter” \( h \) at all nodes.

In order to see how this model develops when moving forward in the binomial lattice, consider first \( T = 2 \):

\[
D(2; 0) = D(1; 0) [(1 - q)D(2; 1, 0) + qD(2; 1, 1)]
\]

\[
= D(1; 0) [1 - q +qh] D(2; 1, 0)
\]

(20)

This results in

\[
D(2; 1, 0) = \frac{D(2; 0)}{D(1; 0)} \frac{1}{1 - q +qh}
\]

(21)

\[
D(2; 1, 1) = \frac{D(2; 0)}{D(1; 0)} \frac{h}{1 - q +qh}
\]

(22)

In this expression the term \( D(2; 0)/D(1; 0) \) is recognized as the forward price for a two period zero-coupon bond with delivery at time 1.

The basic structure of the model requires \( D(t + 1; t, s + 1)/D(t + 1; t, s) = h \). Now assume – as an induction hypothesis – that it has been proved that \( D(T; t, s + 1)/D(T; t, s) = h^{T - t} \) for \( T - t = 1, 2, \ldots, p - 1 \). Inserting this induction hypothesis in the general equation (16) for the nodes \((t, s)\) and \((t, s + 1)\), respectively, one obtains

\[
D(T; t, s) = D(t + 1; t, s) \left[1 - q + qh^{T - (t + 1)}\right] D(T; t + 1, s)
\]

(23)

\[
D(T; t, s + 1) = D(t + 1; t, s + 1) \left[1 - q + qh^{T - (t + 1)}\right] D(T; t + 1, s + 1)
\]

(24)

\(^3\)Conventionally, \( q \) is taken to be 0.5, but this is not in any way part of the “no arbitrage” assumption.

\(^4\)Ho and Lee (1986).
By dividing both sides of equation (24) for \( T - t = p \) by the corresponding sides of (23) it follows that

\[
\frac{D(T;t,s+1)}{D(T;t,s)} = h \frac{D(T;t+1,s+1)}{D(T;t+1,s)} = h \cdot h^{T-(t+1)} = h^{T-t}
\]  

(25)

By the principle of induction the “volatility structure” is determined for all maturities from the postulated volatility structure (19) for zero-coupon bonds with one period to maturity.

This knowledge enables the unraveling of the lattice in order to find \( D(T;1,0) \) and \( D(T;1,1) \) for arbitrary maturity dates \( T \):

\[
D(T;0) = D(1;0) \left[ (1-q)D(T;1,0) + qD(T;1,1) \right]
\]  

(26)

\[
= D(1;0) \left[ 1 - q + qh^{T-1} \right] D(T;1,0)
\]  

(27)

\[
D(T;1,0) = \frac{1}{1 - q + qh^{T-1}} D(T;0)
\]  

(28)

\[
D(T;1,1) = \frac{h^{T-1}}{(1 - q + qh^{T-1})} D(T;0)
\]  

(29)

Given the assumed volatility structure – and, of course, the assumption about constant risk neutral probabilities - the entire lattice can be unraveled by finding the model prices on the lower edge. Since \((q,1-q)\) as well as \(h\) are independent of time and state the model structure will repeat itself – there is no calendar time effect per se in the structure of the model.

Expressed in formulas the model prices on the lower boundary must fulfill the relation

\[
D(T;t,0) = D(t+1;t,0) \left[ 1 - q + qh^{T-(t+1)} \right] D(T;t+1,0) = D(t+1;t,0)D(t+2;t+1,0) \left[ 1 - q + qh^{T-(t+1)} \right] \left[ 1 - q + qh^{T-(t+2)} \right] D(T;t+2,0)
\]  

(30)

By continuing until the expiration date, where \( D(T;T,0) \equiv 1 \), the following relation is obtained:

\[
D(T;t,0) = \prod_{k=t}^{T-1} D(k+1;k,0) \left[ 1 - q + qh^{T-(k+1)} \right]
\]  

(31)

Naturally, this relation is also valid for \( t = 0 \) and for the two zero-coupon bonds with maturity dates equal to the dates \( T \) and \( t \) applied in (31):

\[
D(T;0) = \prod_{k=0}^{T-1} D(k+1;k,0) \left[ 1 - q + qh^{T-(k+1)} \right]
\]  

(32)

\[
D(t;0) = \prod_{k=0}^{t-1} D(k+1;k,0) \left[ 1 - q + qh^{T-(k+1)} \right]
\]  

(33)

By combining these three expressions the entire structure of the Ho-Lee model can be expressed as the following theorem:
Theorem 1 (The Ho-Lee model). The arbitrage free bond prices in the Ho-Lee model are uniquely determined by

\[ D(T; t, s) = h^{(T-t)} \left[ \prod_{k=0}^{t-1} \frac{1 - q + qh^{T-(k+1)}}{1 - q + qh^{T-(k+1)}} \right] \frac{D(T; 0)}{D(t; 0)} \]  

(34)

The term \( D(T; 0)/D(t; 0) \) is the forward price for a zero-coupon bond with maturity \( T \) and delivery date \( t \). The corresponding forward rate, \( f(T, t) \), is by definition given by

\[ \frac{D(T; 0)}{D(t; 0)} \equiv \exp(-f(T, t)(T - t)) \]  

(35)

The other two terms in (34) taken together are called “perturbation functions”.

When zero-coupon bond prices comprise a geometric series, the relevant term structures, \( T \rightarrow y(T; t, s) \), \( T > t \), are parallel curves, and the distance between these possible term structure curves are \(|\log h|\). The latter property follows from

\[ D(T; t, s) \equiv \exp(-y(T; t, s)(T - t)) \Rightarrow \]  

(36)

\[ y(T; t, s) = \frac{-1}{T - t} \log(D(T; t, s)) \Rightarrow \]  

(37)

\[ y(T; t, s) = f(T, t) + \frac{1}{T - t} \sum_{k=0}^{t-1} \log \left[ \frac{1 - q + qh^{T-(k+1)}}{1 - q + qh^{T-(k+1)}} \right] - s \log h \]  

(38)

Due to the time and state independent parameters the Ho-Lee model has very little flexibility. The volatility of the individual zero-coupon rates are governed by the last term in (38), reflecting the parallel term structure curves. It follows that all zero-coupon rates have the same volatility.

Furthermore, the Ho-Lee model will inevitably produce negative interest rates as output, unless some entirely unrealistic restrictions are imposed on the initial term structure.

Example 1

The one-period rates are easily derived from (38) by setting \( T = t + 1 \):

\[ y(t + 1; t, s) = f(t + 1, t) + \log(1 - q + qh) - s \log h \]  

(39)

Assume that the lattice has been arranged so that the smallest interest rates \( -\) equivalent to the largest of the discount factors \( D(t + 1; t, s) \) \( -\) are located at the lower edge. Then \( h < 1 \) and hence \( \log h < 0 \). The smallest one period rate at time \( t \) is

\[ y(t + 1; t, 0) = f(t + 1, t) + \log(1 - q + qh) = \log(1 - q + qh) - \log \left( \frac{D(t + 1; 0)}{D(t; 0)} \right) \]  

(40)

For a given value of \( q \) it is a consequence of the assumed volatility structure and the assumption that \( h < 1 \) that in order to avoid negative forward rates the relation

\[ \frac{D(t + 1; 0)}{D(t; 0)} \frac{1}{1 - q + qh^t} \leq 1 \ \forall t \]  

(41)

must be satisfied. For very large values of \( t \) this means that the initial forward rates must be so large that

\[ f(t + 1, t) > -\log(1 - q) \]  

(42)

For \( q = 1/2 \), e.g., the lower limit amounts to 69-70% per period, which hardly leaves any options open for practical applicability of the model over “longer” periods.
Example 2

Let \( q = 0.5 \) and \( h = 0.997 \). The initial term structure is assumed given by

\[
D(1; 0) = 0.9826, \ D(2; 0) = 0.9651, \ D(3; 0) = 0.9474, \ D(4; 0) = 0.9296, \ D(5; 0) = 0.9119
\]

The entire development of these zero-coupon bond prices is shown for the first three periods in figure 8:

\[
\begin{array}{c}
\left( \begin{array}{c}
0.9826 \\
0.9651 \\
0.9474 \\
0.9296 \\
0.9119
\end{array} \right) \\
\left( \begin{array}{c}
0.9418 \\
0.9225 \\
0.9807 \\
0.9613 \\
0.9119
\end{array} \right) \\
\left( \begin{array}{c}
1.0000 \\
0.9787 \\
0.9574 \\
0.9364 \\
0.9119
\end{array} \right) \\
\left( \begin{array}{c}
1.0000 \\
0.9787 \\
0.9574 \\
0.9364 \\
0.9119
\end{array} \right)
\end{array}
\]

The zero coupon bond expiring at time 5 reaches the following prices at time 4:

\[
D(5; 4, 0) = 0.9869, \ D(5; 4, 1) = 0.9839, \ D(5; 4, 2) = 0.9809,
\]

\[
D(5; 4, 3) = 0.9780, \ D(5; 4, 4) = 0.9751
\]

Example 3

Let \( q = 0.8 \) and \( h = 0.93979 \). The initial term structure is assumed given by

\[
D(1; 0) = 0.9259, \ D(2; 0) = 0.8534, \ D(3; 0) = 0.7836
\]

Tracing out the price development of the zero-coupon bond expiring at time 3 gives the result shown in figure 9.

This choice of parameters leads to a negative interest rate already after two periods, because \( D(3; 2, 0) = 1.0128 \). Since the bond in question has one time period to maturity at \( t = 2 \) its yield is equal to the spot rate at the node \((t, s) = (2, 0)\): \( y(3; 2, 0) = -\log D(3; 2, 0) \simeq -1.27\% \).
Example 4

Consider a horizontal initial term structure curve, i.e.

\[ D(t;0) = \exp(-rt) \]
\[ y(t;0) = r \quad \forall t \] (43)

All initial forward rates are also equal to \( r \). The two possible term structures succeeding at \( t = 1 \) are two parallel curves given by

\[ y(T;1,s) = r + \frac{1}{T-1} \log \left[ 1 - q + qh^{T-1} \right] - s \log h \] (44)

However, these two curves are not horizontal. For \( h > 1 \) the curves become increasing, for \( h < 1 \) decreasing.

Given parameter values \( r = 0.1, h = 1.1, q = 0.5 \) the upper curve varies from 0.1488 to 0.1953, the lower one from 0.05348 to 0.10.

It is possible to prevent the occurrence of negative interest rates by adjusting the Ho-Lee model in such a way that the volatility is a time-dependent variable, which must be chosen within certain upper and lower bounds dictated by the initial term structure. Such a variant of the Ho-Lee model, known as the PST model, has been proposed in Pedersen, Shiu, and Thorlacius (1989), but has not attracted much interest in the literature.

6 The forward algorithm, one-dimensional parameterization

Other known models in the literature are not analytically solvable in the same manner as the Ho-Lee model. That does not invalidate their applicability. A very efficient routine as to determine the possible term structures numerically exists and will be described in some detail here.

The price at time 0 for the Arrow-Debreu security, which pays 1 at node \((T,s)\) and zero in all other situations, is denoted by the symbol \( L(T;0) \). In the same way the symbol \( L(T,s;t,v) \) denotes the price of the same Arrow-Debreu security at node \((t,v)\).

The idea of the forward algorithm is illustrated in figure 10. Any path passing the node \((t,s) = (9,5)\) must necessarily pass through either node \((t,s) = (8,4)\) or \((t,s) = (8,5)\). And any path passing the node \((t,s) = (2,1)\) must continue to one of the nodes \((3,2)\) or \((3,3)\). This observation can be formalized more generally.
If it is required to end at node \((T, s)\) it is necessary to pass a node at time \(T - 1\). Only two possibilities exist: \((T - 1, s)\) or \((T - 1, s - 1)\). On the edges this is not even true, since there is only one way to end at \(T, T\), namely by passing \((T - 1, T - 1)\). Similarly there is only one way to end at \((T, 0)\), namely by passing the node \((T - 1, 0)\).

An Arrow-Debreu security paying 1 at node \((T, s)\) can be constructed as a dynamic portfolio policy at time 0. The advantage of looking at this portfolio policy is its ability to produce numerical results. By denoting the price of this Arrow-Debreu security at time \(T - 1\) and at the two possible preceeding nodes by \(\mathcal{L}(T, s; T - 1, s)\) and \(\mathcal{L}(T, s; T - 1, s - 1)\), respectively, the following dynamic portfolio strategy, starting at time 0, produces the desired Arrow-Debreu security:

1. buy \(\mathcal{L}(T, s; T - 1, s)\) units of the Arrow-Debreu security, which pays exactly at node \((T - 1, s)\) and which by definition has the unit price \(\mathcal{L}(T - 1, s; 0)\)

2. buy \(\mathcal{L}(T, s; T - 1, s - 1)\) units of the Arrow-Debreu security, which pays exactly at node \((T - 1, s - 1)\) and which by definition has the unit price \(\mathcal{L}(T - 1, s - 1; 0)\)

This portfolio has exactly the same payments as the Arrow-Debreu security with a unit payment at node \((T, s)\) and the price \(\mathcal{L}(T, s; 0)\). By bypassing node \((T - 1, s - 1)\) as well as \((T - 1, s)\) the portfolio ends up worthless as it should. If node \((T - 1, s)\) is reached the portfolio pays off exactly \(\mathcal{L}(T, s; T - 1, s)\), which is reinvested in one unit of the appropriate one-period Arrow-Debreu security over the last period. Analogously, if node \((T - 1, s - 1)\) is reached the portfolio pays off exactly \(\mathcal{L}(T, s; T - 1, s - 1)\), which is reinvested in one unit of the appropriate one-period Arrow-Debreu security over the last period.

The price of this dynamic portfolio strategy must correspond to the price of directly acquiring the Arrow-Debreu security with one unit of payment at node \((T, s)\). I.e.

\[
\mathcal{L}(T, s; 0) = \mathcal{L}(T, s; T - 1, s)\mathcal{L}(T - 1, s; 0) + \mathcal{L}(T, s; T - 1, s - 1)\mathcal{L}(T - 1, s - 1; 0)
\]  (45)
This equation is also valid at the edges, but in that case the rhs will be reduced to a single relevant term. Alternatively, any price expression for securities that cannot exist can be read as zero.

The initial Arrow-Debreu prices $L(1,1;0)$ and $L(1,0;0)$, respectively, are known. They are the product of the relevant risk neutral probability and the discount factor $D(1,0)$.

**Theorem 2 (Forward algorithm).** The arbitrage free zero-coupon bond prices derived from a binomial model with

- risk neutral probabilities $(q, 1−q)$
- a given initial term structure $T → D(T;0)$ and
- a given, one-dimensional parameterization of the one-period discount factors

satisfies the following set of recursive equations for the determination – possibly through a numerical solution – of the one-period discount factors and Arrow-Debreu prices:

**Step 0 (Initialization)**

$$L(1,1;0) = qD(1;0) \quad (46)$$
$$L(1,0;0) = (1−q)D(1;0) \quad (47)$$
$$t = 2 \quad (48)$$

**Step 1**

$$D(t;0) = \sum_{s=0}^{t-1} L(t−1,s;0)D(t−1,s) \quad (49)$$

is solved for the zero coupon bond prices $D(t; t−1,s)$, which are tied together in such a way that only one free variable remains to be determined.

Go to step 2.

**Step 2**

$$L(t,s; t−1,s−1) = qD(t; t−1,s−1) \quad \text{for } 1 ≤ s ≤ t \quad (50)$$
$$L(t,s; t−1,s) = (1−q)D(t; t−1,s) \quad \text{for } 0 ≤ s ≤ t−1 \quad (51)$$
$$L(t,s;0) = L(t,s; t−1,s)L(t−1,s;0) + L(t,s; t−1,s−1)L(t−1,s−1;0) \quad (52)$$
$$t = t + 1 \quad (53)$$

Continue with step 1.

Step 1 and the relation (49) express the fact that a zero coupon bond is equivalent to an equally weighted portfolio of one unit of each of the Arrow-Debreu securities with unit payments at nodes $(t,s)$, $s = 0,1,\ldots,t$. This portfolio is duplicable through a portfolio policy that

- at time 0 is composed of $D(t; t−1,s)$ units of the Arrow-Debreu security with unit payment at node $(t−1,s)$, $s = 0,1,\ldots,t−1$.

---

5 Arrow-Debreu prices with meaningless arguments are 0 by definition. This is the case, e.g., for $L(t,t; t−1,t)$. 

13
• in the realized node \((t - 1, s)\) the proceeds \(D(t; t - 1, s)\) from the established portfolio is reinvested in the relevant zero-coupon bond with one time period to maturity

Step 2 and the relations (50)-(51) are identical to the initialization. The one-period Arrow-Debreu securities are always the product of the relevant, risk neutral probability and corresponding one-period discount factor. The updating (52) is just a restatement of (45).

The solution of (49) usually relies on a numerical search - no algorithm is needed if it is possible to solve the model with analytical solutions. Usually this is a relatively simple task, because the rhs will vary \textit{monotonically} with the parameter. Implementing these calculations in a spreadsheet with a “solver”-facility\(^6\) a numerical solution is readily available. For larger problems with a very high number of time periods an appropriate programming language must be employed.

\textit{Example 5: The BDT model}

In the \textit{Ho-Lee} model it was assumed that the one-period zero-coupon bond prices were geometric series. In the \textit{Black-Derman-Toy} model it is assumed that the one-period interest rates are geometric series with a ratio \(v(t)\) specified outside the model:

\[
y(t + 1; t, s) = v(t)y(t + 1; t, s - 1) \iff \log y(t + 1; t, s) = \log v(t) + \log y(t + 1; t, s - 1)
\]  

(54)

Under these circumstances (49) looks like

\[
D(t; 0) = \sum_{s=0}^{t-1} L(t - 1, s; 0) \exp\{-v(t)y(t, t - 1, 0)\}
\]  

(55)

with one unknown: \(y(t; t - 1, 0)\). The rhs is again a monotone function of this unknown, and the forward algorithm is applicable. However, the fact that the volatility \(v(t)\) is \textit{inside} the exponential function prevents us from deriving an analytical expression for the solution.

\section{The forward algorithm, two-dimensional parameterization}

The models described so far take the actual term structure together with an exogenously postulated structure of the model with a given parameterization as their point of departure. E.g. by a volatility parameter for either the interest rates or the bond prices. The output of the models match exactly the term structure used as input.

It is sometimes desirable also to match the existing \textit{volatility structure}, which is defined below.

Assume, e.g., that apart from the zero-coupon term structure \(t \rightarrow y(t; 0)\) a volatility structure is also given for the zero-coupon rates. This volatility structure can be supplied in different ways, but the end result is that the individual points on the two possible, succeeding term structures:

\[
t \rightarrow y(t; 1, 0) \quad t \rightarrow y(t; 1, 1)
\]

are tied together:

\[
y(t; 1, 1) = c(t)y(t; 1, 0) \quad t = 2, 3, \ldots
\]  

(56)

through \(c(t)\) – an input in principle determined by the market.\(^7\) The output from the model must now determine the future \textit{level} as well as the future \textit{volatility structure} as output. The forward-algorithm can do this job.

\(^6\)This goes for standard spreadsheets like Excel, e.g.

\(^7\)This statement should not be interpreted as claiming that the volatility structure can be deduced easily from market data.
Theorem 3. The arbitrage free zero-coupon bond prices derived from a binomial model with

- risk neutral probabilities \((q, 1-q)\)
- a given initial term structure \(T \rightarrow D(T; 0)\) and initial volatility structure \(T \rightarrow c(T)\)
- a given, two-dimensional parameterization of the one-period discount factors and the one-period volatility structure

satisfies the following set of recursive equations for the determination – possibly through a numerical solution – of the one-period discount factors, future volatility of one-period interest rates and Arrow-Debreu prices:

**Step 0 (Initialization)**

\[
\begin{align*}
\mathcal{L}(1, 0; 0) &= (1 - q)D(1; 0) & (57) \\
\mathcal{L}(1, 1; 0) &= qD(1; 0) & (58)
\end{align*}
\]

Solve for \(D(T; 1, 0)\) and \(D(T; 1, 1)\) for \(T = 2, 3, \ldots\) until the maximal maturity under consideration in the model is reached:

\[
D(T; 0) = D(1; 0) [(1 - q)D(T; 1, 0) + qD(T; 1, 1)]
\]  
(59)

With the given volatility structure, linking \(D(T; 1, 1)\) with \(D(T; 1, 0)\), there is only one variable to determine for each \(T\).

\[
\begin{align*}
\mathcal{L}(2, 0; 1, 0) &= (1 - q)D(2; 1, 0) & (60) \\
\mathcal{L}(2, 1; 1, 0) &= qD(2; 1, 0) & (61) \\
\mathcal{L}(2, 1; 1, 1) &= (1 - q)D(2; 1, 1) & (62) \\
\mathcal{L}(2, 2; 1, 1) &= qD(2; 1, 1) & (63)
\end{align*}
\]

\[t = 3\]  
(64)

**Step 1**

The two equations in two unknown parameters:

\[
\begin{align*}
D(t; 1, 0) &= \sum_{s=0}^{t-2} \mathcal{L}(t - 1, s; 1, 0)D(t; t - 1, s) & (65) \\
D(t; 1, 1) &= \sum_{s=1}^{t-1} \mathcal{L}(t - 1, s; 1, 1)D(t; t - 1, s) & (66)
\end{align*}
\]

are solved simultaneously for the bond prices \(D(t; t - 1, s)\) and the volatility \(v(t)\) of the short term interest rate.

Continue with step 2.
Step 2

\[ \mathcal{L}(t, s; t-1, s-1) = qD(t; t-1, s-1) \quad \text{for} \quad 1 \leq s \leq t \]  
\[ \mathcal{L}(t, s; t-1, s) = (1-q)D(t; t-1, s) \quad \text{for} \quad 0 \leq s \leq t-1 \]  
\[ \mathcal{L}(t, s; 1, 0) = \mathcal{L}(t, s; t-1, s-1)\mathcal{L}(t-1, s-1; 1, 0) \]  
\[ \mathcal{L}(t, s; 1, 1) = \mathcal{L}(t, s; t-1, s)\mathcal{L}(t-1, s; 1, 1) + \mathcal{L}(t, s; t-1, s-1)\mathcal{L}(t-1, s-1; 1, 1) \]

\[ t = t + 1 \]

Continue with step 1.

Depending on the specific structure of the model it may be the case that the input consists of the zero-coupon bond prices as well as their volatility, the zero-coupon yields as well as their volatility like in the BDT model. Other possibilities exist. The requirement is that the magnitudes are tied together structurally in such a way that there are exactly two free variables in the two equations (65)-(66).

Again it is the case that a numerical search is easy to implement in a spreadsheet with a “solver”-facility. The algorithm is known as computationally very efficient.

Example 6

Assume a flat initial term structure and a flat volatility structure:

\[ y(t; 0) = 0.1 \quad c(t) = 1.1 \]

Setting \( q = 0.5 \) leads to\(^8\)

\[ \mathcal{L}(1, 0; 0) = \mathcal{L}(1, 1; 0) = 0.5\exp(-0.1) = 0.45241871 \]  
\[ D(T; 0) = \exp(-0.17T) = 0.45241871 \cdot [\exp(-y(T; 1, 0)(T-1)) + \exp(-1.1 \cdot y(T; 1, 0)(T-1))] \]

The solution for \( T = 2 \) and \( T = 3 \) is:

\[ y(2; 1, 0) = 0.095249 \quad y(3; 1, 0) = 0.095260 \]

since \( v(2) \) is known from the input as \( v(2) = c(2) = 1.1 \).

The initialization process leads to:

\[ \mathcal{L}(2, 2; 1, 1) = \mathcal{L}(2, 1; 1, 1) = 0.5D(2; 1, 1) = 0.45026406 \]  
\[ \mathcal{L}(2, 1; 1, 0) = \mathcal{L}(2, 0; 1, 0) = 0.5D(2; 1, 0) = 0.45457327 \]

The first time step 1 is applied the equations that determine the two variables \( D(3; 2, 0) \) and \( v(3) \) look like:

\[ D(3; 1, 0) = \mathcal{L}(2, 0; 1, 0)D(3; 2, 0) + \mathcal{L}(2, 1; 1, 0)D(3; 2, 1) \]  
\[ D(3; 1, 1) = \mathcal{L}(2, 1; 1, 1)D(3; 2, 1) + \mathcal{L}(2, 2; 1, 1)D(3; 2, 2) \]

The result is found through numerical search as \( y(3; 2, 0) = 0.0907425 \) and \( v(3) \simeq 1.1 \).  
\[^8\]This is valid for \( T = 2, 3, \ldots \) because it is assumed that the volatility structure is flat \( c(t) = 1.1 \). In general the volatility is depending on the maturity \( T \).
Upon continuing the algorithm the derived volatilities \( v(t) \) will increase, but this effect disappears by rounding off the results for \( T = 3 \).

**Example 7**

The BDT model is described in an original article\(^9\) that is not easy to follow in detail. The following input values are given:

<table>
<thead>
<tr>
<th>Maturity ( t )</th>
<th>Yield (%)</th>
<th>Yield volatility (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.0</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>11.0</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>12.0</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>12.5</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>13.0</td>
<td>16</td>
</tr>
</tbody>
</table>

The “yield” is calculated as an ordinary discretely calculated “yield” and not as a “continuously compounded rate of return”. This is a convention of no real consequence. The relation (55) can obviously also be formulated as

\[
D(t; 0) = \sum_{s=0}^{t-1} L(t-1, s; 0) (1 + v(t)^s y(t; t-1, 0))^{-1}
\]  

(78)

Furthermore, the magnitudes denoted as “yield volatility” must be transformed in order to fit the notation applied in this note:

\[
c(t) \equiv \exp(2 \cdot (“yield volatility”))
\]  

(79)

The initial term structure is reflected in the zero-coupon bond prices:

\[
\begin{align*}
D(1; 0) &= (1.100)^{-1} = 0.90909091 \\
D(2; 0) &= (1.110)^{-2} = 0.81162243 \\
D(3; 0) &= (1.120)^{-3} = 0.71178025 \\
D(4; 0) &= (1.125)^{-4} = 0.62429508 \\
D(5; 0) &= (1.130)^{-5} = 0.54275994
\end{align*}
\]

Assuming \( q = 0.5 \)\(^10\) we first work through the initialization process:

\[
L(1, 1; 0) = L(1, 0; 0) = 0.5D(1; 0) = 0.45454545.
\]  

(80)

\[
D(2; 0) = 0.5D(1; 0) \left[ (1 + y(2; 1, 0))^{-1} + (1 + \exp(0.38) y(2; 1, 0))^{-1} \right]
\]  

(81)

The suggested solution with \( y(2; 1, 0) = 0.0979 \), \( y(2; 1, 1) = 0.1432 \) is in accordance with equation (81), allowing for an insignificant rounding error.

Next we solve the other initialization equations in order to derive the two possible term structures for \( t = 1 \):

\[
D(3; 0) = 0.5D(1; 0) \left[ (1 + y(3; 1, 0))^{-2} + (1 + \exp(0.36) y(3; 1, 0))^{-2} \right]
\]  

(82)

\[
D(4; 0) = 0.5D(1; 0) \left[ (1 + y(4; 1, 0))^{-3} + (1 + \exp(0.34) y(4; 1, 0))^{-3} \right]
\]  

(83)

\[
D(5; 0) = 0.5D(1; 0) \left[ (1 + y(5; 1, 0))^{-4} + (1 + \exp(0.32) y(5; 1, 0))^{-4} \right]
\]  

(84)

\(^9\)Black, Derman, and Toy (1990).

\(^10\)Which as a parenthetical remark is often assumed in these models, but in no way required by their structure.
The solutions are not explicitly reported in the article. Allowing for marginal rounding errors, which inevitably occur when (only) four decimals’ precision is used, the solutions to the equations (82)-(84) become:

\[ y(3; 1, 0) = 0.1076y(3; 1, 1) = 0.1542 \]
\[ y(4; 1, 0) = 0.1117y(4; 1, 1) = 0.1570 \]
\[ y(5; 1, 0) = 0.1167y(5; 1, 1) = 0.1607 \]

(85)

The initialization is concluded by:

\[ L(2, 0; 1, 0) = L(2, 1; 1, 0) = 0.5 \cdot 0.91082977 = 0.45541488 \]
\[ L(2, 1; 1, 1) = L(2, 2; 1, 1) = 0.5 \cdot 0.87473758 = 0.43736879 \]

(86) (87)

Next we try step 1 for \( t = 3 \). The solutions in the article are “guessed” as \( y(3; 2, 0) = 0.0976, y(3; 2, 1) = 0.1377 \) and \( y(3; 2, 2) = 0.1942 \), respectively. It corresponds to \( v(3) \approx 1.41 \) and a derived volatility of approximately 17.2% in the way this magnitude is reported in the article. We do a test:

\[ D(3; 1, 0) = 0.81514357 = 0.45541488 \cdot [0.91107872 + 0.87896634] \]
\[ D(3; 1, 1) = 0.75065064 = 0.43736879 \cdot [0.87896634 + 0.83738067] \]

(88) (89)

Within a 3 decimal precision the guess in the article is correct.

The next calculation is step 2 for \( t = 3 \). We only report the final results for the Arrow-Debreu securities:

\[ L(3, 0; 1, 0) = 0.20745940 \quad L(3, 1; 1, 0) = 0.40760658 \quad L(3, 2; 1, 0) = 0.20014718 \]
\[ L(3, 1; 1, 1) = 0.19221622 \quad L(3, 2; 1, 1) = 0.37533831 \quad L(3, 3; 1, 1) = 0.18312207 \]

(90)

The algorithm is repeated through step 1 for \( t = 4 \) etc.

It is a characteristic of the BDT model — not very transparent without going through the calculations — that a downward sloping “yield volatility” normally induces a yet more downward sloping derived volatility of one-period rates. This property is not without exceptions\(^\text{11}\). Analogously, an assumed constant \( v(t) \) in the one-dimensional algorithm will normally induce an increasing volatility of the long rates in the model.

## 8 Types of interest rate derivatives

With the notation used here we can characterize a number of different interest rate derivatives.

**Swaps** are assumed to be known. The payments \( P_t \) from a plain vanilla interest rate swap, exchanging fixed rate for floating rate for the time horizon \( t = 1, 2, \ldots, T \), are given as

\[ P_t = K \cdot (Y(t; t - 1, s) - Y_f(T)) \]

(91)

where \( K \) is the notional principle and \( Y_f(T) \) must be determined at the initiation of the contract in order to make the value of the contract equal to zero. One way to do this is to add the notional principle to both payments at the final date \( T \):

\[ P_T = K \cdot (Y(T; T - 1, s) - Y_f(T)) \]
\[ P_T = K \cdot (1 + Y(T; T - 1, s) - (1 + Y_f(T))) \]

(92)

\(^\text{11}\)Through a suitably manipulation with the parameters and the initial term structure the derived volatility may first decrease and then increase afterwards.
Viewed in this way the swap in nothing but a long position in a variable rate bond and a short position in a fixed rate bond with the same maturity. Since a variable rate bond is always priced at par the contract rate \( Y_f(T) \) is the coupon rate necessary on a fixed rate bond with maturity \( T \) in order to be priced at par — sometimes referred to as the *par yield*. Swap rates for different maturities form a term structure of their own, \( T \rightarrow Y_f(T) \), referred to as the *swap curve*.

Options on zero-coupon bonds are treated in the usual way for options in a binomial model, once the interest rates are found in the lattice. However, nice closed-form solutions are not so easy to get in the discrete time models treated here because of the changing nature of the discounting.

Other contracts are often characterized as having multiple payments, i.e. they consist of several options each of which can be exercised independent of the other options. Some of the best known are caps and floors.

A **caplet** with cap rate \( Y_c \) is a contract that pays a multiple of \( \max\{0, Y(t;t-1,s) - Y_c\} \), i.e. the same kind of asymmetry as in an ordinary call option, but the variable governing the payments is not the price of a traded asset.

A **cap** is a portfolio of caplets with identical cap rate \( Y_c \). Such contracts are used as hedging instruments in variable loan contracts with multiple reset dates for the variable rate. This frequency is known as the **tenor** of the contract.

Similarly, a **floor** is a portfolio of contracts that pay \( \max\{0, Y_c - Y(t;t-1,s)\} \), i.e. the same kind of asymmetry as in an ordinary put option. Since

\[
\max\{0, Y(t;t-1,s) - Y_c\} - \max\{0, Y_c - Y(t;t-1,s)\} = Y(t;t-1,s) - Y_c
\]

the two contracts obey a parity relation similar to the put-call parity. The rhs of (93) is the payment stream from a plain vanilla swap contract that pays variable rate \( Y(t;t-1,s) \) against fixed rate \( Y_c \).

A **swaption** is an option to enter into a swap at some future point in time and for a specified future time period. Since a swap essentially is

- a long position in a variable rate bond, which by construction is sold at par and
- a simultaneous short position in a fixed rate bond

a swaption is an option to buy a fixed rate bond — with a prespecified fixed rate — at par.

A **spread option** is an option that pays in accordance with the spread between two interest rates or two prices. E.g. the spread between futures prices on the same underlying, but with different maturities; e.g. oil futures with varying maturity dates. Interest based spread options are options that pay in accordance with e.g. the difference between specified government bond yields, e.g. between Danish and Euro government bond yields at some agreed upon maturity. A variant known as a SYCURVE option pays in accordance with the spread between different yields (at different maturities) on the yield curve.

### 9 The levelling algorithm

Many term structure derivatives are inherently path-dependent securities. In addition to the derivatives mentioned in section 8 there are, of course, variants of the usual barrier options etc. But mortgage products constitute an important class of interest rate derivatives in themselves; in contrast to standard swaps, caps, floors etc. mortgage products are usually associated with a declining notional principal on which the payments on the derivatives are based.

\[12\] Or reverse; the contract has two sides that are necessarily mirror images.
A variable rate mortgage loan, where the principal is being gradually repaid during the term of loan, may in itself be a path-dependent construction; e.g. when the repayments are depending on the actual development in the variable interest rate. We consider here the case of a variable rate annuity with maturity $T$ periods, where the payments are reset in accordance with the annuity formula. If we denote the remaining principal after payment at time $t-1$ as $RP_{t-1}$, then by routine annuity calculations the following relations describe the amortization of the loan:

$$P_t = RP_{t-1} \cdot \alpha_{t-1}^{(T-t+1)Y(t-1,-)}$$  \hfill (94)

$$RP_t = (1 + Y(t; t-1; \cdot)RP_{t-1} - P_t = RP_{t-1} - (1 + Y(t; t-1; \cdot))^{-((T-t+1))}$$  \hfill (95)

$$Z_t \equiv RP_{t-1} - RP_t = RP_{t-1}((1 + Y(t; t-1; \cdot))^{-((T-t+1))} - \frac{Y(t; t-1; \cdot)}{1 - (1 + Y(t; t-1; \cdot))^{-((T-t+1))}}$$  \hfill (96)

Path-dependency is a very bad property in a computational sense. The number of nodes grow exponentially as $2^T$ instead of linearly as $t+1$ for a path-independent derivative. Unless some tricks can be used price computations for a derivative involves calculations along each of these $2^T$ paths, which quickly leads to undesirable computation times. Hence, there is a clear interest in making shortcuts by either finding efficient approximations or by exploiting special structures that can reduce this computational complexity.

Observe from (94)-(96) that the magnitudes $P_t$ and $RP_t$ are proportional to the principal, $RP_0$, through any possible path that the interest rate $Y(t; t-1; \cdot)$ may take. This is typical of many interest rate related calculations, in particular loans and derivatives associated herewith, e.g. caps for variable interest rate loans and similar contracts.

For this reason let $\hat{C}(t,s)$ denote the normalized price of the contingent claim under consideration. In the case of a loan this corresponds to the value of the derivative for a principal of 1 or 100; the actual value can then always be found by scaling with the actual level of the (outstanding) principal. Furthermore, the relative amount of repayment at either of the following nodes, $(t+1, s)$ and $(t+1, s+1)$, respectively, is known at node $(t, s)$ and independent of which node is being realized. We denote this scaling factor as $X(t, s)$.

In terms of formulas, following the usual backwards recursive calculation in the binomial model, this proportionality property can now be expressed as

$$\hat{C}(t, s) = D(t+1; t, s) \left[ P(t+1; t, s) + X(t, s) \left\{ q\hat{C}(t+1, s+1) + (1-q)\hat{C}(t+1, s) \right\} \right]$$

where

- $P(t+1; t, s)$ is the direct payment from the derivative at either of the two nodes $(t+1, s)$ and $(t+1, s+1)$, respectively, given a normalized level at $(t, s)$

- $X(t, s)$ is the fraction of this normalized unit that remains in the continuation of the claim after the payment $P(t+1; t, s)$ at time $t+1$.

From $\hat{C}(0,0)$ we can now find the value of the derivative by scaling with the appropriate amount of principal.

\[13\]In technical terms such a random variable is called predictable; its value is relevant at time $t+1$, but it is known at time $t$. 

20
**Example 8**

Consider the following development of the short term rate of interest, \( Y(t; t - 1, \cdot) \):

![Diagram of interest rate development](image_url)

Figur 11

Assume that we have a variable rate loan with principal 100 and a repayment profile that repays in accordance with the annuity principle. This means that in any period the next payment is fixed “as if” the loan was an annuity – with the remaining principal as the principal value and the time to maturity as the maturity.

For a maturity of four periods there are 8 different paths:

<table>
<thead>
<tr>
<th>Path no.</th>
<th>( t = 0 )</th>
<th>( t = 1 )</th>
<th>( t = 2 )</th>
<th>( t = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.00%</td>
<td>6.50%</td>
<td>7.00%</td>
<td>7.50%</td>
</tr>
<tr>
<td>2</td>
<td>6.00%</td>
<td>6.50%</td>
<td>7.00%</td>
<td>6.75%</td>
</tr>
<tr>
<td>3</td>
<td>6.00%</td>
<td>6.50%</td>
<td>6.25%</td>
<td>6.75%</td>
</tr>
<tr>
<td>4</td>
<td>6.00%</td>
<td>6.50%</td>
<td>6.25%</td>
<td>6.00%</td>
</tr>
<tr>
<td>5</td>
<td>6.00%</td>
<td>5.75%</td>
<td>6.25%</td>
<td>6.75%</td>
</tr>
<tr>
<td>6</td>
<td>6.00%</td>
<td>5.75%</td>
<td>6.25%</td>
<td>6.00%</td>
</tr>
<tr>
<td>7</td>
<td>6.00%</td>
<td>5.75%</td>
<td>5.50%</td>
<td>6.00%</td>
</tr>
<tr>
<td>8</td>
<td>6.00%</td>
<td>5.75%</td>
<td>5.50%</td>
<td>5.25%</td>
</tr>
</tbody>
</table>

Table 1

The sequence of remaining principles is path-dependent. Along the 8 different paths they develop in accordance with (94)-(96) as shown in table 2:

<table>
<thead>
<tr>
<th>Path no.</th>
<th>( t = 1 )</th>
<th>( t = 2 )</th>
<th>( t = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>77.140851</td>
<td>53.028495</td>
<td>27.410865</td>
</tr>
<tr>
<td>2</td>
<td>77.140851</td>
<td>53.028495</td>
<td>27.410865</td>
</tr>
<tr>
<td>3</td>
<td>77.140851</td>
<td>53.028495</td>
<td>27.317710</td>
</tr>
<tr>
<td>4</td>
<td>77.140851</td>
<td>53.028495</td>
<td>27.317710</td>
</tr>
<tr>
<td>5</td>
<td>77.140851</td>
<td>52.850688</td>
<td>27.226112</td>
</tr>
<tr>
<td>6</td>
<td>77.140851</td>
<td>52.850688</td>
<td>27.226112</td>
</tr>
<tr>
<td>7</td>
<td>77.140851</td>
<td>52.850688</td>
<td>27.132592</td>
</tr>
<tr>
<td>8</td>
<td>77.140851</td>
<td>52.850688</td>
<td>27.132592</td>
</tr>
</tbody>
</table>

Table 2

**Remark 1**

In the first period the payment is \( 100 \cdot \alpha_{64}^{-1} = 28.859149 \). The interest payment is 6, so the repayment of principal is 22.859149 leaving a remaining principal of 100 – 22.859149 = 77.140851. Along path 3, e.g., the next payment is found as 77.140851 – 5.01415 = 29.126511. The interest payment is 0.065 – 77.140851 = 5.01415 and the remaining principal becomes 77.140851 + 5.01415 – 29.126511 = 53.028495. And so on.
Assume that debtor considers buying a cap to limit the borrowing rate upwards. If the contract rate on this cap is set at 6%, equal to the initial value of the interest rate, what would be the fair price of this cap?\textsuperscript{14}

Consider first the following revised version of table, where we have marked the nodes where the cap is active by an asterisk (*):

<table>
<thead>
<tr>
<th>Path no.</th>
<th>$t = 0$</th>
<th>$t = 1$</th>
<th>$t = 2$</th>
<th>$t = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.00%</td>
<td>6.50%*</td>
<td>7.00%*</td>
<td>7.50%*</td>
</tr>
<tr>
<td>2</td>
<td>6.00%</td>
<td>6.50%*</td>
<td>7.00%*</td>
<td>6.75%*</td>
</tr>
<tr>
<td>3</td>
<td>6.00%</td>
<td>6.50%*</td>
<td>6.25%*</td>
<td>6.75%*</td>
</tr>
<tr>
<td>4</td>
<td>6.00%</td>
<td>6.50%*</td>
<td>6.25%*</td>
<td>6.00%</td>
</tr>
<tr>
<td>5</td>
<td>6.00%</td>
<td>5.75%</td>
<td>6.25%*</td>
<td>6.75%*</td>
</tr>
<tr>
<td>6</td>
<td>6.00%</td>
<td>5.75%</td>
<td>6.25%*</td>
<td>6.00%</td>
</tr>
<tr>
<td>7</td>
<td>6.00%</td>
<td>5.75%</td>
<td>5.50%</td>
<td>6.00%</td>
</tr>
<tr>
<td>8</td>
<td>6.00%</td>
<td>5.75%</td>
<td>5.50%</td>
<td>5.25%</td>
</tr>
</tbody>
</table>

| Tabel 3 |

There are (at least) two different ways to go about calculating the price of this derivative.

One possibility is to trace the payments along each individual path. This would result in the following table of savings:

<table>
<thead>
<tr>
<th>Path no.</th>
<th>$t = 2$</th>
<th>$t = 3$</th>
<th>$t = 4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.385704</td>
<td>0.530285</td>
<td>0.411163</td>
</tr>
<tr>
<td>2</td>
<td>0.385704</td>
<td>0.530285</td>
<td>0.205581</td>
</tr>
<tr>
<td>3</td>
<td>0.385704</td>
<td>0.132571</td>
<td>0.204883</td>
</tr>
<tr>
<td>4</td>
<td>0.385704</td>
<td>0.132571</td>
<td>0.000000</td>
</tr>
<tr>
<td>5</td>
<td>0.000000</td>
<td>0.132127</td>
<td>0.204196</td>
</tr>
<tr>
<td>6</td>
<td>0.000000</td>
<td>0.132127</td>
<td>0.000000</td>
</tr>
<tr>
<td>7</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
<tr>
<td>8</td>
<td>0.000000</td>
<td>0.000000</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

| Table 4 |

Remark 2

By construction the cap gives no payout at time 1. At time 2 there is only a payout if the interest rate 6.5% (the “up case”) is realized. The cap covers the additional 0.5% over and above the contract rate 6%. For all paths, the remaining principal is 77.140851, so the coverage from the cap is 0.005 \cdot 77.140851, i.e. 0.385704. For paths 5-8 the interest rate is below 6% at $t = 1$.

Along path 6, e.g., there is only one entry where the cap become active, namely for $t = 3$. Since the remaining principal is 52.850688 and the cap covers 0.25% of this amount the payment at $t = 4$ along path 6 is 0.025 \cdot 52.850688 = 0.132127.

To find the value of the cap it is necessary to discount these payments by the appropriate path-dependent sequence of discount factors. The result is shown in table 5.

\textsuperscript{14}To be precise, the cap is assumed only to guarantee against the interest payments themselves. When the cap is activated the payment from the annuity calculation is unchanged, but the debtor gets a reimbursement of the amount of interest payment that exceeds what should be paid at an interest rate of 6%.
\begin{table*}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Path no. & \(t = 2\) & \(t = 3\) & \(t = 4\) & Path value \\
\hline
1 & 0.385704 & 0.530285 & 0.411163 & 1.097310 \\
2 & 0.385704 & 0.530285 & 0.205581 & 0.940102 \\
3 & 0.385704 & 0.132571 & 0.204883 & 0.452190 \\
4 & 0.385704 & 0.132571 & 0.000000 & 0.452190 \\
5 & 0.000000 & 0.132127 & 0.204196 & 0.271544 \\
6 & 0.000000 & 0.132127 & 0.000000 & 0.110937 \\
7 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\
8 & 0.000000 & 0.000000 & 0.000000 & 0.000000 \\
\hline
\textbf{Sum} & & & & \textbf{3.484284} \\
\hline
\end{tabular}
\caption{Table 5}
\end{table*}

Remark 3

Consider path no. 1. The backward recursive calculation along this path is shown in the calculations below:

\begin{align*}
1.075^{-1} \cdot 0.411163 &= 0.382477 \\
1.070^{-1} \cdot (0.382477 + 0.530285) &= 0.853049 \\
1.065^{-1} \cdot (0.853049 + 0.385704) &= 1.163148 \\
1.060^{-1} \cdot 1.163148 &= 1.097310
\end{align*}

The other paths are treated the same way. Finally, each path must be weighted by \(1/8\), given that the value for \(q\) is assumed to be \(1/2\), and the result must be summed. The result in a cap price of \(3.484284/8 = 0.435536\).

This procedure is obviously cumbersome as the number of periods grows. The other possibility is to use the levelling algorithm. The calculations are shown in figure 12.

\begin{figure*}
\centering
\includegraphics[width=\textwidth]{figure12.png}
\caption{The levelling algorithm}
\end{figure*}

Remark 3

The numbers are calculated as follows. For \(t = 3\) the cap is inactive at nodes \((3,0)\) and \((3,1)\). At \((3,3)\) the cap covers 0.75\% of the principal, which is paid at \(t = 4\). In a normalized calculation with principal 100 this payments is 0.75 in each of the two succeeding scenarios; the discounted value of this is \(1.0675^{-1} \cdot 0.75 = 0.702576\). At \((3,4)\) the analogous calculation is \(1.075^{-1} \cdot 1.5 = 1.395349\).

The value at node \((2,0)\) is obviously 0, since both succeeding nodes have value 0. The value at node \((2,1)\) is determined as follows:

\[
\hat{C}(2,1) = 1.0625^{-1} \cdot (0.25 + 0.515152 \cdot 0.5 \cdot (0.7026 + 0)) = 0.405622
\]

where 0.515151... is the fraction of the principal on a two period annuity loan remaining at time 3 when issued at time 2 at the interest rate 6.25\%.
Similarly, the value at node \((2, 2)\) is determined as

\[
\hat{C}(2, 2) = 1.070^{-1} \cdot [1 + 0.516908 \cdot 0.5 \cdot (1.395349 + 0.702576)] = 1.441324
\]

where 0.516908... is the fraction of the principal on a two period annuity loan remaining at time 3 when issued at time 2 at the interest rate 7.0%.

As the last detailed calculation consider the value at node \((1, 1)\):

\[
\hat{C}(1, 1) = 1.065^{-1} \cdot [0.5 + 0.687424 \cdot 0.5 \cdot (1.441324 + 0.405622)] = 1.065556
\]

where 0.687424... is the fraction of the principal on a three period annuity loan remaining at time 2 when issued at time 1 at the interest rate 6.5%.

The value \(\hat{C}(0, 0)\) is the price of the cap for a principal value of 100. The value of the cap for any principal can be found by scaling.

The levelling algorithm produces a price for such path-dependent derivatives by a calculation that is computationally identical to the backwards recursive calculation for a path-independent derivative. The only requirement is that it must be possible to do the scaling at any node.

This is the case for many contracts and it produces a very efficient numerical algorithm for such contracts. But it is not the case for all contracts. One can interpret the scaling requirement by noting that when we know the remaining principal at a given node it must be immaterial how past history got us there.

If the contract has features similar to Asian options, where the payments depend on accumulated observations over the paths, this might not work. Contracts that depend on averages of past observed interest rates, the maximum or the minimum over past realized interest rates or lock the interest rate after certain backwards looking criteria do not satisfy the scaling requirement.

### References


