Term Structure Lattice Models

1 The Term-Structure of Interest Rates

If a bank lends you money for one year and lends money to someone else for ten years, it is very likely that the rate of interest charged for the one-year loan will differ from that charged for the ten-year loan. Term-structure theory has as its basis the idea that loans of different maturities should incur different rates of interest. This basis is grounded in reality and allows for a much richer and more realistic theory than that provided by the *yield-to-maturity* (YTM) framework[1]. We first describe some of the basic concepts and notation that we need for studying term-structure models. In these notes we will often assume that there are m compounding periods per year, but it should be clear what changes need to be made for continuous-time models and different compounding conventions. Time can be measured in periods or years, but it should be clear from the context what convention we are using.

Spot Rates: Spot rates are the basic interest rates that define the term structure. Defined on an annual basis, the spot rate, s_t, is the rate of interest charged for lending money from today ($t = 0$) until time t. In particular, this implies that if you lend A dollars for t years[2] today, you will receive $A(1 + s_t/m)^{mt}$ dollars when the t years have elapsed. The *term structure of interest rates* may be defined to constitute the sequence of spot rates, $\{s_k : \ k = 1, \ldots, n\}$, if we have a discrete-time model with n periods. Alternatively, in a continuous-time model the set $\{s_t : \ t \in [0, T]\}$ may be defined to constitute the term-structure. The *spot rate curve* is defined to be a graph of the spot rates plotted against time. In practice, it is usually upwards sloping in which case $s_{t_1} < s_{t_2}$ whenever $t_1 < t_2$.

Discount Factors: As before, there are discount factors corresponding to interest rates, one for each time, t. The discount factor, d_t, for period t is given by

$$d_t \ := \ \frac{1}{(1 + s_t/m)^{mt}}.$$

Using these discount factors we can compute the present value, P, of any *deterministic* cash flow stream, $(x_0, x_1, \ldots, x_n)$. It is given by

$$P \ = \ x_0 + d_1 x_1 + d_2 x_2 + \ \ldots \ + d_n x_n.$$

Example 1 In practice it is quite easy to determine the spot rate by observing the price of U.S. government bonds. Government bonds should be used as they do not bear *default risk* and so the contracted payments are sure to take place. For example the price, P, of a 2-year zero-coupon government bond with face value $\$100$, satisfies $P = 100/(1 + s_2)^2$ where we have assumed an annual compounding convention. ∎

Forward Rates: A *forward rate*, f_{t_1,t_2}, is a rate of interest[3] that is agreed upon today for lending money from dates t_1 to t_2 where t_1 and t_2 are *future* dates. It is easy using arbitrage arguments to compute forward rates given the set of spot interest rates. For example, if we express time in periods, we have that

$$(1 + s_j/m)^j \ = \ (1 + s_i/m)^i \ (1 + f_{i,j}/m)^{j-i}$$

[1]Nor does the YTM framework preclude arbitrage opportunities.
[2]We assume that t is a multiple of $1/m$ both here and in the definition of the discount factor, d_t, above.
[3]It is quoted on an annual basis unless otherwise stated.

where $i < j$.

Forward Discount Factors: We can also discount a cash flow that occurs at time j back to time $i < j$. The correct discount factor is

$$d_{i,j} \;:=\; \frac{1}{(1 + f_{i,j}/m)^{j-i}}$$

where time is measured in periods and there are m periods per year. In particular, the present value at date i of a cashflow, x_j, that occurs at date $j > i$, is given by $d_{i,j}\ x_j$. It is also easy to see that these discount factors satisfy $d_{i,k} = d_{i,j}\ d_{j,k}$ for $i < j < k$ and they are consistent with earlier definitions.

Short Forward Rates: The term structure of interest rates may equivalently be defined to be the set of forward rates. There is no inconsistency in this definition as the forward rates define the spot rates and the spot rates define the forward rates. We also remark that in an n-period model, there are n spot rates and $n(n+1)/2$ forward rates. The set[4] of short forward rates, $\{r_k^f :\ k = 1, \ldots, n\}$, is a particular subset of the forward rates that also defines the term structure. The short forward rates are defined by $r_k^f := f_{k,k+1}$ and may easily be shown to satisfy

$$(1 + s_k)^k \;=\; (1 + r_0^f)(1 + r_1^f)\ldots(1 + r_{k-1}^f)$$

if time is measured in years and we assume $m = 1$.

Term-Structure Explanations

There are three well known hypotheses that are commonly used for explaining the observed term structure of interest rates: the *expectations* hypothesis, the *liquidity* hypothesis and the *market segmentation* hypothesis.

Expectations Hypothesis: The expectations hypothesis states that the forward rates, $f_{i,j}$, are simply the spot rates, s_{j-i}, that are expected to prevail at time i. While this has some intuitive appeal, if the hypothesis was true then the fact that the spot rate curve is almost always upwards sloping would mean (why?) that the market is almost always expecting spot interest rates to rise. This is not the case.

Liquidity Preference Hypothesis: This hypothesis states that investors generally prefer shorter maturity bonds to longer maturity bonds. This is because longer-maturity bonds are generally more sensitive to changes in the general level of interest rates and are therefore riskier. In order to persuade risk-averse individuals to hold these bonds, they need to be sold at a discount, which is equivalent to having higher interest rates at longer maturities.

Market Segmentation Hypothesis: This states that interest rates at date t_1 have nothing to do with interest rates at date t_2 for $t_1 \neq t_2$. The rationale for this is that short-term securities might be of interest to one group of investors, while longer term securities might be of interest to an altogether different group. Since these investors have nothing in common, the markets for short- and long-term securities should be independent of one another and therefore the interest rates that are set by the market forces of supply and demand, should also be independent. This explanation is not very satisfactory and explains very little about the term structures that are observed in practice.

In practice, the term structure is reasonably well explained by a combination of the expectations and liquidity preference hypotheses.

[4] We generally reserve the term "r_t" to denote the short rate prevailing at time t which, in a stochastic model, will not be known until t.

Example 2 Constructing a Zero-Coupon Bond

Two bonds, A and B both mature in ten years time. Bond A has a 7% coupon and currently sells for $\$97$, while bond B has a 9% coupon and currently sells for $\$103$. The face value of both bonds is $\$100$. Compute the price of a ten-year zero-coupon bond that has a face value of $\$100$.

Solution: Consider a portfolio that buys negative (i.e. is short) seven bonds of type B and buys nine of type A. The coupon payments in this portfolio cancel and the terminal value at $t = 10$ is $\$200$. The initial cost is $-7 \times 103 \ + 9 \times 97 = 152$. The cost of a zero with face value equal to $\$100$ is therefore $\$76$. (The 10-year spot rate, s_{10}, is then equal to 2.78%. ∎

Before developing term-structure models, we should first identify some desirable properties that any such model should possess. A model under consideration should

(i) preclude arbitrage possibilities

(ii) be *tractable* in the sense that *relevant* security prices in the model are computable

(iii) provide an adequate fit to *relevant* empirical data.

The models we will develop in this course will, by construction, usually satisfy (i) automatically. We will spend a lot of time showing that they also satisfy (ii) though this will of course depend on the particular security prices we are seeking to compute. We will not focus on (iii) though we will sometimes discuss the issue of *calibration*.

---

2 Review of Martingale Pricing Theory

We now develop some further notation and review some of the important concepts and results of financial engineering that will be used throughout the course. These definitions are initially given in the context of discrete time models. (We will give their continuous-time analogues later in the course.)

Consider a financial market with $N + 1$ securities and true probability measure, P. We assume that the investment horizon is $[0, T]$ and that there are a total of T trading periods. Securities may therefore be purchased or sold at any date t for $t = 0, 1, \ldots, T - 1$.

A trading strategy is a vector, $\theta_t = (\theta_t^{(0)}(\omega), \ldots, \theta_t^{(N)}(\omega))$ of stochastic processes that describes the number of units of each security held just *before* trading at time t, as a function of t and ω. For example, $\theta_t^{(i)}(\omega)$ is the number of units of the ith security held[5] between times $t - 1$ and t in state ω. We will sometimes write $\theta_t^{(i)}$, omitting the explicit dependence on ω. Note that θ_t is known at date $t - 1$.

In order to respect the evolution of information as time elapses, it is necessary that θ_t be a *predictable* stochastic process. In this context, 'predictable' means that θ_t cannot depend on information that is not yet available at time $t - 1$. We also say that θ_t is $\mathcal{F}_{t-1}$-*measurable* where we use $\mathcal{F}_{t-1}$ to denote all the information in the financial market that is know at date $t - 1$.

Definition 1 *The* **value process**, $V_t(\theta)$, *associated with a trading strategy,* θ_t, *is defined by*

$$V_t = \begin{cases} \sum_{i=0}^{N} \theta_1^{(i)} S_0^{(i)} & \text{for } t = 0 \\ \\ \sum_{i=0}^{N} \theta_t^{(i)} S_t^{(i)} & \text{for } t \geq 1. \end{cases}$$

[5] If $\theta_t^{(i)}$ is negative then it corresponds to the number of units sold short.

Definition 2 *A* **self-financing** *trading strategy is a strategy, θ_t, where changes in V_t are due entirely to trading gains or losses, rather than the addition or withdrawal of cash funds. In particular, a self-financing strategy satisfies*

$$V_t \;=\; \sum_{i=0}^{N} \theta_{t+1}^{(i)} S_t^{(i)} \quad \text{for } t = 1, \ldots, T-1.$$

This condition states that the value of a self-financing portfolio just before trading or *re-balancing* is equal to the value of the portfolio just after trading, i.e., no additional funds have been deposited or withdrawn. In particular we see that

$$V_{t+1} - V_t \;=\; \sum_{i=0}^{N} \theta_{t+1}^{(i)} \left(S_{t+1}^{(i)} - S_t^{(i)} \right) \tag{1}$$

and (1) may also be taken as the definition of a self-financing trading strategy.

Definitions (1) and (2) apply to the case when the securities do not pay dividends, coupons or other intermediate cash flows. When one or more of the securities does pay dividends we can easily adapt these definitions. In particular, let $C_t^{(i)}$ be the dividend paid by the i^{th} security just before trading at time t and let $S_t^{(i)}$ be the ex-dividend price. Then we say θ_t is a self-financing trading strategy if

$$V_{t+1} - V_t \;=\; \sum_{i=0}^{N} \theta_{t+1}^{(i)} \left(S_{t+1}^{(i)} + C_{t+1}^{(i)} - S_t^{(i)} \right). \tag{2}$$

Again the interpretation is that there is no addition or withdrawal of cash funds from the portfolio.

Exercise 1 *Derive equation (1). In particular we see that $\Delta V_t = \theta_{t+1}^T \, \Delta S_t$ thereby emphasizing our interpretation of a self-financing trading strategy as one where changes in the value of the portfolio are due entirely to capital gains and losses.*

Definition 3 *A type A arbitrage opportunity is a self-financing trading strategy, θ_t, such that $V_0(\theta) < 0$ and $V_T(\theta) = 0$. Similarly, a type B arbitrage opportunity is a self-financing trading strategy, θ_t, such that $V_0(\theta) = 0$, $V_T(\theta) \geq 0$ and $\mathrm{E}^P[V_T(\theta)] > 0$.*

Definition 4 *A contingent claim, C, is a random variable whose value is known by time T, i.e., C is $\mathcal{F}_T$-measurable.*

Definition 5 *We say that a contingent claim C is* **attainable** *if there exists a self-financing trading strategy, θ_t, whose value process satisfies $V_T = C$. The value of the claim, C, must equal the initial value of the replicating portfolio if there are no arbitrage opportunities available.*

Definition 6 *We say that the market is* **complete** *if every contingent claim is attainable. Otherwise the market is said to be* **incomplete**.

Definition 7 *A* **numeraire** *security is a security with a strictly* positive *price process.*

Definition 8 *Let N_t be the time t price of a chosen numeraire security and S_t the time t price of any other security. Then S_t/N_t is the* **deflated** *security price.*

Remark 1 *Note that the deflated price of the numeraire security is identically 1.*

Remark 2 *The default numeraire is the* cash account *so whenever we are using a different numeraire this will be stated explicitly.*

Definition 9 *An* **Equivalent Martingale Measure** *(EMM), Q, is a probability measure that is equivalent*[6] *to P under which all deflated security* gains *processes are martingales. Note that an EMM, Q, is specific to the chosen numeraire. In particular, an an EMM-numeraire pair, (Q, N), satisfies*

$$\frac{S_t}{N_t} \;=\; \mathrm{E}_t^Q\left[\frac{S_{t+s}}{N_{t+s}}\right]$$

for any security price process, S_t, that does not pay dividends between s and $s+t$. For a security that pays intermediate cash-flows we have

$$\frac{S_t}{N_t} \;=\; \mathrm{E}_t^Q\left[\sum_{i=t+1}^{t+s}\frac{C_i}{N_i} \;+\; \frac{S_{t+s}}{N_{t+s}}\right]$$

where C_i is the dividend paid by the security at time i.

We have the following results.

Theorem 1 *There is no arbitrage if and only if there exists an EMM, Q.*

A consequence of Theorem 1 is that in the absence of arbitrage, the *deflated* value process, V_t/N_t, of any self-financing trading strategy is a Q-martingale. This implies that the deflated price of any attainable security can be computed as the Q-expectation of the terminal deflated value of the security.

Theorem 2 *Assume there exists a security with strictly positive price process and that there are no arbitrage opportunities. Then the market is complete if and only if there exists exactly one risk-neutral martingale measure, Q.*

---

3 Modelling Philosophy for Term-Structure Models

The modelling philosophy for term-structure models is somewhat different to the modelling philosophy for equity models. In the latter case, stock price dynamics are usually specified under the physical probability measure, P, before their dynamics under an EMM, Q, are determined. For example, in the binomial Black-Scholes framework a unique Q is easily determined after the P-dynamics of the stock-price are given. Moreover, it is easy to check that the model does not allow any arbitrage: we just need $d < R < u$.

In contrast, with term-structure models we often assume that zero-coupon bonds of every maturity exists and it is not always easy to directly specify their P-dynamics in an arbitrage-free manner that it is economically satisfactory. For example, in a T-period binomial model there are $O(T)$ zero-coupon bond prices that we need to specify at each node. Checking that the model is arbitrage-free and that bond price processes have suitable properties (e.g. implied interest rates are always non-negative) can be a cumbersome task. As a result, we usually work with term structure models where we directly specify an EMM, Q, and price all securities using this EMM. By construction, such a model is arbitrage free. Moreover, by leaving some parameters initially unspecified (e.g. short-rate values at nodes or Q-probabilities along branches in a lattice model) we can then *calibrate* them so that security prices in the model coincide[7] with security prices observed in the market.

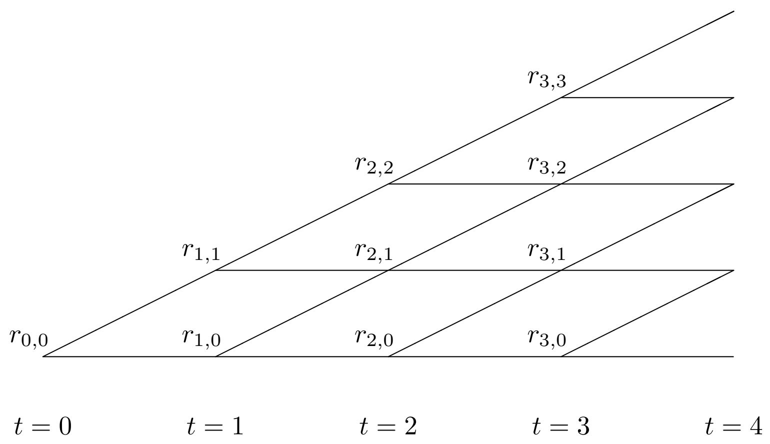
[6] We say P and Q are equivalent if $P(A) = 0 \Leftrightarrow Q(A) = 0$.
[7] The problem of calibrating models to market data can also make use of the true probability measure, P.

4 Binomial-Lattice Models

We begin with binomial-lattice models of the short rate. These models may be viewed as models in their own right or as approximations to more sophisticated models. We will take the latter approach later in these notes when we will explicitly construct trinomial models as approximations to continuous-time short-rate models.

Constructing an Arbitrage-Free Lattice

Consider the binomial lattice below where we specify the short rate, $r_{i,j}$, that will apply for the single period beginning at node $N(i,j)$. This means for example that if $\$1$ is deposited in the cash account at $t = i$, state j, (i.e. node $N(i,j)$, then this deposit will be worth $\$(1 + r_{i,j})$ at time $t + 1$ regardless of the successor node to $N(i,j)$.



We use martingale pricing on this lattice to compute security prices. For example, if $S_i(j)$ is the value of a security at time i and state j, then we *insist* that

$$S_i(j) \;=\; \frac{1}{1 + r_{i,j}} \left[q_u \; S_{i+1}(j+1) + q_d \; S_{i+1}(j) \right].$$

where q_u and q_d are the probabilities of up- and down-moves, respectively. As mentioned earlier, such a model is arbitrage-free by construction.

Computing the Term-Structure from the Lattice

It is easy to compute the price of a zero-coupon bond once the EMM, Q, (with the cash account as numeraire) and the short-rate lattice are specified. In the short rate-lattice below (where the short rate increases by a factor of $u = 1.25$ or decreases by a factor of $d = .9$ in each period), we assume that the Q-probability of each branch is .5 and node-independent. We can then use martingale-pricing to compute the prices of zero-coupon bonds.

Short Rate Lattice					
					0.183
				0.146	0.132
			0.117	0.105	0.095
		0.094	0.084	0.076	0.068
	0.075	0.068	0.061	0.055	0.049
0.060	0.054	0.049	0.044	0.039	0.035
t=0	t=1	t=2	t=3	t=4	t=5

Example 3 **(Pricing a Zero-Coupon Bond)**

We compute the price of a 4-period zero-coupon bond with face value 100 that expires at $t = 4$. Assuming the short-rate lattice is as given above, we see, for example, that the bond price at node $(2, 2)$ is given by

$$83.08 \;=\; \frac{1}{1 + .094}\left[\frac{1}{2}\,89.51 + \frac{1}{2}\,92.22\right].$$

Iterating backwards, we find that the zero-coupon bond is worth 77.22 at $t = 0$.

4-Year Zero				
				100.00
			89.51	100.00
		83.08	92.22	100.00
	79.27	87.35	94.27	100.00
77.22	84.43	90.64	95.81	100.00
t=0	t=1	t=2	t=3	t=4

∎

Note that given the price of the 4-period zero-coupon bond, we can now find the 4-period spot rate, s_4. It satisfies $77.22 = 1/(1 + s_4)^4$ if we quote spot rates on a per-period basis. In this manner we can construct the entire term-structure by evaluating zero-coupon bond prices for all maturities.

Pricing Interest Rate Derivatives

We now introduce and price several interest-rate derivatives using the straightforward martingale pricing methodology. (After we derive the *forward equations* in a later section, we will see an even easier and more efficient[8] method for pricing derivatives.) The following examples will be based on the short-rate lattice and the corresponding zero-coupon bond of Example 3.

Example 4 **(Pricing a European Call Option on a Zero-Coupon Bond)**

We want to compute the price of a European call option on the zero-coupon bond of Example 3 that expires at $t = 2$ and has strike $\$84$. The option price of $\$2.97$ is computed by backwards induction on the lattice below.

[8] Of course the forward equations are themselves derived using martingale pricing so they really add nothing new to the theory.

		0.00	European Call Option
	1.56	3.35	Strike = $84
2.97	4.74	6.64	
t=0	t=1	t=2	

Example 5 **(Pricing an American Put Option on a Zero-Coupon Bond)**

We want to compute the price of an American put option on the same zero-coupon bond. The expiration date is $t = 3$ and the strike is $88. Again the price is computed by backwards induction on the lattice below, where the maximum of the *continuation value* and exercise value is equal to the option value at each node.

			0.00	
		4.92	0.00	American Put Option
	8.73	0.65	0.00	Strike = $88
10.78	3.57	0.00	0.00	
t=0	t=1	t=2	t=3	

Futures Contracts on Bonds

Let F_k be the date k price of a futures contract written on a particular underlying security in a complete[9] market model. We assume that the contract expires after n periods and we let S_k denote the time k price of the security. Then we know that $F_n = S_n$, i.e., at expiration the futures price and the security price must coincide. We can compute the futures price at $t = n - 1$ by recalling that anytime we enter a futures contract, the initial value of the contract is 0. Therefore the futures price, F_{n-1}, at date $t = n - 1$ must satisfy

$$0 = \mathrm{E}_{n-1}^{Q}\left[\frac{F_n - F_{n-1}}{B_n}\right]$$

where we will assume that the numeraire security is the cash account[10] with value B_n at date n. Since B_n and F_{n-1} are both known at date $t = n - 1$, we therefore have $F_{n-1} = \mathrm{E}_{n-1}^{Q}[F_n]$. By the same argument, we also have more generally that $F_k = \mathrm{E}_k^Q[F_{k+1}]$ for $0 \leq k < n$. We can then use the law of iterated expectations to see that $F_0 = \mathrm{E}_0^Q[F_n]$, implying in particular that the futures price process is a martingale. Since $F_n = S_n$ we have

$$\boxed{F_0 = \mathrm{E}_0^Q[S_n].} \tag{3}$$

Remark 3 *Note that the above argument holds regardless of whether or not the underlying security pays dividends or coupons as long as the settlement price, S_n, is ex-dividend. In particular, we can use (3) to price futures on zero-coupon and coupon bearing bonds.*

[9] A complete market is assumed so that we can uniquely price any security and more generally, the futures price process. We could, however, have assumed markets were incomplete and still compute the futures price process as long as certain securities were replicable.

[10] We assume without loss of generality that B_n is the value of the cash account at $t = n$ when $1 was deposited there at $t = 0$.

Remark 4 *It is important to note that (3) applies only when the EMM, Q, is the EMM corresponding to when the cash account is numeraire.*

Forward Contracts on Bonds

Now let us consider the date 0 price, G_0, of a forward contract for delivery of the same security at the same date, $t = n$. We recall that G_0 is chosen in such a way that the contract is initially worth zero. In particular, martingale pricing implies

$$0 \quad = \quad \mathrm{E}_0^Q \left[\frac{S_n - G_0}{B_n} \right].$$

Rearranging terms and using the fact that G_0 is known at date $t = 0$ we obtain

$$G_0 = \frac{\mathrm{E}_0^Q \left[S_n / B_n \right]}{\mathrm{E}_0^Q \left[1 / B_n \right]} \tag{4}$$

Remark 5 *If B_n and S_n are Q-independent, then $G_0 = F_0$. In particular, if interest rates are deterministic, we have $G_0 = F_0$.*

Remark 6 *If the underlying security does not pay dividends or coupons (storage costs may be viewed as negative dividends), then we obtain $G_0 = S_0 / \mathrm{E}_0^Q [1/B_n] = S_0 / d(0, n)$.*

Example 6 (Pricing a Forward Contract on a Coupon-Bearing Bond)

We now price a forward contract for delivery at $t = 4$ of a 2-year 10% coupon-bearing bond where we assume that delivery takes place just after a coupon has been paid. In the lattice below we use backwards induction to compute the $t = 4$ price (ex-coupon) of the bond. We then use (4) to price the contract, with the numerator given by the value at node $(0, 0)$ of the lattice and the denominator equal to the 4-year discount factor. Note that between $t = 0$ and $t = 4$ in the lattice below, coupons are (correctly) ignored.

```
                                                         110.00
                                               102.98    110.00
                                     | 91.66   107.19    110.00
                             85.08   | 98.44   110.46    110.00
                    81.53    93.27   |103.83   112.96    110.00
           79.99    90.45    99.85   |108.00   114.84    110.00
  79.83    89.24    97.67   104.99   |111.16   116.24    110.00

    t=0      t=1      t=2      t=3      t=4      t=5       t=6

  Forward Price = 100*79.83 / 77.22  = 103.38
```

∎

Example 7 (Pricing a Futures Contract on a Coupon-Bearing Bond)

The $t = 0$ price of a futures contract expiring at $t = 4$ on the same coupon-bearing bond is given at node $(0, 0)$ in the lattice below. This lattice is constructed using (3). Note that the forward and futures price are close but not equal.

Futures Price				
				91.66
			95.05	98.44
		98.09	101.14	103.83
	100.81	103.52	105.91	108.00
103.22	105.64	107.75	109.58	111.16
t=0	t=1	t=2	t=3	t=4

∎

Example 8 **(Pricing a Caplet)**

A *caplet* is similar to a European call option on the interest rate r_t. They are usually settled *in arrears* but they may also be settled in advance. If the maturity is τ and the strike is c then the payoff of a caplet (settled in arrears) at time τ is $(r_{\tau-1} - c)^+$. That is, the caplet is a call option on the short rate prevailing at time $\tau - 1$, settled at time τ.

In the lattice below we price a caplet that expires at $t = 6$ with strike $= 2\%$. Note, however, that it is easier to record the time 6 cash flows at their time 5 predecessor nodes, discounting them appropriately. For example, the value at node $N(4, 0)$ is given by

$$0.021 \;=\; \frac{1}{1.039}\left[\frac{1}{2}\frac{\max(0, .049 - .02)}{1.049} + \frac{1}{2}\frac{\max(0, .035 - .02)}{1.035}\right].$$

Caplet Strike = 2%					
					0.138
				0.103	0.099
			0.080	0.076	0.068
		0.064	0.059	0.053	0.045
	0.052	0.047	0.041	0.035	0.028
0.042	0.038	0.032	0.026	0.021	0.015
t=0	t=1	t=2	t=3	t=4	t=5

∎

Remark 7 *In practice caplets are usually based on* LIBOR *(London Inter-Bank Offered Rates).*

Caps: A *cap* is a string of caplets, one for each time period in a fixed interval of time and with each caplet having the same strike price, c.

Floorlets: A *floorlet* is similar to a caplet, except it has payoff $(c - r_{\tau-1})^+$ and is usually settled in arrears at time τ.

Floors: A *floor* is a string of floorlets, one for each time period in a fixed interval of time and with each floorlet having the same strike price, c.

The Forward Equations

Definition 10 *Let $P_e(i,j)$ denote the time 0 price of a security that pays $\$1$ at time i, state j and 0 at every other time and state. We say such a security is an* **elementary security** *and we refer to the $P_e(i,j)$'s as* state prices.

It is easily seen that the elementary security prices satisfy the **forward equations**:

$$\begin{aligned} P_e(k+1,s) &= \frac{P_e(k,s-1)}{2(1+r_{k,s-1})} + \frac{P_e(k,s)}{2(1+r_{k,s})}, \quad 0 < s < k+1 \\ P_e(k+1,0) &= \frac{1}{2}\frac{P_e(k,0)}{(1+r_{k,0})} \\ P_e(k+1,k+1) &= \frac{1}{2}\frac{P_e(k,k)}{(1+r_{k,k})} \end{aligned}$$

Exercise 2 *Derive the forward equations.*

Since we know $P_e(0,0) = 1$ we can use the above equations to evaluate all the state prices by working *forward* from node $N(0,0)$. Working with state prices has a number of advantages:

1. Once you compute the state prices they can be stored and used repeatedly for pricing derivative securities without needing to use backwards iteration each time.
2. In particular, the term structure can be computed using $O(T^2)$ operations by first computing the state prices. In contrast, computing the term structure by first computing all zero-coupon bond prices using backwards iteration takes $O(T^3)$ operations. While this makes little difference given today's computing power, writing code to price derivatives in lattice models is often made easier when we price with the state prices.
3. They can be useful for calibrating lattice models to observed market data.

It should be noted that when pricing securities with an early exercise feature, state prices are of little help and it is still necessary to use backwards iteration.

Hedging in the Binomial Lattice

It is very easy to hedge in a binomial-lattice model. For example, if we wish to hedge a payoff that occurs at $t = s$, then we can use any two securities (that do not expire before $t = s$) to do this. In particular, we could use the cash account and a zero-coupon bond with maturity $t > s$ as our hedging instruments. At each node and time, we simply choose our holdings in the cash account and zero-coupon bond so that their combined value one period later exactly matches the value of the payoff to be hedged. This of course is exactly analogous to hedging in the classic binomial model for stocks.

That we can hedge with just two securities betrays a weakness of the binomial lattice model and short-rate models more generally. In practice, the value of an interest-rate derivative will be sensitive to changes in the entire term-structure and it is unlikely that such changes can be perfectly hedged using just two securities. It is possible in some circumstances, however, that good *approximate* hedges can be found using just two securities. In the example below, we hedge using elementary securities.

Example 9 (a) Consider the short rate lattice of Figure 4.1 where interest rates are expressed as percentages. You may assume the risk-neutral probabilities are given by $q = 1 - q = 1/2$ at each node.

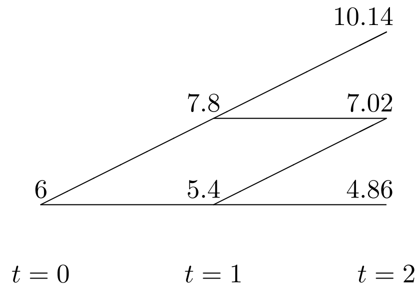


Figure 4.1

Use the forward equation to compute the elementary prices at times $t = 0$, $t = 1$ and $t = 2$.

(b) Consider the different short rate lattice of Figure 4.2 where again you may assume $q = 1 - q = 1/2$. The short rates *and* elementary prices are given at each node. Using only the elementary prices, find the price of a zero-coupon bond with face value \$100 that matures at $t = 3$.

(c) Compute the price of a European call option on the zero-coupon bond of part (b) with strike $= \$93$ and expiration date $t = 2$.

(d) Again referring to Figure 4.2, suppose you have an obligation whose value today is \$100 and whose value at node $(1, 1)$ (i.e. when $r = 7.2\%$) is \$95. Explain how you would immunize this obligation using date 1 elementary securities?

(e) Consider a forward-start swap that begins at $t = 1$ and ends at $t = 3$. The notional principal is \$1 million, the fixed rate in the swap is 5%, and payments at $t = i$ (i=2,3) are based on the fixed rate minus the floating rate that prevailed at $t = i - 1$. Compute the value of this forward swap.

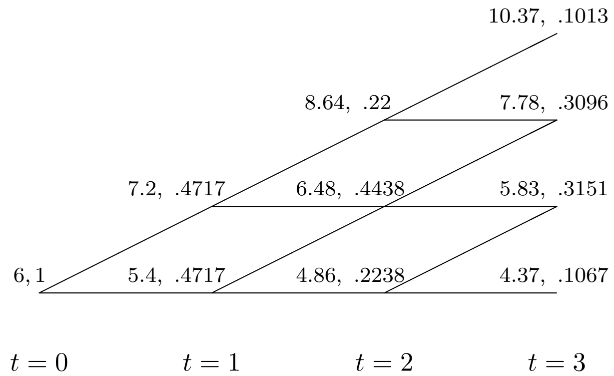


Figure 4.2

Solution

(a) $P_e(0, 0) = 1$, $P_e(1, 0) = .4717$, $P_e(1, 1) = .4717$, $P_e(2, 0) = 0.2238$, $P_e(2, 1) = 0.4426$, $P_e(2, 2) = 0.2188$.

(b) $(.1013 + .3096 + .3151 + .1067) \times 100 = 83.27$.

(c) The value of the zero coupon bond at nodes $(2, 0)$, $(2, 1)$ and $(2, 2)$ are 95.3652, 93.9144 and 92.0471 respectively. This means the payoffs of the option are 2.3652, .9144 and 0 respectively. Now use the time 2 elementary prices to find that the option price equals $2.3652 \times .2238 + .9144 \times .4438 = .9351$.

(d) First find the value, x, of the obligation at node $(1, 0)$. It must be that x satisfies

$$100 = \frac{1}{2}\frac{95}{1.06} + \frac{1}{2}\frac{x}{1.06}$$

which implies $x = 117$. So to hedge the obligation, buy 95 node $(1, 1)$ state securities and 117 node $(1, 0)$ state securities. The cost of this is $117 \times .4717 + 95 \times .4717 = 100$, of course!

(e) The value of the swap today is (why?)

$$(.022/1.072 \times .4717) + (.004/1.054 \times .4717) + (.0364/1.0864 \times .22) + (.0148/1.0648 \times .4438) - (.0014/1.0486 \times .2238) = 0.0247 \text{ million.}$$

∎

---

5 Ho-Lee, Black-Derman-Toy and their Calibration

If we want to use a lattice model in practice then we need to *calibrate* the model to market data. The most basic requirement is that the term-structure (equivalently zero-coupon bond prices) in the market should also match the term-structure of the lattice model. We now show how this can be achieved in the context of two specific models, the Ho-Lee model and the Black-Derman-Toy (BDT) model. More generally, we could also calibrate the lattice to interest-rate volatilities and other security prices.

We will assume[11] again that the Q-probabilities are equal to $.5$ on each branch and are node independent. First we describe the Ho-Lee and Black-Derman-Toy models.

The Ho-Lee Model

The Ho-Lee model assumes that the interest rate at node $N(i,j)$ is given by

$$r_{i,j} = a_i + b_i j.$$

Note that a_i is a *drift* parameter while b_i is a volatility parameter. In particular, the standard deviation of the short-rate at node $N(i,j)$ is equal to $b_i/2$. The continuous-time version of the Ho-Lee model assumes that the short-rate dynamics satisfy

$$dr_t = \mu_t \, dt + \sigma_t \, dB_t.$$

The principal strength of the Ho-Lee model is its tractability.

Exercise 3 *What are the obvious weaknesses of the Ho-Lee model?*

The Black-Derman-Toy Model

The BDT model assumes that the interest rate at node $N(i,j)$ is given by

$$r_{i,j} = a_i e^{b_i s}.$$

Note that $\log(a_i)$ is a *drift* parameter while b_i is a volatility parameter (for $\log(r)$). The continuous-time version of the BDT model assumes that the short-rate has dynamics of the form

$$dY_t = \left(a_t + \frac{1}{\sigma_t}\frac{\partial \sigma}{\partial t} Y_t\right) dt + \sigma_t \, dB_t$$

where $Y_t := \log(r_t)$.

Exercise 4 *What are some of the strengths and weaknesses of the BDT model?*

[11] Later when we approximate diffusion models with trinomial lattices we will not set the Q-probabilities so arbitrarily.

Calibration to the Observed Term-Structure

Consider an n-period lattice and let $(s_1,\ldots,s_n)$ be the term-structure observed in the market place assuming that spot rates are compounded per period. Suppose we want to calibrate the BDT model to the observed term-structure. Assuming $b_i = b$, a constant for all i, we first note that

$$\begin{aligned}\frac{1}{(1+s_i)^i} &= \sum_{j=0}^{i} P_e(i,j) \\ &= \frac{P_e(i-1,0)}{2(1+a_{i-1})} + \sum_{j=1}^{i-1}\left(\frac{P_e(i-1,j)}{2(1+a_{i-1}e^{bj})} + \frac{P_e(i-1,j-1)}{2(1+a_{i-1}e^{b(j-1)})}\right) + \frac{P_e(i-1,i-1)}{2(1+a_{i-1}e^{b(i-1)})}. \qquad (5)\end{aligned}$$

Note that Equation (5) can now be used to solve iteratively for the a_i's as follows:

- Set $i=1$ in (5) and note that $P_e(0,0)=1$ to see that $a_0 = s_0$.
- Now use the forward equations to find $P_e(1,0)$ and $P_e(1,1)$.
- Now set $i=2$ in (5) and solve for a_1.
- Continue to iterate forward until all a_i's have been found.

By construction, this algorithm will match the observed term structure to the term structure in the lattice.

Example 10 (Pricing a Payer Swaption in a BDT Model)

We would like to price a $2-8$ *payer swaption* in a BDT model that has been calibrated to match the observed term-structure of interest rates in the market place. The $2-8$ terminology means that the swaption is an option that expires in 2 months to enter an 8-month swap. The swap is settled in arrears so that payments would take place in months 3 through 10 based on the prevailing short-rate of the previous months. The "payer" feature of the option means that if the option is exercised, the exerciser "pays fixed and receives floating". (The owner of a *receiver swaption* would "receive fixed and pay floating".) For this problem the fixed rate is set at 11.65%.

We use a 10-period lattice for our BDT model, thereby assuming that a single period corresponds to 1 month. The spot-rate curve in the market place has been observed to be

$$\begin{aligned} s_1 = 7.3,\quad s_2 = 7.62,\quad s_3 = 8.1,\quad s_4 = 8.45,\quad s_5 = 9.2 \\ s_6 = 9.64,\quad s_7 = 10.12,\quad s_8 = 10.45,\quad s_9 = 10.75,\quad s_{10} = 11.22 \end{aligned}$$

with a monthly compounding convention. We will also assume[12] that $b_i = b = .005$ for all i.

The first step is to choose the a_i's so that the term-structure in the lattice matches the observed term-structure (spot-rate curve) given above. For a 10-period problem, this is easy to do in *Excel* using the *Solver* add-in. For problems with many periods, however, it would be preferable to use the algorithm outlined above. The short-rate lattice below gives the values of $a_i,\quad i=0,\ldots,9$, so that the model spot-rate curve matches the market spot-rate curve. (Note also that the short-rates at each node $N(i,j)$ do indeed satisfy $r_{i,j} = a_i e^{.005j}$.)

We are now in a position to price the swaption, assuming a notional principal of \$1. Let S_2 denote the value of the swap at $t=2$. We can compute S_2 by discounting the cash-flows back from $t=10$ to $t=2$. It is important to note that it is much easier to record time t cash flows (for $t=3,\ldots,10$) at their predecessor nodes at time $t-1$, discounting them appropriately. This is why in the swaption lattice above, there are no payments recorded at $t=10$ despite the fact that payments do actually take place then.

Exercise 5 *Why is it more convenient to record those cashflows at their predecessor nodes?*

[12] We could, however, also calibrate various term-structure volatilities and / or other security prices by not fixing the b_i's in advance.

The option is then exercised if and only if $S_2 > 0$. In particular, the value of the swaption at maturity $(t = 2)$ is $\max(0, S_2)$. This is then discounted backwards to $t = 0$ to find[13] a swaption value today of $.0013$.

Short-Rate Lattice

										15.90
									13.45	15.82
								13.01	13.38	15.74
							13.24	12.94	13.32	15.66
						12.02	13.18	12.88	13.25	15.58
					12.38	11.96	13.11	12.81	13.18	15.50
				9.58	12.31	11.90	13.05	12.75	13.12	15.43
			9.11	9.53	12.25	11.84	12.98	12.69	13.05	15.35
		7.96	9.07	9.48	12.19	11.78	12.92	12.62	12.99	15.27
	7.30	7.92	9.02	9.44	12.13	11.72	12.85	12.56	12.92	15.20
Year	0	1	2	3	4	5	6	7	8	9
Spot	7.3	7.62	8.1	8.45	9.2	9.64	10.12	10.45	10.75	11.22
a	7.30	7.92	9.02	9.44	12.13	11.72	12.85	12.56	12.92	15.20

Pricing a 2-8 Payer Swaption

										0.0366
									0.0479	0.0360
								0.0539	0.0467	0.0353
							0.0610	0.0523	0.0456	0.0347
						0.0568	0.0590	0.0508	0.0444	0.0340
					0.0560	0.0546	0.0571	0.0492	0.0433	0.0334
				|0.0311	0.0535	0.0524	0.0552	0.0477	0.0422	0.0327
			0.0040	|0.0284	0.0511	0.0502	0.0533	0.0461	0.0411	0.0321
		0.0024	0.0011	|0.0257	0.0486	0.0480	0.0514	0.0446	0.0399	0.0314
	0.0013	0.0005	0.0000	|0.0230	0.0461	0.0458	0.0495	0.0431	0.0388	0.0308
Year	0	1	2	3	4	5	6	7	8	9

∎

A related derivative instrument that is commonly traded is a *Bermudan swaption* (payer or receiver). This is the same as a swaption except now the option can be exercised at any of a predetermined set of times, $\mathcal{T} = (t_1, \ldots, t_m)$.

[13] An exercise in Assignment 1 discusses whether or not the BDT model together with our method of calibration, is a good approach to pricing swaptions.

6 Trinomial Lattice Approximations for Short-Rate Models

We now briefly describe how to approximate one-dimensional models of the short-rate with a trinomial lattice. We prefer to use a trinomial lattice as it is considerably easier to *match moments* in a regular trinomial lattice than it is in a binomial lattice. Our method of constructing the trinomial lattice will follow[14] the method of Hull and White (1993) but it should be mentioned that there are generally many ways[15] of doing this. Derivatives pricing proceeds using backwards induction in the usual straightforward manner. Alternatively, forward equations can be used to compute state prices which in turn may be used to price securities.

It is worth mentioning that computing security prices through lattice approximations is equivalent to using an explicit finite-difference scheme for solving the pricing PDE. The difficulties associated with the stability of finite-difference schemes therefore also apply to lattice approximations. In particular, it is important to choose the right balance between the length of a time step, Δt, and the size of a state step, Δx, when constructing lattices as approximations to diffusions.

Finally, we adopt the usual term-structure modelling philosophy of modelling directly under an EMM, Q, which is chosen by calibrating parameters to observed market prices. We therefore assume that the short-rate dynamics are given by

$$dr_t \;=\; \alpha(t, r_t)\; dt + \beta(t, r_t)\; dW_t \tag{6}$$

where W_t is a Q-Brownian motion, and that we wish to construct a trinomial lattice that will approximate these dynamics. Towards this end, consider the trinomial lattice in Figure 6.1 which may be interpreted as follows:

1. We use $r_{i,j}$ to denote the short-rate that applies at node $N(i, j)$ (time $= i$, state $= j$) for borrowing or lending during the period that immediately follows. At each node, there are three successor nodes that correspond to an up-move, central-move and down-move, respectively. This therefore implies that there are a total of $2n + 1$ nodes or states at time $t = n$. Unlike our convention with binomial lattices, we assume the states at $t = n$ run from $-n$ to $+n$. Therefore the short-rate is given by $r_{n,-n}$ at the lowermost node when $t = n$, and by $r_{n,n}$ at the uppermost node. The central node's short-rate is then $r_{n,0}$.

2. There are 3 risk-neutral probabilities that we need to specify at each node. We will use $q^u_{i,j}$, $q^c_{i,j}$ and $q^d_{i,j}$ to denote the probabilities of an up-move, central-move and down-move, respectively, at node $N(i, j)$. More generally, we will use $q_{i,j}(k)$ to denote the probability of going from node $N(i, j)$ to node $N(i + 1, k)$. For example, we therefore have $q_{i,j}(j - 1) = q^d_{i,j}$.

3. We assume that in any period the short-rate either stays the same, moves up by a fixed amount, Δr, or moves down by the same fixed amount, Δr. This fixed quantity, Δr, does not vary with time.

One of the problems with constructing a lattice in this manner is that the spread of possible short rates for large t will be much greater than required. A related problem is that it may not be possible to choose the $q_{i,j}(k)$'s at some of the extreme nodes in a manner that is consistent with (6).

Exercise 6 *Can you guess why it might indeed be difficult to choose the $q_{i,j}(k)$'s at some of the extreme nodes in a manner that is consistent with (6)?*

[14]This method is also described in *Interest Rate Models: An Introduction*, by Cairns.

[15]See, for example, Appendix C of *Interest Rate Models: Theory and Practice*, by Brigo and Mercurio for an alternative construction.

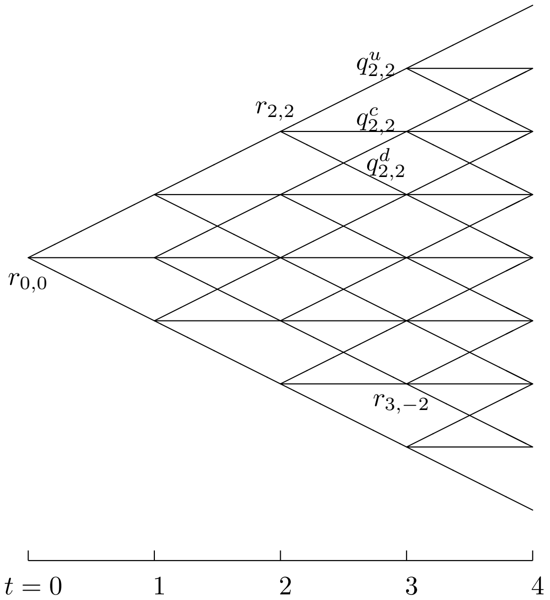


Figure 6.1

For this reason, Hull and White (1993) suggested using a restricted lattice of the type depicted in Figure 6.2. Note that at the extreme nodes in the upper part of the lattice we no longer allow up-moves to take place. Instead we allow for an extra down-move where the short-rate falls by $2\Delta r$. Similarly at the extreme nodes in the lower part of the lattice, we no longer allow down-moves to take place and instead allow for an extra up-move where the short-rate increases by $2\Delta r$.

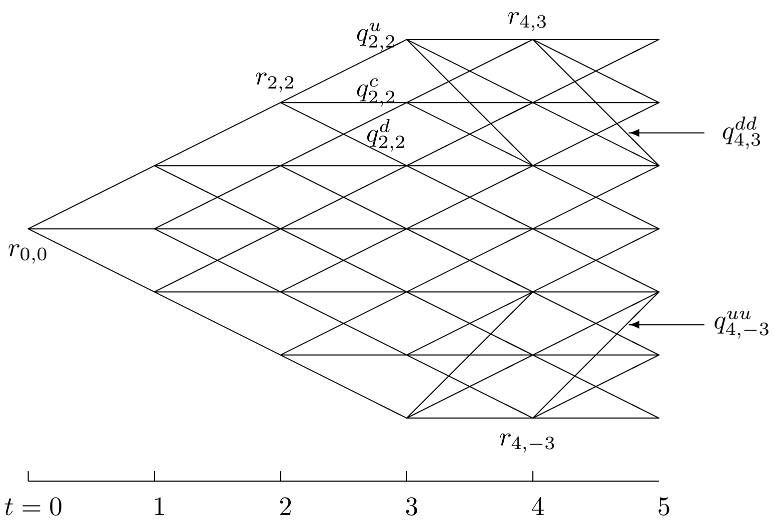


Figure 6.2

We now describe how to construct such a restricted lattice that is consistent with (6). The idea is simply to choose the probabilities, $q_{i,j}(k)$, in such a way as to approximately match the first and second methods of changes in the short rate. That is, at node $N(i,j)$ we choose $Q_{i,j} := (q^{uu}_{i,j}, q^u_{i,j}, q^c_{i,j}, q^d_{i,j}, q^{dd}_{i,j})$ so as to satisfy

$$\begin{aligned} \mathrm{E}^Q\left[r(i+1) - r(i) | \text{state } = (i,j)\right] &\approx \alpha(i, r_{i,j})\Delta t \\ &= q^{uu}_{i,j} \times 2\Delta r \; + \; q^u_{i,j} \times \Delta r \; + \; q^d_{i,j} \times (-\Delta r) \; + \; q^{dd}_{i,j} \times (-2\Delta r) \quad (7) \end{aligned}$$

$$\mathrm{E}^Q\left[(r(i+1) - r(i))^2 | \text{state } = (i,j)\right] \approx \alpha(i, r_{i,j})^2 \Delta t^2 + \beta(i, r_{i,j})^2 \Delta t$$

$$= \quad q^{uu}_{i,j} \times 4\Delta r^2 \ + \ q^{u}_{i,j} \times \Delta r^2 \ + \ q^{d}_{i,j} \times \Delta r^2 \ + \ q^{dd}_{i,j} \times 4\Delta r^2 \qquad (8)$$

$$1 \quad = \quad q^{uu}_{i,j} + q^{u}_{i,j} + q^{c}_{i,j} + q^{d}_{i,j} + q^{dd}_{i,j} \quad \text{and each} \quad q^{k}_{i,j} \geq 0 \qquad (9)$$

where Δt is length of a time period in the lattice. Note that at any node only three of the $q^k_{i,j}$'s will actually be positive, with the particular three depending on whether or not we are at a regular node (up- and down-moves possible), an extreme upper node, or an extreme lower node.

Remark 8 *If we cannot find a solution to (7), (8) and (9) at a regular node (i.e. one of the q's is negative) then it will be necessary to designate this node as an extreme upper node or as an extreme lower node, depending on the nature of the problem. In particular, we do not necessarily have control over the nodes that are designated as extreme upper or extreme lower nodes.*

Transforming the Diffusion

When constructing lattice approximations, it is generally desirable for computational purposes to approximate a diffusion with a constant[16] volatility coefficient. To achieve this, we define $X_t := f(t, r_t)$ where for t fixed, $f(t, \cdot)$ is monotonic in r. If $f(\cdot, \cdot)$ is sufficiently smooth then we may apply Itô's Lemma to obtain

$$dX_t \ = \ \left(\frac{\partial f}{\partial t} + \alpha(t, r_t) \frac{\partial f}{\partial r} + \frac{1}{2} \beta(t, r_t)^2 \frac{\partial^2 f}{\partial r^2} \right) \ dt \ + \ \beta(t, r_t) \frac{\partial f}{\partial r} \ dW_t \qquad (10)$$

If we can therefore find a function, $f(\cdot, \cdot)$, such that

$$\frac{\partial f}{\partial r} \ = \ \frac{1}{\beta(t, r_t)}$$

we will have constructed a diffusion with unit volatility, as desired.

Exercise 7 *What would you take $f(\cdot, \cdot)$ to be in the CIR model where*

$$dr_t \ = \ \alpha(\mu - r_t) \ dt \ + \ \sigma \sqrt{r_t} \ dW_t \ ; \qquad r_0 > 0.$$

The Algorithm

The lattice approximation then proceeds as follows:

1. Find a function $f(\cdot, \cdot)$ with the properties given above and define $X_t = f(t, r_t)$.
2. Use Itô's Lemma to find the Q-dynamics of X_t.
3. Construct a lattice to approximate the dynamics of X_t. This requires us to solve the equations that are analogous to (7), (8) and (9). We can find closed form expressions for these solutions but see Remark 8 above.
4. Price securities using our approximating lattice. Note that since $f(t, \cdot)$ is monotonic we can invert it to find the appropriate short rate at any node, $N(i, j)$.

Remark 9 *For reasons concerning numerical stability, we generally choose $\Delta x = \sqrt{k \Delta t}$ for some $k > 0$, when the diffusion has unit volatility. In practice it is common to choose $k \approx 3$.*

[16] Without loss of generality, we will assume that we want to approximate a diffusion with unit volatility.

Pricing Derivative Securities

As usual, derivative securities can be priced using backwards induction. In particular, if we use $V(i,j)$ to denote the ex-dividend price of a security at node $N(i,j)$ and $C(i,j)$ its dividend payout at that node, then we have

$$V(i,j) \;=\; \frac{1}{1+r_{i,j}} \sum_{k=-n}^{n} q_{i,j}(k)\left[C(i+1,k)+V(i+1,k)\right].$$

American options may be priced in a similar manner except at each node we must also decide whether or not it is optimal to exercise at that point. Forward equations for the trinomial lattice may also be derived and they can then be solved forwards for the state prices.

Exercise 8 *How would you consider the general issue of calibration in the context of using lattice models to approximate continuous-time diffusion models?*

Approximating Multi-Dimensional Diffusions with Lattices

It is also possible to approximate multi-dimensional processes with lattices though the constructions do become more involved. Hull and White (1994) describe methods for doing this in the context of 2-factor models of the short-rate. See also Appendix C of Brigo and Mercurio's *Interest Rate Models: Theory and Practice* for more in this regard.