CHAPTER 20

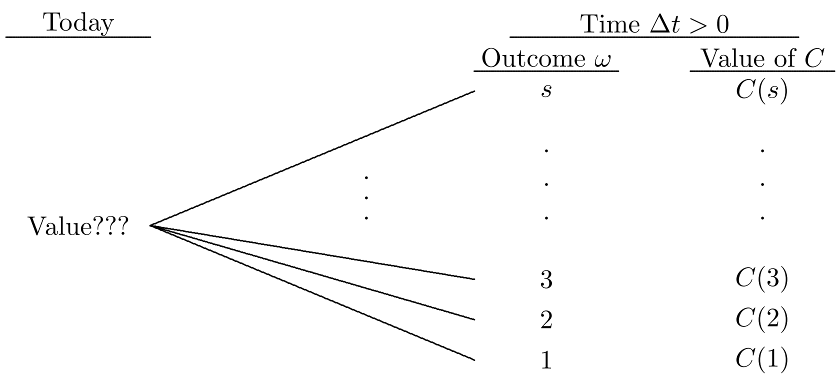
ARBITRAGE-FREE LATTICE MODELS

Here we develop a framework for lattice based arbitrage-free valuation models with stochastic interest rates. We begin with a one period model.

§1. The Basic One-Period Setup

Basic Axioms. Suppose that $\Omega = \{1, 2, \dots, s\}$ represents all possible *states of the world* at some future moment $\Delta t > 0$ (which we'll sometimes refer to as "tomorrow"). Our convention is that $t = 0$ corresponds to today. We will use the expressions "state of the world", or simply "state", or "outcome" interchangeably. For now we leave the meaning of state of the world as an abstraction — in practice it will vary from application to application. But suppose the notion of state of the world contains enough information so that the value of some specified security, call it C, is known (i.e., can be computed) in each such state. Let $C(1), C(2), \dots, C(s)$ denote the value of C in each of the different outcomes. Assume that the security generates no cash flow between today and time Δt. The problem is to calculate the value of C *today*.

Figure 1. The one-period valuation problem.



We begin with a few assumptions. First, the value of C depends only on the numbers $C(1)$, $C(2), \dots, C(s)$. Therefore we may write

$$\textit{value of } C \textit{ today} = \ell(C) = \ell(C(1), C(2), \dots, C(s)). \tag{1}$$

A consequence of this assumption is that if two securities, or portfolios of securities, have the same payoff in every state of the world tomorrow then they have the same value today. In the literature this is sometimes called the *law of one price*. Second, we assume that value is *linear* in the sense that:

$$\ell(aC_1 + bC_2) \;=\; a\ell(C_1) + b\ell(C_2) \tag{2}$$

where a and b are real numbers and the security $aC_1 + bC_2$ pays $aC_1(\omega) + bC_2(\omega)$ in state ω. This says that a portfolio may be valued by aggregating the values of its constituent pieces. For example, the portfolio of 100 shares of GE and 50 shares of IBM is worth 100 times the value of one share of GE plus 50 times the value of one share of IBM. Finally, we have the *arbitrage-free* condition: If the payoff of a security C is positive in every state of the world tomorrow, then its value today is positive. Mathematically,

$$C(\omega) > 0 \text{ for all } \omega \in \Omega \;\;\Longrightarrow\;\; \ell(C) > 0. \tag{3}$$

This is sometimes called the *no free lunch* condition because it implies that something that is guaranteed to be of value (tomorrow) cannot be obtained without cost (today).

Consequences of the Axioms. Let R denote the *risk-free security*; it pays \$1 in every state of the world: $R(\omega) = 1$ for all $\omega \in \Omega$. By (3) we have $\ell(R) > 0$, so we may define the continuously compounding risk-free rate r by the relation $\ell(R) = e^{-r\Delta t}$, so also $e^{r\Delta t}\ell(R) = 1$.

Let S_i denote the i^{th} *state security*. It pays \$1 in state i and 0 in all other states:

$$S_i(\omega) \;=\; \begin{cases} 1 & \text{if } \omega = i \\ \\ 0 & \text{otherwise.} \end{cases}$$

Figure 2 illustrates the payoff of these securities and an illustrative security C in a situation where $s = 3$. Clearly any security may be represented as a portfolio of various amounts of the individual state securities. Referring to Figure 2 for example, $R = S_1 + S_2 + S_3$ and $C = 2S_1 + 4S_2 - S_3$. In general, for arbitrary C (and s),

$$C \;=\; \sum_{i\in\Omega} C(i) S_i. \tag{4}$$

In linear algebra terminology, the state securities form a basis for the vector space of all securities.

Figure 2. Some security payoffs.

Outcome ω (State of the world)	Security Payoffs S_1	S_2	S_3	R	C
3	0	0	1	1	-1
2	0	1	0	1	4
1	1	0	0	1	2

By repeated application of (2) to (4), we get that

$$\begin{aligned}\ell(C) &= \sum_{i\in\Omega} C(i)\ell(S_i) \\ &= e^{-r\Delta t}\sum_{i\in\Omega} e^{r\Delta t}\ell(S_i)C(i) \\ &= e^{-r\Delta t}\sum_{i\in\Omega} p_i C(i), \end{aligned} \tag{5}$$

where we define $p_i = e^{r\Delta t}\ell(S_i)$. The numbers p_i are referred to as *arbitrage probabilities.*

Our first order of business is to show that the numbers p_i do indeed look like probabilities, i.e., that

$$\sum_{i\in\Omega} p_i = 1 \tag{6}$$

and

$$0 \le p_i \le 1 \quad \text{for all} \quad i \in \Omega. \tag{7}$$

To see (6), take $C = R$ in (5). Then

$$\ell(R) = e^{-r\Delta t}\sum_{i\in\Omega} p_i R(i) = e^{-r\Delta t}\sum_{i\in\Omega} p_i,$$

since each $R(i) = 1$, and $e^{r\Delta t}\ell(R) = \sum_{i\in\Omega} p_i$. But $e^{r\Delta t}\ell(R) = 1$. ✓

In view of (6), it suffices to show that $0 \le p_i$ for all i to establish (7). Here we take $C = S_i + \epsilon R$ in (5), where $\epsilon > 0$. Then

$$C(\omega) = S_i(\omega) + \epsilon R(\omega) \ge \epsilon R(\omega) = \epsilon > 0$$

so, by the arbitrage free condition (3), we must have $\ell(C) > 0$. Hence

$$0 < \ell(C) = \ell(S_i) + \epsilon\ell(R)$$

and, multiplying by $e^{r\Delta t}$ (which is positive),

$$0 \ < \ e^{r\Delta t}\ell(S_i) + \epsilon e^{r\Delta t}\ell(R) \ = \ p_i + \epsilon.$$

This holds for any positive ϵ; by letting $\epsilon \to 0$ we see that $0 \le p_i$. ✓

If we think of Ω as a probability outcome space, we may define a probability measure P on Ω by the relation $P\{\omega = i\} = p_i$. In this framework, a security C becomes a random variable on Ω and we have that

$$\ell(C) \ = \ e^{-r\Delta t}\sum_{i\in\Omega} p_i C(i) \ = \ e^{-r\Delta t}E[C]. \tag{8}$$

In other words, our arbitrage-free pricing model ℓ calculates the expected value of the security at time Δt (with respect to the arbitrage probabilities) and then discounts it back to today at the risk-free rate r. This structure is imposed on $\ell(\cdot)$ by our three axioms: (1) law of one price; (2) linearity; and (3) no free lunch.

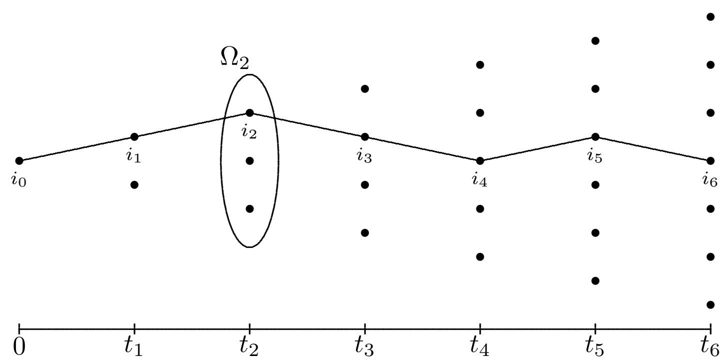
§2. The Multi-Period Generalization

To generalize to a multiple period situation, consider the following abstract model. Suppose we are attempting to value a security C that generates cash flow at various future times $t_1 < t_2 < \cdots < t_N$. We think of t_N as the *horizon time* in the sense that the security generates no cash flow after that time. The cash flow at each time is uncertain but depends only on the state of the world at that time. Let Ω_n denote the possible states of the world at time t_n. At time $0 = t_0$ there is only one such state (the state of the world today) which we'll call state 0, so $\Omega_0 = \{0\}$. Then put

$$\Omega \ = \ \big\{\omega = (i_0, i_1, i_2, \ldots, i_N) : \text{each } i_n \in \Omega_n\big\},$$

so each outcome $\omega \in \Omega$ represents a possible unfolding of the states of the world as time elapses. One such path is illustrated in Figure 3 for a *binomial* lattice.

Figure 3. A sample path through the states.



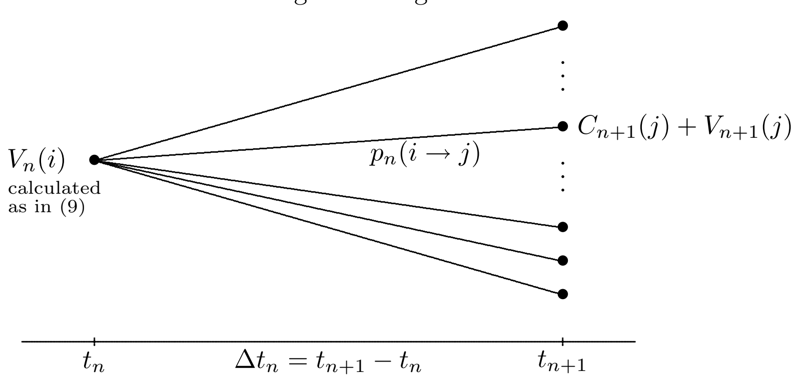
We refer to the locations in Figure 3 corresponding to each time and state (the $\bullet$s) as "nodes". Suppose that corresponding to each node (time t_n, state i) we have a set of arbitrage transition probabilities $p_n(i \to j)$ representing the probability of moving to state j at time t_{n+1} given that you are in state i at time t_n. (The process as described here is Markovian. Any process that is not Markovian can be made so by carrying along sufficient information in the description of a state. This can be tricky — we will see an example of this in the next chapter.) Suppose further that at each such node a time- and state-dependent risk free rate $r_n(i)$ is specified. Then each node, equipped with this information, represents a single-period arbitrage-free valuation model, as we presently illustrate.

We let $C_n(\omega)$ denote the cash flow generated by the security C at time t_n under the outcome (path) ω. We have assumed that the model is specified so that $C_n(\omega)$ actually depends only on i_n — the state of the world at the time of the cash flow. Reflecting this, we use the notation $C_n(i_n)$ and $C_n(\omega)$ interchangeably. Note also that at time t_n, the security will also generate future cash flows at times $t_{n+1}, \ldots, t_N$. We let $V_n(i_n)$ denote the value in state i_n at time t_n of those *remaining* cash flows. So, for example, $V_N(i) = 0$ for all $i \in \Omega_N$, since at time t_N (the horizon date) there is no subsequent cash flow. Suppose we have calculated $V_{n+1}(j)$ for all $j \in \Omega_{n+1}$ for period n (we have this for $n = N-1$). Observe that at time t_{n+1} in state j the security C generates cash flow $C_{n+1}(j)$ and the remaining cash flows have value $V_{n+1}(j)$ generating a total value at that time and state of $C_{n+1}(j) + V_{n+1}(j)$. By our single-period analysis we know that we must calculate the $V_n(i)$ as

$$V_n(i) \ = \ e^{-r_n(i)\Delta t_n} \sum_{j \in \Omega_{n+1}} p_n(i \to j)[C_{n+1}(j) + V_{n+1}(j)], \tag{9}$$

where $\Delta t_n = t_{n+1} - t_n$ (see Figure 4 for a generic node). In this fashion we may move recursively backward through the lattice until we have eventually calculated $V_0(0)$, i.e., today's value.

Figure 4. A generic node.



We remark that the familiar stock price model for option valuation falls into this rubric. In that model, the state dependent risk-free rates $r_n(i)$ are held constant throughout the lattice. What varies from state to state is the value of some underlying stock. In our models it will be the interest rate that varies from state to state. Another difference is that the stock price model is generally used to value options that expire in less than six months. There the term structure of rates is of minimal importance, as discounting takes place over a short time period. In our fixed income applications, we will be valuing securities that generate cash flow over long periods of time — up to 30 years. In this setting the discounting process is of substantial importance.

Henceforth we assume, for convenience only, that the temporal increment is a constant Δt so that $t_n = n\Delta t$. In all of our applications we will have $\Delta t = 0.5$ years.

The Single-State Multi-Period Model. It is instructive to study this valuation framework when, for each period, there is only one state of the world — call it state 0. This is a valuation model with no uncertainty — there is only one path along which the state of the world can unfold. In this case, the i on the left hand side of (9) is always 0 and the summation on the right hand side reduces to a single term corresponding to $j = 0$. Additionally, we clearly have $p_n(0 \to 0) = 1$. Putting this together gives

$$V_n(0) \;=\; e^{-r_n(0)\Delta t}[C_{n+1}(0) + V_{n+1}(0)] \tag{10}$$

Let's apply (10) to a zero coupon bond paying \$1 at the end of period m (with no other cash flow). Then $V_m(0) = 0$ (as there is no cash flow subsequent to period m), also $C_m(0) = 1$, and $C_n(0) = 0$ for $n < m$. Therefore, when $n = m - 1$, (10) gives

$$\begin{aligned} V_{m-1}(0) \;&=\; e^{-r_{m-1}(0)\Delta t}[C_m(0) + V_m(0)] \\ &=\; e^{-r_{m-1}(0)\Delta t}[1 + 0] \;=\; e^{-r_{m-1}(0)\Delta t}, \end{aligned} \tag{11}$$

and when $n < m - 1$, (10) gives

$$V_n(0) \;=\; e^{-r_n(0)\Delta t}[0 + V_{n+1}(0)] \;=\; e^{-r_n(0)\Delta t}V_{n+1}(0) \tag{12}$$

Combining (11) with repeated applications of (12) yields that

$$V_0(0) \;=\; e^{-r_0(0)\Delta t}e^{-r_1(0)\Delta t}\cdots e^{-r_{m-1}(0)\Delta t}.$$

If the period-by-period risk free rates $r_n(0)$ are chosen to explain today's term structure, we must have $V_0(0) = d(t_m)$ — the value today of one dollar to be paid with certainty at time t_m. Referring to equation (5) of Chapter 19 (where $\Delta t = 0.5$ years), one obtains that

$$e^{-r_0(0)\Delta t}e^{-r_1(0)\Delta t}\cdots e^{-r_{m-1}(0)\Delta t} \;=\; e^{-f(0)\Delta t}e^{-f(t_1)\Delta t}\cdots e^{-f(t_{m-1})\Delta t},$$

where each $f(t_i)$ is the single-period forward rate t_i years forward. If the valuation model fully explains today's term structure, this equality holds for all m. Taking $m = 1$ yields that $e^{-r_0(0)\Delta t} = e^{-f(0)\Delta t}$, or $r_0(0) = f(0)$. Taking $m = 2$ gives that $e^{-r_0(0)\Delta t}e^{-r_1(0)\Delta t} = e^{-f(0)\Delta t}e^{-f(t_1)\Delta t}$, or $r_1(0) = f(t_1)$. Moving forward period by period we see that $r_n(0) = f(t_n)$ for all $n \geq 0$. In other words, in the case with no uncertainty, the state-dependent risk-free rate precisely follows the single-period forward rates.

§3. Another Way to Compute $V_0(0)$

Here we describe a different computation which, as we shall prove, also yields the value $V_0(0)$. Once a path $\omega = (i_0, i_1, \ldots, i_N)$ through the lattice is specified, we may compute the cash flow generated by a security C if the future unfolds according to the path ω, namely $C_1(i_1), C_2(i_2), \ldots, C_N(i_N)$. We have formulated the model so the cash flow in period n only depends on the state i_n in period n, yet we may think of C_n as depending on the path ω. That is, take $C_n(\omega) = C_n(i_n)$, where $\omega = (i_0, i_1, \ldots, i_n, \ldots, i_N)$. By assigning probabilities to paths via

$$P\{\omega\} \;=\; \prod_{n=0}^{N-1} p_n(i_n \to i_{n+1}), \tag{13}$$

we turn $\Omega = \{\text{all paths through the lattice}\}$ into a probability space. In this context, the $C_n(\omega)$, $1 \leq n \leq N$, are random variables.

We may also compute path-dependent discount factors corresponding to each period in time via

$$D_n(\omega) \;=\; \prod_{m=0}^{n-1} e^{-r_m(i_m)\Delta t}, \tag{14}$$

which is analogous to (5) in Chapter 19. The D_ns are also random variables. This allows us to compute the present value of the security's path dependent cash flow with a discounting process that also depends on the path:

$$X(\omega) \;=\; \sum_{n=1}^{N} D_n(\omega)C_n(\omega), \tag{15}$$

which is analogous to (1) in Chapter 19. Finally, we compute the expectation of X by summing across all paths through the lattice:

$$EX \;=\; \sum_{\omega\in\Omega} P\{\omega\}X(\omega) \;=\; \sum_{\omega\in\Omega} P\{\omega\} \sum_{n=1}^{N} D_n(\omega)C_n(\omega). \tag{16}$$

The Expected Value of X is $V_0(0)$. Here we show that $EX = V_0(0)$. (Feel free to skip to Concluding Observations!) We proceed by mathematical induction on the length N of the lattice. If $N = 1$, using (9),

$$
\begin{aligned}
V_0(0) &= e^{-r_0(0)\Delta t} \sum_{i_1 \in \Omega_1} p_0(0 \to i_1)[C_1(i_1) + V_1(i_1)] \\
&= \sum_{i_1} p_0(0 \to i_1) e^{-r_0(0)\Delta t} C_1(i_1) \\
&= \sum_{\omega} P\{\omega\} D_1(\omega) C_1(\omega) \;=\; \sum_{\omega} P\{\omega\} \sum_{n=1}^{N} D_n(\omega) C_n(\omega).
\end{aligned}
$$

The second equality above holds because each $V_1(i_1) = 0$ (t_1 is the time horizon if $N = 1$). The third equality holds because summing over all paths through a lattice of length 1 is the same as summing over all states at time t_1. The fourth equality holds because $N = 1$. We have established that the statement $EX = V_0(0)$ is true for all lattices of length 1.

Now suppose that statement is true for all lattices of length $N - 1$. We will show that it must also be true for all lattices of length N. Then, again using (9), for a lattice of length N:

$$
V_0(0) \;=\; e^{-r_0(0)\Delta t} \sum_{i_1} p_0(0 \to i_1)[C_1(i_1) + V_1(i_1)]. \tag{17}
$$

Now, each of the $V_1(i_1)$ is the end result of a recursive calculation on a lattice of length $N - 1$, one period shorter. It follows from the induction hypothesis that, for each state i_1 at time t_1,

$$
V_1(i_i) \;=\; \sum_{\tilde{\omega}}^{(i_1)} P\{\tilde{\omega}\} \sum_{n=2}^{N} \tilde{D}_n(\tilde{\omega}) C_n(\tilde{\omega}). \tag{18}
$$

The notation $\sum_{\tilde{\omega}}^{(i_1)}$ means sum over all paths $\tilde{\omega} = (i_1, i_2, \ldots, i_N)$ starting at state i_1 at time t_1. Here $\tilde{D}_n$ is the path dependent discount factor for bringing cash flow back from time t_n to time t_1, i.e.,

$$
\tilde{D}_n(\tilde{\omega}) \;=\; \prod_{m=1}^{n-1} e^{-r_m(i_m)\Delta t}.
$$

Setting (18) into (17) yields that

$$\begin{aligned}
&V_0(0)\\
&= e^{-r_0(0)\Delta t_0}\sum_{i_1} p_0(0\to i_1)\Big[C_1(i_1)+\sum_{\tilde\omega}^{(i_1)} P\{\tilde\omega\}\sum_{n=2}^{N}\tilde D_n(\tilde\omega)C_n(\tilde\omega)\Big]\\
&= e^{-r_0(0)\Delta t_0}\sum_{i_1} p_0(0\to i_1)\sum_{\tilde\omega}^{(i_1)} P\{\tilde\omega\}\Big[C_1(i_1)+\sum_{n=2}^{N}\tilde D_n(\tilde\omega)C_n(\tilde\omega)\Big] \qquad (19)\\
&= \sum_{i_1}\sum_{\tilde\omega}^{(i_1)} p_0(0\to i_1)P\{\tilde\omega\}\Big[e^{-r_0(0)\Delta t_0}C_1(i_1)+\sum_{n=2}^{N} e^{-r_0(0)\Delta t_0}\tilde D_n(\tilde\omega)C_n(\tilde\omega)\Big].
\end{aligned}$$

In the second equality in (19) we have used that $\sum_{\tilde\omega}^{(i_1)} P\{\tilde\omega\}=1$.

Summing over all states i_1 at time t_1 and then over all paths $\tilde\omega$ of length $N-1$ at starting in state i_1 at time t_1 is equivalent to summing over all paths ω starting at state 0 time 0. With this correspondence, $p_0(0\to i_1)P\{\tilde\omega\}=P\{\omega\}$ where $\tilde\omega=(i_1,\ldots,i_N)$ and $\omega=(0,i_1,\ldots,i_N)$ — see (13). Furthermore, with this correspondence, $C_1(i_1)=C_1(\omega)$, $e^{-r_0(0)\Delta t_0}=D_1(\omega)$, and $e^{-r_0(0)\Delta t_0}\tilde D_n(\tilde\omega)=D_n(\omega)$ — see (14). Setting this into (19) gives

$$\begin{aligned}
V_0(0) &= \sum_{\omega} P\{\omega\}\Big[D_1(\omega)C_1(\omega)+\sum_{n=2}^{N} D_n(\omega)C_n(\omega)\Big]\\
&= \sum_{\omega} P\{\omega\}\sum_{n=1}^{N} D_n(\omega)C_n(\omega).\ \checkmark
\end{aligned}$$

Therefore the statement $V_0(0)=EX$ holds for all lattices of length N and the induction argument is complete.

Concluding Observations. It is not obvious that one should discount path dependent cash flows using path dependent discount factors. For example, one could compute the expected cash flow in any given period, EC_n, and then discount these expected flows using today's discount function, as in

$$\widehat{V} = \sum_{n=1}^{N} d(t_n)EC_n, \qquad (20)$$

which seems to be a natural analog of (1) in Chapter 19. As we have seen, generally this computation is not correct.

There is one situation, however, when the computations in (16) and (20) agree. That is when the cash flows, thought of as random variables, are independent of the interest rate process. In this situation, referring to (15), we see that

$$EX = \sum_{n=1}^{N} E[D_nC_n] = \sum_{n=1}^{N} ED_n\, EC_n, \qquad (21)$$

where the second equality uses the assumed independence of D_n and C_n. Consider a zero coupon bond, call it $C^{(\text{zero})}$, paying \$1 at time t_n. Then

$$C_k^{(\text{zero})}(\omega) \;=\; \begin{cases} 1 & \text{if } k = n \\ 0 & \text{otherwise} \end{cases}$$

for all paths ω. Using (15) and (16) to value $C^{(\text{zero})}$ we get

$$V_0(0) \;=\; EX \;=\; E\Big[\sum_{k=1}^{N} C_k^{(\text{zero})} D_k\Big] \;=\; ED_n,$$

since the only non-zero term in the sum is when $k = n$ and that term is $1 \cdot D_n$. If the lattice is constructed to explain today's term structure, we will have $V_0(0) = d(t_n)$, yielding $ED_n = d(t_n)$. Setting this into (21) gives that

$$EX \;=\; \sum_{n=1}^{N} d(t_n) EC_n,$$

showing that $\widehat{V} = V_0(0)$. In general, cash flows are not independent of the interest rate process. Indeed, for fixed income securities the evolution of interest rates is usually what *determines* their cash flow!

Finally, we note that the computations of (15) and (16), and the fact that $EX = V_0(0)$, are the justification for the Monte Carlo simulation methodology that we will be discussing in the next chapter. In particular, we will randomly sample paths $\omega_1, \omega_2, \ldots, \omega_m$ through the lattice in a manner consistent with their probabilities and compute the numbers $X_k = \sum_n D_n(\omega_k) C_n(\omega_k)$. The resulting sample mean,

$$\overline{X} \;=\; \frac{X_1 + \cdots + X_m}{m},$$

where m is large, will serve as an estimate of EX.