VII

An Actuarial Layman's Guide to Building Stochastic Interest Rate Generators

James A. Tilley

Abstract

A stochastic interest rate generator is a valuable actuarial tool. The parameters that specify a stochastic model of interest rates can be adjusted to make the model arbitrage-free, or they can be adjusted to accommodate an individual investor's subjective views. The arbitrage-free settings of the parameters must be used when pricing streams of interest-rate-contingent cash flows, for example, when establishing the risk-neutral position for asset-liability management. The real-world settings of the parameters should be used when evaluating the risk-reward tradeoffs inherent in deviating from the risk-neutral position.

Without relying on formulas, this paper presents the important concepts underlying the theory of arbitrage-free pricing of interest-rate-contingent cash flows: absence of opportunities for riskless arbitrage; completeness of markets; relative prices that do not depend on individual investors' subjective views or risk preferences; and expected-value pricing in the risk-neutral world. Using these concepts, the paper then describes the steps required to build continuous stochastic models of interest rates, including models that are either partially or fully arbitrage free. After studying the paper, all actuaries should be able to comprehend better some of the literature in this important subject area. Then, after studying some of the technical references, many actuaries should be in a position to begin to build their own practical models.

1. Introduction

In recent years, the literature of financial economics has featured papers on how to value interest-rate-contingent claims by means of option-pricing models (for example, [2], [8], [11], [14], and [18]). The most important applications include the pricing of fixed-income instruments with embedded options: callable corporate bonds, mortgage-backed securities subject to prepayment risk, collateralized mortgage obligations (CMOs) created by allocating the cash flows arising from pools of mortgages to different classes of bonds, floating-rate and other indexed bonds, and various hedging instruments, such as futures, options, and interest rate swaps, caps, and floors. Life insurers have begun to use option-pricing models to value complicated interest-rate-contingent liabilities that contain embedded options, such as the insurer's right to reset periodically the interest rate credited to a policyholder's account, or the policyholder's right to take loans at below-market interest rates or to surrender a policy for a cash amount that does not fully take into account the level of interest rates prevailing at the time of surrender. In the U.S., the capital adequacy of depository institutions (banks and thrifts) is now measured against risk-based capital guidelines that include an interest rate risk component, for which an option-pricing model is needed to value mortgage-related assets properly.

Unfortunately, the papers about option pricing are often very technical, leaving almost all actuaries frustrated, because they recognize the importance of utilizing option-based models, but they do not understand the theory well enough to be able to write computer programs to implement it. I have made no attempt in this paper to review the literature on the subjects of interest rate models and option pricing. That would have diluted my efforts in achieving the paper's objective of bringing the actuary who is not an expert in either financial economics or in the mathematics of stochastic processes (martingales and stochastic calculus, in particular) comfortably to the point of understanding how a useful model for valuing streams of interest-rate-contingent cash flows can be built. Several routes could have been followed to achieve this objective. After much consideration, I decided to develop the paper without formulas, with one exception: in offering an example of a continuous stochastic process for interest rates, it seemed easier to write down a few equations than to write elaborately around them. After reading this paper, and perhaps relying to some extent on the references cited, mathematically inclined actuaries will likely be able to construct stochastic interest rate generators appropriate to their needs. Other actuaries, if unable to build such generators themselves, should at least be able to apply the generators in solving asset and liability valuation problems. The principal goal of this paper is to discuss thoroughly the concepts underlying the valuation of interest-rate-contingent streams of cash flows, not to provide a set of mathematical recipes that can be programmed into an option-pricing model on a computer.

Central to the problem of valuing interest-rate-contingent cash flows is the creation of an appropriate set of interest rate paths or scenarios. In fact, once a theoretically sound stochastic interest rate generator has been constructed, all the applications described above can be handled. Each application involves projecting the relevant cash flows along a path, then discounting the projected cash flows for the path, using the short-term interest rates along the path, to a present value number for the path, and finally averaging the present value numbers for all paths to obtain the arbitrage-free value of the cash-flow stream. The rigorous proof that such a simple procedure works is highly mathematical (see, for example, the texts [12] and [16]). However, one can develop an intuitive feel for the validity of the approach without having to face intimidating mathematics. In this paper, I offer some explanation that serves to build such intuition, but not so much as to distract us from the main goal of laying the foundation for constructing arbitrage-free stochastic interest rate generators.

Section 2 introduces the concepts of current-coupon yields, spot rates of interest, and forward rates of interest, and describes the relationships among them. Section 3 briefly describes both discrete-state and continuous-state models of interest rates and debates the strengths and weaknesses of each. Section 4 introduces several key concepts from financial economics, and then indicates how the assumptions of complete markets and the lack of riskless arbitrage opportunities allow one to move into a special equilibrium world characterized as *risk neutral,* in which the valuation of interest-rate-contingent cash flows becomes a straightforward expected-value problem. Section 5 fulfills the purpose of the paper by documenting how to construct a path generator based on a continuous process, and Section 6 then indicates how such a generator can be used. Section 7 lists the key conclusions of the paper.

2. Yield Curve and Term Structure

This paper focuses on interest rates for instruments free from default and call risk, which, in the financial markets in the U.S., means U.S. Treasury bills, notes, and bonds. All other investment-grade fixed-income financial assets are priced relative to U.S. Treasury obligations. There are several equivalent ways to express the set of yields applying to risk-free debt obligations of various maturities. The most common, because it is the basis on which traders make quotations, is the concept of the *yield curve.* The yield curve is a graph that depicts the yields of hypothetical U.S. Treasury obligations that trade at a price of par as a function of their terms to maturity in years. By convention, the yields on such *par* bonds are expressed as annual rates of interest payable semiannually, referred to as *bond-equivalent yields,* because bonds issued in the U.S. usually pay coupons semiannually. The hypothetical bonds trading at a price of par that constitute the yield curve are said to have *current coupons.*

Another way to express the information contained in the yield curve is to compute the yields of *zero-coupon* bonds of various maturities from the yields of all current-coupon bonds. A unit par value zero-coupon bond having a maturity of n years pays its holder $1 at the end of n years and nothing before then. A zero-coupon bond is sometimes referred to as a *pure discount* bond, because it must always trade at a price less than par, that is, at a discount to par. The yield of a zero-coupon bond with maturity n years is referred to as the n-year *spot yield* or *spot rate*. The graph that depicts default-free spot rates as a function of term to maturity is known as the *term structure* of interest rates. The prices of the zero-coupon bonds that define the term structure are often referred to as *spot prices* and, as already stated, are always less than par.

Yet a third way to express the information contained in either the yield curve or the term structure is to compute the yields for *forward* loans. For example, an investor might agree to lend a borrower money in m years and to be repaid in full (principal plus all accumulated interest) in n years from that point in time—that is, at the end of $m + n$ years from today. Such an arrangement is known as an m-year forward n-year loan. The rate of interest for such a loan is referred to as the (m,n) year *forward rate*. More generally, the $(m,n;t)$ year forward rate refers to the interest rate on a loan that will be arranged t years from today, under which an investor will lend a borrower money $m + t$ years from today and will be repaid in full $m + n + t$ years from today.

Using the terms defined above, it can be shown that the n-year spot price is equal to the product of n positive discount factors. The first factor involves only the (0,1) year forward rate; the second factor involves only the (1,1) year forward rate; and the n-th factor involves only the $(n - 1,1)$ year forward rate. Thus, it follows that the $(n + 1)$-year spot price divided by the n-year spot price is equal to a positive number that depends only on the $(n,1)$ year forward rate. This number will be less than or equal to 1 (in other words, it will be a "discount" factor) if, and only if, the $(n,1)$ year forward rate is non-negative.

The information contained in the sets of current-coupon yields, spot rates, and forward rates is equivalent. (Further material on this subject can be found in the text by Sharpe and Alexander [22].) Any one set of yields or rates is sufficient to derive the other two sets. Depending on the situation, there may be a natural set to use, but all carry identical information. For example, when speaking with traders or portfolio managers to obtain interest rate assumptions for pricing an annuity product, an actuary would likely ask about the yield curve. When discounting a stream of fixed and certain cash flows arising from structured settlement annuity liabilities to obtain a current market value, an actuary would naturally use spot rates. When constructing an arbitrage-free theory of interest rate dynamics, most financial economists would use forward rates as the starting point.

3. Discrete versus Continuous Models

Throughout this paper, the term *state of the world* refers to the yield curve prevailing at a particular time or epoch. In a model of interest rates, the adjectives *discrete* and *continuous,* without any modifiers, are used best to describe the type of states of the world represented, not the type of time interval used. In practical applications, regardless of the model used, cash flows are assumed to occur at discrete time intervals: monthly for typical mortgages; quarterly for CMOs, preferred stocks, and some floating-rate bonds; and semiannually for typical bonds. In asset-liability cash-flow analyses, quarter-year periods typically are used. So the basic issue is not whether discrete-time models are to be preferred to continuous-time models; rather, it is whether discrete-state models are to be preferred to continuous-state models.

Most of the recent literature describes discrete models, in which the states of the world are represented by nodes on a lattice (refer to the papers [2], [11], and [18] cited earlier). The vast majority of such models utilize binomial lattices, on which the world evolves from any given state at one epoch to one of two different states at the next epoch. These two states at the end of a time interval are usually referred to as the *up* state and the *down* state with respect to the state at the beginning of the interval. For reasons of computational efficiency, *connected* lattices are almost always used. From any node in a connected lattice, the two-period evolution of states *up first, then down* and the two-period evolution of states *down first, then up* must lead to the same ending node. In a connected lattice, the world can evolve from a single initial state at epoch 0 to one of two states at epoch 1, to one of three states at epoch 2, and so on, to one of $H + 1$ states at epoch H. In a connected binomial lattice model, it is unlikely that the possible states of the *real world* will be sampled sufficiently finely at the early epochs. To remedy this problem, the time interval can be reduced. For example, with daily intervals, there are about 30 states at the end of any one-month period, but the computational demands of creating and using such a model can be enormous, especially for long-term assets or liabilities. Moreover, it is unnatural (and should be unnecessary) to choose a time interval much shorter than the shortest period between cash flows for typical assets and liabilities. Thus, the coarseness-of-sampling difficulty of connected lattice models remains in many practical situations. Continuous models do not suffer this weakness.

Continuous-state models are described in the academic literature by means of differential equations that represent continuous-time stochastic processes (refer to the papers [8] and [14] cited earlier, and also to the text by Hull [13]). For practical applications, though, the continuous-time process needs to be sampled only at regular time intervals, and the models are reformulated as stochastic difference equations. The time interval is often chosen to equal the shortest period between the cash flows for the assets and liabilities under study. In continuous models, one samples *paths* of interest rates by iterating the difference equation. If P interest rate paths are used, there are P states of the world represented at every epoch. Because the sample of P states at each epoch is drawn from a continuous distribution, the resulting paths of interest rates do not appear to have been constructed artificially. Stated a little differently, it is difficult for an experienced portfolio manager to tell whether an interest rate path was generated from a good continuous model or was constructed from segments of actual interest rate history. The same claim cannot be made for interest rate paths sampled from a lattice.

A connected lattice model has a significant weakness that can be overcome by using a continuous model. For a connected lattice to be *arbitrage-free* (defined in Section 4), severe constraints have to be placed on how it is constructed. These constraints greatly limit the possible yield curve dynamics, and for most models, the resulting evolution of yield curves does not correspond adequately to real-world behavior. The problem arises in simple lattice models because a *single* stochastic factor—the short-term rate of interest—drives the dynamics of the *entire* yield curve, resulting in perfect correlation of yield movements across the curve. In the real world, the movements of neighboring segments of the yield curve may be highly correlated, but they are not perfectly correlated. Arbitrary correlation can be accommodated in a continuous model, because different parts of the yield curve can be assumed to follow correlated stochastic processes.

Discrete and continuous models can also be compared for computational efficiency, which depends on the type of problem to be solved. In the case of interest-rate-contingent, but *path-independent,* cash flows, as are usually associated with pure options, callable bonds, and optional sinking fund bonds, *backward induction* algorithms can be used on a lattice to determine the optimal exercise strategies. Such algorithms are processed backward in time from the latest epoch to the earliest epoch, and such algorithms need to evaluate conditions occurring only at all states in the lattice, not along all paths through the lattice. From epoch 0 to epoch H, there are 2^H paths through a binomial lattice, but only $(H + 1)(H + 2)/2$ total states, if the lattice is connected. Thus, many option-pricing problems can be solved efficiently and accurately on a connected lattice. Without a lattice (whether connected or not), backward induction is not possible. From a purely mathematical viewpoint, it is difficult to construct *optimal* exercise strategies for many option problems by doing calculations on interest rate paths sampled from a continuous model. From a practical viewpoint, note that real-world options are exercised by people who manage portfolios or trading positions, or who run corporations or other businesses. The *behavior* of these people, as to their strategies for rational (if not mathematically optimal) exercise of the options they hold, can be modeled sufficiently accurately that the options are valued properly by way of calculations performed on paths sampled from a continuous model.

Many important problems involve *path-dependent* cash flows, for example, the pricing of prepayable mortgages and instruments derived from them, and the valuation of interest-sensitive insurance liabilities. For such problems, the possible paths of interest rates must be considered, not merely the possible states of the world. A connected lattice offers no special computational advantages in these situations. In fact, when path-dependent cash flows are involved and a lattice model is used, P paths of interest rates will have to be sampled, just as if a continuous model were being used.

In summary, several compelling factors favor the use of continuous-state models over discrete-state models: (i) a discrete model's lack of computational advantage in the common case of path-dependent cash flows; (ii) the need to use the same model consistently for all assets and liabilities, whether their cash flows are path independent or path dependent; and (iii) a continuous model's ability to sample states of the world sufficiently densely and to accommodate realistic yield curve dynamics.

4. Riskless Arbitrage, Complete Markets, and the Risk-Neutral World

4.1 Basic Concepts

This section is shorter than it could be, so that we can proceed to the main subject of the paper. The underlying mathematics are generally presented in an imposing manner and have been the subject of numerous lengthy seminal papers on the application of stochastic process theory to financial economics. The topic of riskless arbitrage is dealt with well in the paper by Pedersen, Shiu, and Thorlacius [18], and brief comments on the role of the risk-neutral world in option-pricing calculations can be found in the text by Cox and Rubinstein [6].

The concept of a *riskless arbitrage opportunity* is not difficult. If one asset or portfolio of assets can be sold and the proceeds of the sale can be used to purchase a different asset or portfolio of assets whose performance will be superior to that of the original asset or portfolio over a specified holding period (infinitesimal or finite depending on the situation), regardless of the states of the world during and at the end of the holding period, then a riskless arbitrage opportunity is said to exist. One need merely sell the first asset or portfolio and purchase the second to be guaranteed of having more wealth at the end of the holding period without having incurred greater risk. The reason that such an opportunity is said to be *riskless is* that wealth can be created without investing any capital at all by selling the first asset *short* (that is, selling it before purchasing it), and using the proceeds of the short sale to purchase the second asset. In this situation, there is no net outlay of funds, but there is a guarantee of positive wealth at the end of the holding period, because the second asset can then be sold for more than is then needed to cover (close out) the short position by purchasing the first asset.

Financial economists and other reasonable people assume that no riskless arbitrage opportunities exist in an equilibrium world. In other words, prices of assets are assumed to adjust continuously to eliminate opportunities for riskless arbitrage. For this to occur, a number of assumptions must be made: assets are perfectly divisible, unlimited short sales are possible, trading takes place continuously without transaction costs, investors act rationally and prefer more wealth to less wealth, and there are no taxes. Although these assumptions are quite stringent, one should not debate too strenuously whether small arbitrages can exist in the real world for brief periods because the assumptions are only approximations to reality. Instead, one should regard the concept of an equilibrium world in which riskless arbitrage opportunities do not exist as fundamental to constructing a sound financial theory for pricing assets.

To see how the concept of riskless arbitrage can lead to a theory for establishing the *relative* prices of assets, consider again the situation described above, modified slightly. Suppose that an asset for which one wants to establish the arbitrage-free price is equivalent to a portfolio of different assets for which one knows the prices. *Equivalence* is used in the sense that the performance of the single asset and that of the portfolio are *identical* over a specified holding period. Then it follows that the single asset and the portfolio of assets must have the same prices, or else there would be a riskless arbitrage opportunity, wherein the more expensive one could be sold short and the less expensive one purchased, guaranteeing a profit without taking any risk. Thus, one establishes the arbitrage-free price of the single asset as equal to the *known* price of the portfolio of assets. For this approach to be generally applicable and therefore lead to a pricing theory, it is necessary to assume that the financial markets are complete, meaning that any given asset is equivalent to some portfolio of fundamental assets.[1] This *replicating portfolio* might not be equivalent to the given asset over all holding periods. The portfolio's holdings might have to be adjusted from time to time, perhaps continuously, to maintain the equivalence. Having to rebalance the replicating portfolio is of no consequence, however, because the ability to trade continuously absent transaction costs, as assumed earlier, enables equivalence to be maintained without having to inject additional money into the portfolio; the replicating strategy is said to be *self-financing*.

If the financial markets are complete and no opportunities for riskless arbitrage exist, then the prices of all assets can be determined relative to the prices of their replicating portfolios. Under these assumptions, the relative prices of assets cannot depend on individual investors' *preferences,* which include their differing subjective views on the probabilities of occurrence of various future states of the world and their differing degrees of aversion to risk. Otherwise, riskless arbitrage opportunities would arise. Because relative asset prices must be *preference-free,* we can choose a frame of reference in which the pricing of assets is particularly straightforward, namely, the *risk-neutral world.* It does not mean that we must adopt such a setting, only that we are permitted to do so, and that we will obtain the correct relative prices for assets if we do. Black and Scholes [3] derived their now-famous formula for the price of a call option on a share of non-dividend-paying stock in terms of the price of the underlying stock by applying the no-riskless-arbitrage condition to a combined position of buying the call option and selling short its replicating portfolio. They solved the resulting differential equation for the price of the call option after establishing appropriate boundary conditions. Only later did others show that a simpler derivation is possible by moving into the risk-neutral world and performing the pricing calculation there (for example, refer to [5]).

What is the risk-neutral world, and why are pricing calculations simpler there? In the risk-neutral world, investors do not require a premium for assuming risk. Thus, assets are priced at their expected present values. In other words, risk-neutral investors behave like traditional actuaries. When pricing assets, they project cash flows along interest rate paths, then discount the cash flows at the one-period interest rates occurring along the paths, and finally calculate the *expected present value* by weighting the present value for each path by that path's probability of occurrence.[2] Moreover, in the risk-neutral world, the probabilities of occurrence of the various paths do not depend on investors' subjective views of the likelihood that different future states of the world will arise. I now show how this description of the risk-neutral world can be used to construct an arbitrage-free model of interest rates.

In a binomial model, the states of the world represented at the nodes of the lattice can be determined from the assumed stochastic process for the one-period interest rate, for example, a discrete geometric Brownian motion. Then, the *risk-neutral probabilities* of up and down transitions at each node can be established to ensure that all zero-coupon bonds are priced properly by the expected-present-value algorithm described above. Alternatively, the no-riskless-arbitrage conditions can be used to establish the possible states of the world, given assumed risk-neutral probabilities for up and down transitions at each node; for example, $\theta_{i,t}$ for the up transition and $1 - \theta_{i,t}$ for the down transition at the i-th node at epoch t, with $0 < \theta_{i,t} < 1$. This is the approach used by Pedersen, Shiu, and Thorlacius [18]. In a continuous model, it is convenient to adopt the approach of assuming that the risk-neutral probabilities are given, and then generating a *finite* number of interest rate paths appropriately. It is usual to generate *equal-probability* paths of interest rates by randomly sampling, epoch to epoch, from an assumed stochastic process, and to adjust, epoch by epoch, the distribution of interest rates to ensure that the no-riskless- arbitrage conditions hold.

~~*4.2 Example: A One-Factor Lognormal Model of Short-Term Interest Rates*~~

~~The rest of this section is devoted to an example in which the natural logarithm of the ratio of the one-period rate of interest at epoch t to the one-period rate of interest at epoch $t - 1$ is normally distributed with mean μ and standard deviation σ. It is conventional to refer to this example as a lognormal stochastic process for the one-period (spot) rate. The initial term structure of interest rates (all the spot rates or all the spot prices) is assumed to be specified exogenously. The objective is to generate an arbitrage-free set of P equal-probability paths of one-period interest rates out to epoch H, the assumed horizon for the desired application. In practical applications, limitations on computer memory and execution time usually constrain the choice of P to between 100 and 1000.~~

~~A single path of one-period interest rates can be created by starting from the given initial one-period rate, then randomly sampling from the assumed lognormal distribution a one-period rate at epoch 1 and using it as the starting one-period rate for randomly sampling from the assumed lognormal distribution a one-period rate at epoch 2, and so on, out to epoch H. Independently repeating this entire sequence of computations P times gives rise to P equal-probability paths of one-period interest rates. Unfortunately, the set of paths is *not* arbitrage free. To obtain an arbitrage-free set of paths, all P one-period rates at each epoch must be multiplied by an appropriate adjustment factor that is the same for all P rates, but that differs from epoch to epoch.[3] The proper approach involves generating and adjusting the one-period rates at epoch 1, which evolve from the given initial one-period rate at epoch 0; then generating and adjusting the one-period rates at epoch 2, which evolve from the adjusted one-period rates at epoch 1; and so on; and finally generating and adjusting the one-period rates at epoch H, which evolve from the adjusted~~