Notes to Financial Engineering: Fixed Income Derivatives

Matthew Cocci

Contents

1. The Fundamental Theorem of Asset Pricing

1.1. Introduction

Derivatives require special pricing techniques aside from the traditional discounted cash flow (DCF) approach, as DCF requires an estimate of the appropriate risk-adjusted rate of return. However, the risk of a derivative varies over time, which makes it difficult to estimate the derivative's risk-adjusted return.

As a result, derivatives pricing turns to the no-arbitrage approach (NA), which eliminates the need to build risk into the model.

1.2. Trading Strategy and Derivative Pricing Definitions

A **trading strategy** is a dynamically-rebalanced portfolio.

A trading strategy is **self-financing** if it generates no intermediate cash inflows and requires no intermediate outflows between the time the portfolio is initiated and the time it is liquidated. This implies that

i. All dividends are reinvestd.

ii. Value of the assets sold at a rebalance time must equal the value of the assets bought.

A trading strategy is **strictly positive** if the value of the traded portfolio can never become zero or negative.

Let N be the value of a *strictly positive*, *self-financing* trading strategy. Then N is a **numeraire process** or, simply, a **numeraire**. Here are a few examples:

- Price of Dividend paying asset: NO, as there are intermediate cash outflows, violating self-financing condition.
- The price of a forward contract: NO, as it can go negative, violating the strictly positive condition.
- Price of a Foreign Currency: NO, as it is equivalent to a dividend paying asset because you think of it as an investment in an interest-bearing account.
- Price of a non-defaultable zero-coupon bond: YES.
- Value of a money market account earning the risk free rate, where there are no interim deposits or withdrawals: YES.

1.3. Martingales and Change of Measure

A **martingale** is a stochastic process X with the property

$$E_t[X(T) - X(t)] = 0 \;\Leftrightarrow\; E_t[X(T)] = X(t), \qquad T > t.$$

A **probability measure** is a specification of the probabilities of all the possible states of the words, mapping states to real numbers.

Suppose that ξ is a nonnegative random variable on $(\Omega, \mathcal{F}, P)$ with $E_P[\xi] = 1$. (The subscript P highlights that the last expectation is with respect to measure P.) Then define a new measure

$$Q : \mathcal{F} \to [0, 1]$$

$$Q(A) = E\left[1_A \xi\right] = \int_A \xi(\omega) dP(\omega), \qquad A \in \mathcal{F} \tag{1}$$

Clearly, Q is a probability measure on $(\Omega, \mathcal{F})$ and it is absolutely continuous with respect to P—i.e. we have

$$Q(A) > 0 \Rightarrow P(A) > 0.$$

Note that it is common to write the random variable ξ as

$$\xi = \frac{dQ}{dP},$$

and we often refer to ξ as the *Radon-Nikodym derivative* or the *likelihood ratio* of Q with respect to P.

Radon-Nikodym Theorem If P and Q are two probability measures on $(\Omega, \mathcal{F})$, then there *will exist* such a random variable ξ so that Expression 1 holds.

1.4. Fundamental Theorem of Asset Pricing (No Dividends)

1.4.1. Statement of Theorem

Suppose we have n non-dividend-paying assets with price processes $S_1, S_2, \ldots, S_n$. Let N be some numeraire process. Then, barring market imperfection, there are no arbitrage opportunities among these assets if and only if there exists a strictly positive probability measure Q_N (so it's dependent upon the numeraire, N) under which each of the processes S_i/N is a martingale.

- Note that S_i/N is the price of asset i in units of the numeraire N. Therefore, we call S_i/N the **normalized price process**.

- The probability measure Q_N will, in general depend on the numeraire. Therefore, we call Q_N the **martingale measure** or **pricing measure** associated with the numeraire N.

- We can paraphrase FTAP by saying that, if the is no arbitrage or market imperfections, then given *any* numeraire process N, there must exist a corresponding martingale measure Q_N under which the normalized price of any non-dividend paying asset is a martingale:

$$\frac{S_i(t)}{N(t)} = E_t^{Q_N}\left[\frac{S_i(T)}{N(T)}\right], \qquad T > t.$$

  From this, we see that changing N will generally change Q_N as well.

1.4.2. Consequences for Derivative Pricing

Let V denote the price of a derivative with payoff $V(T)$ at time T. Then we can apply FTAP to get

$$V(t) = N(t)E_t^{Q_N}\left[\frac{S_i(T)}{N(T)}\right], \qquad T > t.$$

Note, the price we get for a deriative is *invariant* to the choice of the numeraire.

1.4.3. Special Numeraires and Martingale Measures

T-forward measure Let $P(t,T)$ be the price of a non-defaultable zero-coupon bond with unit face value. Then

$$N(t) = P(t,T)$$

is our numeraire. The associated martingale measure, denoted Q_T, is called the *T-forward martingale measure*. This yields a derivative price of

$$V(t) = E_t^{Q_B}\left[e^{-\int_t^T r(s)ds}V(T)\right]$$

Risk Neutral Measure Let's consider the value of a money market account with unit initial value as our numeraire. Then

$$N(t) = B(t) = e^{\int_0^t r(s)ds}$$

where r is the instantaneous risk-free rate. The associated martingale measure, denoted by Q_B, is called the *risk-neutral martingale measure*. This yields a derivative price of

$$V(t) = P(t,T)\, E_t^{Q_T}[V(T)] \;=\; e^{-r(t,T)(T-t)} E_t^{Q_T}[V(T)]$$
$$r(t,T) = -\ln P(t,T)/(T-t)$$

If interest rates are stochastic (and they probably are), then this measure isn't as convenient a tthe T-forward measure.

1.5. FTAP for Dividend-Paying Assets

Consider an asset with price process S and let $D(t)$ denote the cumulative dividend paid by the asset from time 0 up to time t. We can consider the undiscounted cash flows from holding an asset from t to T:

$$S(T) - S(t) + D(T) - D(t) = GP(T) - GP(t)$$

where $GP(t) = S(t) + D(t)$ is the asset's gain process.

Given a numeraire N, the asset's *normalized gain process*, denoted NGP, measures the gains from holding the assets in units of N:

$$NGP(t) = \frac{S(t)}{N(t)} + \int_0^t \frac{dD(s)}{N(s)}$$

where $dD(s)$ is the dividend paid by the asset at time s.

Theorem Now, let's restate the fundamental theorem of asset pricing allowing for dividend paying assets. So again, consider assets with price processes $S_1, \ldots, S_n$ and cumulative dividend processes $D_1, \ldots, D_n$, letting N be any numeraire process. Then there are no arbitrage opportunities across these assets if and only if there exists a strictly positive probability measure Q_N under which each

$$\frac{S_i(t)}{N(t)} + \int_0^t \frac{dD_i(s)}{N(s)}$$

is a martingale. This implies

$$\frac{S_i(t)}{N(t)} = E_t^{Q_N}\left[\frac{S_i(T)}{N(T)} + \int_t^T \frac{dD_i(s)}{N(s)}\right]$$

And so the normalized price of any asset is equalt to the conditional expectation under the martingale measure of the assets normalized payoffs (including future dividends).

2. Continuous Time Stochastic Processes

Here, we develop the necessary machinery in continuous time stochastic processes to model asset price evolution properly and with sufficient richness and generality. In particular, we discuss a heirarchy of model classes that includes Brownian Motion $\subset$ Generalized Brownian Motion $\subset$ Diffusions $\subset$ Ito Processes.

2.1. Introduction

A *stochastic process* X is a collection of random variables indexed by time: $X = \{X_t : t \in \mathcal{T}\}$.

- *Discrete Time*: $\mathcal{T}$ countable, and process changes only at discrete time intervals.
- *Continuous Time*: $\mathcal{T}$ uncountable.

Definition A process X has stationary increments if $X_T - X_t$ has the same distribution as $X_{T'} - X_{t'}$ provided that $T - t = T' - t'$.

2.2. Brownian Motion

Definition The most basic continuous-time process is *Brownian Motion* (or the *Wiener Process*). It has three defining properties:

i. $W(0) = O$.

ii. $W(t)$ is continuous, so no jumps.

iii. Given any two times, $T > t$, the increment $W(T) - W(t)$ is independent of all previous history and normally distributed with mean 0 and variance $T - t$.

A few consequences of the definition of $W(t)$:

- Brownian motion has independent stationary increments.
- $W(t)$ is normally distributed with $\mu = 0$, $\sigma^2 = t$.

2.3. Generalized Brownian Motion

Definition A *generalized Brownian motion* is a continuous-time process X with the following property:

$$X(t) = X(0) + \mu t + \sigma W(t)$$

where μ is the *drift*, σ is the *volatility*, and W is simple Brownian motion. The differential equation equivalent is written:

$$dX(t) = \mu\, dt + \sigma\, dW(t).$$

It follows immediately from the definition that

- $X(t)$ is continuous, so no jumps.

- $X(t)$ is normally distributed with mean $X(0) + \mu t$, variance $\sigma^2 t$.

- Given any two times, $T > t$, the increment $X(T) - X(t)$ is independent of all previous history and normally distributed with mean $\mu(T-t)$ and variance $\sigma^2(T-t)$.

- X is a martingale if and only if $\mu = 0$.

Theorem It also happens that Generalized Brownian motions are the only continuous time processes with continuous sample paths and stationary increments.

2.4. Ito Processes

Even more general than Brownian Motion (which is retained as a special case), an *Ito Process* is a stochastic process X defined by one of two equivalent formulations:

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dW(t)$$

$$X_t = X_0 + \int_0^t \mu(s)\,ds + \int_0^t \sigma(s)\,dW(s)$$

for any arbitrary stochastic processes μ (the drift) and σ (the volatility) along with some Brownian Motion $W(t)$. Here are some properties

- Has continuous sample paths and is a martingale if and only if $\mu(t) = 0$.

- Increments are not necessarily stationary, as μ and σ can change *randomly* with time.

Definition If the drift and volatility of an Ito process depend only upon the current value of the process and time, then X is a *diffusion*. Mathematically, X is a *diffusion* if

$$dX(t) = \mu(X(t), t)\,dt + \sigma(X(t), t)\,dW(t)$$

for some functions μ and σ.

2.5. Ito's Lemma

Suppose that X is an Ito process defined by

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dW(t) \tag{2}$$

and we define a new process $Y(t) = f(X(t), t)$ where f is some function that's twice differentiable in X and once in t. Then we have that

$$dY(t) = f_X(X(t), t)\,dX(t) + f_t(X(t), t)\,dt + \frac{1}{2} f_{XX}(X(t), t)\sigma(t)^2\,dt. \tag{3}$$

where subscripts on f denote the partial derivatives.[1] Subbing Equation 2 into Equation 3, we get that

$$dY(t) = \left(f_X(X(t),t)\mu(t) + f_t(X(t),t) + \frac{1}{2} f_{XX}(X(t),t)\sigma(t)^2 \right) dt + f_X(X(t),t)\sigma(t)\, dW(t)$$

Thus, it is clear that Y is also an Ito Process by the statement above with the drift and volatility given by the coefficients on dt and $dW(t)$ as always.

Using Ito's Lemma In practice, we use Ito's Lemma whenever we have a (typically complicated) Ito Process that we want to solve. Given the process X and its corresponding Ito Process, we posit a function f that could could help. Then we write a new Ito Process using Ito's lemma with $dY(t) = df(X(t),t)$ on the LHS. From there, hopefully we can integrate $dY(t)$ easily on the left and solve out for $X(t)$.

2.6. Multi-dimensional Ito's Lemma

For the sake of completeness, let's generalize Ito's Lemma to consider the case of a finite number of Ito processes, $X_1, X_2, \ldots, X_n$,

$$dX_i(t) = \mu_i(t)\, dt + \sigma_i(t)\, dW_i(t).$$

Next, define $Y(t) = f(X_1(t), \ldots, X_n(t), t)$ for some differentiable function f. Then multi-dimensional Ito's Lemma says

$$\begin{aligned} dY(t) = &\sum_{i=1}^{n} f_{X_i}(X_1(t), \ldots, X_n(t), t)\, dX_i(t) \\ &+ f_t(X_1(t), \ldots, X_n(t), t)\, dt \\ &+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} f_{X_i,X_j}(X_1(t), \ldots, X_n(t), t)\, \rho_{ij}\sigma_i(t)\sigma_j(t)\, dt \end{aligned}$$

where ρ_{ij} is the correlation coefficient between dW_i and dW_j. Note that you'll have to plug back in for the dX_i in the first sum.

We'll mostly consider with the two-dimensional case for the two specific instances below:

i. $Y(t) = X_1(t)X_2(t)$, which gives us

$$dY(t) = X_2(t)\, dX_1(t) + X_1(t)\, dX_2(t) + \rho_{12}\sigma_1(t)\sigma_2(t)\, dt$$

Note that you'll have to plug back in for dX_1 and dX_2. This is a type of integration-by-parts formula because (after rearranging terms) it relates $X_1\, dX_2$ to $X_2\, dX_1$.

[1] Note that Equation 3 almost looks like the chain rule from traditional calculus, except for that extra term with f_{XX} partial derivative. That arises from the additional variability due to the inclusion of stochastic factors like $W(t)$ in the original Ito Process.

ii. $Y(t) = X_1(t)/X_2(t)$, which gives us

$$dY(t) = \frac{1}{X_2(t)} dX_1(t) - \frac{X_1(t)}{X_2(t)^2} dX_2(t) + \frac{X_1(t)}{X_2(t)^3} \sigma_2(t)^2 \, dt - \frac{1}{X_2(t)^2} \, \rho_{12} \sigma_1(t) \sigma_2(t) \, dt.$$

Note that you'll have to plug back in for dX_1 and dX_2. This is a type of integration-by-parts formula because (after rearranging terms) it relates $X_1 \, dX_2$ to $X_2 \, dX_1$.

2.7. Geometric Brownian Motion

Let's consider the process X governed by

$$dX(t) = X(t)\mu(t) \, dt + X(t)\sigma(t) dW(t).$$

To solve, let us consider the process $Y(t) = \log X(t)$. We compute the partials and apply Ito's Lemma:

$$f_X = \frac{1}{X(t)}, \qquad f_{XX} = -\frac{1}{X(t)^2}, \qquad f_t = 0$$

$$dY(t) = \frac{1}{X(t)} dX(t) - \frac{1}{2} \frac{1}{X(t)^2} (\sigma(t) X(t))^2) \, dt$$

which simplifies (after subbing in for $dX(t)$) into the expression

$$dY(t) = \left(\mu(t) - \frac{1}{2} \sigma(t)^2 \right) + \sigma(t) \, dW(t).$$

Next, integrating both sides and substituting back in with $Y(t) = \log X(t)$, we get

$$Y(t) = Y(0) + \int_0^t \left(\mu(s) - \frac{1}{2} \sigma(s)^2 \right) ds + \int_0^t \sigma(s) \, dW(s)$$

$$X(t) = X(0) e^{\int_0^t \left(\mu(s) - \frac{1}{2} \sigma(s)^2 \right) ds + \int_0^t \sigma(s) \, dW(s)}$$

where it also follows that X is strictly positive.

Special Case Suppose that μ and σ are constant, in which case X follows a *geometric Brownian motion.* Then it follows that

$$\log X(t) \sim N \left(\log X(0) + \left(\mu - \frac{1}{2} \sigma^2 t \right), \sigma^2 t \right)$$

so that we say $X(t)$ is *lognormally distributed.*

2.8. Girsanov's Theorem

Theorem Suppose that X is an Ito process

$$dX(t) = X(t)\mu(t)\ dt + X(t)\sigma(t)\ dW(t),$$

where μ an σ are stochastic processes and W is Brownian Motion under some probability measure P—like maybe the real world measure. Then if Q is any other strictly positive probability measure, then X is also an Ito process under Q—i.e., there exist processes $\hat{\mu}$, $\hat{\sigma}$, and $\hat{W}$ with the property that

$$dX(t) = X(t)\hat{\mu}(t)\ dt + X(t)\hat{\sigma}(t)\ d\hat{W}(t). \tag{4}$$

Even better, $\hat{\sigma} = \sigma$.

This is particularly useful because, in general, we will have to work with two different probability measures: the true/historical P and the martingale probability measure Q_N.

3. Interest Rate Derivatives

Just a quick word about terminology. We set terms at the *contracting date* first about the *expiration*—the date at which an option-holder must make a decision about whether to exercise an option like a cap, floor, or swaption. The *maturity* is when the option would stop paying (assuming you chose to exercise). Finally, suppose we have a simple one-period derivative whose payoff at T_{i+1} depends on the rate at T_i. Then T_i is the so-called *fixing date*, while T_{i+1} is the maturity.

3.1. Basics and Notation

Before we begin, let's define a few important terms, assets, and concepts. Throughout, we will use the convention that lowercase rates denote *continuous compounding* while uppercase rates represents *simple interest*.

- *Zero Coupon Bond*: Defined by $P(t,T)$, it represents the price at time t of a non-defaultable zero with unit face value.

- *Spot Interest Rates*: Available at time t for maturity at time T,

$$r(t,T) = -\frac{\ln P(t,T)}{T-t}, \qquad R(t,T) = \frac{1 - P(t,T)}{(T-t)P(t,T)}$$

- *Forward Interest Rates*: These are forward rates available at time t for a loan starting at time T with maturity at time $T+\tau$,

$$f(t,T,T+\tau) = \frac{\ln P(t,T) - \ln P(t,T+\tau)}{\tau}, \quad F(t,T,T+\tau) = \frac{P(t,T) - P(t,T+\tau)}{\tau P(t,T+\tau)}$$

Now that we have the basics, we'll refer explictly to some schedule or set of maturities, $\{T_0, T_1, \ldots, T_{n-1}, T_n\}$, just so that we can simplify notation. We'll assume that the dates are evenly spaced so that $T_i - T_{i-1} = \tau$.

$$\begin{aligned} P_i(t) &= P(t,T_i) \\ R_i(t) = R(t,T_i) &= \frac{1 - P_i(t)}{(T_i - t)P_i(t)} \\ F_i(t) = F(t,T_{i-1},T_i) &= \frac{P_{i-1}(t) - P_i(t)}{\tau P_i(t)} \end{aligned}$$

We also have a couple of identities that result from our specification:

$$\begin{aligned} F_i(T_{i-1}) &= R_i(T_{i-1}) \\ \frac{P_i(t)}{P_{i-1}(t)} &= \frac{1}{1 + F_i(t)\tau} \end{aligned}$$

Finally, the last thing to mention is that we write the forward martingale measure (which we will primarily use in our analysis) as $Q_i = Q_{T_i}$ with P_i as the associated numeraire.

3.2. Forward Rate Agreements

Payoff: A $T_{i-1} \times T_i$ forward rate agreement (FRA) settled in arrears is a forward contract with payoff at T_i equal to the difference between interest on a notional N at the floating rate $R_i(T_{i-1})$ and interest on the same notional at a pre-specified fixed rate K:

$$N \cdot (R_i(T_{i-1}) - K)\tau$$

But since we know the payoff to an FRA already at T_{i-1}, we typically settle FRA's in advance at the fixing time by taking the present value of the payoff:

$$N \cdot P_i(T_{i-1}) \cdot (R_i(T_{i-1}) - K)\tau = N \cdot \frac{(R_i(T_{i-1}) - K)\tau}{1 + R_i(T_{i-1})\tau} \tag{5}$$

Terminology: Everything is from the perspective of the fixed:

- The person who pays fixed is *long*, or has a *payer* FRA.

- The person who receives fixed is *short*, or has a *receiver* FRA.

Pricing: It can be shown that at any time $t \leq T_{i-1}$, an agreement for a loan from T_{i-1} to T_i is priced as follows:

$$\begin{aligned} FRA(t, T_i, K) &= P_{i-1}(t) - (1 + K\tau)P_i(t) \\ &= P_i(t)(F_i(t) - K)\tau \end{aligned} \tag{6}$$

where K is the strike. Typically we want to set K so that the initial value is 0, which implies that at time 0

$$K = F_i(0)$$

Applying FTAP: By the definition of FTAP, we know that the value of the FRA should be the discounted future payoff we defined above in Equation 5, so under the T-forward martingale measure:

$$\begin{aligned} FRA(t, T_i, K) &= P_i(t) E_t^{Q_i} \left[\frac{P_i(T_{i-1})(R_i(T_{i-1}) - K)\tau}{P_i(T_{i-1})} \right] \\ &= P_i(t) \left(E_t^{Q_i} \left[R_i(T_{i-1}) \right] - K \right) \tau \end{aligned} \tag{7}$$

Now if we compare the value today of the FRA, expressed two different ways in Equations 6 and 7, we see that we must have

$$F_i(t) = E_t^{Q_i} \left[R_i(T_{i-1}) \right] = E_t^{Q_i} \left[F_i(T_{i-1}) \right]$$

implying that the forward rate, F_i is a martingale under Q_i.[2]

[2]The third equality follows from an identitity we saw in the previous section:

$$F_i(T_{i-1}) = R_i(T_{i-1}).$$

3.3. Interest Rate Swaps

Spot Interest-Rate Swap (IRS): A contract that exchanges the fixed interest payment $N \cdot K \cdot \tau$—where K is the fixed interest rate called the *swap rate*—for the variable interest rate payment $N \cdot R_i(T_{i-1}) \cdot \tau$ at each settlement date, $\{T_1, \ldots, T_n\}$.[3] Clearly an IRS is simply equivalent to a portfolio of FRA's settled in arrears, with the swap rate K usually chosen so that the value of the IRS at T_0 is zero.[4]

Forward IRS Starting at T_j: A contract that exchanges the fixed interest payment $N \cdot K \cdot \tau$ for the variable interest payment $N \cdot R_i(T_{i-1}) \cdot \tau$ on each settlement date, $\{T_{j+1}, \ldots, T_n\}$.[5]

Pricing Forward IRS: First we start by pricing the floating leg using FTAP then the fixed leg for unit face value:

$$\begin{aligned}
&\text{Floating} \qquad \sum_{i=j+1}^{n} P_i(t) E_t^{Q_i}\left[R_i(T_{i-1})\tau\right] = \sum_{i=j+1}^{n} P_i(t) F_i(t) \tau = \sum_{i=j+1}^{n} \left(P_{i-1}(t) - P_i(t)\right) \\
&\qquad\qquad\qquad = P_j(t) - P_n(t) \\
&\text{Fixed} \qquad \sum_{i=j+1}^{n} P_i(t) K\tau = K\tau \sum_{i=j+1}^{n} P_i(t) = K\tau A(t, T_j, T_n) \\
&\text{Total value at } t \leq T_j \qquad IRS(t, T_j, T_n, K) = P_j(t) - P_n(t) - K\tau A(t, T_j, T_n)
\end{aligned}$$

Again, we typically want the value of the swap to have zero initial value, so the previous line implies that what we define as the *forward swap rate* at time t is[6]

$$FSR(t, T_j, T_n) = \frac{P_j(t) - P_n(t)}{\tau A(t, T_j, T_n)} = K$$

Notice that if the swap has a single payment date, the swap collapses to an FRA and the forward swap rate collapses to the forward rate:

$$IRS(t, T_{i-1}, T_i, K) = FRA(t, T_i, K), \qquad FSR(t, T_{i-1}, T_i) = F_i(t)$$

In fact, a little fiddling shows that the forward swap rate is a weighted average of forward rates,

$$FSR(t, T_j, T_n) = \frac{P_j(t) - P_n(t)}{\tau A(t, T_j, T_n)} = \sum_{i=j+1}^{n} w_i F_i(t), \qquad w_i = \frac{P_i}{A(t, T_j, T_n)}$$

Swaps are quoted in terms of the swap rate, FSR(0,0,T_n) with varying tenors. They are typically against $\tau = 6$ mo. LIBOR for Euro swaps but $\tau = 3$ mo. LIBOR for USD.

[3] Just a bit of terminology, the length of time $T_n - T_0$ is called the *tenor* of the swap.

[4] It's also a convention that the party who payes fixed is long, or has a *payer* IRS.

[5] In other words, the first fixing date, T_j, does not coincide with the contract date, T_0, as we have in the standard spot IRS.

[6] If we get to time $t = T_j$, however, then this simply becomes the swap rate.

3.4. Caps and Floors

Interest Rate Cap: Like a payer IRS, but cash flows are only settled if positive. Thus the net cash flow on a generic settlement date, T_i, is

$$N \cdot (R_i(T_{i-1}) - K)^+ \tau$$

The buyer of the cap pays the seller an up-front premium, and all subsequent cash flows are paid by the seller to the buyer.[7] Note that caps are typically *forward-starting*, so that the first fixing date is not $T_0 = 0$ (the contract date), but $T_1 = \tau$, where τ is the lag between settlement dates.

Interest Rate Floor: Like a receiver IRS, but cash flows are only settled if positive. Thus the net cash flow on a generic settlement date, T_i, is

$$N \cdot (K - R_i(T_{i-1})^+ \tau$$

Caplets and Floorlets: Similar to breaking up an IRS into a portfolio of FRAs, caps and floors can be thought of as a portfolio of options on the individual FRAs that compose the swap. We call these individual options "caplets" and "floorlets." In this way, caplets are call options on the interest rate, R_i, while floorlets are puts on the interest rate.

So we let $\text{Cap}(t, T_n, K)$ denote the value at $t < T_1$ of a cap with first fixing date T_1 and last settlement date T_n, and we also let $\text{Cpt}(t, T_i, K)$ denote the value of a caplet with fixing data T_{i-1} and settlement date T_i. Similar notation holds for floors and floorlets. This allows us to write

$$\text{Cap}(t, T_n, K) = \sum_{i=2}^{n} \text{Cpt}(t, T_i, K), \qquad \text{Flr}(t, T_n, K) = \sum_{i=2}^{n} \text{Flt}(t, T_i, K)$$

3.5. Cap-Floor Parity

Next, it's clear that the combination of long a cap and short a floor will generate the same cash flows as a forward IRS starting at T_1 so that we arrive at the formula for *cap-floor parity*:

$$\text{Cap}(t, T_n, K) - \text{Flr}(t, T_n, K) = IRS(t, T_1, T_n, K)$$

$$\text{Cpt}(t, T_i, K) - \text{Flt}(t, T_i, K) = FRA(t, T_i, K)$$

In the special case where $K = FSR(t, T_1, T_n$ in the first expression, then the value of the IRS is zero and the cap has the same value as the corresonding floor. Thus the forward swap rate (FSR) is the at-the-money strike for caps and floors.

[7] The reason it's called a cap is because it compensates you if floating interest rates rise above a certain level, allowing people with floating liabilities to hedge. Obviously, caps only make sense for those with floating liabilities, where there is some uncertainty in future cash flows. Contrast this with those who have fixed liabilities, who know *exactly* what they will have to pay.

3.6. Black's Formula for Caps and Floors

3.6.1. Pricing

A simple variation of the Black formula allows us to value a caplet under certain assumptions. FTAP and an identity give us

$$\begin{aligned}\mathrm{Cpt}(t,T_i,K) &= P_i(t)E_t^{Q_i}\left[(R_i(T_{i-1})-K)^+\tau\right]\\ &= P_i(t)E_t^{Q_i}\left[(F_i(T_{i-1})-K)^+\right]\tau\end{aligned}$$

This expression makes clear that a caplet is simply a call option on the forward rate, F_i. From there, we recall that F_i is a martingale under Q_i so that

$$dF_i(t) = F_i(t)\sigma_i(t)d\hat{W}_i(t)$$

where σ_i is the proportional volatility of the forward rate and $\hat{W}_i$ is a Brownian motion under Q_i. By Ito's Lemma, we get that

$$\ln F_i(T_{i-1}) = \ln F_i(t) - \frac{1}{2}\int_t^{T_{i-1}} \sigma_i(s)^2\, ds + \int_t^{T_{i-1}} \sigma_i(s)\, d\hat{W}_i(s)$$

If we assume that $\sigma(s)$ is deterministic, then

$$F_i(T_{i-1}) \sim N\left(-\frac{1}{2}\bar{\sigma}_i^2(T_{i-1}-t),\;\; \bar{\sigma}_i^2(T_{i-1}-t)\right), \qquad \bar{\sigma}_i^2 = \frac{1}{T_{i-1}-t}\int_t^{T_{i-1}} \sigma_i(s)^2\, ds$$

$$\begin{aligned}\mathrm{Cpt}_B(t,T_i,K) &= P_i(t)\left[F_i(t)N(y_i) - KN\left(y_i - \bar{\sigma}_i\sqrt{T_{i-1}-t}\right)\right]\tau\\ \text{where}\quad y_i &= \frac{1}{\bar{\sigma}_i\sqrt{T_{i-1}-t}}\left[\ln\left(\frac{F_i(t)}{K}\right)\right] + \frac{1}{2}\bar{\sigma}_i\sqrt{T_{i-1}-t}\end{aligned} \tag{8}$$

Under the assumptions of the Black formula, the caplet implied volatility is equal to the average proportional volatility of the underlying *forward* rate. The formula for a floorlet is similar

$$\mathrm{Flt}_B(t,T_i,K) = P_i(t)\left[KN\left(-y_i + \bar{\sigma}_i\sqrt{T_{i-1}-t}\right) - F_i(t)N(-y_i)\right]\tau \tag{9}$$

where y_i is as above. Putting everything together, we get that

$$\mathrm{Cap}_B(t,T_n,K) = \sum_{i=2}^{n}\mathrm{Cpt}_B(t,T_i,), \qquad \mathrm{Flr}_B(t,T_n,K) = \sum_{i=2}^{n}\mathrm{Flt}_B(t,T_i,) \tag{10}$$

3.6.2. Implied Volatilies

As with equities, market quotes for caps and floors are typically provided in terms of Black implied volatilities. So if Cap_M is the market price, the cap implied volatility, σ_n^{Cap} for a cap with maturity T_n will be defined so that

$$\text{Cap}_M(t, T_n, K) = \sum_{i=2}^{n} \text{Cpt}(t, T_i, K; \sigma_n^{\text{Cap}})$$

In this way, cap implied volatilities are averages of the implied volatilities of the different caplets that compose the cap.

If we were to actually look at a plot of the Black Implied Volatilities for European at the money caps, we'd see a hump shape with noticeable right skew. This comes about for a few reasons:

1. Cap IVs are averages of the IVs of the component caplets.
2. Caplet IVs are related to the percentage volatilities of forward rates.
3. Empirically, these percentage volatilities of forward rates also display a hump figure similar to what we'd see in the IVs.

Why do we see this empirical hump in the volatilities of forward rates? Well, it makes a lot of sense.

- Forward rates with a short time to maturity are closely tied to spot rates over the short term. As a result, they don't vary much as the Federal Reserve sets the interest rate.
- Forward rates over the long term build in effects over several business cycles, so the result is a sort of averaging. This means lower observed volatility.
- Finally, in the medium term, business cycles can have an impact and the time horizon is long enough for Federal Reserve policy to change. This generates the hump shape.

3.7. Swaptions

So far, we've seen forward rate agreements, swaps on interest rates, along with caps and floors, which are a type of option on interest rates. Now, let's look at *swaptions*, which are options on interest rate *swaps*. There are two basic types:

- *Payer Swaption*: Gives the right to enter into a payer IRS upon exercise.
- *Receiver Swaption*: Gives the right to enter into a receiver IRS upon exercise.

In either case, the fixed rate in the IRS is equal to the strike of the swaption, and *no* price is paid upon exercise. Typically, the swap starts immediately upon exercise so that the first reset date of the underlying IRS coincides with the swaptions exercise date.

Now, if we let $\mathrm{PSO}(t, T_j, T_n, K)$ denote the value at time t of a a payer swaption with expiration T_j on a swap with unit notional and tenor $T_n - T_j$ (so that it's a $T_j - t$ years into $T_n - T_j$ years swaption), we can write the value of a payor swaption on the expiration date (using identities and results derived above) as follows:

$$\begin{aligned}\mathrm{PSO}(T_j, T_j, T_n, K) &= \mathrm{IRS}(T_j, T_j, T_n, K)^+ \\ &= (P_j(T_j) - P_n(T_j) - K\tau A(T_j, T_j, T_n))^+ \\ &= A(T_j, T_j, T_n)\left(\mathrm{FSR}(T_j, T_j, T_n) - K\right)^+ \tau\end{aligned}$$

If we consider the payoff from a swaption on a swap with a single settlement date T_n, we get the same payoff as a caplet settled on the fixing date:

$$\begin{aligned}\mathrm{PSO}(T_{n-1}, T_{n-1}, T_n, K) &= P_n(T_{n-1})(R_n(T_{n-1}) - K)^+\tau \\ \Rightarrow \quad \mathrm{PSO}(t, T_{n-1}, T_n, K) &= \mathrm{Cpt}(t, T_n, K)\end{aligned}$$

Similarly for receiver swaptons:

$$\begin{aligned}\mathrm{RSO}(T_j, T_j, T_n, K) &= (-\mathrm{IRS}(T_j, T_j, T_n, K))^+ \\ &= (K\tau A(T_j, T_j, T_n) + P_n(T_j) - P_j(T_j))^+ \\ &= A(T_j, T_j, T_n)\left(K - \mathrm{FSR}(T_j, T_j, T_n)\right)^+ \tau\end{aligned}$$

$$\mathrm{RSO}(t, T_{n-1}, T_n, K) = \mathrm{Flt}(t, T_n, K)$$

3.8. Swaption Parity

Clearly, long a payer swaption and short the corresponding receiver swaption will be equivalent to a long position in a forward IRS:

$$\mathrm{PSO}(t, T_j, T_n, K) - \mathrm{RSO}(t, T_j, T_n, K) = \mathrm{IRS}(t, T_j, T_n, K)$$

This relationship is called *swaption parity*. Note that if $K = \mathrm{FSR}(t, T_j, T_n)$, the value of the payer swaption is the same as the value of the receiver swaptions. This means that the forward swap rate is the at-the-money strike for swaptions—just as for caps and floors.

3.9. Black's Formula for Swaptions

3.9.1. Pricing

Just like for caps and floors, we can price swaptions using a version of the Black Formula. To do so, we'll use annuities as our numeraire. So recall that $A(t, T_j, T_n)$ is the value at time t of an annuity paying \$1 at times $\{T_{j+1}, T_{j+2}, \ldots, T_n\}$. Because there are no intermediate payments before T_{j+1}, we can take this as our numeraire, which we'll denote Q_A and call the *forward swap measure*. This allows us to write

$$\begin{aligned}
\mathrm{PSO}(t, T_j, T_n, K) &= A(t, T_j, T_n) E_t^{Q_A} \left[\frac{A(T_j, T_j, T_n)(\mathrm{FSR}(T_j, T_j, T_n) - K)^+ \tau}{A(T_j, T_j, T_n)} \right] \\
&= A(t, T_j, T_n) E_t^{Q_A} \left[(\mathrm{FSR}(T_j, T_j, T_n) - K)^+ \right] \tau
\end{aligned}$$

Now its clear that

$$\mathrm{FSR}(t, T_j, T_n) = \frac{P_j(t) - P_n(t)}{\tau A(t, T_j, T_n)}$$

must be a martingale under Q_A because it represents the discounted price process of assets available in the market. So assuming the FSR is an Ito process, we have

$$d\mathrm{FSR}(t, T_j, T_n) = \mathrm{FSR}(t, T_j, T_n) \sigma_{\mathrm{FSR}}(t)\, d\hat{W}(t)$$

where σ_{FSR} is the percentage volatility of the forward swap rate and $\hat{W}(t)$ is a Brownian motion under Q_A. If we assume that σ_{FSR} is deterministice, it follows that $\mathrm{FSR}(T_j, T_j, T_n)$ is lognormally distributed under Q_A and the value of payer and receiver options are

$$\begin{aligned}
\mathrm{PSO}_B(t, T_j, T_n, K) &= A(t, T_j, T_n) \Big(\mathrm{FSR}(t, T_j, T_n) N(y) - K N(y - \bar{\sigma}_{\mathrm{FSR}} \sqrt{T_j - t}) \Big) \tau \\
\mathrm{RSO}_B(t, T_j, T_n, K) &= A(t, T_j, T_n) \Big(K N(-y + \bar{\sigma}_{\mathrm{FSR}} \sqrt{T_j - t}) - \mathrm{FSR}(t, T_j, T_n) N(-y) \Big) \tau \\
y &= \frac{1}{\bar{\sigma}_{\mathrm{FSR}} \sqrt{T_j - t}} \left[\ln \left(\frac{\mathrm{FSR}(t, T_j, T_n)}{K} \right) \right] + \frac{1}{2} \bar{\sigma}_{\mathrm{FSR}} \sqrt{T_j - t} \\
\bar{\sigma}_{\mathrm{FSR}}^2 &= \frac{1}{T_j - t} \int_t^{T_j} \sigma_{\mathrm{FSR}}(s)^2\, ds
\end{aligned}$$

3.9.2. Implied Volatilities

As with caps and floors, swaption quotes are typicall provided in terms of Black Implied Volatilities. Quotes are typically arranged in a volatility matrix in which rows correspond to different expirations and columns correspond to different swap tenors.

3.10. Derivatives on Bonds, Rather than Interest Rates

In previous sections, we derived everything in terms of derivatives on interest rates (including forward or forward swap rates). As we go on, however, it will be more useful to think of caps, floors, and swaptions as options on *bonds* rather than on *interest rates*. This section examines how we can transform payoffs of these derivatives into payoffs of standard vanilla options on bonds.

Caplets and Floorlets: For these particular options, there is a relation to the prices of Zero Coupon Bonds (ZCBs):

$$\begin{aligned}\mathrm{Cpt}(T_{i-1}, T_i, K) &= P_i(T_{i-1})(R_i(T_{i-1}) - K)^+\tau \\ &= P_i(T_{i-1})\left(\frac{1}{P_i(T_{i-1})} - 1 - K\tau\right)^+ \\ &= (1 + K\tau)\left(\frac{1}{1+K\tau} - P_i(T_{i-1})\right)^+\end{aligned}$$

This is the payoff of $1 + K\tau$ puts with strike $1/(1 + K\tau)$ and expiration T_{i-1} on the Zero Coupon Bond (ZCB) with maturity T_i.

Similarly, the value of a floorlet with fixing date T_{i-1}, settlement date T_i, and strike K is equal to the value of $1 + K\tau$ calls with strike $1/(1 + K\tau)$ and expiration T_{i-1} on the ZCB with maturity T_i.

Swaptions: Rather than ZCBs, swaptions can be expressed as options on *coupon* bonds. Recall the payoff of a payer swaption on the expiration date T_j:

$$\begin{aligned}\mathrm{PSO}(T_j, T_j, T_n, K) &= (P_j(T_j) - P_n(T_j) - K\tau A(T_j, T_j, T_n))^+ \\ &= (1 - [P_n(T_j) + K\tau A(T_j, T_j, T_n)])^+\end{aligned}$$

which is the payoff of a put with strike 1 and expiration T_j on a coupon bond with coupon rate K and maturity T_n. This means the value of a payor swaption with expiration T_j and strike K on a swap with tenor $T_n - T_j$ is equal to the value of a put with expiration T_j and unit strike on a coupon bond with coupon rate K and maturity T_n.

4. One-Factor Spot Rate Models

4.1. Uses of Interest Rate Models

Interest rate models serve one or more of three main purposes:

1. Forecasting future interest rates.
2. Pricing bonds.
3. Pricing interest rates derivatives.

Overall, there's a fair amount of overlap, but also some extremely important differences. In particular, the first task operates under the *real-world* probability measure, while the second two operate under various *martingale* measures.

In addition, while the second two both require operations under martingale measure, they differ in their goals. In the case of pricing bonds, it's not crucial that the model generates prices for zero-coupon bonds (ZCBs) which perfectly match the current term structure. In fact, you typically want to capture mispricing, so you'll want some discrepancy between your model prices and the observeable market prices.

On the other hand, when pricing interest rate derivatives, it doesn't matter whether bonds are valued *correctly*.[8] Rather, we just care about the price at which we can buy and sell the asset so that we can form a replicating portfolio. With this in mind, we would therefore like our models to match the initial term structure—which determines the prices at which we can buy and sell bonds—*exactly*.

[8]Considering an analagous situation, recall that When we valued equity derivatives, we didn't check whether the stock price equaled the discounted expected future cash flows—which is the "correct" price.

4.2. General Form of One-Factor Models

One-factor spot-rate models assume that the instantaneous spot rate $r(t) = \lim_{T \to t} r(t,T)$ follow a diffusion:

$$dr(t) = \mu(r(t), t)\ dt + \sigma(r(t), t)\ d\hat{w}(t) \tag{11}$$

where $\hat{w}(t)$ is Brownian motion under the risk neutral measure, Q_B. By FTAP, we know that the price of a riskless ZCB should be

$$P(t,T) = B(t) E_t^{Q_B}\left[\frac{1}{B(T)}\right] = E_t^{Q_B}\left[\exp\left(-\int_t^T r(s)\ ds\right)\right] \tag{12}$$

In this way, we can model the *entire* term structure just based on the short term spot rate.

In general, we will want $P(t,T)$ as expressed in Equation 12 to have an explicit, closed form solution, which becomes clear if we consider how we would price a call on a ZCB with expiration $t < T$ and maturity T. If today is t_0, suppose that we have an option on ZCB that expires at time t. Then we want to simulate the price $P(t,T)$ given what we know today, $P(t_0, T)$. With a closed form solution, it's easy to take draws for what's inside expectation operator in Equation 12 and average. Otherwise, we'd have to simulate interest rate paths form t_0 to t, then from t to T in order to compute the price $P(t,T)$. This would lead to exponentialy growth in the number of simulations.

4.3. General Form of Affine One-Factor Models

Most one-factor models assume an affine process to satisfy the requirement that ZCB prices be available in closed form,[9] Models of the following form are called *affine one-factor models*:

$$dr(t) = (\theta(t) - \kappa(t) r(t))\ dt + \sqrt{\sigma_0(t)^2 + \sigma_1(t)^2 r(t)}\ d\hat{w}(t) \tag{13}$$

It can also be shown that in any affine one-factor model, ZCB bond prices are given by

$$P(t,T) = e^{a(t,T) - b(t,T) r(t)} \tag{14}$$

where the functions a and b solve the two first-order ODEs

$$\begin{aligned} b_t(t,T) &= b(t,T)\kappa(t) + \frac{1}{2} b(t,T)^2 \sigma_1(t)^2 - 1 \\ a_t(t,T) &= b(t,T)\theta(t) - \frac{1}{2} b(t,T)^2 \sigma_0(t)^2 \end{aligned}$$

with boundary conditions $b(T,T) = 0$ and $a(T,T) = 0$.

[9] Recall the definition of an affine process, which stipulates a process whose drift and variance rate (which is volatility squared) are *affine* (linear) functions of the level of the process. While no general solution exists to the SDE that defines an affine process, the Fourier transform of the probability density of the value of the process at any time t is known.

4.4. Vasicek Model

The most basic model, the *Vasicek Model*, is a homogenous affine one-factor spot rate model of the form

$$dr(t) = \kappa(\theta - r(t))\, dt + \sigma\, d\hat{w}(t), \qquad \kappa \geq 0, \quad \theta, \sigma > 0 \tag{15}$$

It has a few desireable properties:

- The spot rate, $r(t)$, is mean reverting if $\kappa > 0$.
- The functions a and b in Equation 14 can be computed explicitly.
- Explicit formulas can be derived for vanilla bond options and more complext derivatives.

It does, however have important disadvantages in that the volatility is not proportional to the level of interest rates, and negative interest rates are not ruled out.

Solution to the Vasicek Model: Now let's solve the SDE to get a closed form solution for the interest rate. To do so, we set $z(t) = r(t)e^{\kappa t}$, get the partials, and apply Ito's Lemma:

$$\frac{\partial z}{\partial r} = e^{\kappa t}, \qquad \frac{\partial z^2}{\partial r^2} = 0, \qquad \frac{\partial z}{\partial t} = \kappa r(t) e^{\kappa t}$$

$$\begin{aligned} dz(t) &= \frac{\partial z}{\partial t} dt + \frac{\partial z}{\partial r} dr + \frac{1}{2}\frac{\partial^2 z}{\partial r^2}\sigma^2 dt \\ &= \kappa r(t) e^{\kappa t}\, dt + e^{\kappa t}\left[\kappa\left(\theta - r(t)\right)\, dt + \sigma\, d\hat{w}(t)\right] + 0 \\ &= \kappa\theta e^{\kappa t}\, dt - e^{\kappa t}\sigma\, d\hat{w}(t) \\ \Rightarrow d\left(r(t)e^{\kappa t}\right) &= \kappa\theta e^{\kappa t}\, dt - e^{\kappa t}\sigma\, d\hat{w}(t) \end{aligned}$$

Since the right hand side does not contain $r(t)$, we can take Ito-integrate everything:

$$\begin{aligned} \int_0^t d\left(r(s)e^{\kappa s}\right) &= \int_0^t \kappa\theta e^{\kappa s}\, ds + \int_0^t e^{\kappa s}\sigma\, d\hat{w}(s) \\ e^{\kappa t} r(t) - e^{\kappa \cdot 0} r(0) &= \theta(e^{\kappa t} - 1) + \sigma \int_0^t e^{\kappa s} d\hat{w}(s) \\ \Rightarrow \quad r(t) &= r(0)e^{-\kappa t} + \theta(1 - e^{-\kappa t}) + \sigma \int_0^t e^{-\kappa(t-s)} d\hat{w}(s) \end{aligned}$$

This implies that given $r(0)$, $r(t)$ is normally distributed with[10]

$$\begin{aligned} Er(t) &= r(0)e^{-\kappa t} + \theta(1 - e^{-\kappa t}) \\ Var[r(t)] &= Var\left(\sigma \int_0^t e^{-\kappa(t-s)} d\hat{w}(s)\right) = \sigma^2 \int_0^t e^{-2\kappa(t-s)}\, ds \\ &= \frac{\sigma^2}{2\alpha}\left(1 - e^{-2\kappa t}\right) \end{aligned}$$

[10]The variance term uses a standard Ito Calculus result $Var\left(\int_0^t f(s)\, dW(s)\right) = \int_0^t f^2(s)\, ds$.

4.5. Cox-Ingersoll-Ross (CIR) Model

An alternative to the Vasicek model, the CIR model is another homogeneous affine one-factor model:

$$dr(t) = \kappa(\theta - r(t))\, dt + \sigma\sqrt{r(t)}\, d\hat{w}(t), \qquad \kappa, \theta, \sigma > 0 \tag{16}$$

It retains the mean reversion feature along with the explicit formula result while eliminating the possibility of negative interest rates.

The chief disadvantage, however, of both the Vasicek and CIR models is their inability to match the current term structure—a sympton of the relatively small amount of free parameters.

4.6. Ho-Lee Model

The Ho-Lee model was an attempt to develop a model that fits the initial term structure better. It has characterization

$$dr(t) = \theta(t)\ dt + \sigma\ d\hat{w}(t), \qquad \sigma > 0 \tag{17}$$

In this process, θ is *time dependent*. This has a number of intractible features, namely that interest rates are not mean-reverting and that there is a high probability of negative interest rates.

Solution to the Ho-Lee Model:

4.7. Extended Vasicek Model

4.7.1. Governing Process

The *Extended Vasicek* (or *Hull-White*) model is a non-homogeneous affine one-factor model that assumes

$$dr(t) = \kappa(\theta(t) - r(t))\, dt + \sigma d\hat{w}(t) \tag{18}$$

Clearly, this is identical to the traditional Vasicek model with the crucial exception that $\theta(t)$ is time dependent.[11] Using Ito's Lemma, with $z(t) = r(t)e^{\kappa t}$ it has solution

$$\begin{aligned} r(t) &= e^{-\kappa t}\left(r(0) + \kappa \int_0^t e^{\kappa s}\theta(s)\, ds + \sigma \int_0^t e^{\kappa s}\, d\hat{w}(s)\right) \\ \Rightarrow \quad r(t) &\sim N\left(e^{-\kappa t}\left[r(0) + \kappa \int_0^t e^{\kappa s}\theta(s)\, ds\right],\ \frac{1-e^{-2\kappa t}}{2\kappa}\sigma^2\right) \end{aligned} \tag{19}$$

That's one characterization, but we could express the EV model another way by examining the soltuion and letting

$$\begin{aligned} \varphi(t) &= \kappa \int_0^t e^{-\kappa(t-s)}\theta(s)\, ds \\ x(t) &= e^{-\kappa t}\left(r(0) + \sigma \int_0^t e^{\kappa s}\, d\hat{w}(s)\right) \\ \Rightarrow r(t) &= \varphi(t) + x(t) \end{aligned}$$

It's clear that $\varphi(t)$ is entirely deterministic; only $x(t)$ is random. Moreover, by applying Ito's lemma,[12] we can express

$$dx(t) = -\kappa x(t)\, dt + \sigma\, d\hat{w}(t), \qquad x(0) = r(0)$$

which is an affine process. This formulation is often more convenient to work with, and we'll use it to price ZCBs.

[11] As with the Vasicek model, the interest rate is mean-reverting and also has the unfortunate possibility of becoming negative. However, as we see in the solution, the probability of negative interest rates is small provided that θ and κ are large while σ is small.

[12] To do so, we rewrite the the expression for $x(t)$ as

$$\begin{aligned} e^{\kappa t}x(t) - e^{\kappa 0}x(0) &= \sigma \int_0^t e^{\kappa s}\, d\hat{w}(s), \qquad x(0) = r(0) \\ \int_0^t d\left(e^{\kappa s}x(s)\right) &= \sigma \int_0^t e^{\kappa s}\, d\hat{w}(s), \quad \Leftrightarrow \quad d\left(e^{\kappa t}x(t)\right) = \sigma e^{\kappa t}\, d\hat{w}(t) \end{aligned}$$

From there, we define $y(t) = e^{\kappa t}x(t)$, implying $x(t) = e^{-\kappa t}y(t)$. Since we have $dy(t)$, we can use Ito's Lemma to get $dx(t)$.

4.7.2. Pricing Bonds under the EV Model

Now that we have a reasonable model of interest rate dynamics, let's price zero-coupon bonds. Recall how we do that:

$$\begin{aligned} P(t,T) &= E\left[e^{\int_t^T r(s)\,ds}\right] = E\left[e^{\int_t^T \varphi(s)+x(s)\,ds}\right] \\ &= e^{\int_t^T \varphi(s)\,ds} E\left[e^{\int_t^T x(s)\,ds}\right] \end{aligned}$$

where we could pull the exponential of $\varphi(\cdot)$ in front because it is entirely deterministic. To compute the expectation, let's do a couple intermediate steps:

1. We can compute the expectation of $x(t)$ simply and in closed form by using Ito's lemma similarly to how we used it for the complete $r(t)$. If we do so, we get:

$$\begin{aligned} & x(t) = e^{-\kappa t}\left(x(0) + \sigma \int_0^t e^{\kappa s}\, d\hat{w}(s)\right) \\ \Rightarrow \quad & E[x(t)] = x(0)e^{-\kappa t} \end{aligned}$$

2. From there, we can use that to compute the expectation of the integral, which will be useful later on:

$$\begin{aligned} E\left[\int_0^t x(s)\,ds\right] &= E\left[\int_0^t e^{-\kappa s}\left(x(0) + \sigma \int_0^s e^{\kappa u}\, d\hat{w}(u)\right)\, ds\right] \\ &= E\left[x(0)\int_0^t e^{-\kappa s}\right] + E\left[\int_0^t e^{-\kappa s}\left(\sigma\int_0^s e^{\kappa u}\, d\hat{w}(u)\right)\, ds\right] \\ &= x(0)E\left[\int_0^t e^{-\kappa s}\right] + \sigma E\left[\int_0^t \int_0^s e^{\kappa(u-s)}\, d\hat{w}(u)\, ds\right] \\ &= \frac{x(0)}{\kappa}\left[1 - e^{-\kappa t}\right] \end{aligned}$$

The second term drops because it's an expecation of Brownian Increments which all have expectation 0.[13]

3. Next, we want the variance of $x(t)$.

FINISH DERIVATION
At the very end, we see that ZCB prices are

$$\begin{aligned} P(t,T) &= E_t^{Q_B}\left[e^{-\int_t^T r(s)\,ds}\right] \\ &= E_t^{Q_B}\left[e^{-\int_t^T \varphi(s)+x(s)\,ds}\right] = e^{-\int_t^T \varphi(s)\,ds} E_t^{Q_B}\left[e^{-\int_t^T x(s)\,ds}\right] \\ &= e^{-\int_t^T \varphi(s)\,ds - a(T-t) - b(T-t)x(s)} \end{aligned} \tag{20}$$

$$\text{where} \quad a(\tau) = \left(\frac{b(\tau)-\tau}{\kappa} + \frac{b(\tau)^2}{2}\right)\frac{\sigma^2}{2\kappa}, \qquad b(\tau) = \frac{1-e^{-\kappa\tau}}{\kappa}$$

[13] You see this if you switch the order of integration, integrate out s, then break it up into two Ito integrals, which will have expecation zero.

Equation 20 follows from the fact that $x(t)$ is an affine process and can be expressed as in Equation 14.

4.7.3. Bond Price Dynamics

4.7.4. Calibration

To calibrate the model, we'll need to specify φ so that the model matches the term structure on the calibration date, time $t = 0$. This means we choose φ so that

$$
\begin{aligned}
P(0,T) = e^{-\int_0^T \varphi(s)\,ds - a(T) - b(T)x(0)} &= P_M(0,T) \\
\Leftrightarrow \quad e^{-\int_0^T \varphi(s)\,ds} &= e^{a(T)+b(T)x(0)} P_M(0,T)
\end{aligned}
$$

for all T, so that our calibration matches the *entire* term structure. If φ is, in fact, chosen so that this is true for all T, then it follows that

$$
\begin{aligned}
e^{-\int_t^T \varphi(s)\,ds} &= \frac{e^{-\int_0^T \varphi(s)\,ds}}{e^{-\int_0^t \varphi(s)\,ds}} \\
&= \frac{e^{a(T)+b(T)x(0)} P_M(0,T)}{e^{a(t)+b(t)x(0)} P_M(0,t)}
\end{aligned}
$$

This means that we can can plug this into Equation 20 and use the fact that $x(0) = r(0)$ to get

$$
\begin{aligned}
P(t,T) &= A(t,T) e^{-b(T-t)x(t)} \\
A(t,T) &= \frac{e^{a(T)+b(T)r(0)} P_M(0,T)}{e^{a(t)+b(t)r(0)} P_M(0,t)} e^{-a(T-t)}
\end{aligned}
$$

We now have the price of ZCBs expressed as a function of the initital term structure, a stochastic factor $x(t)$, and two constant parameters—κ and σ.

Since $x(t)$ is normally distributed, it follows that $P(t,T)$ are log-normally distributed, while simply compounded spot rates, which are expressed

$$
R(t,T) = \frac{1}{T-t}\left(\frac{1}{P(t,T)} - 1\right),
$$

will have a shifted log-normal distribution, while continuously compounded rates are normally distributed.

4.7.5. Pricing Coupon Bonds and Payer Swaptions

4.8. Cox-Ingersoll Ross Extension

4.8.1. CIR++ versus Extended CIR

4.8.2. Pricing ZCB's under the CIR++ Model

5. Multi-Factor Models

A. Principal Component Analysis

A.1. Motivation and Setup

Principal component analysis is relevant for Fixed Income in order to determine how many factors a multi-factor spot rate model might need. Recall that we suppose our zero-coupon bonds can be priced by

$$\begin{aligned} P(t,T) &= A(t,T)e^{-\sum_{i=1}^{n} b_i(T-t)x_i(t)} \\ \Rightarrow \quad r(t,T) &= -\frac{\ln A(t,T)}{T-t} + \sum_{i=1}^{n} \frac{b_i(T-t)}{T-t} x_i(t) \end{aligned} \tag{21}$$

We start by assuming we have a time series of interest rates with constant times to maturity: $\tau_1, \ldots, \tau_N$. So we might observe the 1-year rate over time, the 2-year rate over time, etc.

Next, we let $\mathbf{r}(t)$ denote the N-dimensional vector of observed rates at time t. It has ith entry $\mathbf{r}_i(t) = r(t,\ t+\tau_i)$. Then if the observed rates are determined by an n factor model, Equation 21 tells us that

$$\begin{aligned} &\mathbf{r}(t) = \mathbf{a}(t) + \mathbf{B}\mathbf{x}(t) & \\ \text{where} \quad &\mathbf{a}_i(t) = -\ln\left(A(t,t+\tau_i)/\tau_i\right), \qquad & \mathbf{a}_i(t) \in \mathbf{a}(t) \subset \mathbb{R}^N \\ &\mathbf{x}_j(t) = x_j(t), & \mathbf{x}_i(t) \in \mathbf{x}(t) \subset \mathbb{R}^n \\ &\mathbf{B}_{iJ} = b_j(\tau_i)/\tau_i & \mathbf{B}_{iJ} \in \mathbf{B} \subset \mathbb{R}^N \times \mathbb{R}^n \end{aligned}$$

Then we let $\Delta\mathbf{r}(t)$ and $\Delta\mathbf{x}(t)$ denote the demeaned change $\mathbf{r}(t)$ and $\mathbf{x}(t)$, respectively, so that

$$\Delta\mathbf{r}(t) = \mathbf{B}\Delta\mathbf{x}(t) \tag{22}$$

If we assume that the factors are not linearly dependent, then the rank of the $N \times N$ covariance matrix of $\Delta\mathbf{r}$ (denoted by V) will be $\min\{n, N\}$. So if we could measure the covariance matrix, V, without error, we could estimate the number of factors by simply obtaining data on $\Delta\mathbf{r}$ with N sufficiently large so that $N > n$ and then computing the rank of V.

Measurement Error: However, because there will be measurement/sampling error, this won't work. Even if $\text{rank}(V) = n < N$, measurement errors will distort linear independence between rows and columns. So we'll try to find the "best" possible approximation of the observed covariance matrix, V, by a covariance matrix of rank n, denoted V_n. Then, we'll determine how large n has to be in order to obtain an acceptable approximation.

A.2. Approximating the Covariance Matrix

We'll want to approximate the covariance matrix with a matrix of lower rank. In doing so, we'll need to use eigenvectors. Here's the main process: If V, an $N \times N$ covariance matrix, then

1. V has exactly N distinct eigenvectors and the rank of V is equal to the number of non-zero eigenvalues.

2. Let Λ_N be the $N \times N$ diagonal matrix whose diagonal elements are the eigenvalues of V. Also, let Q_N be the $N \times N$ matrix whose columns are the eigenvectors of V. Then we have $V = Q_N \Lambda_N Q_N^T$.

3. The best best rank $n \leq N$ approximation of V is

$$V_n = Q_n \Lambda_n Q_n^T$$

   where Λ_n is an $n \times n$ diagonal matrix whose diagonal elements are the n largest eigenvalues of V and Q_n is the $N \times n$ matrix whose columns are the eigenvectors corresonding to the n largest eigenvalues of V.

The best rank n approximation V_n to covariance matrix V is the covariance matrix corresponding to a n-factor model, and the factors in this n-factor model are called the first n principal components of $\mathbf{r}$.

Note that it also follows that

$$\Delta \mathbf{x}(t) = Q_n^T \Delta \mathbf{r}(t)$$

$$\mathbf{B} = Q_n$$

Moreover, the n factors jointly explain a fraction of the total variance of the interest rates included in $\Delta \mathbf{r}$:

$$\frac{\sum_{i=1}^{n} \lambda_i}{\sum_{i=1}^{N} \lambda_i}$$

where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of V sorted in decreasing order.

B. Antithetic Variates: A Variance Reduction Technique

The technique of *antithetic variates* reduces the sample size needed to achieve a given precision (in the form of standard error) in Monte Carlo simulations. Let's go through the steps:

1. Suppose that we want to estimate EY using MC simulation. Instead of simulating a single sample of size N, we simulate $N/2$ pairs, $(Y_1, Y_2)_i$, where each pair is i.i.d. for all pairs $i = 1, \dots, N/2$.

2. We can get an unbiased estimator of EY by first computing the average value $\bar{Y}_i = \frac{1}{2}(Y_{i1} + Y_{i2})$ then averaging across the $\bar{Y}_i$.

3. If we simulate the $N/2$ pairs independently, then

$$\begin{aligned} Var(\bar{Y}_i) = Var\left(\frac{1}{2}(Y_{i1} + Y_{i2})\right) &= \frac{1}{4}\left[Var(Y_{i1}) + Var(Y_{i2}) + 2Cov(Y_{i1}, Y_{i2})\right] \\ &= \frac{(1+\rho)}{2} Var(Y), \quad \rho \text{ is the correlation between } Y_{i1},\, Y_{i2} \end{aligned}$$

   So then if we want the total variance of Y over our entire simulation, then we get

$$Var(\bar{Y}) = \frac{(1+\rho)Var(Y)}{2(N/2)} = \frac{(1+\rho)Var(Y)}{N}$$

4. Clearly, by making the correlation negative between the terms Y_1 and Y_2 in each given pair, we can lower the variance of our estimator for EY as a whole. (Note that the pairs themselves are still independent from pair to pair.)

Note that some of the negative correlation will wash out and we won't get perfect -1 correlation when we jump from random draws to prices, as typically the prices are some convoluted *function* of the random draws.

Also, the gain—while significant—for path *dependent* derivatives is not quite as large due to the dependence of the payoffs upon the entire path.