Continuous-Time Short Rate Models

These notes provide an overview of single- and multi-factor models of the short rate. We will begin with a generic single-factor model where the dynamics of r_t under the physical measure, P, are given. Following the approach of Vasicek, we then derive the PDE that must be satisfied by derivative security prices. We then use the martingale approach to give an alternative (and familiar) expression for derivative security prices. The consistency of the two approaches is then demonstrated using the Feynman-Kac PDE representation. We show how the Martingale Representation theorem can be used to construct hedging strategies and then discuss some specific single-factor models. These examples include the Vasicek and CIR models, and more generally, affine models. We then conclude with multi-factor models and describe some specific examples.

Once we have moved on to discussing specific models, we will generally describe the dynamics of r_t (and other factors in multi-factor models) under the EMM, Q, and not bother with the P-dynamics of r_t. This is the standard approach in term-structure modelling and was the approach we followed when we used binomial and trinomial lattices to model the short rate.

1 The PDE Approach

Let us assume that the P-dynamics of the short rate are given by the SDE

$$dr_t \;=\; \alpha_p(t, r_t)\; dt \;+\; \beta_p(t, r_t)\; dW_t^P. \tag{1}$$

where W_t^P is a P-Brownian motion. In particular, r_t is a Markov process. Consider now the time t price, Z_t^T, of a zero-coupon bond maturing at time T. We assume that Z_t^T is a sufficiently smooth function of (t, r_t) so that Itô's Lemma implies

$$dZ_t^T \;=\; \left[\frac{\partial Z^T}{\partial t} + \alpha_p(t,r_t)\frac{\partial Z^T}{\partial r} + \frac{1}{2}\beta_p(t,r_t)^2\frac{\partial^2 Z^T}{\partial r^2}\right]\; dt \;+\; \beta_p(t,r_t)\frac{\partial Z^T}{\partial r}\; dW_t^P \tag{2}$$

$$\;=\; Z_t^T\left[m(t,r;T)\; dt + S(t,r;T)\; dW_t^P\right] \tag{3}$$

where $m(t,r;T)$ and $S(t,r;T)$ are implicitly defined[1] by (3). Now consider times $t < T_1 < T_2$ where T_1 and T_2 are fixed. We will first construct a self-financing portfolio (portfolio A) involving[2] the cash account, B_t, and $Z_t^{T_2}$ that replicates $Z_t^{T_1}$ (portfolio B).

Portfolio A: Hold a_t units of $Z_t^{T_2}$ and b_t units of B_t at time t.

Portfolio B: Hold 1 unit of $Z_t^{T_1}$ at time t.

We choose a_t and b_t such that portfolio A is self-financing[3] and the two portfolios have equal value, i.e., $V_t^A = V_t^B$ where $V_t^A := a_t Z_t^{T_2} + b_t B_t$ and $V_t^B := Z_t^{T_1}$. We then have

$$a_t\; dZ_t^{T_2} + b_t\; dB_t \;=\; dZ_t^{T_1}. \tag{4}$$

Exercise 1 *Why does (4) imply that portfolio A is self-financing?*

[1]We will also use $m_t^{T_i}$ and $S_t^{T_i}$ to denote $m(t,r;T_i)$ and $S(t,r;T_i)$, respectively.

[2]It is a slight abuse of notation to use $Z_t^{T_i}$ to refer to both the zero-coupon bond price and the zero-coupon bond itself but we do so to avoid introducing further notation!

[3]Note that portfolio B is clearly self-financing.

Substituting (3) into (4) we obtain

$$a_t\, Z_t^{T_2}\left[m_t^{T_2}\; dt + S_t^{T_2}\; dW_t^P\right] + b_t\; r_t B_t\; dt \;=\; Z_t^{T_1}\left[m_t^{T_1}\; dt + S_t^{T_1}\; dW_t^P\right]. \tag{5}$$

Equating drift and volatility terms then implies

$$\begin{aligned} a_t &= \frac{S_t^{T_1} Z_t^{T_1}}{S_t^{T_2} Z_t^{T_2}} \\ b_t &= \frac{1}{r_t B_t}\left[m_t^{T_1} Z_t^{T_1} - \frac{m_t^{T_2} Z_t^{T_1} S_t^{T_1}}{S_t^{T_2}}\right]. \end{aligned}$$

Since $V_t^A = V_t^B$, we then have

$$\frac{S_t^{T_1} Z_t^{T_1}}{S_t^{T_2} Z_t^{T_2}} Z_t^{T_2} + \frac{1}{r_t B_t}\left[m_t^{T_1} Z_t^{T_1} - \frac{m_t^{T_2} Z_t^{T_1} S_t^{T_1}}{S_t^{T_2}}\right] B_t \;=\; Z_t^{T_1}. \tag{6}$$

Simplifying we obtain

$$\begin{aligned} \frac{S_t^{T_1}}{S_t^{T_2}} + \frac{1}{r_t}\left[m_t^{T_1} - \frac{m_t^{T_2} S_t^{T_1}}{S_t^{T_2}}\right] &= 1 \\ \Rightarrow\; \frac{m_t^{T_1} - r_t}{S_t^{T_1}} &= \frac{m_t^{T_2} - r_t}{S_t^{T_2}} \;=:\; \lambda(t, r_t) \end{aligned} \tag{7}$$

independently of T. Note that in deriving (7) we only used the fact that $Z_t^{T_i}$ was the price of a non-dividend paying security and did not use any other properties of zero-coupon bonds. We therefore obtain that for any derivative security with time t price, P_t, and dynamics given by

$$dP_t \;=\; P_t\left[\mu(t, r_t)\; dt + \;\sigma(t, r_t)\; dW_t^P\right]$$

we must have $\lambda(t, r_t) = (\mu(t, r_t) - r)/\sigma(t, r_t)$.

Remark 1 *$\lambda(t, r_t)$ is usually called the* market-price-of risk*. Since r_t is stochastic*[4]*, we sometimes refer to the market-price-of-risk* process*.*

Using (2), (3) and (7) we obtain[5] the PDE that any derivative security price, P_t, must satisfy

$$(\alpha_{p,t} - \lambda_t \beta_{p,t})\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} + \frac{1}{2}\beta_{p,t}^2 \frac{\partial^2 P}{\partial r^2} - r_t P \;=\; 0. \tag{8}$$

In order to solve (8) we also need to specify boundary conditions. These conditions will depend on the nature of the derivative security. For example, if P_t is the price of a zero-coupon bond that matures at time $T > t$ then the boundary condition is $P_T \equiv 1$.

Remark 2 *Note that if we know the dynamics of just one traded security, i.e. we know $\mu(t, r_t)$ and $\sigma(t, r_t)$ for that security, then we know $\lambda(t, r_t)$ and we can then use (8) to compute the price of any other security.*

Remark 3 *We sometimes call (8) the* Fundamental PDE *for 1-factor models. For multi-factor models we can easily derive an analogous PDE.*

[4]In the Black-Scholes model for equities λ_t is a constant since each of μ, σ and r is also constant.
[5]We could also derive a similar PDE for dividend paying securities.

2 The Martingale Approach

In the martingale approach we begin with an EMM, Q, under which all discounted security prices are martingales. Let us assume that that the Q-dynamics of r_t are given by

$$dr_t \;=\; \alpha(t, r_t)\; dt \;+\; \beta(t, r_t)\; dW_t \tag{9}$$

where W_t is a Q-Brownian motion. Then the time t price, P_t, of a traded security (that does not pay any intermediate cash-flows) is given by

$$P_t \;=\; B_t \mathrm{E}_t^Q\left[\frac{P_T}{B_T}\right] \tag{10}$$

where $B_t := \exp\left(\int_0^t r_s\; ds\right)$ is the time t value of the cash account. For example, we have

$$Z_t^T \;=\; \mathrm{E}_t^Q\left[e^{-\int_t^T r_s\; ds}\right]. \tag{11}$$

An obvious question that comes to mind is how the PDE and martingale approaches are related. We now discuss this issue and it will not be surprising to see that the Feynman-Kac representation and Girsanov's Theorem provide the link.

Connection to PDE Approach

We consider the case of a non-dividend paying security with time t price, P_t, that satisfies (10). The Feynman-Kac formula then states that P_t also satisfies the following PDE

$$\alpha(t, r_t)\frac{\partial P}{\partial r} + \frac{\partial P}{\partial t} + \frac{1}{2}\beta(t, r_t)^2\frac{\partial^2 P}{\partial r^2} - r_t P \quad = \quad 0 \tag{12}$$

Since P_t satisfies both (8) and (12) it must be the case that

$$\beta_p(t, r_t) \quad = \quad \beta(t, r_t) \;\; \text{and}$$

$$\lambda(t, r_t) \quad = \quad \frac{\alpha_p(t, r_t) - \alpha(t, r_t)}{\beta_p(t, r_t)}. \tag{13}$$

Comparing (1) and (9) we also see that

$$dW_t^P \;=\; dW_t - \lambda_t\; dt$$

so that in particular, we have

$$\frac{dQ}{dP} \;=\; \exp\left(-\int_0^t \lambda_s\; dW_s^P - \frac{1}{2}\int_0^t \lambda_s^2\; ds\right). \tag{14}$$

by Girsanov's Theorem.

Remark 4 *While we used non-dividend paying securities to derive the relationships (13) and (14), we could*[6] *also have done so using dividend paying securities. In particular, the EMM Q may be used to price all securities (dividend or non-dividend paying) in the economy.*

Remark 5 *We mention once again that it is standard in term-structure modelling to model interest rates and security price processes directly under Q without any direct reference to the physical measure, P.*

[6] The PDEs in (8) and (12) would change slightly to reflect dividend payments.

3 Specific Models

The Vasicek Model

The Vasicek model assumes that the short rate, r_t, follows a Gaussian model where

$$dr_t \;=\; \alpha(\mu - r_t)\;dt \;+\; \sigma\; dW_t \tag{15}$$

and where as usual, W_t is a Q-Brownian motion. We can then solve (15) by applying Itô's Lemma to $Y_t := \exp(\alpha t) r_t$. We obtain

$$Y_t \;=\; Y_0 + \alpha\mu \int_0^t e^{\alpha u}\; du + \sigma \int_0^t e^{\alpha u}\; dW_u$$

so that

$$r_t \;=\; e^{-\alpha t} r_0 + \mu\left(1 - e^{-\alpha t}\right) + \sigma e^{-\alpha t} \int_0^t e^{\alpha u}\; dW_u. \tag{16}$$

In order to compute the term-structure in Vasicek's model we have to compute

$$Z_0^T \;=\; \mathrm{E}_0^T\left[e^{-\int_0^T r_s\; ds}\right]. \tag{17}$$

We therefore need to find the distribution of the random variable $X := \int_0^T r_s\; ds$. To do this we apply Theorem 7 from the lecture notes, *Overview of Stochastic Calculus*, to see that $X \sim \mathsf{N}(a, b^2)$ where

$$\begin{aligned} a &= \frac{(\mu - r_0)}{\alpha}\left(e^{-\alpha T} - 1\right) + \mu T \\ b^2 &= \int_0^T \sigma^2 e^{2\alpha v}\left(\int_v^T e^{-\alpha t}\; dt\right)^2 dv \;=\; \frac{\sigma^2}{2\alpha^2}\left(T + \frac{4e^{-\alpha T} - e^{-2\alpha T} - 3}{2\alpha}\right). \end{aligned}$$

If we now observe that $\mathrm{E}[\exp(\phi X)] = \exp(\phi a + \phi^2 b^2/2)$ when $X \sim \mathsf{N}(a, b^2)$ we then find

$$Z_0^T \;=\; \exp(A(0,T) - B(0,T) r_0) \tag{18}$$

where

$$\begin{aligned} B(t,T) &= \frac{1 - e^{-\alpha(T-t)}}{\alpha} \\ A(t,T) &= \left(\mu - \frac{\sigma^2}{2\alpha^2}\right)\left[B(t,T) - (T-t)\right] - \frac{\sigma^2}{4\alpha} B(t,T)^2. \end{aligned}$$

More generally, it is easy to see (either by repeating the above analysis or using the fact that r_t is a Markov process) that

$$Z_t^T \;=\; \exp(A(t,T) - B(t,T) r_t). \tag{19}$$

Remark 6 *Note that Z_t^T depends on t and T only through their difference, $T - t$.*

Exercise 2 *Find the SDE satisfied by Z_t^T. What do you notice about its behavior as $t \to T$?*

Example 1 (Derivatives with ZCB as Underlying Security in Vasicek's Model)
Suppose we wish to price a security that expires at time T with payoff $C_T := C(Z_T^U)$ where Z_T^U is the value at time T of a zero-coupon bond maturing at time $U > T$. For example, if the security in question is a call option then $C(x) = (x - K)^+$. Martingale pricing tell us that the time t security price, C_t, satisfies

$$C_t = B_t \mathrm{E}_t^Q \left[\frac{C_T}{B_T} \right]. \tag{20}$$

In order to evaluate the expectation on the right-hand-side of (20) it would appear that we need to find the joint distribution of Z_T^U and B_T. While this is feasible, we can also compute C_t by changing to the *forward measure*, P^T. Equation (17) in the lecture notes, *Overview of Stochastic Calculus*, tells us that we can also represent C_t as

$$C_t = Z_t^T \mathrm{E}_t^{P^T} \left[C(Z_T^U) \right]. \tag{21}$$

To take advantage of (21), we need to find the P^T-dynamics of Z_t^U. Towards this end, we take the following steps:

1. Apply Itô's Lemma to (19) and find that Z_t^T satisfies

$$\begin{aligned} dZ_t^T &= r_t Z_t^T \, dt - Z_t^T B(t,T) \sigma dW_t & (22) \\ &= r_t Z_t^T \, dt - Z_t^T S(t,T) dW_t & (23) \end{aligned}$$

   where W_t is a Q-Brownian motion and $S(t,T) := \sigma B(t,T)$.

2. Now define $Y_t^U := Z_t^U / Z_t^T$ and again apply Itô's Lemma to find that

$$dY_t^U = Y_t^U S(t,T) \left[S(t,T) - S(t,U) \right] \, dt + Y_t^U \left[S(t,T) - S(t,U) \right] \, dW_t \tag{24}$$

3. Note that by definition of P^T, it must be that Y_t^U is a P^T-martingale. Girsanov's Theorem then tells us that $W_t^{P^T}$ is a P^T Brownian motion where

$$W_t^{P^T} := W_t + \int_0^t S(v,T) \, dv.$$

   Note also that we then have

$$dY_t^U = Y_t^U \left[S(t,T) - S(t,U) \right] \, dW_t^{P^T} \tag{25}$$

4. Note that since $Z_T^T \equiv 1$ we may write (21) as

$$C_t = Z_t^T \mathrm{E}_t^{P^T} \left[C(Y_T^U) \right]. \tag{26}$$

   Therefore to compute C_t we need the P^T-distribution of Y_T^U.

Exercise 3 *What is the P^T-distribution of Y_T^U? (Hint: Note that the volatility term, $S(t,U) - S(t,T)$, is a deterministic function of t!)*

Exercise 4 *Give a Black-Scholes-like expression for the time t price of a European call option that expires at time T with payoff $(Z_T^U - K)^+$.*

∎

Remark 7 *Note that the analysis of Example 1 is not specific to the Vasicek model. In fact, whenever the volatility process, $S(t,T)$, for the time T-maturing zero-coupon bond is deterministic, we can obtain a similar Black-Scholes-like formula.*

Remark 8 *We could also have solved the problem of Example 1 without switching to the measure, P^T, by showing that $(\int_0^T r_s\ ds,\ r_T)$ has a bivariate normal distribution under Q . (See Cairns for a proof based on computing the joint moment generating function of $(\int_0^T r_s\ ds,\ r_T)$.)*

The Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross (CIRR) model assumes that the short-rate, r_t, satisfies

$$dr_t \ = \ \alpha(\mu - r_t)\ dt \ + \ \sigma\sqrt{r_t}\ dW_t\ ; \qquad r_0 > 0. \tag{27}$$

where α and μ are positive constants. The principal advantage of the CIR model over the Vasicek model is that the short rate is guaranteed to remain non-negative.[7] Unlike the Vasicek model, however, the CIR model is not Gaussian and is therefore considerably more difficult[8] to analyze. Perhaps surprisingly, zero-coupon bond prices can still be computed analytically and are given by

$$Z_t^T \ = \mathrm{E}_t^Q\left[e^{-\int_t^T r_s\ ds}\right] \ = \ \exp(A(T-t) - B(T-t)r_t). \tag{28}$$

where

$$\begin{aligned} A(\tau) &= \frac{2\alpha\mu}{\sigma^2}\log\left(\frac{2\gamma e^{(\gamma+\alpha)\tau/2}}{(\gamma+\alpha)(e^{\gamma\tau}-1)+2\gamma}\right) \\ B(\tau) &= \frac{2(e^{\gamma\tau}-1)}{(\gamma+\alpha)(e^{\gamma\tau}-1)+2\gamma} \\ \gamma &= \sqrt{\alpha^2+2\sigma^2} \end{aligned}$$

Remark 9 *The expression for Z_t^T in (28) can be verified by checking that it satisfies the PDE (12) with $P_T \equiv 1$, $\alpha(t,r_t) = \alpha(\mu - r_t)$ and $\beta(t,r_t) = \sigma\sqrt{r_t}$.*

Closed form solutions for option prices on zero-coupon bonds can also be found in this model. In general, derivatives prices can be estimated by either numerically solving the PDE in (12) with appropriate boundary conditions or by using Monte-Carlo methods.

Exercise 5 *What is the time 0 forward price for delivery at time τ of a zero-coupon bond that matures at time $T > \tau$ in the CIR model?*

---

[7] Indeed, it can be shown that if $2\alpha\mu > \sigma^2$ then the short rate will remain strictly positive. See Cairns for a proof of this fact.

[8] See Cairns or Shreve for detailed treatments of the CIR model.

4 Generalization to Time-Varying Parameters

The ability to calibrate models to market data is a desirable feature of any model and this points to one of the main drawbacks[9] of the Vasicek and CIR models. These models only have a finite number of free parameters[10] and so it is not possible to specify these parameter values in such a way that model prices (e.g. for zero-coupon bonds of different maturities) coincide with observed market prices. This problem is overcome[11] by allowing the parameters to vary deterministically with time. Some of the more well-known models with time-varying parameters are described below.

Example 2 (The Ho-Lee Model)

We assume that the short-rate satisfies

$$dr_t = \theta(t)\ dt + \sigma dW_t$$

where $\theta(t)$ is a deterministic function of time. It is possible to extend the Ho-Lee model and also make $\sigma(t)$ a deterministic function of time. ∎

Example 3 (The Black-Derman-Toy Model)

This model assumes that $Y_t := \log(r_t)$ satisfies

$$dY_t = \theta(t)\ dt + \sigma dW_t$$

where again $\theta(t)$ is a deterministic function of time. Itô's Lemma may be applied to see that r_t is a geometric Brownian motion. This model was originally specified as a lattice model. ∎

Example 4 (The Black-Karansinski Model)

The Black-Karansinski model is a generalization of the Black-Derman-Toy model where we assume that $Y_t := \log(r_t)$ satisfies

$$dY_t = \alpha(t)(\log\mu(t) - Y_t)\ dt + \sigma(t) dW_t$$

where $\alpha(t)$, $\mu(t)$ and $\sigma(t)$ are deterministic functions of time. Itô's Lemma can be applied to find the dynamics of r_t. ∎

Example 5 (Hull and White)

The Hull-White model generalizes the Vasicek model and assumes

$$dr_t = \alpha(\mu(t) - r_t)\ dt + \sigma dW_t$$

where $\mu(t)$ is a deterministic function of time that may be interpreted as a local mean-reversion level. It is also possible for α and σ to be deterministic functions of time. The Hull-White model is a Gaussian model and so it is straightforward to price derivatives using the same techniques we used for the Vasicek model. ∎

Remark 10 *Question* 1 *of Assignment* 3 *shows how to calibrate a Ho-Lee model with time-varying drift to a given term structure.*

[9]Of course models with too many free parameters are often guilty of over-fitting.
[10]Three, to be specific.
[11]Though new problems associated with over-fitting can then arise!

Single-Factor Affine Term-Structure Models

Some of the models we have described thus far (e.g. Vasicek, CIR, Hull and White) have term structures of the form

$$Z_t^T = \exp\left(a(t,T) + b(t,T)r_t\right). \tag{29}$$

If we define the *yield* as $Y_t^T := -\log(Z_t^T)/(T-t)$ so that $Z_t^T = \exp\left(-(T-t)Y_t^T\right)$, then we see that these models have yields that are *affine* in r_t. A model with this property is called an affine term-structure model. We have the following result.

Theorem 1 *Consider a* 1*-factor model of the form* $dr_t = \alpha(t, r_t)\ dt + \beta(t, r_t)\ dW_t$ *where* W_t *is a* Q*-Brownian motion. Then for all* $t \in [0, T]$*,*

$$Z_t^T = \exp\left(a(t,T) + b(t,T)r_t\right) \tag{30}$$

if and only if $\alpha(r_t, t)$ *and* $\beta(r_t, t)^2$ *are affine in* r_t*, i.e.*

$$\alpha(t,r) = \alpha_1(t) + \alpha_2(t)r \tag{31}$$

$$\beta(t,r)^2 = \beta_1(t) + \beta_2(t)r \tag{32}$$

Proof:

(i) Suppose (30) holds. Then we may apply Itô's Lemma to obtain

$$dZ_t^T = Z_t^T \left[\frac{\partial a}{\partial t} + r_t \frac{\partial b}{\partial t} + b(t,T)\alpha(t,r_t) + \frac{1}{2} b(t,T)^2 \beta(t,r_t)^2\right] dt + Z_t^T \beta(t,r_t)\ dW_t. \tag{33}$$

However under Q it must also be the case that the instantaneous drift of dZ_t^T is given by $r_t Z_t^T\ dt$. This implies

$$\frac{\partial a}{\partial t} + r_t \frac{\partial b}{\partial t} + b(t,T)\alpha(t,r_t) + \frac{1}{2} b(t,T)^2 \beta(t,r_t)^2 - r_t = 0. \tag{34}$$

Now if we differentiate twice with respect to r_t across (34) we see that

$$b(t,T)\ \frac{\partial^2}{\partial r^2}\alpha(t,r_t) + \frac{1}{2} b(t,T)^2\ \frac{\partial^2}{\partial r^2}(\beta(t,r_t)^2) = 0. \tag{35}$$

Since (35) must hold for all T, it must be the case that

$$\frac{\partial^2}{\partial r^2}\alpha(t,r_t) = 0 \qquad \text{and} \qquad \frac{\partial^2}{\partial r^2}(\beta(t,r_t)^2) = 0$$

implying (31) and (32) as desired.

(ii) For the opposite direction, suppose now that (31) and (32) hold. We first substitute (31) and (32) into (12) to obtain

$$\begin{aligned} \left[\alpha_1(t) + \alpha_2(t)r\right] \frac{\partial Z_t^T}{\partial r} + \frac{\partial Z_t^T}{\partial t} + \frac{1}{2}\left[\beta_1(t) + \beta_2(t)r\right] \frac{\partial^2 Z_t^T}{\partial r^2} - r_t P &= 0 \qquad (36) \\ Z_T^T &= 1 \end{aligned}$$

and then use (30) to substitute expressions for the partial derivatives in (36). The resulting PDE may then be reduced to a pair of ODE's in $a(t,T)$ and $b(t,T)$ for which solutions may be shown to exist subject to technical conditions. ∎

Other Models

As we saw in (9) a generic single-factor model for r_t is given by

$$dr_t \;=\; \alpha(t,r_t)\;dt \;+\; \beta(t,r_t)\;dW_t \tag{37}$$

where W_t is a Q-Brownian motion. Many single-factor models can be therefore be analyzed once we specify the functional forms of $\alpha(t,r_t)$ and $\beta(t,r_t)$ in (37). In general, care is needed in specifying $\alpha(t,r_t)$ and $\beta(t,r_t)$ so that: (i) (37) has a solution and (ii) the solution implies behavior for r_t that is satisfactory. Once these conditions are satisfied, the principal concerns are whether or not the model is tractable and can easily be calibrated to market data, and whether or not the empirical properties of the model are consistent with those observed in the market-place.

---

5 Hedging in Single-Factor Models

We now describe how to hedge a derivative security in a generic 1-factor model where

$$dr_t \;=\; \alpha(t,r_t)\;dt \;+\; \beta(t,r_t)\;dW_t.$$

and where W_t is a Q-Brownian motion. We assume the derivative expires at some time $\tau > 0$ and that it does not pay intermediate cash-flows between 0 and τ. Its time t value is denoted by C_t so that its payout upon expiration is C_τ. Martingale pricing then tells us that C_t/B_t is a Q-martingale. The Martingale Representation Theorem then gives the existence of an adapted process, ϕ_t, such that

$$\frac{C_t}{B_t} \;=\; C_0 + \int_0^t \phi_s\;dW_s, \tag{38}$$

assuming as usual that $B_0 = 1$. In particular, we have $C_t \;=\; B_t\left[C_0 + \int_0^t \phi_s\;dW_s\right]$ so Itô's Lemma then implies

$$\begin{aligned} dC_t &= \left[C_0 + \int_0^t \phi_s\;dW_s\right] r_t B_t\;dt + B_t\phi_t\;dW_t \\ &= r_t C_t\;dt + B_t\phi_t\;dW_t. \end{aligned} \tag{39}$$

In order to hedge the derivative, we need to choose our *hedging securities*. We will use the cash account, B_t, and some other[12] security, P_t say, with Q-dynamics given by

$$dP_t \;=\; P_t\left[r_t\;dt + \sigma_t\;dW_t\right] \tag{40}$$

where σ_t is an adapted process that is strictly greater than 0. We can now rewrite (39) as

$$dC_t \;=\; r_t C_t\;dt \;+\; \frac{B_t\phi_t}{\sigma_t C_t}\;\sigma_t C_t\;dW_t. \tag{41}$$

Recalling equation (14) in the *Overview of Stochastic Calculus* lecture notes, we see that the Q-dynamics of the wealth process, V_t, associated with a self-financing trading strategy are given by

$$dV_t \;=\; r_t V_t\;dt \;+\; \theta_t\sigma_t V_t\;dW_t \tag{42}$$

where θ_t and $(1-\theta_t)$ are the *fractions* of time t wealth, V_t, invested in the risky security, P_t, and cash account, respectively, at time t.

Now if we compare (41) and (42), we see that the self-financing strategy that replicates C_τ is a portfolio that at time t invests $\theta_t := B_t\phi_t/\sigma_t C_t$ in P_t and $\$(1-\theta_t)C_t$ in the cash account.

[12] P_t might, for example, represent the price of a particular zero-coupon bond. Note also that the Q-dynamics of any traded security must have a drift equal to r_t.

Exercise 6 *Suppose we choose the cash account and a zero-coupon bond maturing at time T as our hedging instruments. What must the relationship between τ and T be?*

Exercise 7 *How would you go about hedging a European option on a zero-coupon bond in the CIR model?*

---

6 Multi-Factor Models

In the single-factor affine term-structure models we saw that the zero-coupon bonds prices were given by

$$Z_t^T = \exp\left(a(t,T) + b(t,T)r_t\right) \tag{43}$$

for some deterministic functions $a(t,T)$ and $b(t,T)$. This then implies that

$$\frac{dZ_t^T}{Z_t^T} = r_t\ dt + b(t,T)\sigma(t,r_t)\ dW_t \tag{44}$$

where $\sigma(t,r_t)$ is the volatility coefficient for r_t. Equation (44) implies that the returns on zero-coupon bonds of different maturities are *instantaneously* perfectly correlated since

$$\frac{b(t,T_1)\sigma(t,r_t)\cdot b(t,T_2)\sigma(t,r_t)}{\sqrt{b(t,T_1)^2\sigma(t,r_t)^2}\sqrt{b(t,T_2)^2\sigma(t,r_t)^2}} = 1.$$

Indeed, this is the reason why we only need the cash account and one other security to hedge derivative securities in these models. Depending on the application of interest, this can be a very unsatisfactory property of single-factor models. For example, a single factor model would be entirely inappropriate for pricing a *slope-of-the-yield-curve* option with a time T payoff given by

$$\begin{aligned} h(r_T,T) &:= \max\left(S(r_T,T) - K,\ 0\right) \\ \text{where } S(r_T,T) &= \frac{Y_T^{T_2} - Y_T^{T_1}}{T_2 - T_1}; \quad T < T_1 < T_2. \end{aligned}$$

More generally, it would be unwise to use a single-factor model for pricing fixed-income derivatives that do not mature in the relatively near future, e.g. 1 or 2 years. A similar comment applies to a derivative written on an underlying security that does not mature in the near future, regardless of whether or not the derivative[13] itself matures in the near future. The simple reason for this is the well-recognized fact that one factor can not adequately explain movements in the entire term structure. For example, we often see yields at opposite ends of the term structure move in opposite directions. This behavior is more easily explained with multi-factor models. Multi-factor models were introduced primarily to overcome these problems. For example, in our brief discussion of Gaussian multi-factor models below we will see that less than perfect instantaneous correlations between bond returns are possible.

Gaussian Multi-Factor Models

We can specify a Gaussian multi-factor model for the short rate by setting $r_t = \sum_i^n X_t^{(i)}$ and assuming that the Q-dynamics of $X_t \in \mathbf{R}^n$ are given by

$$dX_t = A(\mu - X_t)\ dt + C\ dW_t \tag{45}$$

where W_t is an n-dimensional Q-Brownian motion, $\mu \in \mathbf{R}^n$ and A and C are $n \times n$ matrices. It can then be shown that the solution to (45) is a Gaussian process given by

$$X_t = e^{-At}X_0 + \int_0^t e^{-A(t-s)}\mu\ ds + \int_0^t e^{-A(t-s)}C\ dW_s. \tag{46}$$

[13] A 1-year-10-year swaption is such an example.

Zero-coupon bond prices can again be shown to be affine and have the form

$$Z_t^T = \exp\left(\overline{a}(t,T) + \overline{b}(t,T)X_t\right)$$

where $\overline{a}(t,T)$ is a scalar function and $\overline{b}(t,T)$ is an $\mathbf{R}^n$-valued function. Itô's Lemma then implies that the Q-dynamics of Z_t^T are given by

$$\frac{dZ_t^T}{Z_t^T} = r_t\ dt + \overline{b}(t,T)\sigma\ dW_t \tag{47}$$

where σ is the $n \times n$ volatility matrix for r_t.

Exercise 8 *Compute the instantaneous correlation coefficient, ρ, between returns on zero-coupon bonds of maturities T_1 and T_2. Note that ρ need not equal 1.*

The Multi-Factor CIR Model

The multi-factor CIR model[14] builds upon the single-factor CIR model by assuming that each of n factors, $X_t^{(i)}$ for $i = 1, \ldots, n$, follows CIR-type processes that are Q-independent. If the short rate, r_t, then satisfies $r_t = \sum_i^n X_t^{(i)}$ we can use the results from the single-factor CIR model to determine the term structure.

For each $i = 1, \ldots, n$ we therefore assume that $X_t^{(i)}$ satisfies

$$dX_t^{(i)} = a_i \left(\mu_i - X_t^{(i)}\right)\ dt + c_i \sqrt{X_t^{(i)}}\ dW_t^{(i)};\quad X_0^{(i)} > 0 \tag{48}$$

where the $W_t^{(i)}$'s are Q-independent Brownian motions and a_i, c_i and μ_i are positive constants. Since each $X_t^{(i)}$ is a single-factor CIR process, we also know that there exist functions $A_i(t)$ and $B_i(t)$ such that

$$\mathrm{E}_t^Q\left[\exp\left(-\int_t^T X_u^{(i)}\ du\right)\right] = \exp\left(A_i(T-t) + B_i(T-t)X_t^{(i)}\right). \tag{49}$$

If $r_t = R(t, X_t) := \sum_i X_t^{(i)}$ then we have enough information to compute the term structure. In particular

$$\begin{aligned}
Z_t^T &= \mathrm{E}_t^Q\left[\exp\left(-\int_t^T r_u\ du\right)\right] \\
&= \mathrm{E}_t^Q\left[\exp\left(-\int_t^T \sum_{i=1}^n X_u^{(i)}\ du\right)\right] \\
&= \mathrm{E}_t^Q\left[\exp\left(-\sum_{i=1}^n \int_t^T X_u^{(i)}\ du\right)\right] \\
&= \mathrm{E}_t^Q\left[\prod_{i=1}^n \exp\left(-\int_t^T X_u^{(i)}\ du\right)\right] \\
&= \prod_{i=1}^n \mathrm{E}_t^Q\left[\exp\left(-\int_t^T X_u^{(i)}\ du\right)\right] \\
&= \prod_{i=1}^n \exp\left(A_i(T-t) + B_i(T-t)X_t^{(i)}\right) \\
&= \exp\left(\sum_{i=1}^n \left(A_i(T-t) + B_i(T-t)X_t^{(i)}\right)\right) \qquad (50)
\end{aligned}$$

[14]See Duffie's *Dynamic Asset Pricing* for a more complete discussion.

Remark 11 *There are two methods by which we can make the multi-factor CIR model operable:*

(i) *We identify the factors, $X_t^{(i)}$ for $i = 1, \ldots, n$, as economically relevant and measurable variables such as inflation rates, bond yields, economic indicators etc. We then calibrate the model parameters to these variables and the prices of interest-rate dependent securities. This would appear to be a challenging task.*

(ii) *As in (i), we also assume that the factors, $X_t^{(i)}$ for $i = 1, \ldots, n$, represent economically relevant variables. However, instead of actually trying to identify and measure these variables we make a change-of-variable substitution so that bond yields of n different maturities now make up the state variables. Note that bond yields at time t can be expressed in terms of the $X_t^{(i)}$'s. This is clear from (50) and Itô's Lemma then enables us to write the dynamics satisfied by our new state variables, i.e. the n bond yields. This model, sometimes called a* yield-factor *model, is now much easier to calibrate.*

Multi-Factor Affine Models

More generally, a multi-factor affine model[15] for the short-rate is specified by assuming that $r_t = a + b \cdot X_t$ and that the Q-dynamics of $X_t \in \mathbf{R}^n$ are given by

$$dX_t \;=\; \mu(X_t)\; dt + \sigma(X_t)\; dW_t \tag{51}$$

where W_t is an n-dimensional Q-Brownian motion, and $\mu \in \mathbf{R}^n$ and $\sigma(X_t) \in \mathbf{R}^{(n \times n)}$ are affine functions of X_t. That is,

$$\begin{aligned} \mu(X_t) &= c + dX_t \quad \text{and} \\ \sigma(X_t)\sigma(X_t)^T &= e + fX_t \end{aligned}$$

where c and d are n-dimensional vectors, and e and f are $n \times n$ matrices. Duffie and Kan (1996) showed that if the term structure has the form

$$Z_t^T \;=\; \exp\left(A(t,T) + B(t,T)X_t\right)$$

then it must be the case that X_t satisfies[16] (51).

Exercise 9 *Is the multi-factor CIR model an affine model?*

Derivatives Pricing and Hedging in Multi-Factor Models

Pricing and hedging in multi-factor proceeds along the same lines as in single-factor models. If closed form solutions are not available for derivative prices, then they may be computed numerically either using Monte Carlo simulation or by solving the associated Feynman-Kac PDE. For example, suppose $X_t \in \mathbf{R}^n$ satisfies

$$dX_t \;=\; \alpha(t, X_t)\; dt + \beta(t, X_t)\; dW_t \tag{52}$$

where W_t is an n-dimensional Q-Brownian motion. Then the time t price, C_t, of a security that has a dividend rate function, $h(X_t, t)$, and terminal payment function, $g(X_T)$, is given by

$$C_t \;=\; \mathrm{E}_t^Q\left[\int_t^T e^{-\int_t^s r_u\, du}\; h(X_s, s)\; ds \;+\; e^{-\int_t^T r_u\, du} g(X_T)\right].$$

The associated Feynman-Kac PDE is then

$$\begin{aligned} \frac{\partial C}{\partial t} + \frac{\partial C}{\partial x}\;\mu(t,x) + \frac{1}{2} tr\left[\sigma(t,x)\sigma(t,x)^T \frac{\partial^2 C}{\partial x^2}\right] - r(x,t)C + h(x,t) &= 0, \qquad (x,t) \in \mathbf{R} \times [0,T) \\ C(x,T) &= g(x), \qquad x \in \mathbf{R} \end{aligned}$$

[15] For further details on affine models see James and Webber who devote an entire chapter to the subject.

[16] Additional restrictions need to be placed on the parameters to ensure that r_t remains non-negative.

where $\frac{\partial^2 C}{\partial x^2}$ is the matrix of second derivatives and $\frac{\partial C}{\partial x}$ is the vector of first derivatives.

In order to hedge a derivative security in a model driven by n Brownian motions it is necessary, in general, to use $n+1$ *hedging* securities. The hedging securities are often taken to include the cash account and n zero-coupon bonds of different maturities. The replicating portfolio may be constructed in a manner analogous to that of Section 5.

---

7 Strengths and Weaknesses of Short Rate Models

Short-rate models (including lattice models of the short rate) have a number of strengths and weaknesses:

Strengths

- The models are generally tractable and very amenable to numerical and Monte-Carlo simulation methods.
- Derivatives prices can be computed quickly. This is very important for risk-management purposes when many securities need to be priced frequently.
- They are parsimonious and can provide "sanity checks" on more sophisticated models that can often be calibrated to "fit everything". Models that are calibrated to "fit everything" can be unreliable due to the problems associated with over-fitting. Of course short rate models with deterministic time-varying parameters (e.g. Ho-Lee, Hull and White) are also susceptible to over-fitting.

Weaknesses

- The one-factor short-rate models imply that movements in the entire term-structure can be hedged with only two securities. Equivalently, instantaneous returns on zero-coupon bonds of different maturities are perfectly correlated in single-factor models. Neither of these features is realistic but these problems can be overcome by using multi-factor models. In fact models with just 2 or 3 factors can afford considerably more modelling flexibility. Moreover, they generally retain their numerical tractability. Models with 3 or more factors, however, tend to suffer from the curse-of-dimensionality in which case Monte-Carlo simulation becomes[17] the only practical pricing technique.
- They are not as "close to reality" as the LIBOR market models. The latter class of models directly model observable market quantities, i.e., LIBOR rates, and this feature makes these models relatively straightforward to calibrate. Moreover, many securities of interest can easily be priced in these models by making appropriate distributional assumptions about the evolution of particular LIBOR rates.

In practice, short-rate models are often used as a complement and "sanity check" for more sophisticated models. Moreover, their tractability means that they will continue to be used for risk-management purposes.

[17] Unless analytic solutions are available.