Review of Continuous-Time Term-Structure Models

Part I: equilibrium models

Jesper Lund*

Email: mail@jesperjund.com
Homepage: www.jesperlund.com

First draft: April 1997
Current draft: November, 2001

*These lecture notes were originally written for a M.Sc. course in Fixed Income Analysis at the Aarhus School of Business (1997 and 1998). Currently, they are used for the Ph.D. course in Financial Econometrics at the Aarhus School of Business. I am grateful to Peter Honoré for helpful comments on a earlier version of the lecture notes.

1 Introduction

This paper contains a survey of continuous-time term-structure models. The general idea is introducing the reader (student) to the subject, so that he or she hopefully will be able to read journal articles on the subject. The technical and mathematical nature of this paper reflects that objective.

It is (now) customary to divide term-structure models into two groups, called arbitrage-free and equilibrium models, respectively. The former group contains models which fit the initial yield exactly, and prominent examples are the Heath, Jarrow and Morton (1992) model and the Hull and White (1990) extended Vasicek model. On the other hand, the equilibrium models do not (per construction) fit the yield curve exactly. Sometimes, the models in equilibrium framework are called "classical" models since this approach dates back earlier than the more recent arbitrage-free models.[1]

To a large extent, the choice between arbitrage-free and equilibrium models is dictated by the purpose of the analysis. If we are interested in identifying bonds that are mispriced relative to other bonds, we can only use equilibrium models. On the other hand, when pricing fixed-income derivatives it is generally preferable to use an arbitrage-free model, see, for example, chapter 9 in Tuckman (1995) for an excellent discussion. An important topic in financial econometrics studies of the yield curve is to learn something about risk premia. Clearly, this cannot be done with models that take the initial yield curve as given and use relative-pricing techniques (valuation under the risk-neutral measure) to price derivatives. In the present paper, we only discuss equilibrium models with either a single factor or multiple factors. However, in a sequel to these lecture notes [Part II, Lund (1998)], we describe two types of arbitrage-free models: equilibrium-type models with time-dependent parameters (calibrated models) and models in the Heath, Jarrow and Morton framework.

The outline of the present paper is as follows: Section 2 contains a brief summary of the definition of the term structure using continuous compounding. Section 3 presents a general analysis of one-factor models, including a detailed treatment of the stochastic discount factor approach in continuous time. Section 4 describes three examples of one-factor models, all belonging to the linear (affine) class. In Section 5, we extend the modeling framework to multiple factors, and in Section 6 we discuss the exponential-affine class in a multi-factor setup. Most multi-factor models with known analytical solutions for the term structure belong to this class.

[1]Unfortunately, the literature is not consistent with respect to these definitions. In some papers, especially pre-1990 papers, a model is only called an equilibrium model if derived from the utility function of the representative agent. The Vasicek (1977) model is sometimes referred to as an arbitrage-free (or partial equilibrium) model because an "absence of arbitrage" argument is used in the model derivation, cf. also Section 3 in the present paper. In summary, there is bound to be some confusion, and the reader should be careful when encountering definitions along these lines in, especially, the "older" literature.

2 Definitions

There are three different ways to represent the term structure of interest rates:

$P(t,T)$ is the price, at time t, of a zero-coupon bond[2] maturing at time T (the *maturity date*). The *time to maturity* of this bond is $\tau = T - t$. It is important to note the distinction between the maturity date and the time to maturity — they are only identical when $t = 0$. In general, we assume that $P(t,T)$ exists for all $T > t$.

$R(t,T)$ The yield-to-maturity with continuous compounding at time t, for a zero-coupon bond maturing at time T.

$f(t,T)$ The instantaneous forward rate at time t, for a zero-coupon bond maturing at time T.

The yield-to-maturity $R(t,T)$ and forward rate $f(t,T)$ are defined as follows:

$$R(t,T) = \frac{-\log P(t,T)}{T-t} \tag{1}$$

$$f(t,T) = \frac{-\partial P(t,T)/\partial T}{P(t,T)} = \frac{-\partial \log P(t,T)}{\partial T}. \tag{2}$$

Note that log denotes the natural (base e) logarithm. The inverse relationship expresses the bond price, $P(t,T)$, in terms of either $R(t,T)$ or $f(t,T)$:

$$P(t,T) = e^{-R(t,T)(T-t)} \tag{3}$$

$$P(t,T) = e^{-\int_t^T f(t,s)ds}. \tag{4}$$

The first formula, equation (3), follows simply by rearranging the definition of $R(t,T)$ in (1). To derive (4), first note that

$$\log P(t,T) - \log P(t,t) = \int_t^T \frac{\partial \log P(t,s)}{\partial s} ds = -\int_t^T f(t,s)ds, \tag{5}$$

and since $P(t,t) = 1$, we get (4).

Furthermore, by equating the terms in the exponents in (3) and (4), we get the following relationship between yield-to-maturity and forward rates:

$$R(t,T) = \frac{1}{T-t}\int_t^T f(t,s)ds, \tag{6}$$

which may be interpreted as the average forward rate over the (remaining) time to maturity of the bond.

[2]Unless we state otherwise in the text, all bonds are assumed to be zero-coupon bonds which have a single payment of one "unit of account" at time T. Bonds with more than one remaining payment, for example bullets, are called *coupon bonds*.

3 A general one-factor model

In this section, we carefully explain the common mathematical structure of term-structure models with a single factor, a framework that encompasses the Merton (1973), Vasicek (1977) and Cox, Ingersoll and Ross (CIR) (1985) models. The reader is assumed to be familiar with stochastic differential equations (SDEs), including Ito's lemma, at a level comparable to Hull (1997), Luenberger (1997), or a similar (non-mathematically oriented) text.

We make the following assumptions:

A–1 Standard economic assumptions for a continuous-time model: Frictionless bond markets with continuous trading, no distorting taxes, no short-sale restrictions, and no divisibility problems. Investors always prefer more wealth to less, i.e. the marginal utility of wealth is positive at all levels of wealth.

A–2 The short rate (instantaneous interest rate) follows the general SDE:

$$dr_t = \mu(r_t)dt + \sigma(r_t)dW_t, \tag{7}$$

where $\mu(r)$ and $\sigma(r)$ are the drift and volatility functions, respectively, and W_t a Brownian motion (Wiener) process.

A–3 The market price of risk for the term structure, $\lambda(\cdot)$, only depends on the short rate, r_t.

A–4 All bond prices, i.e. $P(t,T)$ for all $T > t$, are functions of a single state variable, the short rate r_t (in addition to t and T). It can be shown that this assumption follows from the previous two assumptions about the short-rate process (A–2) and the market price of risk (A–3). By implication, changes in the yield curve at different maturities are perfectly correlated.[3]

It is important to realize that we do not assume that the relationship between $P(t,T)$ and the short rate, r_t, is known. On the contrary, the entire purpose of the following is deriving that function endogenously from the above assumptions.

In the following, we present two different derivations of the equilibrium yield curve, as represented through $P(t,T)$. In the first approach (Section 3.1), we make direct assumptions about the *stochastic discount factor* (sometimes called the pricing kernel). This approach is used in chapter 11 of Campbell, Lo and MacKinlay (1997) and in Backus et al. (1998) for discrete-time models, and in chapter 19 of Cochrane (2001) for continuous-time models. In the second approach (Section 3.2), we use the classical arbitrage argument of Black and Scholes (1973) to form a riskless bond portfolio, from which we derive equilibrium bond prices as the solution of a partial differential equation. The two approaches complement each other, and the "absence of arbitrage" derivation can be taken as evidence that the stochastic discount factor in Section 3.1 must have the assumed functional form. The *risk neutral valuation* approach, discussed in Section 3.3, is used to tie the two derivations together.

[3]Of course, this is highly restrictive, but later we relax the assumption with the so-called multi-factor models. Right now we want to keep things simple!

3.1 Equilibrium bond prices using stochastic discount factors

In addition to the above assumptions, A–1 to A–4, we assume that the stochastic discount factor, Λ_t, is governed by the following stochastic differential equation:

$$d\Lambda_t = -r_t \Lambda_t dt - \Lambda_t \lambda(r_t) dW_t, \tag{8}$$

where $\lambda(r_t)$ is the market price of risk, and W_t is the same Brownian motion as the one driving innovations to the short rate, cf. equation (7).

We can obtain equilibrium bond prices, $P(t,T)$, by solving the stochastic discount factor SDE (8) forward, and then take the conditional expectation. Specifically, since the payoff for a zero-coupon bond is one unit of account (say, dollar) at time T, we know that the current (time t) price is

$$P(t,T) = E_t\left(\frac{\Lambda_T}{\Lambda_t}\right), \tag{9}$$

where the expectation is taken conditional on time-t information. In general, however, it is not convenient to calculate this conditional expectation directly although a change-of-measure transformation may simplify the problem, as we discuss below in Section 3.3. Instead, utilize that equation (9) implies that the product of the bond price, $P(t,T)$, and the stochastic discount factor, Λ_t, is a martingale. With a slight abuse of notation,

$$E_t\Big(\, d\left\{P(t,T)\Lambda_t\right\} \Big) = 0, \tag{10}$$

which has the more precise mathematical meaning that the drift in the SDE for the product $P(t,T)\Lambda_t$ is zero (hence, the product is a martingale).

The basic idea in the following is to use Ito's lemma to derive the drift in the SDE for the product of $P(t,T)$ and Λ_t. The SDE for Λ_t is given above, but with respect to $P(t,T)$ we only know (by assumption A–4) that the bond price is a function of calendar time t and the short rate r_t. However, this means that we can invoke Ito's lemma to show that

$$dP(t,T) = \mu_P(t,T)P(t,T)dt \;+\; \sigma_P(t,T)P(t,T)dW_t, \tag{11}$$

where

$$\mu_P(t,T)P(t,T) \;=\; \frac{\partial P}{\partial r}\mu(r) + \frac{\partial P}{\partial t} + \frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2(r) \tag{12}$$

$$\sigma_P(t,T)P(t,T) \;=\; \frac{\partial P}{\partial r}\sigma(r). \tag{13}$$

In equation (11), $\mu_P(t,T)$ is the expected instantaneous return of the bond with maturity date T, and $\sigma_P(t,T)$ is the volatility (standard deviation) of the bond return. The expected return and the volatility depend on the short rate, r_t, but in order to simplify the notation, this dependence is suppressed here.

The next step is to combine the SDEs for $P(t,T)$ and Λ_t, and use Ito's lemma to obtain an expression for the drift of $P(t,T)\Lambda_t$. The result is

$$\begin{aligned} E_t\Big(d\left\{P(t,T)\Lambda_t\right\}/\,dt\Big) &= \\ &= E_t\left(dP(t,T)\Lambda_t\right) + E_t\left(d\Lambda_t P(t,T)\right) + E_t\left(d\Lambda_t dP(t,T)\right) \\ &= \mu_P(t,T)P(t,T)\Lambda_t - r_t\Lambda_t P(t,T) - \Lambda_t\lambda(r_t)\sigma_P(t,T)P(t,T), \end{aligned} \tag{14}$$

where the third term in the second line of equation (14) is the extra second-order term in Ito's lemma (the phrase "extra" means compared to a standard first-order Taylor series expansion).

We obtain the economic restriction on bond prices by setting the drift (14) equal to zero. Note that since $\Lambda_t > 0$, we can simplify the expression by dividing all terms with Λ_t, and after substituting in the expression for $\mu_P(t,T)$ and $\sigma_P(t,T)$ from (12) and (13) and re-arranging the terms, we get the following *partial differential equation* for the bond price:

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2(r) \;+\; \frac{\partial P}{\partial r}\left[\mu(r) - \lambda(r)\sigma(r)\right] \;+\; \frac{\partial P}{\partial t} \;-\; rP = 0, \tag{15}$$

with boundary condition $P(T,T) = 1$. Analytical solutions to this PDE exist for several one-factor models, including the Vasicek (1977), Merton (1973) and CIR (1985) models. They are described in detail in section 4 below.

3.2 The classical "absence of arbitrage" derivation

The starting point for this derivation is that the bond price, $P(t,T)$, is governed by the SDE (11), which we have obtained from assumption A–4 and Ito's lemma (note: we have not yet made any assumptions about the stochastic discount factor). The problem is that equilibrium expected returns $\mu_P(t,T)$ for different T's are unknown, so a general expression for the bond price $P(t,T)$ cannot be determined at this stage. The intermediate goal in the following is developing some form of equilibrium model for the expected returns, $\mu_P(t,T)$, for all T. Concretely, we use the principle of no-arbitrage to reduce this problem to specifying a single market price risk (preference) parameter.

Suppose we construct a portfolio consisting of w_1 bonds with maturity date T_1 and w_2 bonds with maturity date T_2.[4] We require $T_1 \neq T_2$, but apart from that T_1 and T_2 can be arbitrary. The value of the resulting portfolio, at time t, is denoted by

$$\Pi_t = w_1 P(t,T_1) + w_2 P(t,T_2), \tag{16}$$

and the value, Π_t, satisfies the SDE:

$$\begin{aligned} d\Pi_t = {} & \left[w_1\mu_P(t,T_1)P(t,T_1) \;+\; w_2\mu_P(t,T_2)P(t,T_2)\right]dt \;+ \\ & \left[w_1\sigma_P(t,T_1)P(t,T_1) \;+\; w_2\sigma_P(t,T_2)P(t,T_2)\right]dW_t\,. \end{aligned} \tag{17}$$

[4]Vasicek (1977) uses the same technique. It is very similar to the method used to derive the Black-Scholes (stock) option-pricing formula, see Hull (1997).

Since there are two bonds and only one source of risk, it must be possible to eliminate the risky part of the portfolio by choosing w_1 and w_2 such that

$$w_1\sigma_P(t,T_1)P(t,T_1) + w_2\sigma_P(t,T_2)P(t,T_2) = 0\,. \tag{18}$$

In general, this requires continuous adjustment of the portfolio (which can be done costlessly since we have assumed away transactions costs). If w_1 and w_2 are continuously readjusted according to (18), the portfolio SDE (17) reduces to:

$$d\Pi_t = [w_1\mu_P(t,T_1)P(t,T_1) \;+ w_2\mu_P(t,T_2)P(t,T_2)]\,dt, \tag{19}$$

which is locally deterministic (riskless). To prevent arbitrage opportunities, the excess return above the short rate r_t must be zero:

$$w_1\left(\mu_P(t,T_1) - r_t\right)P(t,T_1) + w_2\left(\mu_P(t,T_2) - r_t\right)P(t,T_2) = 0\,. \tag{20}$$

To summarize, we have now shown that if the 2×1 vector $w = [w_1\ w_2]'$ solves the equation

$$\begin{bmatrix} \sigma_P(t,T_1)P(t,T_1) & \sigma_P(t,T_2)P(t,T_2) \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \equiv A_1 w = 0\,, \tag{21}$$

the same w is also a solution to the homogeneous system of equations:

$$\begin{bmatrix} \sigma_P(t,T_1)P(t,T_1) & \sigma_P(t,T_2)P(t,T_2) \\ (\mu_P(t,T_1) - r_t)P(t,T_1) & (\mu_P(t,T_2) - r_t)P(t,T_2) \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \equiv A_2 w = 0\,. \tag{22}$$

This is only possible if the rank of the matrix A_2 is 1. To see this, note that if A_2 has full rank, A_2^{-1} exists, and the solution of (22) is $w = 0$ — which is obviously a contradiction. This argument proves that the rank of A_2 must be 1.

Since A_2 has less than full rank, it is possible to write the last (second) row as a linear combination of the other rows — in this case a scalar (function) $\lambda(r)$ times the first row. Hence, we get

$$\mu_P(t,T_j) = r_t + \lambda(r_t)\sigma_P(t,T_j), \quad j = 1,2 \tag{23}$$

where $\lambda(r_t)$ is the so-called *market price of risk*. Further, note that the particular choices of T_1 and T_2 play no role in the above derivation, so (23) must hold for all T, and $\lambda(r)$ must be independent of T. In summary, we have reduced the problem of determining $\mu_P(t,T)$, for all possible T, to that of specifying a single market price of risk parameter, $\lambda(r)$, which is at most a function of the short rate r. Of course, this requires an additional assumption about market preferences, and different models (Vasicek, Merton, CIR) use different specifications of $\lambda(r)$.

Finally, we substitute (23) into (12), and after rearranging terms (and using the definition of $\sigma_P(t,T)$ from equation (13)), we get the following partial differential equation:

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2(r) \;+\; \frac{\partial P}{\partial r}\left[\mu(r) - \lambda(r)\sigma(r)\right] \;+\; \frac{\partial P}{\partial t} \;-\; rP = 0, \tag{24}$$

with boundary condition $P(T,T) = 1$. Note that this is exactly the same partial differential equation as the one we derived in Section 3.1 through a direct specification of the stochastic discount factor. This equivalence suggests that the stochastic discount factor process must have the form (8) in order to rule out arbitrage opportunities. Indeed, this intuition is correct, as we verify in the next section on risk neutral pricing.

3.3 Bond pricing through risk-neutral valuation

A general representation of the solution to (24) is furnished by the *Feynman-Kac* formula:

$$P(t,T) = E_t^Q \left[e^{-\int_t^T r_s ds} \right], \tag{25}$$

where the expectation is taken under the probability measure (probability distribution) corresponding to the drift-adjusted stochastic process:

$$dr_t = \{\mu(r_t) - \lambda(r_t)\sigma(r_t)\}\ dt\ +\ \sigma(r_t) dW_t^Q\,, \tag{26}$$

where W_t^Q is a Brownian motion under the Q-measure, or risk-neutral distribution.

We can also obtain the risk-neutral valuation equation (25), and the risk-neutral process (26), directly from the stochastic discount factor process. First, note that the logarithm of Λ_t follows the SDE

$$d\log\Lambda_t = \left(-r_t - \frac{1}{2}\lambda(r_t)^2\right) dt - \lambda(r_t) dW_t. \tag{27}$$

If we substitute this expression into (9), and integrate from $u = t$ to $u = T$, we get

$$\begin{aligned} P(t,T) &= E_t \left(e^{\log \Lambda_T - \log \Lambda_t} \right) \\ &= E_t \left(e^{-\int_t^T r_u du - \frac{1}{2}\int_t^T \lambda(r_u)^2 du - \int_t^T \lambda(r_u) dW_u} \right). \end{aligned} \tag{28}$$

The only remaining problem is to evaluate the expectation in the second line of (28), and the most straightforward way of doing this is through a change-of-measure transformation. We start by noting that

$$\frac{dQ}{dP} = e^{-\frac{1}{2}\int_0^T \lambda(u)^2 du - \int_0^T \lambda(u) dW_u} \tag{29}$$

is the Radon-Nikodym derivative[5] between the true probability measure P [not to be confused with the bond price $P(t,T)$] and another probability measure, which

[5]See Duffie (1996) or Musiela and Rutkowski (1997) for a thorough discussion of the Radon-Nikodym derivative and the related Girsanov theorem (finance applications). A complete mathematical treatment of these topics is given by, e.g., Karatzas and Shreve (1991).

we denote Q. With the Radon-Nikodym derivative dQ/dP at hand, we can rewrite equation (28) as

$$\begin{aligned} P(t,T) &= E_t^P\left[\frac{dQ}{dP}e^{-\int_t^T r_u du}\right]\Big/ E_t^P\left(\frac{dQ}{dP}\right) \\ &= E_t^Q\left[e^{-\int_t^T r_u du}\right], \end{aligned} \tag{30}$$

where the last line in (30) corresponds to equation (25). Hence, we conclude that Q is the risk-neutral probability distribution.

From the Girsanov theorem [see, e.g., Duffie (1996)] we know that

$$W_t^Q = W_t + \int_0^t \lambda(u)du \tag{31}$$

is a Brownian motion under the Q-measure, and by substituting this into (7), we obtain the Q-dynamics for the short rate:

$$dr_t = \{\mu(r_t) - \sigma(r_t)\lambda(r_t)\}\, dt + \sigma(r_t)dW_t^Q. \tag{32}$$

This is exactly we same SDE as (26), which we obtained from the "fundamental" PDE for the bond price (24), combined with an application of the Feynman-Kac theorem. If we go through the last derivation in the opposite order, that is starting with the risk-neutral valuation equation (30) and applying a change-of-measure transformation from Q to P (the original, or "true", probability measure), the resulting Radon-Nikodym derivation, dQ/dP, becomes the risk premium part of the stochastic discount factor. Hence, we have shown that the stochastic discount factor process must have the form (8) in order to be consistent with market equilibrium (more precisely: rule out arbitrage opportunities).

Needless to say, we still need to calculate (25) in order to get a closed-form expression for the bond price $P(t,T)$, and in most cases it is actually simpler to solve the PDE directly. However, equation (25) offers a lot of intuition about the mechanics of arbitrage-free term-structure models. The current price, $P(t,T)$, is obtained by discounting the final payment of one unit of account back to the present (time t), and since the future short-term interest rates are random, we take the expectation, conditional on the current value of the short rate, r_t. Among financial economists, the technique is known as *risk-neutral* valuation, and consequently (26) is called the risk-neutral stochastic process for the short rate. Note, however, that we are not assuming risk-neutrality on behalf of the economic agents. On the contrary, investor preferences enter the bond-pricing formula through the drift adjustment by $\lambda(r_t)\sigma(r_t)$ in the SDE (26).

4 Three examples of one-factor models

In this section we present the bond-pricing formula, i.e. the solution $P(t,T)$ of the fundamental PDE (24) for three different models: Vasicek (1977), Merton (1973) and CIR (1985). The solution technique is very much the same (separation of variables), so we only provide a detailed discussion for the Vasicek model.

4.1 Vasicek (1977) model

The short rate follows the mean-reverting Gaussian process (sometimes called the Ornstein-Uhlenbeck process):

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t, \tag{33}$$

where κ measures the speed of mean reversion (the larger κ, the faster the speed of mean reversion), μ is the unconditional mean, and σ is the instantaneous volatility of the short rate. The Vasicek process (33) is the continuous-time equivalent of a first-order autoregressive process, or AR(1) model. With respect to the market price of risk, we assume that it is a constant $\lambda(r) = \lambda$.

This results in the following PDE:

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma^2 \;+\; \frac{\partial P}{\partial r}\left[\kappa(\mu - r) - \lambda\sigma\right] \;+\; \frac{\partial P}{\partial t} \;-\; rP = 0, \tag{34}$$

with boundary condition $P(T, T) = 1$.

First, we guess that the solution takes the so-called exponential-affine form:

$$P(t, T) = \exp\left[A(\tau) + B(\tau)r_t\right], \quad \tau = T - t\,. \tag{35}$$

Second, we differentiate (35) with respect to r and t:

$$\frac{\partial P}{\partial r} = B(\tau)P(t, T) \tag{36}$$

$$\frac{\partial^2 P}{\partial r^2} = B(\tau)^2 P(t, T) \tag{37}$$

$$\frac{\partial P}{\partial t} = -\frac{\partial P}{\partial \tau} = -\left[A'(\tau) + B'(\tau)r\right] \cdot P(t, T). \tag{38}$$

Note that $A'(\tau) = \frac{d\,A(\tau)}{d\,\tau}$, and $B'(\tau) = \frac{d\,B(\tau)}{d\,\tau}$.

Third, we substitute (36)–(38) into (34). Since all terms contain a factor $P(t, T)$, we move this factor outside the parenthesis (braces) and get:

$$\left\{\frac{1}{2}B^2(\tau)\sigma^2 \;+\; B(\tau)[\kappa(\mu - r) - \lambda\sigma] \;-\; A'(\tau) - B'(\tau)r \;-\; r\right\} \cdot P = 0. \tag{39}$$

Finally, we divide by P in (39), and collect the terms containing the factor r:

$$\left\{\frac{1}{2}B^2(\tau)\sigma^2 + B(\tau)[\kappa\mu - \lambda\sigma] - A'(\tau)\right\} \;-\; \left\{\kappa B(\tau) + 1 + B'(\tau)\right\} r = 0. \tag{40}$$

The PDE (40) should be satisfied for all values of r, and this can only hold if both expressions in braces are zero. Equating each of the two terms in braces with zero, results in two *ordinary differential equations* (ODEs),

$$A'(\tau) = \frac{1}{2}\sigma^2 B^2(\tau) + [\kappa\mu - \lambda\sigma]B(\tau) \tag{41}$$

$$B'(\tau) = -\kappa B(\tau) - 1. \tag{42}$$

If we can find a solution to these ODEs, we have demonstrated that (35) is indeed the solution to (34). As with PDEs, we need boundary conditions to solve ODEs. The boundary condition from the PDE, that is $P(T,T) = 1$, translates into two boundary (initial) conditions for the ODE:[6]

$$A(0) = 0 \quad \text{and} \quad B(0) = 0\,. \tag{43}$$

Generally, a system of ODEs (as we have) needs to be solved simultaneously. In the present case, however, the solution separates into two univariate ODEs with a recursive structure since (42) only involves $B(\tau)$. Therefore, we first solve (42), and after substituting the solution $B(\tau)$ into (41), we determine $A(\tau)$.

In our effort to solve (42), we first rewrite it as:

$$B'(\tau) + \kappa B(\tau) = -1\,, \tag{44}$$

and multiply on both sides by $\exp(\kappa\tau)$:

$$B'(\tau)e^{\kappa\tau} + \kappa B(\tau)e^{\kappa\tau} = -e^{\kappa\tau}\,. \tag{45}$$

By the product rule for differentiation, the left hand side in (45) can also be written as:

$$\frac{d}{d\tau}\left\{e^{\kappa\tau}B(\tau)\right\} = -e^{\kappa\tau}\,. \tag{46}$$

Since $B(\tau)$ does not appear on the right hand side in (46), the solution can be obtained by ordinary integration. By the standard relationship between differentiation and integration

$$e^{\kappa\tau}B(\tau) = B(0) + \int_0^\tau \frac{d}{ds}\left\{e^{\kappa s}B(s)\right\}\,ds = -\int_0^\tau e^{\kappa s}ds, \tag{47}$$

where the last equality is obtained by using the boundary condition $B(0) = 0$, as well as equation (46). Finally, we arrive at the desired solution:

$$\begin{aligned} B(\tau) &= -e^{-\kappa\tau}\int_0^\tau e^{\kappa s}ds = -\int_0^\tau e^{-\kappa(\tau - s)}ds \\ &= -\left[\frac{1}{\kappa}e^{-\kappa(\tau - s)}\right]_0^\tau = \frac{e^{-\kappa\tau} - 1}{\kappa}. \end{aligned} \tag{48}$$

Having found $B(\tau)$, we turn to $A(\tau)$. Again, since the function in question, i.e. $A(\tau)$ does not appear on the right hand side of the ODE (41), the solution can be determined by ordinary integration:

$$A(\tau) = A(0) + \int_0^\tau A'(s)ds = \frac{1}{2}\sigma^2\int_0^\tau B^2(s)ds + [\kappa\mu - \lambda\sigma]\int_0^\tau B(s)ds. \tag{49}$$

[6] Note that $\tau = 0$ when $t = T$ (the bond matures).

Thus, we need to calculate the integral of $B(s)$ and $B^2(s)$. To conserve on space, we state the requisite results rather briefly, leaving most of the details to the reader.

$$
\begin{aligned}
(\kappa\mu - \lambda\sigma)\int_0^\tau B(s)ds &= (\kappa\mu - \lambda\sigma)\,\frac{\left(1 - e^{-\kappa\tau}\right)/\kappa \;-\; \tau}{\kappa} \\
&= -\left(\mu - \lambda\sigma/\kappa\right)\left[\tau \;+\; \frac{e^{-\kappa\tau} - 1}{\kappa}\right] \qquad (50)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{1}{2}\sigma^2\int_0^\tau B^2(s)ds &= \frac{1}{2}\sigma^2\frac{\left(1 - e^{-2\kappa\tau}\right)/2\kappa \;-\; 2\left(1 - e^{-\kappa\tau}\right)/\kappa \;+\; \tau}{\kappa^2} \\
&= \frac{1}{2}\left(\frac{\sigma}{\kappa}\right)^2\left[\frac{1 - e^{-2\kappa\tau} \;-\; 4\left(1 - e^{-\kappa\tau}\right)}{2\kappa} \;+\; \tau\right] \qquad (51)
\end{aligned}
$$

This concludes the derivation of the Vasicek bond-pricing formula. For convenience, we restate the entire formula below (after having worked a bit on the expressions):

$$
R(\infty) = \mu - \frac{\lambda\sigma}{\kappa} - \frac{1}{2}\left(\frac{\sigma}{\kappa}\right)^2 \qquad (52)
$$

$$
B(\tau) = \frac{e^{-\kappa\tau} - 1}{\kappa} \qquad (53)
$$

$$
A(\tau) = -R(\infty)\left(\tau + B(\tau)\right) \;-\; \frac{\sigma^2}{4\kappa}B^2(\tau). \qquad (54)
$$

This corresponds to equation (27) on page 185 of the Vasicek (1977) paper. Note, though, that his notation is different from ours. Among other things, he uses the opposite sign for the market price of risk (which he calls q, instead of "our" λ).

In (53), it straightforward to see that $B(\tau) < 0$, so an increase in r_t lowers bond prices. The reader is encouraged to investigate how different parameter values for κ, μ, σ and λ affect the shape of the term structure.

4.2 The Merton (1973) model

The short rate is governed by the SDE:

$$
dr_t = \mu dt + \sigma dW_t, \qquad (55)
$$

and the market price of risk is a constant λ, as in the Vasicek model. It can be shown that the bond price is given by (35) (see Section 4.1) with the following definitions of $A(\tau)$ and $B(\tau)$:

$$
B(\tau) = -\tau \qquad (56)
$$

$$
A(\tau) = -\frac{1}{2}\left(\mu - \sigma\lambda\right)\tau^2 + \frac{1}{6}\sigma^2\tau^3. \qquad (57)
$$

The proof is left to the reader. Compared to the Vasicek model, it is actually quite simple (and doing it is a very good exercise).

4.3 The CIR (1985) model

The famous CIR, or Cox, Ingersoll and Ross, model uses the so-called square-root SDE (process) for the short rate:

$$dr_t = \kappa(\mu - r_t)\, dt \;+\; \sigma\sqrt{r_t} dW_t. \tag{58}$$

CIR specifies the market price of risk as follows: $\lambda(r) = \lambda\sqrt{r}/\sigma$. The scaling by σ is done only to simplify the subsequent derivations.

The derivations are considerably more tedious than in the Vasicek case, so we simply state the result below, and refer to Ingersoll (1987, pp. 397-399) or Lund (1993, pp. 37–41) for further details. Once again, the bond-pricing formula takes the familiar exponential-affine form (35), with $B(\tau)$ and $A(\tau)$ being defined as follows:

$$B(\tau) \;=\; \frac{-2\left(1 - e^{-\gamma\tau}\right)}{2\gamma + (\kappa + \lambda - \gamma)\left(1 - e^{-\gamma\tau}\right)} \tag{59}$$

$$A(\tau) \;=\; \frac{2\kappa\mu}{\sigma^2} \log\left[\frac{2\gamma e^{(\kappa + \lambda - \gamma)\tau/2}}{2\gamma + (\kappa + \lambda - \gamma)\left(1 - e^{-\gamma\tau}\right)}\right] \tag{60}$$

where

$$\gamma = \sqrt{(\kappa + \lambda)^2 + 2\sigma^2}. \tag{61}$$

The main advantage over the Vasicek model is that r_t is restricted to be non-negative. However, for realistic parameter values, there is rarely much difference between the yield curves obtained from the Vasicek and CIR models, respectively.

5 Multi-factor models

The main advantage of one-factor models is their simplicity as the entire yield curve is a function of just one state variable. Moreover, this state variable is observable — at least in principle (in practice, we use a short-term interest rate as a proxy). However, there are several problems with one-factor models.

First, the model assumes that changes in the yield curve, and hence bond returns, are perfectly correlated across maturities, and not surprisingly this assumption is easily contradicted by the empirical evidence. Apart from that, the assumption of perfect correlation is highly problematic for several "practical" purposes, for example Value-at-Risk calculations, and pricing derivatives on interest-rate spreads. The latter case is discussed by Canabarro (1995). Second, the shape of the yield curve is severely restricted. Specifically, the Vasicek and CIR models can only accommodate yield curves that are monotonic increasing or decreasing and humped. An inversely humped yield curve, for example, cannot be generated with these models. Moreover, with time-invariant parameters, one-factor models tend to provide a very poor fit to the actual yield curves observed in the market.

The latter problem can be solved by *calibration* which is discussed in Lund (1998). By making some parameters time-dependent, we obtain a perfect fit to the current (initial) yield curve, and the calibrated model can only be used to price fixed-income derivatives. If the modeling purpose is identifying bonds that are mispriced, the calibration approach cannot be used. Moreover, since the model is extended with deterministic parameters, yield changes are still perfectly correlated, so for some securities, a calibrated one-factor model may still be inadequate.

For these reasons, we discuss multi-factor models in the following. Specifically, the short rate is still governed by stochastic process with time-invariant parameters, but there are now, say, m sources of innovation, and not just one as in (7). For practical purposes, this means that we get a better (but not perfect) fit to the yield curve, and yield curve changes are no longer perfectly correlated.

5.1 A general framework for multi-factor models

The underlying assumptions of multi-factors models are very similar to the one-factor case, and the modifications relate only to the stochastic process for the short rate and the risk premia. For convenience, however, we restate the full list of assumptions:

1. Standard economic assumptions for a continuous-time model, see Section 3.

2. All bond prices are a function of a $m \times 1$ vector of state variables, denoted X_t. Together with the next assumption, this implies that the market prices of risk at time t, $\lambda(X_t)$, are functions of the m state variables.

3. The short rate is a known function of X_t, i.e. $r_t = r(X_t)$. In most case, r_t is the first element of the vector X_t, that is $r_t = w'X_t$, where w is a vector with one as the first element and zeros elsewhere.

4. The dynamics of the state variables are governed by:

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \tag{62}$$

   where $\mu(X)$ is a $m \times 1$ drift vector, $\sigma(X)$ is a $m \times m$ matrix containing the volatilities coefficients, and W_t an m-dimensional Brownian motion. Unless otherwise noted, we specify $\sigma(X)$ as a diagonal matrix and let the m univariate Brownian motions in the vector W_t be correlated. The requisite correlation coefficients are denoted ρ_{ij}.

The purpose of the following analysis is deriving the functional relationship between the m state variables, X_t, and the prices of zero-coupon bonds, $P(t,T)$, for all T. As in section 3, we start by deriving a stochastic process for bond prices, including an expression for the expected returns on different bonds. Next, we use the economic theory (absence of arbitrage) to impose an APT-like restriction on bond returns, or we use the martingale restriction involving the stochastic discount factor. In either case, we need specific assumptions about the market prices of risk. Finally, we obtain a PDE for bond prices, as well as a risk-neutral process for the short rate (through the m state variables).

By an appropriate multivariate version of Ito's lemma, bond prices can be shown to evolve according to

$$d\,P(t,T) = \mu_P(t,T)P(t,T)dt \ + \ \sum_{i=1}^{m} \sigma_{Pi}(t,T)P(t,T)dW_{it}, \tag{63}$$

where drift is given by

$$\mu_P(t,T)P(t,T) = \sum_{i=1}^{m} \frac{\partial P}{\partial X_i}\mu_i(X) + \frac{\partial P}{\partial t} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 P}{\partial X_i \partial X_j}\sigma_i(X)\sigma_j(X)\rho_{ij}, \tag{64}$$

and the i'th bond volatility is given by

$$\sigma_{Pi}(t,T)P(t,T) = \frac{\partial P}{\partial X_i}\sigma_i(X). \tag{65}$$

Note that $\rho_{ii} = 1$ in (64). The expected return and the bond volatilities depend on the state variables, but to keep the notation manageable, this dependence is suppressed here.

5.2 The stochastic discount factor

The stochastic discount factor is governed by the following stochastic process:

$$d\Lambda_t = -r_t\Lambda_t dt - \Lambda_t \lambda(X_t)' dW_t, \tag{66}$$

where $\lambda(X_t)$ is an $m \times 1$ vector containing the market prices of risk.

We know that the product of the bond price, $P(t,T)$, and the stochastic discount factor, Λ_t, is a martingale. Therefore, we use Ito's lemma to derive an expression for the drift of $P(t,T)\Lambda_t$, and by setting the drift term to zero, we obtain the following PDE:

$$\begin{aligned} &\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 P}{\partial X_i \partial X_j}\sigma_i(X)\sigma_j(X)\rho_{ij} \ + \\ &\quad \sum_{i=1}^{m} \frac{\partial P}{\partial X_i}\left\{\mu_i(X) - \lambda_i(X)\sigma_i(X)\right\} \ + \ \frac{\partial P}{\partial t} \ - \ rP = 0, \end{aligned} \tag{67}$$

where $\lambda_i(X)$ is the i'th element of the vector $\lambda(X)$, and so forth. The details of the PDE derivation are left as an exercise to the reader (alternatively, see Section 5.3).

As in the one-factor case, cf. Section 3, we can use the Feynman-Kac theorem to represent the solution of the PDE (76) as the risk-neutral expectation

$$P(t,T) = E_t^Q\left[e^{-\int_t^T r(X_s)ds}\right], \tag{68}$$

where the Q-measure corresponds to the stochastic process (SDE)

$$dX_t = \{\mu(X_t) - \sigma(X_t)\lambda(X_t)\}\ dt \ + \ \sigma(X_t)dW_t^Q. \tag{69}$$

Note that the drift and volatility functions of (69) are obtained from the coefficients of the first-order and second-order derivatives in (76), respectively. The same result can be obtained by using a change-of-measure transformation of the conditional expectation of the stochastic discount factor. Since the requisite derivations represent a straightforward generalization of Section 3.3, we leave the details to the reader.

5.3 The APT restriction for multi-factor models

The purpose of this section is to present a multi-factor generalization of the classical "absence of arbitrage" arguments, which we used to determine the bond price equation for a general one-factor model in Section 3.2.

In order to derive the appropriate APT restriction on $\mu(t,T)$ for different maturity dates T, we construct a portfolio consisting of $K = m+1$ with distinct maturities. The number of bonds with maturity date T_i, $i = 1,\ldots,K$, is denoted by w_i. The instantaneous changes in the value of this portfolio, Π_t, can be written as:

$$\begin{aligned} d\Pi_t &= \sum_{k=1}^{K} w_k \cdot d\,P(t,T_k) = \\ &\left[\sum_{k=1}^{K} w_k \mu_P(t,T_k) P(t,T_k)\right] dt + \sum_{i=1}^{m}\left[\sum_{k=1}^{K} w_k \sigma_{Pi}(t,T_k) P(t,T_k)\right] dW_{it}, \end{aligned} \tag{70}$$

where we have interchanged the order of summation between i and k in the second line of (70). Since there are more bonds than sources of risk, it must be possible to choose non-zero portfolio weights, w_k, which make the portfolio locally riskless. This means that the weights must satisfy m restrictions of the form

$$\sum_{k=1}^{K} w_k \sigma_{Pi}(t,T_k) P(t,T_k) = 0, \quad i = 1,\ldots,m. \tag{71}$$

By continuously readjusting the portfolio weights, we can ensure that the price dynamics of the portfolio are always riskless, or deterministic. Absence of arbitrage requires that the expected (and realized) return is equal to the short rate r_t — otherwise there is a "free lunch" by either buying or selling the portfolio and taking the opposite position in the money markets (both investments are locally riskless). Stated otherwise, the expected excess return must be zero,

$$\sum_{k=1}^{K} w_k P(t,T_k) \cdot \{\mu_P(t,T_k) - r_t\} = 0 \tag{72}$$

We have shown that, if the vector $z = [P(t,T_1)w_1,\ldots,P(t,T_K)w_K]'$, with $z \neq 0$, solves the system of equations:

$$\begin{bmatrix} \sigma_{P1}(t,T_1) & \ldots & \sigma_{P1}(t,T_K) \\ \sigma_{P2}(t,T_1) & \ldots & \sigma_{P2}(t,T_K) \\ \ldots & \ldots & \ldots \\ \sigma_{Pm}(t,T_1) & \ldots & \sigma_{Pm}(t,T_K) \end{bmatrix} \begin{bmatrix} P(t,T_1)w_1 \\ P(t,T_2)w_2 \\ \ldots \\ P(t,T_K)w_K \end{bmatrix} \equiv A_1 z = 0\,, \tag{73}$$

the same $K \times 1$ vector z also solves the larger system:

$$\begin{bmatrix} \sigma_{P1}(t,T_1) & \dots & \sigma_{P1}(t,T_K) \\ \sigma_{P2}(t,T_1) & \dots & \sigma_{P2}(t,T_K) \\ \dots & \dots & \dots \\ \sigma_{Pm}(t,T_1) & \dots & \sigma_{Pm}(t,T_K) \\ \mu_P(t,T_1)-r_t & \dots & \mu_P(t,T_K)-r_t \end{bmatrix} \begin{bmatrix} P(t,T_1)w_1 \\ P(t,T_2)w_2 \\ \dots \\ P(t,T_K)w_K \end{bmatrix} \equiv A_2 z = 0. \tag{74}$$

Since (74) is a homogeneous system of equations and $z \neq 0$, this is only possible if the rank of A_2 is equal to m. If A_2 is non-singular, the solution to (74) is $z = A_2^{-1} \cdot 0 = 0$, which is obviously a contradiction.[7] Since A_2 has $m+1$ rows but rank m, the last row can be written as a linear combination of the other rows. Moreover, this result does not depend on the specific maturities T_k, so for any T we have

$$\mu_P(t,T) = r_t + \sum_{i=1}^{m} \lambda(X_t)\sigma_{Pi}(t,T), \tag{75}$$

where $\lambda_i(X)$ is the market price of risk for the i'th state variable (factor). Note that the risk premia can only depend on X_t and possibly calendar time t, but not on the maturity dates T (or other characteristics of the K securities, for that matter).

To complete the derivation of bond prices, we substitute (75) into (64). After rearranging terms and using the definition of $\sigma_{Pi}(t,T)$ in (65), we get the following PDE for bond prices:

$$\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 P}{\partial X_i \partial X_j}\sigma_i(X)\sigma_j(X)\rho_{ij} + \sum_{i=1}^{m} \frac{\partial P}{\partial X_i}\left\{\mu_i(X) - \lambda_i(X)\sigma_i(X)\right\} + \frac{\partial P}{\partial t} - rP = 0, \tag{76}$$

with boundary condition $P(T,T) = 1$. Again, this is exactly the same PDE as the one obtained in Section 5.2.

6 The exponential-affine class of models

In this section we consider a general class of models, called *exponential-affine* models, where a general analytical solution is available [Duffie and Kan (1996)]. For some parametric specifications, we can obtain a closed-form expression like in the Vasicek model, but in the worst case we will have to solve a system of ordinary differential equations numerically, and this can be done very efficiently with the Runge-Kutta method. This is especially true for multi-factor models since the complexity of the numerical solution only increases linearly in the number of state variables, whereas

[7] This line of reasoning depends on K being equal to $m+1$. If $K > m+1$, we can show that absence of arbitrage implies $\text{rank}(A_2) = m$ by noting that A_1 and A_2 have the same nullspace. This approach is used in proofs of the APT theory, see Ross (1976). Johnston (1984) contains an introduction to nullspaces.

the complexity of finite-difference PDE solutions generally increases exponentially in the number of state variables.

Duffie and Kan (1996) propose a general class of term-structure models that include Gaussian models as a special case. Under the original (true) probability measure, the m state variables in the vector X_t are governed the process

$$dX_t = \mathcal{K}\left(\Theta - X_t\right)dt + C\sigma(X_t)dW_t, \tag{77}$$

where $\sigma(X_t)$ is a $m \times m$ diagonal matrix whose i'th diagonal element given by

$$\sigma_{ii}(X_t) = \sqrt{\alpha_i + \beta_i' X_t}. \tag{78}$$

In this setup, the m univariate Brownian motions are independent, and the dependence structure between the innovations to X_t is captured by the $m \times m$ matrix C. Loosely speaking, this means that the $m \times m$ variance-covariance matrix for changes in X_t is given by

$$\mathrm{Cov}(dX_t) = C\sigma^2(X_t)C'\,dt, \tag{79}$$

with representative element (i, j)

$$[\mathrm{Cov}(dX_t)]_{ij} = \mathrm{Cov}(dX_{it}, dX_{jt}) = \sum_{k=1}^{m} C_{ik}C_{jk}\sigma_{kk}^2(X_t)dt. \tag{80}$$

We refer to Duffie and Kan (1996) for a thorough discussion of conditions ensuring that (77) is a well-defined stochastic process. The short rate, or instantaneous interest rate, r_t, is specified as an affine function of X_t:

$$r_t = r(X_t) = w_0 + \sum_{i=1}^{m} w_i X_{it} = w_0 + w' X_t. \tag{81}$$

Generally, the vector w consists of either zeros or ones. Finally, to complete the model specification, we make the following assumptions about the market prices of risk:

$$\lambda(X_t) = \sigma(X_t)\lambda \tag{82}$$

With these assumptions, the fundamental PDE can be written as

$$\begin{aligned}
&\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 P}{\partial X_i \partial X_j}\left(\sum_{k=1}^{m} C_{ik}C_{jk}\sigma_{kk}^2(X)\right) + \\
&\sum_{i=1}^{m}\frac{\partial P}{\partial X_i}\left[\sum_{k=1}^{m}\mathcal{K}_{ik}(\Theta_k - X_k) - \sum_{k=1}^{m} C_{ik}\sigma_{kk}^2(X)\lambda_k\right] + \frac{\partial P}{\partial t} - r(X)P \;=\; 0,
\end{aligned} \tag{83}$$

or more compactly using matrix algebra and the trace operator,[8]

$$\frac{1}{2}\mathrm{Tr}\left(\frac{\partial^2 P}{\partial X \partial X'}C\sigma^2(X)C'\right) + \frac{\partial P}{\partial X'}\left[\mathcal{K}(\Theta - X) - C\sigma^2(X)\lambda\right]$$

[8]If A and B are symmetric square $(m \times m)$ matrices,

$$\mathrm{Tr}\,(AB) = \sum_{i=1}^{m}\{AB\}_{ii} = \sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij}B_{ji} = \sum_{i=1}^{m}\sum_{j=1}^{m} A_{ij}B_{ij},$$

that is $\mathrm{Tr}\,(AB)$ is a convenient way to write the sum of the product of all m^2 elements in A and B.

$$+\frac{\partial P}{\partial t} - [w_0 + w'X]\,P \;=\; 0. \tag{84}$$

Following the general idea of Section 4, we guess that the solution takes the following form

$$P(\tau, X) = \exp\left[A(\tau) + B(\tau)'X\right], \tag{85}$$

where $A(\tau)$ is a scalar function, and $B(\tau)$ an $m \times 1$ vector. In order to verify whether the solution is of the form (85), we compute the requisite partial derivatives of (85) and substitute these expressions into the PDE (84). If we can obtain an ODE system defining $A(\tau)$ and $B(\tau)$, we have demonstrated that the solution is of the form (85). Moreover, by solving the ODE system, either analytically or by numerical methods (Runge-Kutta), we obtain the bond-pricing formula.

First, we have after straightforward calculations

$$\frac{\partial P}{\partial X_i} \;=\; B_i(\tau) \cdot P(t,T), \quad i = 1, \ldots, m \tag{86}$$

$$\frac{\partial^2 P}{\partial X_i \partial X_j} \;=\; B_i(\tau)B_j(\tau) \cdot P(t,T), \quad i = 1, \ldots, m \;\; j = 1, \ldots, m \tag{87}$$

$$\frac{\partial P}{\partial t} \;=\; -\frac{\partial P}{\partial \tau} \;=\; -\left[\frac{dA(\tau)}{d\tau} + \frac{dB(\tau)'}{d\tau}X\right] \cdot P(t,T). \tag{88}$$

If we substitute these expressions into the PDE, and collect terms involving X_i, for $i = 1, \ldots, m$ and the constant (one), we obtain the following ODEs — after rearranging several terms and using the property that $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$:

$$\frac{dB(\tau)}{d\tau} \;=\; \frac{1}{2}\sum_{i=1}^{m}[C'B(\tau)]_i^2\,\beta_i - \mathcal{K}'B(\tau) - \sum_{i=1}^{m}\lambda_i[C'B(\tau)]_i\,\beta_i - w \tag{89}$$

$$\frac{dA(\tau)}{d\tau} \;=\; \frac{1}{2}\sum_{i=1}^{m}[C'B(\tau)]_i^2\,\alpha_i + B(\tau)'\mathcal{K}\Theta - \sum_{i=1}^{m}\lambda_i[C'B(\tau)]_i\,\alpha_i - w_0. \tag{90}$$

Here, $[C'B(\tau)]_i$ refers to the i'th element of the $m \times 1$ vector $C'B(\tau)$. Finding a general closed-form solution to this ODE system does not seem to be possible, but many special cases (models) can be solved in closed form.

Having derived the ODEs, it is worth emphasizing the specific restrictions which result in the exponential-affine bond price (85). Duffie and Kan (1996) show that we obtain (85) if the following conditions hold:

- The short rate is an affine function of the state variables, that is $r_t = w_0 + w'X_t$.
- The risk-neutral drift function (vector) is affine in X_t.
- The covariances between dX_i and dX_j for all i, j are affine functions of X_t. For the SDE (77) this holds if $\sigma^2(X_t)$ is affine in X_t, cf. equation (80).

The stochastic process (77) combined with the risk premia (82) is the most general specification satisfying the sufficient conditions in Duffie and Kan (1996). In applications, further restrictions are often needed to obtain a tractable model.

In the above setup, nothing is assumed about the state variables, and accordingly they are taken as unobserved variables. As mentioned during the introductory remarks of Section 5.2, we can always invert the bond-pricing formula and express the m state variables in terms of m (distinct) zero-coupon yields. Still, the starting point of the modeling effort is an unobserved stochastic process, and the identification with m bond yields is only made indirectly. Alternatively, Duffie and Kan (1996) propose taking m "reference" yields as the state variables, that is specifying the stochastic process (under the Q-measure) directly for these m yields. This approach — called *yield factor* models — facilitates direct identification with observable quantities (points on the yield curve), but as the state variables are now traded assets, we need to impose parameter restrictions on (77), such that the m bonds are priced correctly, see Duffie and Kan (1996). Unfortunately, the requisite parameter restrictions are often quite complex and therefore difficult to impose. In essence, there is a tradeoff between direct interpretation of the state variables and the time-invariant model parameters.

6.1 Examples of affine multi-factor models

The three one-factor models in Section 4 belong to the exponential-affine class. For models with multiple factors, there are a lot of different specifications, and we cannot provide an exhaustive list. Instead, we offer a few examples from the multi-factor literature.

6.1.1 Gaussian central tendency model

This model has been proposes by, among others, Beaglehole and Tenney (1991) and Jegadeesh and Pennacchi (1996). The short rate is governed by the two-factor Gaussian process

$$dr_t = \kappa_1(\mu_t - r_t)dt + \sigma_1 dW_{1t} \tag{91}$$

$$d\mu_t = \kappa_2(\theta - \mu_t)dt + \sigma_2 dW_{2t}, \tag{92}$$

and the two Brownian motions may be correlated with correlation coefficient ρ. The market prices of risks are specified as constants, λ_1 and λ_2. The central tendency models generalized the Vasicek model by letting the short rate revert towards a time-varying (stochastic) mean which is governed by a separate process. Sometimes this feature is referred to as a "double decay" model.

The PDE is given by:

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma_1^2 + \frac{1}{2}\frac{\partial^2 P}{\partial \mu^2}\sigma_2^2 + \frac{\partial^2 P}{\partial r \partial \mu}\rho\sigma_1\sigma_2 + \frac{\partial P}{\partial r}\left[\kappa_1(\mu - r) - \lambda_1\sigma_1\right]$$
$$+\frac{\partial P}{\partial \mu}\left[\kappa_2(\theta - \mu) - \lambda_2\sigma_2\right] + \frac{\partial P}{\partial t} - rP = 0, \tag{93}$$

subject to the boundary condition $P(T,T) = 1$. It is straightforward to verify that this model is exponential-affine (since the model is Gaussian). Therefore,

$$P(t, t+\tau) = \exp\left[A(\tau) + B_1(\tau)r_t + B_2(\tau)\mu_t\right]. \tag{94}$$

If we substitute the requisite partial derivatives into (93) and divide by P on both sides of the equation, get

$$\frac{1}{2}B_1^2(\tau)\sigma_1^2 + \frac{1}{2}B_2^2(\tau)\sigma_2^2 + B_1(\tau)B_2(\tau)\rho\sigma_1\sigma_2 + B_1(\tau)\left[\kappa_1(\mu - r) - \lambda_1\sigma_1\right]$$
$$+B_2(\tau)\left[\kappa_2(\theta - \mu) - \lambda_2\sigma_2\right] - A'(\tau) - B_1'(\tau)r - B_2'(\tau)\mu - r = 0. \tag{95}$$

Since (95) must hold for all values of r and μ, we obtain the following ODE systems after collecting terms:

$$B_1'(\tau) = -\kappa_1 B_1(\tau) - 1 \tag{96}$$
$$B_2'(\tau) = \kappa_1 B_1(\tau) - \kappa_2 B_2(\tau) \tag{97}$$
$$\begin{aligned} A_1'(\tau) = {} & \frac{1}{2}\sigma_1^2 B_1^2(\tau) + \frac{1}{2}\sigma_2^2 B_2^2(\tau) + \rho\sigma_1\sigma_2 B_1(\tau)B_2(\tau) \\ & -\lambda_1\sigma_1 B_1(\tau) + (\kappa_2\theta - \lambda_2\sigma_2)B_2(\tau), \end{aligned} \tag{98}$$

with boundary (initial) conditions $B_1(0) = 0$, $B_2(0) = 0$, and $A(0) = 0$ as $P(T,T) = 1$ for all r_t and μ_T. It is possible to solve the entire ODE system in closed form, but for reasons of space we concentrate on $B_1(\tau)$ and $B_2(\tau)$. First, note that the ODE defining $B_1(\tau)$ is exactly the same as in the Vasicek model. This means that

$$B_1(\tau) = \frac{e^{-\kappa_1\tau} - 1}{\kappa_1} \tag{99}$$

If we substitute (99) into (97), we get another linear ODE, which can be solved by the same technique as we used in Section 4.1:

$$B_2(\tau) = \frac{e^{-\kappa_2\tau} - 1}{\kappa_2} - \frac{e^{-\kappa_1\tau} - e^{-\kappa_2\tau}}{\kappa_1 - \kappa_2}. \tag{100}$$

Finally, we can substitute (99) and (100) into (98), and $A(\tau)$ can be calculated by ordinary integration. Since the expression for $A(\tau)$ is rather length, it is omitted here.

6.1.2 Fong-Vasicek stochastic volatility model

Fong and Vasicek (1991) propose another extension of the Vasicek model, where the Ornstein-Uhlenbeck process is augmented with a stochastic volatility factor:

$$dr_t = \kappa_1(\mu - r_t)dt + \sqrt{V_t}dW_{1t} \tag{101}$$
$$dV_t = \kappa_2(\alpha - V_t)dt + \eta\sqrt{V_t}dW_{2t} \tag{102}$$

The correlation coefficient between two Brownian motions is denoted ρ. Fong and Vasicek (1991) specify the market prices of risk as

$$\lambda_i(\cdot) = \lambda_i\sqrt{V}, \quad i = 1, 2 \tag{103}$$

since this is the only specification which preserves the affine property under the Q-measure. With these assumptions, the PDE becomes

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}V + \frac{1}{2}\frac{\partial^2 P}{\partial V^2}\eta^2 V + \frac{\partial^2 P}{\partial r \partial V}\rho\eta V + \frac{\partial P}{\partial V}\left[\kappa_1(\mu - r) - \lambda_1 V\right]$$

$$+\frac{\partial P}{\partial V}\left[\kappa_2(\alpha - V) - \lambda_2\eta V\right] + \frac{\partial P}{\partial t} - rP = 0. \tag{104}$$

The solution is of the form

$$P(t, t+\tau) = \exp\left[A(\tau) + B_1(\tau)r + B_2(\tau)V\right]. \tag{105}$$

After substituting the requisite partial derivatives of (105) into the PDE and collecting terms, we get the following ODE system defining the functions $A(\tau)$, $B_1(\tau)$ and $B_2(\tau)$:

$$B_1'(\tau) = -\kappa_1 B_1(\tau) - 1 \tag{106}$$

$$\begin{aligned} B_2'(\tau) = {} & \frac{1}{2}B_1^2(\tau) + \frac{1}{2}\eta^2 B_2^2(\tau) + \rho\eta B_1(\tau)B_2(\tau) \\ & -\lambda_1 B_1(\tau) - (\kappa_2 + \lambda_2\eta)B_2(\tau) \end{aligned} \tag{107}$$

$$A'(\tau) = \kappa_1\mu B_1(\tau) + \kappa_2\alpha B_2(\tau). \tag{108}$$

The solution for $B_1(\tau)$ is the same as in the Vasicek model. Closed-form expressions for $B_2(\tau)$ and $A(\tau)$ are presented in Selby and Strickland (1995). The expressions are quite complicated — involving infinite-order series expansions — so it might be worthwhile to consider solving the ODEs numerically instead.

6.1.3 Multi-factor CIR models

Neither the Gaussian central-tendency model nor the Fong-Vasicek stochastic volatility model restrict the short rate to be non-negative. A popular multi-factor model with this property is the m-factor CIR model which is obtained by adding m independent square root processes,

$$dr_t = \sum_{i=1}^{m} y_{it} \tag{109}$$

$$dy_{it} = \kappa_i(\mu_i - y_{it})dt + \sigma_i\sqrt{y_{it}}dW_{it}, \tag{110}$$

and the market price of risk for the i'th factor is specified as in the one-factor CIR model, that is

$$\lambda_i(\cdot) = (\lambda_i/\sigma_i)\sqrt{y_{it}}. \tag{111}$$

The easiest way to derive an expression for bond prices is using risk-neutral expectations

$$P(t,T) = E_t^Q\left[e^{-\int_t^T\left(\sum_{i=1}^m y_{is}\right)ds}\right], \tag{112}$$

where y_{it} evolves according to

$$dy_{it} = \{\kappa_i(\mu_i - y_{it}) - \lambda_i y_{it}\} \, dt + \sigma_i \sqrt{y_{it}} dW_{it}^Q \tag{113}$$

under the Q-measure. By interchanging the order of integration and summation, equation (112) may be rewritten as

$$\begin{aligned} P(t,T) &= E_t^Q \left[e^{-\sum_{i=1}^m \left(\int_t^T y_{is} ds \right)} \right] \\ &= E_t^Q \left[\prod_{i=1}^m e^{-\int_t^T y_{is} ds} \right] \\ &= \prod_{i=1}^m E_t^Q \left[e^{-\int_t^T y_{is} ds} \right] \qquad (114) \\ &= \prod_{i=1}^m P_i(t,T), \qquad (115) \end{aligned}$$

where $P_i(t,T)$ is the price formula for a one-factor CIR model with parameters κ_i, μ_i, σ_i and λ_i, as well as "short rate" y_{it}. Note that the third line follows because of independence between the m square-root processes.

6.1.4 The Vasicek model with a stochastic market price of risk

Lund (1999, Appendix B) presents a Gaussian term-structure models, where r_t is driven by the univariate Ornstein-Uhlenbeck process under the original probability measure, but, contrary to the Vasicek model, the bond price $P(t,T)$ also depends on a second factor which is called a stochastic market price of risk. In the following, we derive the Lund (1999) model using the stochastic discount factor approach.

The model is based on two independent Ornstein-Uhlenbeck processes:

$$\begin{aligned} dr_t &= \kappa_1(\mu_1 - r_t)dt + \sigma_1 dW_{1t} \qquad (116) \\ d\lambda_t^* &= \kappa_2(\mu_2 - \lambda_t^*)dt + \sigma_2 dW_{2t}. \qquad (117) \end{aligned}$$

Although Lund (1999) assumes that the Brownian motions W_{1t} and W_{2t} are independent, it is straightforward to incorporate a non-zero correlation coefficient between the two Brownian motions. The stochastic discount factor, Λ_t, is governed by the SDE (66) with $X_t = (r_t, \lambda_t^*)'$ and

$$\lambda(X_t) = \begin{bmatrix} -\lambda_t^* \kappa_1 / \sigma_1 \\ 0 \end{bmatrix}. \tag{118}$$

It is because of this particular specification of $\lambda(X_t)$ that Lund (1999) calls λ_t^* a stochastic market price of risk (apart from the sign and scaling).

The risk-neutral SDE for $X_t = (r_t, \lambda_t^*)'$ is given by

$$\begin{aligned} dr_t &= \{\kappa_1(\mu_1 - r_t) + \kappa_1 \lambda_t^*\} \, dt + \sigma_1 dW_{1t}^Q \qquad (119) \\ d\lambda_t^* &= \kappa_2(\mu_2 - \lambda_t^*)dt + \sigma_2 dW_{2t}^Q, \qquad (120) \end{aligned}$$

and the fundamental PDE for the bond price $P(t,T)$ is

$$\frac{1}{2}\frac{\partial^2 P}{\partial r^2}\sigma_1^2 + \frac{1}{2}\frac{\partial^2 P}{\partial \lambda^{*2}}\sigma_2^2 + \frac{\partial P}{\partial \lambda^*}\kappa_2(\mu_2 - \lambda^*) + \\ \frac{\partial P}{\partial r}\kappa_1(\mu_1 - r + \lambda^*) + \frac{\partial P}{\partial t} - rP(t,T) = 0, \tag{121}$$

subject to the boundary condition $P(T,T) = 1$. Since the model is Gaussian, and hence exponential-affine, the solution of (121) has the well-known form:

$$P(t, t+\tau) = \exp\left[A(\tau) + B(\tau)r_t + C(\tau)\lambda_t^*\right]$$

The scalar functions $A(\tau)$, $B(\tau)$ and $C(\tau)$ satisfy the following system of ordinary differential equations (ODEs):

$$A'(\tau) = \kappa_1\mu_1 B(\tau) + \kappa_2\mu_2 C(\tau) + \frac{1}{2}\sigma_1^2 B^2(\tau) + \frac{1}{2}\sigma_2^2 C^2(\tau) \tag{122}$$

$$B'(\tau) = -\kappa_1 B(\tau) - 1 \tag{123}$$

$$C'(\tau) = -\kappa_2 C(\tau) + \kappa_1 B(\tau), \tag{124}$$

with boundary conditions $A(0) = 0$, $B(0) = 0$ and $C(0) = 0$. First, it is easily verified that

$$B(\tau) = \frac{e^{-\kappa_1\tau} - 1}{\kappa_1} \tag{125}$$

$$C(\tau) = \frac{e^{-\kappa_2\tau} - 1}{\kappa_2} - \frac{e^{-\kappa_1\tau} - e^{-\kappa_2\tau}}{\kappa_1 - \kappa_2} \tag{126}$$

We note, in passing, that (125) and (126) are exactly the same expressions as in the double-decay model, cf. Section 5.4.1. Second, $A(\tau)$ is obtained by integrating the right hand side of (122), see Lund (1999, p. 361) for the full (rather lengthy) expression.

References

Backus, D., S. Foresi and C. Telmer (1998), "Discrete-Time Models of Bond Pricing," Manuscript, Department of Finance, Stern School of Business, New York University. Available at `http://www.stern.nyu.edu/~dbackus/`

Beaglehole, D.R. and M.S. Tenney (1991), "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations," *Journal of Fixed Income*, 1, Sept., 69–83.

Black, F. and M. Scholes (1973), "The Valuation of Options and Corporate Liabilities," *Journal of Political Economy*, 108, 144–172.

Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press.

Canabarro, E. (1995), "Where do One-Factor Models Fail?" *Journal of Fixed Income*, September, 31–52.

Cochrane, J.H. (2001), *Asset Pricing*, Princeton University Press.

Cox, J.C., J.E. Ingersoll and S.A. Ross (1985), "A Theory of the Term Structure of Interest Rates, " *Econometrica*, 53, 385–407.

Duffie, D. (1996), *Dynamic Asset Pricing Theory, 2nd Edition*, Princeton University Press.

Duffie, D. and R. Kan (1996), "A Yield Factor Model of Interest Rates," *Mathematical Finance*, 6, 379–406.

Fong, H.G. and O.A. Vasicek (1991), "Fixed Income Volatility Management," *Journal of Portfolio Management*, Summer, 41–46.

Heath, D., R. Jarrow and A. Morton (1992), "Bond Prices and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation," *Econometrica*, 60, 77–105.

Hull, J.C. (1997), *Options, Futures, and other Derivative Securities*, Prentice-Hall.

Hull, J.C. and A. White (1990), "Pricing Interest Rate Derivative Securities," *Review of Financial Studies*, 3, 573–592.

Ingersoll, J.E. (1987), *Theory of Financial Decision Making*, Rowman and Littlefield, Maryland.

Jegadeesh, N. and G.G. Pennacchi (1996), "The Behavior of Interest Rates Implied by the Term Structure of Euro-Dollar Futures," *Journal of Money, Credit and Banking*, 28, 426–446.

Johnston, J. (1994), *Econometric Methods, 3rd Ed.*, McGraw Hill, New York, NY.

Karatzas, I. and S.E. Shreve (1991), *Brownian Motion and Stochastic Calculus, 2nd Edition*, Springer Verlag.

Luenberger, D. (1998), *Investment Science*, Oxford University Press.

Lund, J. (1993), “Dynamic Models of the Term Structure of Interest Rates,” unpublished Master’s thesis, Department of Finance, Aarhus School of Business. Available at `http://www.JesperLund.com`

Lund, J. (1998), “Review of Continuous-Time Term-Structure Models — Part II: Arbitrage-Free Models,” Lecture Notes, Department of Finance, Aarhus School of Business, April 1998.

Lund, J. (1999), “A Model for Studying the Effect of EMU on European Yield Curves,” *European Finance Review*, 2, 321–363.

Merton, R.C. (1973), “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 4, 141–183.

Musiela, M. and M. Rutkowski (1997), *Martingale Methods in Financial Modelling*, Springer Verlag.

Ross, S.A. (1976), “The Arbitrage Theory of Capital Asset Pricing,” *Journal of Economic Theory*, 13, 341–360.

Selby, M.J.P. and C.R. Strickland (1991), “Computing the Fong and Vasicek Pure Discount Bond Price Formula,” *Journal of Fixed Income*, September, 78–84.

Tuckman, B. (1995), *Fixed Income Securities: Tools for Today’s Markets*, Wiley.

Vasicek, O.A. (1977), “An Equilibrium Characterization of the Term Structure of Interest Rates,” *Journal of Financial Economics*, 5, 177–188.