Fixed Income Analysis
Term-Structure Models in Continuous Time

Multi-factor equilibrium models (general theory)
  The Brennan and Schwartz model
  Exponential-affine models

Jesper Lund
April 14, 1998

Outline

1. One-factor models (review and problems)
2. Calibration vs. multi-factor models
3. Mathematical techniques (multivariate SDEs and Ito’s lemma)
4. A general multi-factor model
5. Examples of multi-factor models
6. The Brennan and Schwartz two-factor model
7. The exponential-affine class of multi-factor models
8. A two-factor central-tendency model (Beaglehole-Tenney)
One-Factor Models – 1

- Key features of one-factor (equilibrium) models:
  - All bond prices are a function of a **single** state variable, the short rate.
  - The short rate evolves according to the univariate SDE:
    \[
    dr_t = \mu(r_t)dt + \sigma(r_t)dW_t.
    \] (1)
  - Using the "absence of arbitrage" assumption and Ito's lemma, we derive a PDE for bond prices:
    \[
    \frac{1}{2} \frac{\partial^2 P}{\partial r^2} \sigma^2(r) + \frac{\partial P}{\partial r} [\mu(r) - \lambda(r)\sigma(r)] + \frac{\partial P}{\partial t} - rP = 0,
    \] (2)
    with boundary condition \( P(T, T) = 1 \).

- Advantages of one-factor models:
  - **Simple** model — with a limited number of parameters
  - The state variable (short rate) is **observable**, at least in principle.
  - Numerical solutions (e.g. binomial trees) can be implemented, if necessary.

One-Factor Models – 2

- Problems with one-factor models:
  - Changes in the yield curve are **perfectly correlated** across different maturities.
  - Shape of the yield curve highly restricted (monotonic increasing and decreasing, and hump-shaped, but **not** inversely hump-shaped).
  - Model unable to fit the actual yield curve (when the model parameters are time-invariant, as they are supposed to be). Cause of concern for pricing derivatives (e.g., mortgage-backed securities).

- Solutions:
  - Calibrated one-factor models with time-dependent parameters (advocated by Hull and White (1990) as modifications of the Vasicek and CIR models).
  - Alternatively: HJM models which fit the initial yield curve per construction.
  - Models with multiple factors (but still with time-invariant parameters).
Solution 1: Calibrated one-factor models

- For example, Vasicek with time-dependent drift

\[ dr_t = \kappa \{ \theta(t) - r_t \} \, dt + \sigma dW_t^Q, \]  

where \( \theta(t) \) is chosen to fit the current yield curve exactly.

- Problems solved:
  1. Perfect fit to the current yield curve (including any bond mispricing).
  2. Any shape of the current yield curve can be accommodated.

- Problems remaining and new problems:
  1. Still a one-factor model with perfect-correlation assumption. Inadequate for certain derivatives, e.g. options on yield spreads [Canabarro (1995)].
  2. The approach is (inherently) useless for detecting mispricing of bonds.
  3. Model will not fit future yield curves, unless parameters are re-calibrated.
  4. Hedging and risk-management applications are problematic — because of the "perfect correlation" assumption.

Solution 2: Multi-Factor Models

- Main assumptions:
  - All bond prices are a function of a \( m \)-dimensional state vector \( X_t \).
  - The short rate is a known function of \( X_t \), that is \( r_t = r(X_t) \).
  - The state variables in \( X_t \) evolve according to the multivariate SDE

\[ dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \]  

where \( W_t \) is an \( m \)-dimensional Brownian motion, and \( \sigma(X_t) \) diagonal.

- Problems with multi-factor models:
  1. Changes in yield curve are no longer perfectly correlated, but they still lie in an \( m \)-dimensional subspace (a great improvement, of course).
  2. We may need "many" factors to fit the entire yield curve.
  3. Factors are, in principle, unobservable. What is \( X_t \) anyway?
  4. Finding an analytical solution for bond prices may be difficult.
  5. Numerical solutions (for derivatives) can be computationally involved.
Multi-Factor Models — How?

- Without loss of generality, the short rate can be taken as one of the \( m \) state variables, since \( r_t = r(X_t) \).

- Under no-arbitrage assumption all bond prices (still) satisfy:

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

where \( Q \) denotes the risk-neutral distribution. **Note:** the risk-neutral process for \( r_t \) has yet to be determined.

- Using “traded assets” as additional state variables?
  - Examples: 30Y yield or the consol yield (Brennan-Schwartz model).
  - We must specify how the state variables affect \( r_t \) under the \( Q \)-measure.
  - Parameter restrictions, since (5) must hold for these assets also.

Multivariate SDEs

- Multivariate SDE:

\[
dx_t = \mu(X_t)dt + \sigma(X_t)dW_t,
\]

where \( X_t = (X_{1t}, \ldots, X_{mt}) \).

- The \( i \)'th row of (6) is a univariate SDE, whose drift and volatility functions depend on all \( m \) state variables:

\[
dx_{it} = \mu_i(X_t)dt + \sigma_i(X_t)dW_{it}.
\]

In this setup, the \( m \) univariate Brownian motions can be correlated, with \( \text{Corr}(dW_{it}, dW_{ij}) = \rho_{ij}dt \).

- Consider a scalar function, \( F(X,t) \), representing a mapping from \( \mathbb{R}^m \times \mathbb{R} \) to the real line, \( \mathbb{R} \). The dynamics of \( F(X,t) \) are obtained by applying a multivariate version of **Ito’s lemma**.
Ito’s lemma (multivariate)

- If $X_t$ evolves according to the vector SDE (6), the function $F$, given by $F = F(X,t)$ follows the univariate SDE:

$$dF_t = \alpha(X_t)dt + \sum_{i=1}^{m} \beta_i(X_t)dW_{it}, \quad (8)$$

- The drift in (8) is given by:

$$\alpha(X) = \sum_{i=1}^{m} \frac{\partial F}{\partial X_i} \mu_i(X) + \frac{\partial F}{\partial t} + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 F}{\partial X_i \partial X_j} \sigma_i(X) \sigma_j(X) \rho_{ij}, \quad (9)$$

where $\rho_{ii} = 1$.

- The $i$’th volatility coefficient in (8) is given by:

$$\beta_i(X) = \frac{\partial F}{\partial X_i} \sigma_i(X). \quad (10)$$

A General Multi-Factor Model – 1

- As in the one-factor case, we determine (endogenously) the relationship between $X_t$ and bond prices, $P(t,T)$.

- Since $P(t,T)$ is a function of $X_t$ and $t$,

$$dP(t,T) = \mu_P(t,T)P(t,T)dt + \sum_{i=1}^{m} \sigma_{P_i}(t,T)P(t,T)dW_{it}, \quad (11)$$

and the drift and volatility coefficients are obtained from Ito’s lemma.

- Absence of arbitrage implies the APT restriction:

$$\mu_P(t,T) = r_t + \sum_{i=1}^{m} \lambda_i(X_t) \sigma_{P_i}(t,T). \quad (12)$$

- In equation (12), $\lambda_i(X_t)$ is the market price of risk for the $i$’th factor, and it is independent of $T$. 


A General Multi-Factor Model – 2

• By Ito’s lemma, \( \mu_P(t, T) \) and \( \sigma_P(t, T) \) can also be written as:

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

\[
\sigma_P(t, T) P(t, T) = \frac{\partial P}{\partial X_i} \sigma_i(X), \quad i = 1, 2, \ldots, m
\]

• After substituting these equations into the APT restriction (12), we get the following PDE:

\[
\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 P}{\partial X_i \partial X_j} \sigma_i(X) \sigma_j(X) \rho_{ij} + \sum_{i=1}^{m} \frac{\partial P}{\partial X_i} [\mu_i(X) - \lambda_i(X) \sigma_i(X)] + \frac{\partial P}{\partial t} - r(X) P = 0 \quad (13)
\]

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

A General Multi-Factor Model – 3

• Feynman-Kac solution:

\[
P(t, T) = E_t^Q \left[ e^{-\int_t^T r(X_s) ds} \right], \quad (14)
\]

• The expectation in (14) is taken under the probability measure corresponding to the risk-neutral (drift-adjusted) process:

\[
dX_t = \{\mu(X_t) - \sigma(X_t) \lambda(X_t)\} \ dt + \sigma(X_t) dW_t^Q. \quad (15)
\]

• The \( i \)'th element of the SDE (15) is

\[
dX_{it} = \{\mu_i(X_t) - \lambda_i(X_t) \sigma_i(X_t)\} \ dt + \sigma_i(X_t) dW_{it}^Q. \quad (16)
\]
Four examples of multi-factor models

1. Double-Decay (Central-Tendency) model:
   \[
   dr_t = \kappa_1(\mu - r_t)dt + \sigma_1 dW_{1t} \\
   d\mu_t = \kappa_2(\theta - \mu_t)dt + \sigma_2 dW_{2t}
   \]  
   (17)  
   (18)

2. Fong-Vasicek stochastic volatility model:
   \[
   dr_t = \kappa_1(\mu - r_t)dt + \sqrt{V_t}dW_{1t} \\
   dV_t = \kappa_2(\alpha - V_t)dt + \eta \sqrt{V_t}dW_{2t}
   \]  
   (19)  
   (20)

3. Brennan-Schwartz model:
   \[
   d\log r_t = [\alpha(l_t - r_t) - \alpha \log p]dt + \sigma_1 dW_{1t} \\
   dl_t = \beta_2(r, l)dt + \sigma_2 l_t dW_{2t},
   \]  
   (21)  
   (22)
   where \(l_t\) is the consol rate (annuity that never matures).

4. Multi-factor CIR model:
   \[
   r_t = \sum_{i=1}^{m} y_{it} \\
   dy_{it} = \kappa_i(\mu_i - y_{it})dt + \sigma_i \sqrt{y_{it}}dW_{it}, \quad i = 1, 2, \ldots, m
   \]  
   (23)  
   (24)
   where the \(m\) Brownian motions are independent.

The Brennan and Schwartz (1979) model

- State variables in the model:
  \(r_t\)  the short rate (instantaneous interest rate).
  \(l_t\)  the yield-to-maturity on a consol bond with a “continuous coupon”.

- General stochastic process:
  \[
  dr_t = \beta_1(r_t, l_t)dt + \eta_1(r_t, l_t)dW_{1t} \\
  dl_t = \beta_2(r_t, l_t)dt + \eta_2(r_t, l_t)dW_{2t}.
  \]  
  (25)  
  (26)

- The particular process used in the paper:
  \[
  d\log r_t = [\alpha(l_t - r_t) - \alpha \log p]dt + \sigma_1 dW_{1t} \\
  dl_t = \beta_2(r, l)dt + \sigma_2 l_t dW_{2t}.
  \]  
  (27)  
  (28)

- For pricing purposes, we do not need to specify \(\beta_2(l, r)\) as the second state variable is a traded asset.
BS consol price dynamics

- A consol is an annuity that never matures. If $V_t$ denotes the price of the consol, we have the following relation:

$$V_t = \int_0^\infty e^{-lt} ds = \left[-\frac{1}{l} e^{-lt}\right]_0^\infty = \frac{1}{l_t}$$  \hspace{1cm} (29)

- Note: the relationship between $X_t = (r_t, l_t)$ and $V_t$ is known.

- Consol price dynamics:

$$\frac{dV_t}{V_t} = \mu_V(r_t, l_t) dt + 0 \cdot dW_{1t} + \sigma_V(r_t, l_t) dW_{2t} \hspace{1cm} (30)$$

where

$$V \cdot \mu_V(r, l) = -l^{-2} \beta_2(l, r) + l^{-3} \eta_2^2(l, r) = l^{-1} \left[-l^{-1} \beta_2(l, r) + l^{-2} \eta_2^2(l, r)\right] \hspace{1cm} (31)$$

$$V \cdot \sigma_V(r, l) = -l^{-2} \eta_2(l, r) = l^{-1} \left[-l^{-1} \eta_2(l, r)\right] \hspace{1cm} (32)$$

since $V_t = l_t^{-1}$ does not depend on $r_t$.

BS fundamental PDE – 1

- The bond price, $P(t, T)$, satisfies the PDE:

$$\frac{1}{2} \frac{\partial^2 P}{\partial r^2} \eta_1^2(r, l) + \frac{1}{2} \frac{\partial^2 P}{\partial l^2} \eta_2^2(r, l) + \frac{\partial^2 P}{\partial r \partial t} \rho \eta_1(r, l) \eta_2(r, l) + \frac{\partial P}{\partial r} \left\{\beta_1(r, l) - \lambda_1(r, l) \eta_1(r, l)\right\} + \frac{\partial P}{\partial l} \left\{\beta_2(r, l) - \lambda_2(r, l) \eta_2(r, l)\right\} + \frac{\partial P}{\partial t} - rP = 0. \hspace{1cm} (33)$$

- Because the $l_t$ is a known function of a traded asset, we can eliminate $\beta_2(r, l)$ and $\lambda_2(r, l)$ from the above PDE.

- First, we substitute the SDE for the consol price dynamics (30) into the APT relationship used to derive the PDE:

$$\mu_V(r, l) + l = r + \lambda_2(r, l) \sigma_V(r, l) \hspace{1cm} (34)$$

Why do we add $l$ on the LHS?
BS fundamental PDE – 2

- Second, we substitute (31) and (32) into (34):

\[ -l^{-1}\beta_2(r,l) + l^{-2}\eta_2^2(r,l) + l = r - \lambda_2(r,l)l^{-1}\eta_2(r,l). \] (35)

- Third, after multiplying by \( l \) on both sides of (35), we get

\[ \beta_2(r,l) - \lambda_2(r,l)\eta_2(r,l) = l^{-1}\eta_2^2(r,l) + l^2 - rl \] (36)

- Finally, we substitute (36) into (33):

\[
\frac{1}{2} \frac{\partial^2 P}{\partial r^2} \eta_1^2(r,l) + \frac{1}{2} \frac{\partial^2 P}{\partial l^2} \eta_2^2(r,l) + \frac{\partial P}{\partial r} \{ \beta_1(r,l) - \lambda_1(r,l)\eta_1(r,l) \} + \frac{\partial P}{\partial l} \{ l^{-1}\eta_2^2(r,l) + l^2 - rl \} + \frac{\partial P}{\partial t} - rP = 0,
\] (37)

which is the BS PDE.

Assessment of the BS model

- Advantages of the Brennan-Schwartz model:
  - State variables are observable (in principle), and they can be interpreted as short and long-run factors.
  - Only one market price of risk (preference) parameter in the model.

- Problems with the Brennan-Schwartz model:
  - No analytical solution for bond prices. The PDE can only be solved with numerical methods — either by finite-difference PDE solutions or Monte Carlo evaluation of the Feynman-Kac formula.
  - In most bond markets, there are no actively traded consol bonds.
  - Technical problems with the BS model: by the definition of \( l_t \),

\[ V_t = l_t^{-1} = \int_t^\infty P(t,s)ds = F(r_t,l_t), \] (38)

but the requisite parameter constraint(s) are not imposed in the BS model.

- This problem is, in fact, an argument against using traded assets (yields) as state variables (not just in the BS model).
Exponential-affine models – 1

- Fundamental PDE for a general multi-factor model:
  \[
  \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 P}{\partial X_i \partial X_j} \sigma_i(X) \sigma_j(X) \rho_{ij} + \\
  \sum_{i=1}^{m} \frac{\partial P}{\partial X_i} [\mu_i(X) - \lambda_i(X) \sigma_i(X)] + \frac{\partial P}{\partial t} - r(X) P = 0
  \]  
  \hspace{1in} (39)

  - The Brennan and Schwartz model with \( X_t = (r_t, l_t) \) does not lead to an analytical solution of (39) for bond prices.

  - There are several term-structure models with an analytical solution for \( P(t, T) \), and for most of these models we get
  \[
  P(t, t + \tau) = \exp \left[ A(\tau) + B(\tau)' X_t \right].
  \]  
  \hspace{1in} (40)

  - Models with bond prices of the form (40) are called exponential-affine models [Duffie and Kan (1996)].

Exponential-affine models – 2

- What are the sufficient conditions for obtaining (40) as the solution to (39)?

  - All bond prices, solutions to (39), depend on:
    1. The mapping from \( X_t \) to \( r_t \), given by \( r_t = r(X_t) \).
    2. \( m \) risk-neutral drifts:
      \[
      \mu_i^*(X) = \mu_i(X) - \lambda_i(X) \sigma_i(X)
      \]  
      \hspace{1in} (41)

      3. \( m(m+1)/2 \) variance-covariance terms: \( \sigma_i(X) \sigma_j(X) \rho_{ij} \).

  - Sufficient conditions for exponential-affine models:
    \[
    r(X) = w_0 + w_1' X \\
    \mu_i^*(X) = a_i + b_i' X, \quad i = 1, \ldots, m \\
    \sigma_i(X) \sigma_j(X) \rho_{ij} = c_{ij} + d_{ij}' X, \quad i, j = 1, \ldots, m
    \]  
    \hspace{1in} (42) \hspace{1in} (43) \hspace{1in} (44)

  - That is, all “coefficients” in the PDE are linear in \( X \).
Exponential-affine models – 3

- The function $A(\tau)$ and the $m \times 1$ vector (of functions) $B(\tau)$ depend on the specific model.
- $A(\tau)$ and $B(\tau)$ are obtained as the solution to an ODE system with dimension $(m + 1)$.
- Same procedure as with one-factor models:
  - First, we guess that the solution is of the form (40).
  - Second, we substitute the requisite partial derivatives in to the PDE.
  - Finally, we collect terms with the factor $X_i$ ($i = 1, 2, \ldots, m$) and a constant (remaining terms).
  - This provides the $m + 1$ ODEs which must be solved somehow (perhaps numerically, using Runge-Kutta integration).
  - Boundary conditions for the ODE: $A(0) = 0$ and $B(0) = 0_{m \times 1}$.

Gaussian central-tendency model – 1

- Stochastic process for the short rate:
  \[
  \begin{align*}
  dr_t &= \kappa_1 (\mu_t - r_t) dt + \sigma_1 dW_{1t} \\
  d\mu_t &= \kappa_2 (\theta - \mu_t) dt + \sigma_2 dW_{2t},
  \end{align*}
  \]
  (45)
  (46)

- The Brownian motions are dependent, $\text{Corr}(dW_{1t}, dW_{2t}) = \rho dt$, and the market prices of risk are constants, $\lambda_1$ and $\lambda_2$.

- PDE:
  \[
  \begin{align*}
  \frac{1}{2} \frac{\partial^2 P}{\partial r^2} \sigma_1^2 + \frac{1}{2} \frac{\partial^2 P}{\partial \mu^2} \sigma_2^2 + \frac{\partial^2 P}{\partial r \partial \mu} \rho \sigma_1 \sigma_2 + \frac{\partial P}{\partial r} [\kappa_1 (\mu - r) - \lambda_1 \sigma_1] \\
  + \frac{\partial P}{\partial \mu} [\kappa_2 (\theta - \mu) - \lambda_2 \sigma_2] - \frac{\partial P}{\partial \tau} - r P &= 0,
  \end{align*}
  \]
  (47)

- We guess that
  \[
  P(t, t + \tau) = \exp \left[ A(\tau) + B_1(\tau) r_t + B_2(\tau) \mu_t \right].
  \]
  (48)
Gaussian central-tendency model – 2

• Substitution of the partial derivatives of the function (48) into the PDE (47) gives

\[
\left\{ \frac{1}{2}B_1^2(\tau)\sigma_1^2 + \frac{1}{2}B_2^2(\tau)\sigma_2^2 + B_1(\tau)B_2(\tau)\rho\sigma_1\sigma_2 + B_1(\tau) [\kappa_1(\mu - r) - \lambda_1\sigma_1] \\
+ B_2(\tau) [\kappa_2(\theta - \mu) - \lambda_2\sigma_2] - A'(\tau) - B_1'(\tau)r - B_2'(\tau)\mu - r \right\} P = 0 \tag{49}
\]

• After dividing by \( P \) and collecting terms we get

\[
\left\{ \frac{1}{2}\sigma_1^2B_1^2(\tau) + \frac{1}{2}\sigma_2^2B_2^2(\tau) + \rho\sigma_1\sigma_2B_1(\tau)B_2(\tau) \\
- \lambda_1\sigma_1B_1(\tau) + (\kappa_2\theta - \lambda_2\sigma_2)B_2(\tau) - A'(\tau) \right\} \\
- \left\{ \kappa_1B_1(\tau) + B_1'(\tau) + 1 \right\} r \\
+ \left\{ \kappa_1B_1(\tau) - \kappa_2B_2(\tau) - B_2'(\tau) \right\} \mu = 0 \tag{50}
\]

Gaussian central-tendency model – 3

• Since (50) must hold for all values of \( r \) and \( \mu \), we have

\[
B_1'(\tau) = -\kappa_1B_1(\tau) - 1 \tag{51}
\]

\[
B_2'(\tau) = \kappa_1B_1(\tau) - \kappa_2B_2(\tau) \tag{52}
\]

\[
A_1'(\tau) = \frac{1}{2}\sigma_1^2B_1^2(\tau) + \frac{1}{2}\sigma_2^2B_2^2(\tau) + \rho\sigma_1\sigma_2B_1(\tau)B_2(\tau) \\
- \lambda_1\sigma_1B_1(\tau) + (\kappa_2\theta - \lambda_2\sigma_2)B_2(\tau). \tag{53}
\]

• ODE solutions:

\[
B_1(\tau) = \frac{e^{-\kappa_1\tau} - 1}{\kappa_1} \tag{54}
\]

\[
B_2(\tau) = \frac{e^{-\kappa_2\tau} - 1}{\kappa_2} - \frac{e^{-\kappa_1\tau} - e^{-\kappa_2\tau}}{\kappa_1 - \kappa_2} \tag{55}
\]

\[
A(\tau) = \int_0^\tau A'(s)ds, \quad \text{where } A'(s) \text{ is the RHS of (53).} \tag{56}
\]