2.3. AFFINE MODELS OF THE TERM STRUCTURE

- Fundamental asset pricing concept - The pricing of any financial asset is based on a very intuitive result - the price corresponds to the present value of the future asset pay-off:

(1) $P_t = E_t[P_{t+1} M_{t+1}]$

being P_t the price of a financial asset providing nominal cash-flows and M_{t+1} the nominal stochastic discount factor (sdf) or pricing kernel, as it is the determining variable of P_t .
In fact, solving equation (1) forward, the asset price may be written solely as a function of the pricing kernel, as:

(2) $P_t = E_t[M_{t+1} \ldots M_{t+n}]$

- Asset prices and returns are related to their risk, i.e., to the asset capacity of offering higher cash-flows when they are more needed and valued.
- Actually, the more an asset helps to smooth income fluctuations, the less risky it is and the higher will be its demand for ensuring against "bad times".
- Considering that

$$E(XY) = E(X)E(Y) + COV(X,Y)$$

- Equation (1) may be written as:

(3) $$P_t = E_t[P_{t+1}]E_t[M_{t+1}] + Cov_t[P_{t+1}, M_{t+1}]$$

- When the asset is riskless, its pay-off in t+1 is known in t with certainty => P_{t+1} may be considered as a constant in t, which implies, from (1):

(4) $$\frac{P_t}{P_{t+1}} = E_t[M_{t+1}]$$

- As the LHS of (4) is the inverse of the risk-free asset's gross return, denoted by $1 + i_{t+1}^f$, replacing in equation (3) $E_t[M_{t+1}]$ by $1/1 + i_{t+1}^f$, it is obtained:

(3) $P_t = E_t[P_{t+1}]E_t[M_{t+1}] + Cov_t[P_{t+1}, M_{t+1}]$

(5) $$P_t = E_t[P_{t+1}]\frac{1}{1 + i_{t+1}^f} + Cov_t[P_{t+1}, M_{t+1}]$$

The asset price is the discounted expected value of its future pay-off or price, adjusted by the covariance of its return with the sdf.

- As it will become clear later, this covariance consists in a risk factor and it is positive for assets that pay higher returns when they are more needed.

- The same result may be obtained for interest rates, instead of prices. Actually, dividing both sides of equation (1) by P_t, one gets:

(1) $P_t = E_t[P_{t+1} M_{t+1}]$

(6) $$1 = E_t\left[(1 + i_{t+1}) M_{t+1}\right]$$

- Applying the already used statistical result

$$E(XY) = E(X)E(Y) + COV(X,Y)$$

to (6) it is obtained

(6) $1 = E_t[(1 + i_{t+1})M_{t+1}]$

(7) $$E_t(1+i_{t+1}) \cdot E_t(M_{t+1}) + Cov(i_{t+1}, M_{t+1}) = 1 \Leftrightarrow E_t(1+i_{t+1}) = \frac{[1 - Cov(i_{t+1}, M_{t+1})]}{E_t(M_{t+1})}$$

- Following equation (4) we obtain:

(4) $\frac{P_t}{P_{t+1}} = E_t[M_{t+1}]$

(8) $$E_t(1+i_{t+1}) = \frac{1}{E_t(M_{t+1})} - \frac{Cov(i_{t+1}, M_{t+1})}{E_t(M_{t+1})} \Leftrightarrow E_t(1+i_{t+1}) = (1+i^f_{t+1}) - \frac{Cov(i_{t+1}, M_{t+1})}{E_t(M_{t+1})}$$

- Therefore, we get:

(9) $$E_t[i_{t+1}] = i^f_{t+1} - \frac{Cov_t[M_{t+1}, i_{t+1}]}{E_t[M_{t+1}]}$$

The interest rate of an asset results from the risk-free rate, adjusted by a risk factor => the lower the covariance, the higher the risk and the interest rate.

- With some additional self-explanatory algebra, the following result is obtained:

(10) $$E_t[i_{t+1}] = i^f_{t+1} + \frac{Cov_t[M_{t+1}, i_{t+1}]}{Var_t[M_{t+1}]} \cdot \left(-\frac{Var_t[M_{t+1}]}{E_t[M_{t+1}]} \right) = i^f_{t+1} + \beta_{i_{t+1}, M_{t+1}} \lambda .$$

- $\beta_{i_{t+1}, M_{t+1}}$ is the coefficient of a regression of i_{t+1} on M_{t+1}

↓

- β measures the correlation between the asset's return and the stochastic discount factor (sdf) - quantity of risk.

- Market price of risk: $\lambda = -\frac{Var_t[M_{t+1}]}{E_t[M_{t+1}]}$

- From equation (8), denoting by $\rho_{M_{t+1},i_{t+1}}$ the correlation coefficient between the sdf and the asset's rate of return and $\sigma_{M_{t+1}}$ and $\sigma_{i_{t+1}}$, the excess return of any asset over the risk-free asset is: **(8)** $E_t(1+i_{t+1}) = \frac{1}{E_t(M_{t+1})} - \frac{Cov(i_{t+1},M_{t+1})}{E_t(M_{t+1})} \Leftrightarrow E_t(1+i_{t+1}) = (1+i^f_{t+1}) - \frac{Cov(i_{t+1},M_{t+1})}{E_t(M_{t+1})}$

(11) $$\Lambda_t = E_t[i_{t+1}] - i^f_{t+1} = -\rho_{M_{t+1},i_{t+1}} \frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]}$$

- <u>Equation (11) illustrates a basic result in finance theory:</u> the excess return of any asset over the risk-free asset depends on the correlation of its rate of return with the sdf => an asset with pay-off negatively correlated to the sdf is riskier.

- The mean-variance frontier will correspond to the limiting values of equation (11) => expected values and standard-deviations must lie in the interval $\left[-\frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]}, \frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]}\right]$.

(11) $\Lambda_t = E_t[i_{t+1}] - i^f_{t+1} = -\rho_{M_{t+1},i_{t+1}} \frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]}$

mean-variance region

minimum risk (frontier): $\rho_{M_{t+1},i_{t+1}} = 1$

- <u>As on the frontier all asset returns are perfectly correlated with the sdf, all asset returns are also perfectly correlated with each other</u> => it is possible to define the return of any asset as a linear combination of the returns of any 2 other assets - market or wealth portfolio and the risk-free asset:

(12) $E_t[i_{t+1}] = \beta_{i_{t+1},i^W_{t+1}} E[i^W_{t+1}] + \left(1-\beta_{i_{t+1},i^W_{t+1}}\right) i^f_{t+1} = i^f_{t+1} + \beta_{i_{t+1},i^W_{t+1}} \left(E[i^W_{t+1}] - i^f_{t+1}\right)$

CAPM

i^W_{t+1} - Rate of return of market portfolio

$$\textbf{(10)}\quad E_t[i_{t+1}] = i^f_{t+1} + \frac{Cov_t[M_{t+1}, i_{t+1}]}{Var_t[M_{t+1}]} \cdot \left(-\frac{Var_t[M_{t+1}]}{E_t[M_{t+1}]} \right) = i^f_{t+1} + \beta_{i_{t+1}, M_{t+1}} \lambda .$$

$$\textbf{(12)}\quad E_t[i_{t+1}] = \beta_{i_{t+1}, i^W_{t+1}} E[i^W_{t+1}] + (1 - \beta_{i_{t+1}, i^W_{t+1}}) i^f_{t+1} = i^f_{t+1} + \beta_{i_{t+1}, i^W_{t+1}} (E[i^W_{t+1}] - i^f_{t+1})$$

- (10) + (12) => CAPM assumes the sdf as a function of the gross rate of return of the wealth market portfolio, while the market price of risk is the spread between the expected market portfolio return and the risk-free asset return.

- A representative investor solves the following optimisation problem:

(13) $$\text{Max } E_t\left[\sum_{j=0}^{\infty}\delta^j U(C_{t+j})\right]$$

δ - time-constant discount factor

C_{t+j} - investor's consumption in the period *t+j*

$U(C_{t+j})$ - utility of consumption in the period *t+j*

- Optimisation problem in a two-period setting:

(14) $$\underset{\zeta}{\text{Max}}\ U(C_t)+\delta E_t\left[U(C_{t+1})\right]$$
$$\text{s.t. } C_t = e_t - P_t^R\zeta$$
$$C_{t+1} = e_{t+1} + P_{t+1}^R\zeta$$

e – income

P^R - price of a financial asset providing real cash-flows

ζ - number of asset units bought

Consumption in t+1 = Income in t+1 + Proceeds of the Investment done at t

Consumption = Income – financial investments (or savings)
t and *t+1* – present and future

- Inserting the constraints in the objective function in (14), the following optimisation problem arises:

(15) $$\underset{\zeta}{\mathrm{Max}}\, U(C_t, C_{t+1}) = U(e_t - P_t^R \zeta) + \delta E_t\left[U(e_{t+1} + P_{t+1}^R \zeta)\right]$$

- Solution:

(16) $$\frac{\partial U(\ )}{\partial \zeta} = 0 \Leftrightarrow P_t^R U'(C_t) = \delta E_t\left[U'(C_{t+1}) P_{t+1}^R\right]$$

- In equilibrium:

- marginal utility of consuming one real monetary unit less at time *t* = discounted expected value of the marginal utility of consuming at time *t+1* the proceeds of an investment of P^R monetary units at time *t* in the financial asset.

- Consumption CAPM (CCAPM) equation (from 16):

(17) $$P_t^R = E_t\left[P_{t+1}^R D_{t+1}\right]$$

being

(18) $$D_{t+1} = \delta \frac{U'(C_{t+1})}{U'(C_t)}$$

intertemporal marginal rate of substitution, stochastic discount factor (sdf) or pricing kernel

- Real assets are usually scarce.

- The most well known are inflation-indexed Government bonds and they exist only in a few countries (UK, US and France are the most liquid, but these securities have also been issued in Germany, Canada, Greece, Australia, Italy, Japan, Russia, Sweden, Spain, Hong-Kong, Iceland, India, Brazil and Mexico).

- Equation (17) is often adapted to nominal assets:

(19) $$\frac{P_t}{Q_t} = E_t\left[\frac{P_{t+1}}{Q_{t+1}} D_{t+1}\right]$$

similar to (1), with $M_{t+1} = \frac{D_{t+1}}{\Pi_{t+1}}$ and $\Pi_{t+1} = \frac{Q_{t+1}}{Q_t}$ => real and nominal assets are priced in the same way.

(11) $$\Lambda_t = E_t\left[i_{t+1}\right] - i_{t+1}^f = -\rho_{M_{t+1},i_{t+1}} \frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t\left[M_{t+1}\right]}$$

(18) $$D_{t+1} = \delta \frac{U'(C_{t+1})}{U'(C_t)}$$

- **CCAPM:** an asset will pay a higher return or is riskier when the correlation of its return with the marginal utility of consumption is lower, i.e. when consumption is higher => the asset is riskier when it pays more when those cash-flows are less needed.

- Affine models: log-linear relationship between asset prices and the sdf, on one side, and the factors or state variables, on the other side.

- These models were originally developed by Duffie and Kan (1996), for the term structure of interest rates.

- They allow for a parsimonious representation of the TSIR dynamics, as a function of a given number of observed or unobserved factors (or state variables).

- Equation (1) in logs:

(20) $$p_t = \log\left(E_t\left[P_{t+1}M_{t+1}\right]\right)$$

- Assuming joint log-normality of asset prices and discount factor => if *log X ~ N(*μ,σ^2*)* then *log E(X)* = $\mu + \sigma^2/2$ (as *X* is lognormally distributed, being its mean *E(X) = exp(*$\mu + \sigma^2/2$*)* => basic equation considered in the affine models:

(21) $$p_t = E_t\left[m_{t+1} + p_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{t+1}\right]$$

μ σ^2

- DK models: multifactor affine models of the term structure, where the pricing kernel is a linear function of several factors

$$z_t^T = \left(z_{1,t} \dots z_{k,t}\right)$$

- **DK models advantages**:

(i) Accommodate the most important term structure models, from Vasicek (1977) and CIR one-factor models to multi-factor models.

(ii) Allow the estimation of the term structure simultaneously on a cross-section and time-series basis.

(iii) Provide a way of computing and estimating simple closed-form expressions for the spot, forward, volatility and term premium curves.

- Discount factors:

(22) $$-m_{t+1} = \xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}$$

V(*Zt*) - variance matrix of the random shocks on the sdf, defined as a diagonal matrix with elements $v_i(z_t) = \alpha_i + \beta_i^T z_t$ and No.rows/columns equal to the No. factors.

et - independent shocks $\varepsilon_t \sim N(0, I)$

λ^T - market prices of risks, as they govern the covariance between the stochastic discount factor and the yield curve factors.

- Higher λs ⇔ higher covariance between the discount factor and the asset return ⇔ lower expected rate of returns or lower risk.

- Another way to write the pricing kernel (from (22)):

(23)

$$-m_{t+1} = \xi + \gamma_1 z_{1t} + \gamma_2 z_{2t} + \cdots + \gamma_k z_{kt} + \lambda_1 \sigma_{1t} \varepsilon_{1,t+1} + \lambda_2 \sigma_{2t} \varepsilon_{2,t+1} + \cdots + \lambda_k \sigma_{kt} \varepsilon_{k,t+1}$$

- The k factors z_t are defined as mean reverting, forming a k-dimensional vector:

(24) $$z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$$

θ - long-run mean of the factors.

Φ - has positive diagonal elements, that determine the speed of convergence of the factors to the long-term mean, ensuring that the factors are stationary;

- From (24), we have the factors as follows:

(25) $z_{i,t+1} = (1-\phi_i)\theta_i + \phi_i z_{i,t} + \sigma_{i,t}\varepsilon_{i,t+1}$, where $\sigma_{it} = \sqrt{\alpha_i + \beta_{i1}z_{1t} + \beta_{i2}z_{2t} + \ldots + \beta_{ik}z_{kt}}$

- <u>Asset prices are also log-linear functions of the factors</u>:

(26) $$-p_{n,t} = A_n + B_n^{\ T} z_t$$

n - term to maturity

A_n and B_n - vectors of parameters to be estimated.

B_n - factor loadings (impact of a random shock on the factors over the log of asset prices).

- The question now is how to relate the parameters of the stochastic factors to the parameters of bond prices and the term structure of interest rates ⇔ identification of the parameters.

- In term structure models, the identification of the parameters is easier assuming that the term structure is modelled using zero-coupon bonds paying 1 monetary unit => the log of the maturing bond price $p_{0,t}$ = 0 => (from (26)) $A_0 = B_0 = 0$

- According to (22) and (26), the 1st term on the RHS of (21) is in (27):

(21) $p_t = E_t\left[m_{t+1} + p_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{t+1}\right]$

(22) $-m_{t+1} = \xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}$

(26) $-p_{n,t} = A_n + B_n{}^T z_t$ => in t+1: $-p_{n-1,t+1} = A_{n-1} + B_{n-1}^T z_{t+1}$

(27) $E_t\left[m_{t+1} + p_{t+1}\right] = E_t\left\{-\left[\xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}\right] - \left(A_{n-1} + B_{n-1}{}^T z_{t+1}\right)\right\}$

- Using the factor definition in (24) $z_{t+1} = (I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$ we get from (27), replacing in ($z_{t+1} = (I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$

(27) $$E_t[m_{t+1}+p_{t+1}] = E_t\left\{-\left[\xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2}\varepsilon_{t+1}\right] - \left(A_{n-1} + B_{n-1}{}^T z_{t+1}\right)\right\}$$

(28) $$E_t[m_{t+1}+p_{t+1}] = -E_t\left\{\begin{array}{l}\xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2}\varepsilon_{t+1} + A_{n-1} \\ +B_{n-1}{}^T\left[(I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}\right]\end{array}\right\}$$

- Computing the expected value and given that the random shocks are assumed to have zero mean => all terms in ε_{t+1} will be cancelled => (28) may be simplified to:

(29) $$\begin{aligned}E_t[m_{t+1}+p_{t+1}] &= -\left\{\xi + \gamma^T z_t + A_{n-1} + B_{n-1}{}^T\left[(I-\Phi)\theta + \Phi z_t\right]\right\} \\ &= -\left[A_{n-1} + \xi + B_{n-1}{}^T(I-\Phi)\theta\right] - \left(\gamma^T + B_{n-1}{}^T\Phi\right)z_t\end{aligned}$$

(22) $-m_{t+1} = \xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}$ + (24) $z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2} \varepsilon_{t+1}$ +

(26) $-p_{n,t} = A_n + B_n^{\ T} z_t$

- To obtain the variance in the 2nd term on the RHS of

(21) $p_t = E_t[m_{t+1} + p_{t+1}] + 0.5 \cdot Var_t[m_{t+1} + p_{t+1}]$, all constant terms will be eliminated:

(30)
$$Var_t\left[\lambda^T V(z_t)^{1/2} + B_{n-1}^{\ T} V(z_t)^{1/2}\right] = Var_t\left[(\lambda^T + B_{n-1}^{\ T}) V(z_t)^{1/2}\right]$$
$$= (\lambda^T + B_{n-1}^{\ T})\left[Var_t(\alpha + \beta^T z_t)\right]^{1/2} (\lambda + B_{n-1})$$

- Evidencing the independent terms and the terms in z_t :

(31)
$$Var_t\left[\lambda^T V(z_t)^{1/2} + B_{n-1}^{\ T} V(z_t)^{1/2}\right] = (\lambda + B_{n-1})^T \alpha (\lambda + B_{n-1}) + (\lambda + B_{n-1})^T \beta^T z_t (\lambda + B_{n-1})$$

- From

(21) $$p_t = E_t\left[m_{t+1} + p_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{t+1}\right]$$

(29) $$E_t\left[m_{t+1} + p_{t+1}\right] = -\left\{\xi + \gamma^T z_t + A_{n-1} + B_{n-1}{}^T\left[(I-\Phi)\theta + \Phi z_t\right]\right\}$$
$$= -\left[A_{n-1} + \xi + B_{n-1}{}^T(I-\Phi)\theta\right] - \left(\gamma^T + B_{n-1}{}^T\Phi\right)z_t$$

(31) $$Var_t\left[\lambda^T V(z_t)^{1/2} + B_{n-1}{}^T V(z_t)^{1/2}\right] = (\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1}) + (\lambda + B_{n-1})^T \beta^T z_t(\lambda + B_{n-1})$$

(32) $$-p_{n,t} = \left\{\left[A_{n-1} + \xi + B_{n-1}{}^T(I-\Phi)\theta\right] + \left(\gamma^T + B_{n-1}{}^T\Phi\right)z_t\right\}$$
$$-\frac{1}{2}\left[(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1}) + (\lambda + B_{n-1})^T \beta^T z_t(\lambda + B_{n-1})\right]$$

- Putting in evidence the independent terms and the terms in z_t , from (32) one obtains:

(33) $$-p_{n,t} = \left\{\left[A_{n-1} + \xi + B_{n-1}{}^T(I-\Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})\right]\right\}$$
$$+\left(\gamma^T + B_{n-1}{}^T\Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^T(\lambda + B_{n-1})\right)z_t$$

- Comparing the coefficients on the RHS of

(26) $-p_{n,t} = A_n + B_n^{\ T} z_t$

to the independent term and the term associated to the factor in

(33) $$-p_{n,t} = \left\{\left[A_{n-1} + \xi + B_{n-1}^{\ T}(I-\Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})\right]\right\} + \left(\gamma^T + B_{n-1}^{\ T}\Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^T(\lambda + B_{n-1})\right) z_t$$

the recursive restrictions in (34) and (35) are obtained:

(34) $$A_n = A_{n-1} + \xi + B_{n-1}^{\ T}(I-\Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})$$

(35) $$B_n^{\ T} = \gamma^T + B_{n-1}^{\ T}\Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^{T'}(\lambda + B_{n-1})$$

- Considering that the continuously compounded yield is:

(36) $$y_{n,t} = -\frac{\log P_{n,t}}{n}$$

- From (36) and

(26) $$-p_{n,t} = A_n + B_n{}^T z_t$$

the yield curve is defined as:

(37) $$y_{n,t} = \frac{1}{n}\left(A_n + B_n{}^T z_t\right)$$

- From equations (34), (35) and (37)

(34) $A_n = A_{n-1} + \xi + B_{n-1}{}^T (I - \Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})$ and

(35) $B_n{}^T = \gamma^T + B_{n-1}{}^T \Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^{T'}(\lambda + B_{n-1})$

(37) $y_{n,t} = \frac{1}{n}\left(A_n + B_n{}^T z_t\right)$

as well as the normalisation $A_0 = B_0 = 0$, it is obtained the short-term rate (as with n=1, A_{n-1} and B_{n-1} will be A_0 and B_0 correspondingly, both equal to 0):

(38) $y_{1,t} = \xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t$

- Correspondingly, using the definition of the factors in

(24) $z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$

- and solving backwards, one gets:

(39) $$E_t(z_{i,t+n}) = (1-\phi_i)\theta_i + \phi_i z_{i,t+n-1} = (1-\phi_i)\theta_i + \phi_i\left[(1-\phi_i)\theta_i + \phi_i z_{i,t+n-2}\right]$$
$$= \ldots = \sum_{j=1}^{n}\left[\phi_i^{j-1}(1-\phi_i)\theta_i\right] + \phi_i^n z_{i,t}$$

- Given that the expression in the sum corresponds to the sum of the first *n*-terms of a geometric progression with rate ϕ and first term equal to $(1-\phi_i)\theta_i$, equivalent to $\left[(1-\phi_i)\theta_i\right]\frac{1-\phi_i^n}{1-\phi_i}$, the following expression is obtained:

(40) $$E_t(z_{i,t+n}) = \left[(1-\phi_i)\theta_i\right]\frac{1-\phi_i^n}{1-\phi_i} + \phi_i^n z_{i,t} = \theta_i(1-\phi_i^n) + \phi_i^n z_{i,t}$$

- To calculate the expected value of future short-term interest rate, one can use

(38) $$y_{1,t} = \xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t$$

- and plug

(40) $$E_t\left(z_{i,t+n}\right) = \left[\left(1-\phi_i\right)\theta_i\right]\frac{1-\phi_i^n}{1-\phi_i} + \phi_i^n z_{i,t} = \theta_i\left(1-\phi_i^n\right) + \phi_i^n z_{i,t}$$

writing in matrix form (as the matrices involved in the computations are diagonal)

(41) $$\begin{aligned} E_t\left(y_{1,t+n}\right) &= E_t\left(\xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_{t+n}\right) \\ &= \xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] E_t\left(z_{t+n}\right) \\ &= \xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right]\left[\left(I-\Phi^n\right)\theta + \Phi^n z_t\right] \end{aligned}$$

- From

(24) $$z_{t+1} = (I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$$

(37) $$y_{n,t} = \frac{1}{n}\left(A_n + B_n{}^T z_t\right)$$

one gets the variance of interest rates:

(42) $$Var_t\left(y_{n,t+1}\right) = \frac{1}{n^2} B_n{}^T V(z_t) B_n$$

- Instantaneous or one-period forward rate = log of the inverse of the gross return =>

(43) $$f_{n,t} = p_{n,t} - p_{n+1,t}$$

- From

(26) $-p_{n,t} = A_n + B_n^{\ T} z_t$

(43) $f_{n,t} = p_{n,t} - p_{n+1,t}$

(34) $A_n = A_{n-1} + \xi + B_{n-1}^{\ T}(I-\Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})$

(35) $B_n^{\ T} = \gamma^T + B_{n-1}^{\ T}\Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^{T}$

one gets the instantaneous or one-period forward curve:

(44)
$$f_{n,t} = \left(A_{n+1} + B_{n+1}^{\ T} z_t\right) - \left(A_n + B_n^{\ T} z_t\right) = \left(A_{n+1} - A_n\right) + \left(B_{n+1}^{\ T} - B_n^{\ T}\right) z_t =$$
$$= \left[\xi + B_n^{\ T}(I-\Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \alpha_i\right] + \left[\gamma^T + B_n^{\ T}(\Phi - I) - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \beta_i^T\right] z_t$$

- From the current and the one-period ahead bond prices in the price equation in

(26) $-p_{n,t} = A_n + B_n{}^T z_t$

and the short-term rate in

(38) $y_{1,t} = \xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t$

it is obtained the term premium as the difference between the one-period expected return and the short-term interest rate:

(45)
$$\begin{aligned}
\Lambda_{n,t} &= E_t p_{n,t+1} - p_{n+1,t} - y_{1,t} \\
&= E_t\left(-A_n - B_n{}^T z_{t+1}\right) + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left(\xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t\right) \\
&= -A_n - B_n{}^T\left[(I-\Phi)\theta + \Phi z_t\right] + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left[\xi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i + \left(\gamma^T - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right) z_t\right] \\
&= \left\{A_{n+1} - \left[A_n + \xi + B_n{}^T (I-\Phi)\theta\right] - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i\right\} + \left[B_{n+1}{}^T - \left(\gamma^T + B_n{}^T \Phi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right)\right] z_t
\end{aligned}$$

- From the recursive restrictions on the factor loadings in

(27) $$A_n = A_{n-1} + \xi + B_{n-1}{}^T (I-\Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})$$

(28) $$B_n{}^T = \gamma^T + B_{n-1}{}^T \Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^{T'} (\lambda + B_{n-1})$$

equation (45) can be simplified as in (46):

(45)
$$\begin{aligned}
\Lambda_{n,t} &= E_t p_{n,t+1} - p_{n+1,t} - y_{1,t} \\
&= E_t\left(-A_n - B_n{}^T z_{t+1}\right) + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left(\xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t\right) \\
&= -A_n - B_n{}^T\left[(I-\Phi)\theta + \Phi z_t\right] + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left[\xi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i + \left(\gamma^T - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right) z_t\right] \\
&= \left\{A_{n+1} - \left[A_n + \xi + B_n{}^T (I-\Phi)\theta\right] - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i\right\} + \left[B_{n+1}{}^T - \left(\gamma^T + B_n{}^T \Phi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right)\right] z_t
\end{aligned}$$

(46) $$\Lambda_{n,t} = -\sum_{i=1}^{k} \lambda_i B_{i,n} \alpha_i - \frac{B_{i,n}^2 \alpha_i}{2} - \left(\sum_{i=1}^{k} \lambda_i B_{i,n} \beta_i - \frac{B_{i,n}^2 \beta_i}{2}\right) z_t$$

- The term premium can alternatively be calculated from the pricing equation

(21) $p_t = E_t[m_{t+1} + p_{t+1}] + 0.5 \cdot Var_t[m_{t+1} + p_{t+1}]$

- Solving in order to $E_t[p_{t+1}]$, we get:

(47) $E_t[p_{t+1}] = p_t - E_t[m_{t+1}] - 0.5 \cdot Var_t[m_{t+1} + p_{t+1}]$

(48) $E_t p_{n,t+1} - p_{n+1,t} = \{p_{n+1,t} - E_t[m_{t+1}] - 0.5 \cdot Var_t[m_{t+1} + p_{n,t+1}]\} - p_{n+1,t}$

$$= -E_t[m_{t+1}] - \frac{Var_t(m_{t+1}) + Var_t(p_{n,t+1}) + 2Cov(m_{t+1}, p_{n,t+1})}{2}$$

- Given that the $Cov(m_{t+1}, p_{n,t+1}) = Cov(m_{t+1}, i_{n,t+1})$, as $p_{n,t+1}$ is the only stochastic component in the rate of return ($i = P_{t+1}/P_t = lnP_{t+1} - lnP_t = p_{t+1} - p_t$), the previous equation is equal to:

(49) $E_t p_{n,t+1} - p_{n+1,t} = -E_t[m_{t+1}] - \frac{Var_t(m_{t+1})}{2} - \frac{Var_t(i_{n,t+1})}{2} - Cov(m_{t+1}, i_{n,t+1})$

- According to

(21) $$p_t = E_t\left[m_{t+1} + p_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{t+1}\right]$$

and considering the assumption $p_{0t} = 0$

- Solving in order to $E_t\left[p_{t+1}\right]$, we get the price of the short-term bond:

(50) $$p_{1,t} = E_t\left[m_{t+1} + p_{0,t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{0,t+1}\right] = E_t\left[m_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1}\right]$$

- From

(49) $$E_t p_{n,t+1} - p_{n+1,t} = -E_t[m_{t+1}] - \frac{Var_t(m_{t+1})}{2} - \frac{Var_t(i_{n,t+1})}{2} - Cov(m_{t+1}, i_{n,t+1})$$

(50) $$p_{1,t} = E_t\left[m_{t+1} + p_{0,t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1} + p_{0,t+1}\right] = E_t\left[m_{t+1}\right] + 0.5 \cdot Var_t\left[m_{t+1}\right]$$

(36) $$y_{n,t} = -\frac{\log P_{n,t}}{n}$$

(38) $$y_{1,t} = \xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t$$ and $$\Lambda_{n,t} = E_t p_{n,t+1} - p_{n+1,t} - y_{1,t}$$

the term premium will be equal to:

(51) $$\Lambda_{n,t} = -COV_t(i_{n,t+1}, m_{t+1}) - Var_t(i_{n,t+1}) / 2$$

Risk premium determined by the covar. of the asset's rate of return with the stochastic discount factor => the lower the covar., the higher the risk premium is.

- As from

(26) $-p_{n,t} = A_n + B_n^{\,T} z_t$

we get

(52) $i_{n,t+1} = p_{n,t+1} - p_{n+1,t} = -A_n - B_n^T z_{t+1} + A_{n+1} + B_{n+1}^T z_t$

the covariance in

(51) $\Lambda_{n,t} = -COV_t\left(i_{n,t+1}, m_{t+1}\right) - Var_t\left(i_{n,t+1}\right)/2$ is $-B_n^T COV_t\left(z_{t+1}, m_{t+1}\right)$

- Consequently, equation (51) for **the term premium** becomes equivalent to:

(53) $\Lambda_{n,t} = B_n^T COV\left(z_{t+1}, m_{t+1}\right) - B_n^T Var_t\left(z_{t+1}\right) B_n / 2$

- From

(22) $-m_{t+1} = \xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}$ and

(24) $z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2} \varepsilon_{t+1}$

the term premium in

(53) $\Lambda_{n,t} = B_n^T COV(z_{t+1}, m_{t+1}) - B_n^T Var_t(z_{t+1}) B_n / 2$

may be written as:

(54) $\Lambda_{n,t} = -\lambda^T V(z_t) B_n - \frac{B_n^T V(z_t) B_n}{2}$

at least one of the market prices of risk must be negative in order to have a positive term premium.

- One-factor models were the first step in modelling the term structure of interest rates.

- These models are grounded on the estimation of bond yields as functions of the short-term interest rate.

- Vasicek (1977) presented the whole term structure as a function of a single factor, the short-term interest rate, whose volatility was assumed to be constant - **Gaussian models**:

K	θ_i	Φ	α_ι	β_ι	ξ	γ_ι
1	0 or θ *	ϕ	σ^2	0	$\delta+\lambda^2/2$	1

* Depending on whether the true values of interest rates or their differences to the mean are considered.

- The Cox *et al*. (1985a) model added the stochastic volatility feature to the Vasicek model, avoiding interest rates to go negative, as in the Vasicek model. Thus, it corresponds to an analogous particular case of the DK model, with $\alpha_i = 0$ and $\beta_i = \sigma_i^2$.

- Affine models may be classified according to:

(i) number of factors considered;

(ii) volatility properties.

- According to Litterman and Scheinkman (1991), the pronounced hump-shape of the US yield curve => 3 factors are required to explain the shifts in the whole term structure of interest rates.
- These factors are usually identified as the level, the slope and the curvature, being the level often responsible for the most important part of interest rate variation.

- Given the stochastic properties of interest rates volatility, Gaussian or constant volatility models are often rejected. Besides, these models impose constant volatility and one-period term premium curves (non-pure version of expectations theory).
- The forward rate also exhibits some shortcomings.
- Nonetheless, Gaussian models are used very often as:

(i) interest rate volatilities don't suffer significant changes during most periods;

(ii) Constant volatility models as much easier to implement, namely with non-observable or latent factor, given that the volatility depends on the square root of the factors in stochastic volatility models => signal restrictions have to be imposed, which is harder to do in iterative econometric processes.

Shortcomings of the forward rates under constant volatility:

- It can be shown that the asymptotic forward rate cannot be simultaneously finite and time-varying.

(37)
$$f_{n,t} = \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left(A_n + B_n{}^T z_t\right) = \left(A_{n+1} - A_n\right) + \left(B_{n+1}{}^T - B_n{}^T\right) z_t =$$
$$= \left[\xi + B_n{}^T (I - \Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \alpha_i\right] + \left[\gamma^T + B_n{}^T(\Phi - I) - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \beta_i^T\right] z_t$$

- β_i=0 in Gaussian models => the forward rate may be written as:

(55) $$f_{n,t} = \xi + B_n{}^T (I - \Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \alpha_i + \left[\gamma^T + B_n{}^T(\Phi - I)\right] z_t$$

- Vasicek models => volatility is constant => last term of the RHS of

(35) $B_n{}^T = \gamma^T + B_{n-1}{}^T \Phi - \frac{1}{2}\left(\lambda + B_{n-1}\right)^T \beta^T z_t \left(\lambda + B_{n-1}\right)$ is zero (β_i=0), while γ_i=1.

- Each factor loading in a multifactor Vasicek model corresponds to:

(56) $$B_{i,n} = 1 + \varphi_i + \varphi_i^{\,2} + \ldots + \varphi_i^{\,n} = \sum_{i=1}^{n} \varphi_i^{\,n-1} = u_1 \times \frac{1-r^n}{1-r} = \frac{1-\varphi_i^{\,n}}{1-\varphi_i}$$

- From

(55) $$f_{n,t} = \xi + B_n^{\,T}(I-\Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2 \alpha_i + \left[\gamma^T + B_n^{\,T}(\Phi - I)\right] z_t$$

the one-period forward rate may thus be written as:

(57) $$f_{n,t} = \xi + B_n^{\,T}(I-\Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i \sigma_i + \frac{1-\varphi_i^n}{1-\varphi_i}\sigma_i\right)^2 + \sum_{i=1}^{k}\left[\varphi_i^n z_{it}\right]$$

- Though this specification of the forward-rate curve accommodates very different shapes, the limiting forward rate cannot be simultaneously finite and time-varying.

- In fact, if $\varphi_i < 1$, the asymptotic value will not depend on the factors, as the limit of the last term on the RHS is zero.

- If $\varphi_i = 1$, the limiting value of the instantaneous forward becomes time-varying but assumes infinite values, as $\frac{1-\varphi_i^n}{1-\varphi_i} = n$ in this case.

- According to the non-pure version of the expectations theory, the forward rate corresponds to:

(58) $$f_{n,t} = E_t\left(y_{1,t+n}\right) + \Lambda_n$$

→ the term premium doesn't have the subscript *t*, as it is constant under this theory

- Law of iterated expectations => (58) corresponds to:

(59) $$f_{n,t} = E_t\left(E_{t+1}\left(y_{1,t+n}\right) + \Lambda_n\right) = E_t\left(f_{n-1,t+1}\right) + \left(\Lambda_n - \Lambda_{n-1}\right)$$

⇓

- The Gaussian model implies that the forward rates are martingales if the term structure of risk premium is constant along time (non-pure version of the expectations theory) and flat.

- Solving (59) in order to the expected future value of the forward, one can assess whether the expectations theory holds, by performing the following regression:

(59) $f_{n,t} = E_t\left(E_{t+1}\left(y_{1,t+n}\right) + \Lambda_n\right) = E_t\left(f_{n-1,t+1}\right) + \left(\Lambda_n - \Lambda_{n-1}\right)$

(60) $f_{n-1,t+1} - y_{1,t} = constant + c_n\left(f_{n,t} - y_{1,t}\right) + residual$

- If expectations theory holds => the best estimate for the difference between the forward and the current short-term rate at a future time will be the current value for that difference + a constant that embeds the slope of the term premium curve => $c_n = 1$

- A rejection of this hypothesis suggests that term premia vary with time, i.e., that the expectations theory does not hold.

(36) $y_{n,t} = -\frac{\log P_{n,t}}{n}$ => $y_{1,t} = -p_{1,t}$

- from (43) $f_{n,t} = p_{n,t} - p_{n+1,t}$

$f_{0,t} = p_{0,t} - p_{1,t} = 0 + y_{1,t} = y_{1,t}$

(61) $f_{0,t} = A_1 + B_1^T z_t$

- The theoretical values of c_n implied by the Gaussian model correspond to :

(62) $$c_n = \frac{Cov\left(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t}\right)}{Var\left(f_{n,t} - f_{0,t}\right)}$$

- Solving (59) in order to the expected future value of the forward, one can assess whether the expectations theory holds, by performing the following regression:

(59) $f_{n,t} = E_t\left(E_{t+1}\left(y_{1,t+n}\right) + \Lambda_n\right) = E_t\left(f_{n-1,t+1}\right) + \left(\Lambda_n - \Lambda_{n-1}\right)$

(60) $f_{n-1,t+1} - y_{1,t} = constant + c_n\left(f_{n,t} - y_{1,t}\right) + residual$

- If expectations theory holds => the best estimate for the difference between the forward and the current short-term rate at a future time will be the current value for that difference + a constant that embeds the slope of the term premium curve => $c_n = 1$

- A rejection of this hypothesis suggests that term premia vary with time, i.e., that the expectations theory does not hold.

(36) $y_{n,t} = -\frac{\log P_{n,t}}{n}$ => $y_{1,t} = -p_{1,t}$

- from (43) $f_{n,t} = p_{n,t} - p_{n+1,t}$

$f_{0,t} = p_{0,t} - p_{1,t} = 0 + y_{1,t} = y_{1,t}$

(61) $f_{0,t} = A_1 + B_1^T z_t$

- The theoretical values of c_n implied by the Gaussian model correspond to :

(62) $$c_n = \frac{Cov\left(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t}\right)}{Var\left(f_{n,t} - f_{0,t}\right)}$$

- Also from (43) $f_{n,t} = p_{n,t} - p_{n+1,t}$ and from (26) $-p_{n,t} = A_n + B_n{}^T z_t$

(63) $f_{n,t} = (A_{n+1} - A_n) + (B_{n+1}^T - B_n^T) z_t$

(64) $f_{n-1,t+1} = (A_n - A_{n-1}) + (B_n^T - B_{n-1}^T) z_{t+1}$

- The first difference in the covariance in (62) $c_n = \dfrac{Cov(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t})}{Var(f_{n,t} - f_{0,t})}$ becomes:

(65) $f_{n-1,t+1} - f_{0,t} = (A_n - A_{n-1} - A_1) + (B_n^T - B_{n-1}^T) z_{t+1} - B_1^T z_t$

(65) $$f_{n-1,t+1} - f_{0,t} = (A_n - A_{n-1} - A_1) + (B_n^T - B_{n-1}^T) z_{t+1} - B_1^T z_t$$

(24) $$z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2} \varepsilon_{t+1}$$

(66) $$f_{n-1,t+1} - f_{0,t} = (A_n - A_{n-1} - A_1) + (B_n^T - B_{n-1}^T)\left[(I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2} \varepsilon_{t+1}\right] - B_1^T z_t$$
$$= (A_n - A_{n-1} - A_1) + (B_n^T - B_{n-1}^T)(I - \Phi)\theta + (B_n^T - B_{n-1}^T)\Phi z_t + (B_n^T - B_{n-1}^T) V(z_t)^{1/2} \varepsilon_{t+1} - B_1^T z_t$$

(67) $$f_{n,t} - f_{0,t} = (A_{n+1} - A_n - A_1) - (B_1 + B_n - B_{n+1})^T z_t$$

- The covariance and the variance in (62) becomes: $c_n = \dfrac{Cov(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t})}{Var(f_{n,t} - f_{0,t})}$

(68) $$Cov(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t}) = (B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 - \Phi^T (B_n - B_{n-1}))$$

(69) $$Var(f_{n,t} - f_{0,t}) = (B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 + B_n - B_{n+1})$$

- To calculate c_n in (62), it is still necessary to calculate Γ_0 - the variance-covariance matrix of *z*, which corresponds to $E\{[z-E(z)][z-E(z)]^T\}$
- Specification of the factors in

(24) $z_{t+1} = (I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$

↓

(70) $z - E(z) = \Phi(z-\theta) + V$

↓

- Γ_0 components will be calculated as:

(71) $vec(\Gamma_0) = vec(\Phi\Gamma_0\Phi^T) + vec(V)$

↓ as $vec(ABC) = (C' \otimes A)vec(B)$

(72) $vec(\Gamma_0) = (I - \Phi \otimes \Phi^T)^{-1}.vec(V)$

- Kronecker product:

If $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{B}$ is a $p \times q$ matrix, then the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is the $mp \times nq$ block matrix:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix},$$

more explicitly:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & \cdots & a_{11}b_{1q} & \cdots & \cdots & a_{1n}b_{11} & a_{1n}b_{12} & \cdots & a_{1n}b_{1q} \\ a_{11}b_{21} & a_{11}b_{22} & \cdots & a_{11}b_{2q} & \cdots & \cdots & a_{1n}b_{21} & a_{1n}b_{22} & \cdots & a_{1n}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{p1} & a_{11}b_{p2} & \cdots & a_{11}b_{pq} & \cdots & \cdots & a_{1n}b_{p1} & a_{1n}b_{p2} & \cdots & a_{1n}b_{pq} \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ a_{m1}b_{11} & a_{m1}b_{12} & \cdots & a_{m1}b_{1q} & \cdots & \cdots & a_{mn}b_{11} & a_{mn}b_{12} & \cdots & a_{mn}b_{1q} \\ a_{m1}b_{21} & a_{m1}b_{22} & \cdots & a_{m1}b_{2q} & \cdots & \cdots & a_{mn}b_{21} & a_{mn}b_{22} & \cdots & a_{mn}b_{2q} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{m1}b_{p1} & a_{m1}b_{p2} & \cdots & a_{m1}b_{pq} & \cdots & \cdots & a_{mn}b_{p1} & a_{mn}b_{p2} & \cdots & a_{mn}b_{pq} \end{bmatrix}$$

- Example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \otimes \begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} = \begin{bmatrix} 1\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 2\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \\ 3\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} & 4\begin{bmatrix} 0 & 5 \\ 6 & 7 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1\times 0 & 1\times 5 & 2\times 0 & 2\times 5 \\ 1\times 6 & 1\times 7 & 2\times 6 & 2\times 7 \\ 3\times 0 & 3\times 5 & 4\times 0 & 4\times 5 \\ 3\times 6 & 3\times 7 & 4\times 6 & 4\times 7 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 & 10 \\ 6 & 7 & 12 & 14 \\ 0 & 15 & 0 & 20 \\ 18 & 21 & 24 & 28 \end{bmatrix}$$

- From

(62) $$c_n = \frac{Cov(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t})}{Var(f_{n,t} - f_{0,t})}$$

(68) $$Cov(f_{n-1,t+1} - f_{0,t}, f_{n,t} - f_{0,t}) = (B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 - \Phi^T (B_n - B_{n-1}))$$

(69) $$Var(f_{n,t} - f_{0,t}) = (B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 + B_n - B_{n+1})$$

$$f_{n-1,t+1} - y_{1,t} = constant + c_n (f_{n,t} - y_{1,t}) + residual$$

$$y_{n,t+1} - y_{n+1,t} = constant + d_n (y_{n+1,t} - y_{1,t}) + residual$$

we get

(73) $$c_n = \frac{(B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 - \Phi^T (B_n - B_{n-1}))}{(B_1 + B_n - B_{n+1})^T \Gamma_0 (B_1 + B_n - B_{n+1})}$$

⇓

(74) $$\lim_{n\to\infty} c_n = \frac{B_1{}^T \Gamma_0 B_1}{B_1{}^T \Gamma_0 B_1} = 1$$, as $B_{n-1} = B_n = B_{n+1}$ when $n \to \infty$

- Most empirical tests of the pure expectations hypothesis involve instead yield regressions of the form:

(62) $$y_{n,t+1} - y_{n+1,t} = constant + d_n\left(y_{n+1,t} - y_{1,t}\right) + residual$$

- By definition

(75) $$d_n = \frac{Cov\left(y_{n,t+1} - y_{n+1,t}, y_{n+1,t} - y_{1,t}\right)}{Var\left(y_{n,t+1} - y_{n+1,t}\right)}$$

- Due to (37) $y_{n,t} = \frac{1}{n}\left(A_n + B_n^{T} z_t\right)$, the 1st and 2nd elements of the covariance in (75) are correspondingly:

(76) $$Y_{n,t+1} - Y_{n+1,t} = \frac{A_n}{n} - \frac{A_{n+1}}{n+1} + \frac{B_n^T z_{t+1}}{n} - \frac{B_{n+1}^T z_t}{n+1}$$

(77) $$Y_{n+1,t} - Y_{1,t} = \frac{A_{n+1}}{n+1} - A_1 + \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t$$

- As all terms in *A* in (75) $d_n = \frac{Cov(y_{n,t+1} - y_{n+1,t}, y_{n+1,t} - y_{1,t})}{Var(y_{n,t+1} - y_{n+1,t})}$ are constant, they don't contribute to the covariance.

(78)

$$Cov(y_{n,t+1} - y_{n+1,t}, y_{n+1,t} - y_{1,t}) = Cov\left(\frac{A_n}{n} - \frac{A_{n+1}}{n+1} + \frac{B_n^T z_{t+1}}{n} - \frac{B_{n+1}^T z_t}{n+1}, \frac{A_{n+1}}{n+1} - A_1 + \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t\right)$$

$$= Cov\left(\frac{B_n^T\left[(I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}\right]}{n} - \frac{B_{n+1}^T z_t}{n+1}, \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t\right)$$

$$= Cov\left(\frac{B_n^T \Phi z_t}{n} - \frac{B_{n+1}^T z_t}{n+1}, \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t\right) = Cov\left[\frac{B_n^T \Phi z_t}{n} - \frac{\frac{n B_{n+1}^T z_t}{n+1}}{n}, \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t\right]$$

$$= Cov\left[\left(\frac{B_n^T \Phi}{n} - \frac{\frac{n B_{n+1}^T}{n+1}}{n}\right) z_t, \left(\frac{B_{n+1}^T}{n+1} - B_1^T\right) z_t\right] = \left(B_1^T - \frac{B_{n+1}^T}{n+1}\right)\Gamma_0\left(\frac{n B_{n+1}}{n+1} - \Phi' B_n\right)$$

- The variance in (75) $d_n = \dfrac{Cov(y_{n,t+1} - y_{n+1,t}, y_{n+1,t} - y_{1,t})}{Var(y_{n,t+1} - y_{n+1,t})}$ is:

(79) $$Var(y_{n,t+1} - y_{n+1,t}) = Cov\left(\frac{A_n}{n} - \frac{A_{n+1}}{n+1} + \frac{B_n^T z_{t+1}}{n} - \frac{B_{n+1}^T z_t}{n+1}, \frac{A_n}{n} - \frac{A_{n+1}}{n+1} + \frac{B_n^T z_{t+1}}{n} - \frac{B_{n+1}^T z_t}{n+1}\right)$$
$$= \left(B_1 - (n+1)^{-1} B_{n+1}\right)^T \Gamma_0 \left(B_1 - (n+1)^{-1} B_{n+1}\right)$$

(80) $$\lim_{n\to\infty} d_n = \frac{B_1^T \Gamma_0 (I - \Phi^T) B}{B_1^T \Gamma_0 B_1}$$, where $B = \lim_{n\to\infty} B_n$

- **2-factor constant volatility (i.e. Vasicek-type) model:**

K	θ_i	Φ	α_i	β_i	ξ	γ_i
2	θ or 0	$\begin{bmatrix}\varphi_1 & \vdots \\ \cdots & \varphi_2\end{bmatrix}$	σ_i^2	0	$\delta + \sum_{1=1}^{2} \frac{\lambda_i^2}{2}\sigma_i^2$	1

Stochastic discount factor: From (22)

(81) $$-m_{t+1} = \delta + \sum_{i=1}^{k}\left(\frac{\lambda_i^2}{2}\sigma_i^2 + z_{it} + \lambda_i \sigma_i \varepsilon_{i,t+1}\right)$$ $$-m_{t+1} = \xi + \gamma^T z_t + \lambda^T V(z_t)^{1/2} \varepsilon_{t+1}$$

Factors - first-order autoregressive processes with zero mean (corresponds to considering the differences between the "true" factors and their means): From (24)

(82) $$z_{i,t+1} = \varphi_i z_{it} + \sigma_i \varepsilon_{i,t+1}$$ $$z_{t+1} = (I - \Phi)\theta + \Phi z_t + V(z_t)^{1/2} \varepsilon_{t+1}$$

Bond prices: From (26)

(83) $$-p_{n,t} = A_n + B_{1,n} z_{1t} + B_{2,n} z_{2t}$$ $$-p_{n,t} = A_n + B_n{}^T z_t$$

Yield curve:

From (37)

(84) $y_{n,t} = \frac{1}{n}\left(A_n + B_{1,n} z_{1t} + B_{2,n} z_{2t}\right)$

$$y_{n,t} = \frac{1}{n}\left(A_n + B_n{}^T z_t\right)$$

Factor loadings:

From (34), (35) and (81)

(85) $A_n = A_{n-1} + \delta + \frac{1}{2}\sum_{i=1}^{k}\left[\lambda_i^2\sigma_i^2 - \left(\lambda_i\sigma_i + B_{i,n-1}\sigma_i\right)^2\right]$

(86) $B_{i,n} = \left(1 + B_{i,n-1}\varphi_i\right)$

$$A_n = A_{n-1} + \xi + B_{n-1}{}^T (I - \Phi)\theta - \frac{1}{2}(\lambda + B_{n-1})^T \alpha(\lambda + B_{n-1})$$

$$B_n{}^T = \gamma^T + B_{n-1}{}^T \Phi - \frac{1}{2}(\lambda + B_{n-1})^T \beta^{T'}$$

$$-m_{t+1} = \delta + \sum_{i=1}^{k}\left(\frac{\lambda_i^2}{2}\sigma_i^2 + z_{it} + \lambda_i\sigma_i\varepsilon_{t+1}\right)$$

Short-term interest rate:

(87) $y_{1,t} = \delta + \sum_{i=1}^{k} z_{it}$

↓

Given the common normalisation

From (38)

$$y_{1,t} = \xi - \frac{1}{2}\lambda^T \alpha\lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t$$

$$p_{0t} = 0\,, \quad A_0 = B_{10} = B_{20} = 0\,,$$

- This model has the appealing feature of the short-term being the sum of 2 factors plus a constant.
- The usual conjecture is that one factor is related to inflation expectations and that the other factor reflects the *ex-ante* real interest rate.

One-period forward curve:

(88)
$$f_{n,t} = \delta + \frac{1}{2}\sum_{i=1}^{2}\left[\lambda_i^2\sigma_i^2 - \left(\lambda_i\sigma_i + \frac{1-\varphi_i^n}{1-\varphi_i}\sigma_i\right)^2\right] + \sum_{i=1}^{2}\left[\varphi_i^n z_{it}\right]$$

From (44)

$$f_{n,t} = \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left(A_n + B_n{}^T z_t\right) = \left(A_{n+1} - A_n\right) + \left(B_{n+1}{}^T - B_n{}^T\right)z_t =$$
$$= \left[\xi + B_n{}^T(I-\Phi)\theta - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2\alpha_i\right] + \left[\gamma^T + B_n{}^T(\Phi - I) - \frac{1}{2}\sum_{i=1}^{k}\left(\lambda_i + B_{i,n}\right)^2\beta_i^T\right]z_t$$

Volatility curve:

(89)
$$Var_t\left(y_{n,t+1}\right) = \frac{1}{n^2}\sum_{i=1}^{k}\left(B_{i,n}^2\sigma_i^2\right)$$

From (42)
$$Var_t\left(y_{n,t+1}\right) = \frac{1}{n^2}B_n{}^T V\left(z_t\right)B_n$$

as the factors have constant volatility, given by $Var_t\left(z_{i,t+1}\right) = \sigma_i^2$, the volatility of yields depends neither on the level of the factors, nor on the level of the short-rate.

Term premium:

(90)

$$\Lambda_{n,t} = E_t p_{n,t+1} - p_{n+1,t} - y_{1,t} = \frac{1}{2}\sum_{i=1}^{k}\left[\lambda_i^2\sigma_i^2 - \left(\lambda_i\sigma_i + \frac{1-\varphi_i^n}{1-\varphi_i}\sigma_i\right)^2\right]$$

$$= \sum_{i=1}^{k}\left[-\lambda_i\sigma_i^{\ 2}B_{i,n} - \frac{B_{i,n}^{\ 2}\sigma_i^{\ 2}}{2}\right]$$

From (46)

$$\Lambda_{n,t} = -\sum_{i=1}^{k}\lambda_i B_{i,n}\alpha_i - \frac{B_{i,n}^2\alpha_i}{2} - \left(\sum_{i=1}^{k}\lambda_i B_{i,n}\beta_i - \frac{B_{i,n}^2\beta_i}{2}\right)z_t$$

- If the factors that determine the dynamics of the yield curve are assumed to be non-observable and the parameters are unknown, a usual estimation methodology is the Kalman filter and a maximum likelihood procedure.

- <u>Kalman Filter</u> - algorithm that computes the optimal estimate for the state variables at *t* using the information available up to *t*-1.

- <u>Maximum likelihood procedure</u> – provides the estimates for the parameters.

- The starting point for the derivation of the Kalman filter is to write the model in state-space form:

- **observation or measurement equation**

(91) $\underset{(r\times 1)}{Y_t} = \underset{(r\times n)}{A} \cdot \underset{(n\times 1)}{X_t} + \underset{(r\times k)}{H} \cdot \underset{(k\times 1)}{Z_t} + \underset{(r\times 1)}{w_t}$

$$\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} \\ \vdots & \vdots \\ b_{1,l} & b_{2,l} \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} w_{1,t} \\ \vdots \\ w_{l,t} \end{bmatrix}$$

where $y_{1,t}$, … , $y_{l,t}$ are the l zero-coupon yields at time t with maturities j = 1, … , u periods and $w_{1,t}$, … ,$w_{l,t}$ are the normally distributed i.i.d. errors, with null mean and standard-deviation equal to e_j^2, of the measurement equation for each interest rate considered, $a_j = A_j / j$, $b_{1,j} = B_{1,j} / j$, $b_{2,j} = B_{2,j} / j$.

- state or transition equation

(92)
$$\underset{(k\times1)}{Z_t} = \underset{(k\times1)}{C} + \underset{(k\times k)}{F} \cdot \underset{(k\times1)}{Z_{t-1}} + G \underset{(k\times1)}{v_t}$$

$$\begin{bmatrix} z_{1,t+1} \\ z_{2,t+1} \end{bmatrix} = \begin{bmatrix} \varphi_1 & 0 \\ 0 & \varphi_2 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \begin{bmatrix} v_{1,t+1} \\ v_{2,t+1} \end{bmatrix}$$

r – No. variables (interest rates) to estimate
n – No. observable exogenous variables (with no observable factors, *n*=1 => A becomes a column vector with the independent terms for each interest rate)
k – No. non-observable or latent exogenous variables (the factors).
ω_t and υ_t **-** i.i.d. residuals, distributed as $w_t \sim N(0,R)$ and $v_t \sim N(0,Q)$
Variance matrices:

$$\underset{(r\times r)}{R} = E(w_t w_t')$$

$$\underset{(k\times k)}{Q} = E(v_{t+1} v_{t+1}')$$

- One may estimate simultaneously the yields and the volatilities, to avoid implausible estimates for the latter:

(93)
$$\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ Var_t(y_{1,t+1}) \\ \vdots \\ Var_t(y_{l,t+1}) \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} \\ \vdots & \vdots \\ b_{1,l} & b_{2,l} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \end{bmatrix}$$

where $a_{l+j,t} = \frac{1}{n^2}\left(B_{1,j}^2\sigma_1^2 + B_{2,j}^2\sigma_2^2\right)$ and $2l$ is the number of variables to estimate

In our model A is a column vector with elements $a_{j,t}$ for the first l rows and $\frac{1}{n^2}\left(B_{1,j}^2\sigma_1^2 + B_{2,j}^2\sigma_2^2\right)$ for the next l rows; X_t is a $2l$-dimension column vector of one's $(n=1)$, C is a column vector of zeros and F is a $k \times k$ diagonal matrix, with typical element $F_{ii} = \varphi_i$ $(k=2)$.

- Contrary to the pioneer interest rate models, such as Vasicek (1977) and Cox *et al*. (1985a), where the short-term interest influenced the whole term structure, the latent factor models do not use explicit determinants of the yield curve.
- As previously referred, one common conjecture is to assume that one factor is related to the *ex-ante* real interest rate and a 2nd factor linked to inflation expectations.
- Therefore, one may start by estimating the factors and at a second stage try to identify how does one of the factors relate to inflation.
- Alternatively, one may specifically relate inflation to the second factor in the model to be estimated, in line with Fung et al. (1999).

Fung, Ben Siu Cheong, Scott Mitnick and Eli Remolona (1999), "Uncovering Inflation Expectations and Risk Premiums from Internationally Integrated Financial Markets", *Bank of Canada Working Paper Series*, No. 99-6.

- Assuming that inflation (π) is an AR(1) process, being $\overline{\pi}$ its mean, $\tilde{\pi}_t$ the deviation of inflation from its mean and ρ a parameter that measures the rate of mean-reversion:

(94) $$(\pi_{t+1} - \overline{\pi}) = \rho(\pi_t - \overline{\pi}) + u_{t+1}$$

- If the short-term interest is the sum of the factors and one of the factors is related to inflation, we may write:

(95) $$z_{2,t} = E_t(\pi_{t+1} - \overline{\pi}) = \rho(\pi_t - \overline{\pi}) = \rho\tilde{\pi}_t$$

- From the 2 previous equations:

(96) $$z_{2,t+1} = E_{t+1}(\tilde{\pi}_{t+2}) = \rho\tilde{\pi}_{t+1} = \rho(\rho\tilde{\pi}_t + u_{t+1}) = \rho z_{2,t} + \rho u_{t+1}$$

- From (82) $z_{i,t+1} = \varphi_i z_{it} + \sigma_i \varepsilon_{i,t+1}$ and (95)

$$\rho = \varphi_2$$

$$\rho u_{t+1} = \sigma_2 \varepsilon_{2,t+1}$$

$$\tilde{\pi}_t = \frac{1}{\varphi_2} z_{2,t}$$

- If the link between inflation and the second factor is considered, the observation equation becomes:

$$(97)\quad \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ Var_t(y_{1,t+1}) \\ \vdots \\ Var_t(y_{l,t+1}) \\ \tilde{\pi}_t \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \\ 0 \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} \\ \vdots & \vdots \\ b_{1,l} & b_{2,l} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & b_\pi \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \\ \upsilon_\pi \end{bmatrix}$$

- **Major drawback:** it implies the 2nd factor to explain simultaneously the inflation and the long term rates, which in some periods may evidence significantly different volatilities.

⬇

(i) In periods of higher volatility of the long-term rates, the estimated inflation tends to present a more irregular behaviour than the true inflation.

(ii) The AR(1) process for inflation is not necessarily the optimal model for forecasting inflation, being too simple concerning its lag structure and not allowing for the inclusion of other macro-economic information that market participants may use to form their expectations of inflation (e.g. monetary aggregates, commodity prices, exchange rates, wages and unit labour costs).

- However, a more complex model would certainly hamper the identification of the factor.
- One way to overcome these problems is by using a joint model for the term structure and the inflation, where the latter still shares a common factor with the interest rates but is also determined by a second specific factor:

(98) $$\pi_t = \frac{1}{n}\left(A_\pi + B_\pi{}^T z_{\pi t}\right)$$

where $z_{\pi t} = \begin{bmatrix} z_{2t} \\ z_{1\pi,t} \end{bmatrix}$ and $z_{1\pi,t+1} = \varphi_{1\pi} z_{1\pi,t} + \sigma_{1\pi}\varepsilon_{1\pi,t+1}$

- In this case, the observation and the state equations become:

(99)
$$\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ Var_t\left(y_{1,t+1}\right) \\ \vdots \\ Var_t\left(y_{l,t+1}\right) \\ \tilde{\pi}_t \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \\ 0 \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} & 0 \\ \vdots & \vdots & \vdots \\ b_{1,l} & b_{2,l} & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & b_{2\pi} & b_{1\pi} \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{1\pi,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \\ \upsilon_{\pi} \end{bmatrix}$$

(100)
$$\begin{bmatrix} z_{1,t+1} \\ z_{2,t+1} \\ z_{1\pi,t+1} \end{bmatrix} = \begin{bmatrix} \varphi_1 & 0 & 0 \\ 0 & \varphi_2 & 0 \\ 0 & 0 & \varphi_{1\pi} \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{1\pi,t} \end{bmatrix} + \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_{1\pi} \end{bmatrix} \begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \\ \varepsilon_{1\pi;t+1} \end{bmatrix}$$

- One may also use the DK framework to model simultaneously the term structures of interest rates of 2 countries.
- A first attempt to model jointly the term structures of 2 countries is found in Fung *et al*. (1999), where a 2-factor stochastic volatility model is used to estimate simultaneously the U.S. and the Canadian term structures.
- In this case, it was assumed that both countries share a common factor related to the real interest rate, following the close trade relationship between those countries.
- As each country pursued its own monetary policy, it was assumed that the U.S. and the Canadian term structures also depended on a specific factor, related to the inflation expectations and, accordingly, to the monetary policy.

- In the Euro area, the opposite happens, i.e., there is a common monetary policy and real interest rates differ among the member countries.

- One can model the joint term structures of 2 Euro Area countries assuming a common factor related to the inflation expectations and a specific factor that is supposed to be related to the real interest rate, modelling the 1st term structure as previously stated and the 2nd as:

(81) $$-m_{t+1} = \delta + \sum_{i=1}^{k}\left(\frac{\lambda_i^2}{2}\sigma_i^2 + z_{it} + \lambda_i\sigma_i\varepsilon_{t+1}\right)$$

(101) $$-m^*_{t+1} = \delta^* + \frac{\lambda_1^{*2}}{2}\sigma_i^{*2} + z^*_{it} + \lambda_1^*\sigma_1^*\varepsilon^*_{t+1} + \frac{\lambda_2^2}{2}\sigma_2^2 + z_{2t} + \lambda_2\sigma_2\varepsilon_{t+1}$$

- Remaining equations:

(102) $-p_{n,t}^{*} = A_n^{*} + B_{1,n}^{*} z_{1t}^{*} + B_{2,n} z_{2t}$

(103) $z_{1,t+1}^{*} = \varphi_1^{*} z_{1t}^{*} + \sigma_1^{*} \varepsilon_{1,t+1}^{*}$

(104) $y_{n,t}^{*} = \frac{1}{n}\left(A_n^{*} + B_{1,n}^{*} z_{1t}^{*} + B_{2,n} z_{2t}\right)$

- 2-country model with (common) inflation:

(105)
$$\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ y_{1,t}^{*} \\ \vdots \\ y_{l,t}^{*} \\ Var_t\left(y_{1,t+1}\right) \\ \vdots \\ Var_t\left(y_{l,t+1}\right) \\ Var_t\left(y_{1,t+1}^{*}\right) \\ \vdots \\ Var_t\left(y_{l,t+1}^{*}\right) \\ \tilde{\pi}_t \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{1,t}^{*} \\ \vdots \\ a_{l,t}^{*} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \\ a_{l+1,t}^{*} \\ \vdots \\ a_{2l}^{*} \\ 0 \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} & 0 \\ \vdots & \vdots & \vdots \\ b_{1,l} & b_{2,l} & 0 \\ 0 & b_{2,1} & b_{1,1}^{*} \\ \vdots & \vdots & \vdots \\ 0 & b_{2,l} & b_{1,l}^{*} \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & b_{\pi} & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z_{1,t}^{*} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{1,t}^{*} \\ \vdots \\ \upsilon_{l,t}^{*} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \\ \upsilon_{l+1,t}^{*} \\ \vdots \\ \upsilon_{2l,t}^{*} \\ \upsilon_{\pi} \end{bmatrix}$$

- The estimation departs from assuming that the starting value of the state vector Z is obtained from a normal distribution with mean $\bar{Z}_0$ and variance P_0 (usually it is assumed that the starting values of the factors are zero).

- $\hat{Z}_0$ can be seen as a guess concerning the value of Z using all information available up to and including $t = 0$.

- Using $\bar{Z}_0$ and P_0 and following (92) $\underset{(k\times1)}{Z_t} = \underset{(k\times1)}{C} + \underset{(k\times k)}{F} \cdot \underset{(k\times1)}{Z_{t-1}} + G \underset{(k\times1)}{v_t}$, the optimal estimator for Z_1 will be given by:

(106) $\hat{Z}_{1|0} = C + F\hat{Z}_0$

- Consequently, the variance matrix of the estimation error of the state vector will correspond to:

(107)

$$\begin{aligned} P_{1|0} &= E\left[\left(Z_1 - \hat{Z}_{1|0}\right)\left(S_1 - \hat{Z}_{1|0}\right)'\right] \\ &= E\left[(C + FZ_0 + Gv_1 - C - FZ_0)(C + FZ_0 + Gv_1 - C - FZ_0)'\right] \\ &= E\left[(Fv_0 + Gv_1)(v_0'F' + v_1'G')\right] \\ &= E(Fv_0v_0'F') + E(Gv_1v_1'G') \\ &= FP_{1|0}F' + GQ_1G' \end{aligned}$$

- Given that $vec(ABC) = (C'\otimes A)\cdot vec(B)$, $P_{1|0}$ may be obtained from:

(108)

$$\begin{aligned} vec\left(P_{1|0}\right) &= vec\left(FP_{1|0}F'\right) + vec(GQ_1G') \\ &= (F\otimes F)\cdot vec\left(P_{1|0}\right) + (G\otimes G)\cdot vec(Q_1) \\ &= \left[\underset{(n^2\times n^2)}{I} - (F\otimes F)\right]^{-1}\left[(G\otimes G)\cdot vec(Q_1)\right] \end{aligned}$$

- As w_t is independent from X_t and from all the prior information on y and x (denoted by ζ_{t-1}), we can obtain the <u>forecast of y_t conditional</u> on X_t and ζ_{t-1} directly from

(109) $$\underset{(r\times1)}{Y_t} = \underset{(r\times n)}{A} \cdot \underset{(n\times1)}{X_t} + \underset{(r\times k)}{H} \cdot \underset{(k\times1)}{Z_t} + \underset{(r\times1)}{w_t} \qquad \text{with} \quad w_t \sim N(0,R)$$

(110) $$E(y_t | X_t, \zeta_{t-1}) = AX_t + H\hat{Z}_{t|t-1} \qquad \text{with} \quad v_t \sim N(0,Q)$$

- Therefore, from (109) and (110), we have the following expression for the <u>forecasting error</u>:

(111) $$Y_t - E(Y_t|X_t, \zeta_{t-1}) = (AX_t + HZ_t + w_t) - \left(AX_t + H\hat{Z}_{t|t-1}\right) = H\left(Z_t - \hat{Z}_{t|t-1}\right) + w_t$$

- From (110), the conditional variance-covariance matrix of the estimation error of the observation vector will be:

(112)
$$E\{[Y_t - E(Y_t|X_t,\zeta_{t-1})][Y_t - E(Y_t|X_t,\zeta_{t-1})]'\} = E\{[H(Z_t - \hat{Z}_{t|t-1}) + w_t][H(Z_t - \hat{Z}_{t|t-1}) + w_t]'\}$$
$$= HE[(Z_t - \hat{Z}_{t|t-1})(Z_t - \hat{Z}_{t|t-1})']H' + E(w_t w_t')$$
$$= HP_{t|t-1}H' + R$$

- After the updates of the mean and variance-covariance matrices of the dependent variables, the log-likelihood function is computed to estimate the parameters:

(113)
$$\log L(Y_T) = \sum_{t=1}^{T} \log f(Y_t|I_{t-1})$$

$$f(Y_t|I_{t-1}) = (2\pi)^{-1/2}\left|HP_{t|t-1}H' + R\right|^{-1/2} \cdot$$
$$\exp\left[-\frac{1}{2}(Y_t - A - H\hat{Z}_{t|t-1})'(HP_{t|t-1}H' + R)^{-1}(y_t - A - H\hat{Z}_{t|t-1})\right]$$

- The maximization of the log-likelihood function is often performed as the minimization of the symmetric of that function.
- In order to characterise the distribution of the observation and state vectors, it is also required to compute the conditional covariance between both forecasting errors.
- From (81) we get:

(114) $$E\{[Y_t - E(Y_t|X_t, \zeta_{t-1})][Z_t - E(Z_t|X_t, \zeta_{t-1})]'\} = E\{[H(Z_t - \hat{Z}_{t|t-1}) + w_t][Z_t - \hat{Z}_{t|t-1}]'\}$$
$$= HE[(Z_t - \hat{Z}_{t|t-1})(Z_t - \hat{Z}_{t|t-1})']$$
$$= HP_{t|t-1}$$

- Therefore, using (109), (111) and (113), the conditional distribution of the vector (Y_t, Z_t) is:

(115) $$\begin{bmatrix} Y_t|X_t, \zeta_{t-1} \\ Z_t|X_t, \zeta_{t-1} \end{bmatrix} \sim N\left(\begin{bmatrix} AX_t + H_{t|t-1} \\ \hat{Z}_{t|t-1} \end{bmatrix}, \begin{bmatrix} HP_{t|t-1}H' + R & HP_{t|t-1} \\ P_{t|t-1}H' & P_{t|t-1} \end{bmatrix}\right)$$

- Consequently, following (114), the distribution of Z_t given Y_t, X_t and ζ_{t-1} is $N\left(\hat{Z}_{t|t}, P_{t|t}\right)$, where $\hat{Z}_{t|t}$ and $P_{t|t}$ are respectively the optimal forecast of Z_t given $P_{t|t}$ and the mean square error of this forecast, corresponding to the following updating equations of the Kalman Filter:

(116) $\hat{Z}_{t|t} = \hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]$

(117) $P_{t|t} = P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}HP_{t|t-1}$

- Following this update, new estimates can be obtained, generalizing (106) $\hat{Z}_{1|0} = C + F\hat{Z}_0$

(118) $\hat{Z}_{t+1|t} = C + F\hat{Z}_{t|t} = C + F\left\{\hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]\right\}$

$= C + F\hat{Z}_{t|t-1} + FP_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]$

(119) $$P_{t+1|t} = FP_{t|t}F' + GQ_tG'$$
$$= F\left[P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}HP_{t|t-1}\right]F' + GQ_tG'$$
$$= FP_{t|t-1}F' - FP_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}HP_{t|t-1}F' + GQ_tG'$$

→ Ricatti equation

- The matrix $FP_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}$ is usually known as the gain matrix, since it determines the update in $\hat{Z}_{t+1|t}$ due to the estimation error of Y_t.

- Concluding, the Kalman Filter may be applied after specifying starting values for $\hat{Z}_{1|0}$ and $P_{1|0}$ using equations (110), (112), (116), and (117) and iterating on equations (118) and (119).

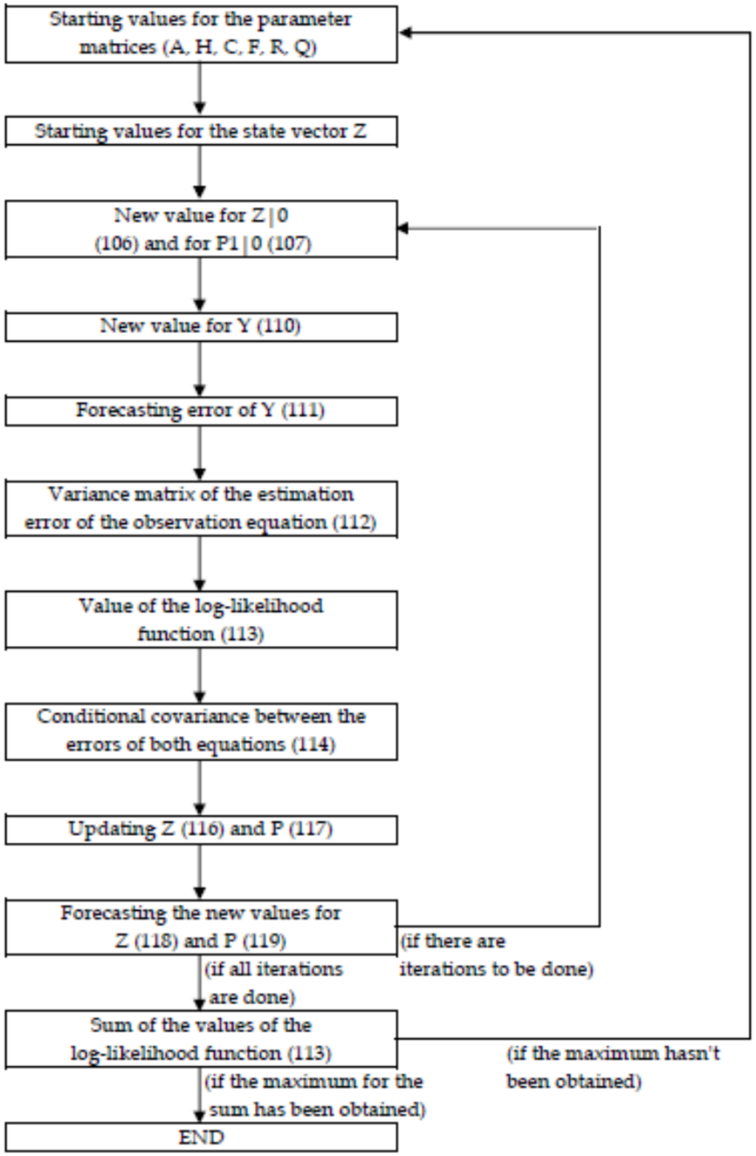
(110) $$E(y_t|X_t,\zeta_{t-1}) = AX_t + H\hat{Z}_{t|t-1}$$

(112) $$E\{[Y_t - E(Y_t|X_t,\zeta_{t-1})][Y_t - E(Y_t|X_t,\zeta_{t-1})]'\} = E\{[H(Z_t - \hat{Z}_{t|t-1}) + w_t][H(Z_t - \hat{Z}_{t|t-1}) + w_t]'\}$$
$$= HE\left[(Z_t - \hat{Z}_{t|t-1})(Z_t - \hat{Z}_{t|t-1})'\right]H' + E(w_t w_t')$$
$$= HP_{t|t-1}H' + R$$

(116) $$\hat{Z}_{t|t} = \hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]$$

(117) $$P_{t|t} = P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}HP_{t|t-1}$$

(118) $$\hat{Z}_{t+1|t} = C + F\hat{Z}_{t|t} = C + F\left\{\hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]\right\}$$
$$= C + F\hat{Z}_{t|t-1} + FP_{t|t-1}H'\left(HP_{t|t-1}H'+R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]$$



(106) $$\hat{Z}_{1|0} = C + F\hat{Z}_0$$

(107) $$\begin{aligned} P_{1|0} &= E\left[\left(Z_1 - \hat{Z}_{1|0}\right)\left(S_1 - \hat{Z}_{1|0}\right)'\right] \\ &= E\left[(C + FZ_0 + Gv_1 - C - FZ_0)(C + FZ_0 + Gv_1 - C - FZ_0)'\right] \\ &= E\left[(Fv_0 + Gv_1)(v_0'F' + v_1'G')\right] \\ &= E(Fv_0v_0'F') + E(Gv_1v_1'G') \\ &= FP_{1|0}F' + GQ_1G' \end{aligned}$$

(110) $$E(y_t | X_t, \zeta_{t-1}) = AX_t + H\hat{Z}_{t|t-1}$$

(111) $$Y_t - E(Y_t | X_t, \zeta_{t-1}) = (AX_t + HZ_t + w_t) - \left(AX_t + H\hat{Z}_{t|t-1}\right) = H\left(Z_t - \hat{Z}_{t|t-1}\right) + w_t$$

(112) $$\begin{aligned} E\{[Y_t - E(Y_t | X_t, \zeta_{t-1})][Y_t - E(Y_t | X_t, \zeta_{t-1})]'\} &= E\left\{\left[H\left(Z_t - \hat{Z}_{t|t-1}\right) + w_t\right]\left[H\left(Z_t - \hat{Z}_{t|t-1}\right) + w_t\right]'\right\} \\ &= HE\left[\left(Z_t - \hat{Z}_{t|t-1}\right)\left(Z_t - \hat{Z}_{t|t-1}\right)'\right]H' + E(w_t w_t') \\ &= HP_{t|t-1}H' + R \end{aligned}$$

(113) $$\log L(Y_T) = \sum_{t=1}^{T} \log f(Y_t | I_{t-1})$$

(114) $$\begin{aligned} E\{[Y_t - E(Y_t | X_t, \zeta_{t-1})][Z_t - E(Z_t | X_t, \zeta_{t-1})]'\} &= E\left\{\left[H\left(Z_t - \hat{Z}_{t|t-1}\right) + w_t\right]\left[Z_t - \hat{Z}_{t|t-1}\right]'\right\} \\ &= HE\left[\left(Z_t - \hat{Z}_{t|t-1}\right)\left(Z_t - \hat{Z}_{t|t-1}\right)'\right] \\ &= HP_{t|t-1} \end{aligned}$$

(116) $$\hat{Z}_{t|t} = \hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]$$

(117) $$P_{t|t} = P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}HP_{t|t-1}$$

(118) $$\begin{aligned} \hat{Z}_{t+1|t} &= C + F\hat{Z}_{t|t} = C + F\left\{\hat{Z}_{t|t-1} + P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right]\right\} \\ &= C + F\hat{Z}_{t|t-1} + FP_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}\left[Y_t - \left(AX_t + H_{t|t-1}\right)\right] \end{aligned}$$

(119) $$\begin{aligned} P_{t+1|t} &= FP_{t|t}F' + GQ_tG' \\ &= F\left[P_{t|t-1} - P_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}HP_{t|t-1}\right]F' + GQ_tG' \\ &= FP_{t|t-1}F' - FP_{t|t-1}H'\left(HP_{t|t-1}H' + R\right)^{-1}HP_{t|t-1}F' + GQ_tG' \end{aligned}$$

2.4. HJM

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

Goal: Model the dynamics of the entire yield curve, assuming there is just one factor in a risk-neutral world.

A zero-coupon bond return will be the risk-free rate

$$dP(t, T) = r(t)P(t, T)\,dt + v(t, T, \Omega_t)P(t, T)\,dz(t)$$

$P(t, T)$: Price at time t of a risk-free zero-coupon bond with principal \$1 maturing at time T

Ω_t: Vector of past and present values of interest rates and bond prices at time t that are relevant for determining bond price volatilities at that time

$v(t, T, \Omega_t)$: Volatility of $P(t, T)$

$f(t, T_1, T_2)$: Forward rate as seen at time t for the period between time T_1 and time T_2

$F(t, T)$: Instantaneous forward rate as seen at time t for a contract maturing at time T

$r(t)$: Short-term risk-free interest rate at time t

$dz(t)$: Wiener process driving term structure movements.

Stochastic process:

(1) $$dP(t,T) = r(t)P(t,T)\,dt + v(t,T,\Omega_t)P(t,T)\,dz(t)$$

Forward rate:

(2) $$f(t,T_1,T_2) = \frac{\ln[P(t,T_1)] - \ln[P(t,T_2)]}{T_2 - T_1}$$

From (1) and Ito's Lemma:

$$d\ln[P(t,T_1)] = \left[r(t) - \frac{v(t,T_1,\Omega_t)^2}{2}\right]dt + v(t,T_1,\Omega_t)\,dz(t)$$

$$d\ln[P(t,T_1)] = \left[r(t) - \frac{v(t,T_1,\Omega_t)}{2}\right]dt + v(t,T_1,\Omega_t)\,dz(t)$$

The risk-neutral process for the forward rate depends solely on the bond price volatility

(3) $$df(t,T_1,T_2) = \frac{v(t,T_2,\Omega_t)^2 - v(t,T_1,\Omega_t)^2}{2(T_2 - T_1)}dt + \frac{v(t,T_1,\Omega_t) - v(t,T_2,\Omega_t)}{T_2 - T_1}dz(t)$$

It is possible to show that:

(4) $$dF(t,T) = v(t,T,\Omega_t)\, v_T(t,T,\Omega_t)\, dt - v_T(t,T,\Omega_t)\, dz(t)$$

⬇

There is a link between the drift and the standard-deviation of the instantaneous forward rate (*F(t,T)*).

Key problem: risk-free interest rate is non-Markov ⇔ the risk-free interest rate process depends on its previous path.

HJM can be extended to several factors:

(5) $$dF(t,T) = m(t,T,\Omega_t)\, dt + \sum_k s_k(t,T,\Omega_t)\, dz_k$$