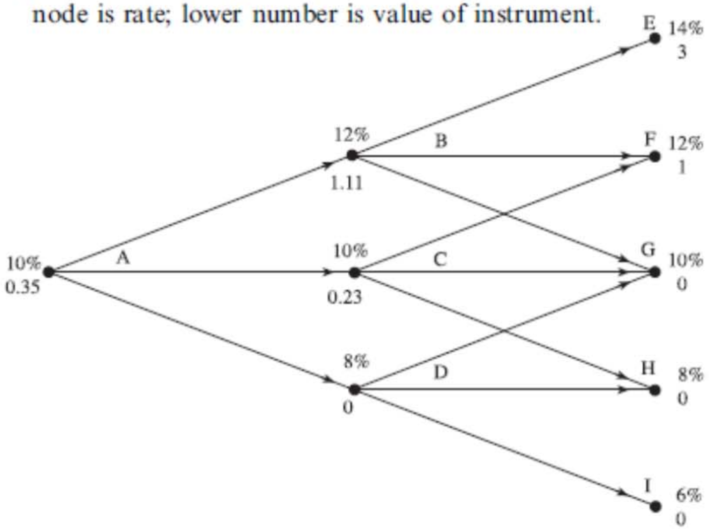
2.2. SHORT RATE MODELS

2.2.1. Interest Rate Trees

- Focus: How to model the term structure by specifying the behavior of the short-term interest rate?
- Why do we use trees? - A tree is a discrete-time representation of the stochastic process.
- Most trees are binomial, event though trinomial trees are also used, namely to value interest rate derivatives.

Figure 31.6 Example of the use of trinomial interest rate trees. Upper number at each node is rate; lower number is value of instrument.



The tree is used to value a derivative that provides a payoff at the end of the second time step of

$$\max[100(R - 0.11), 0]$$

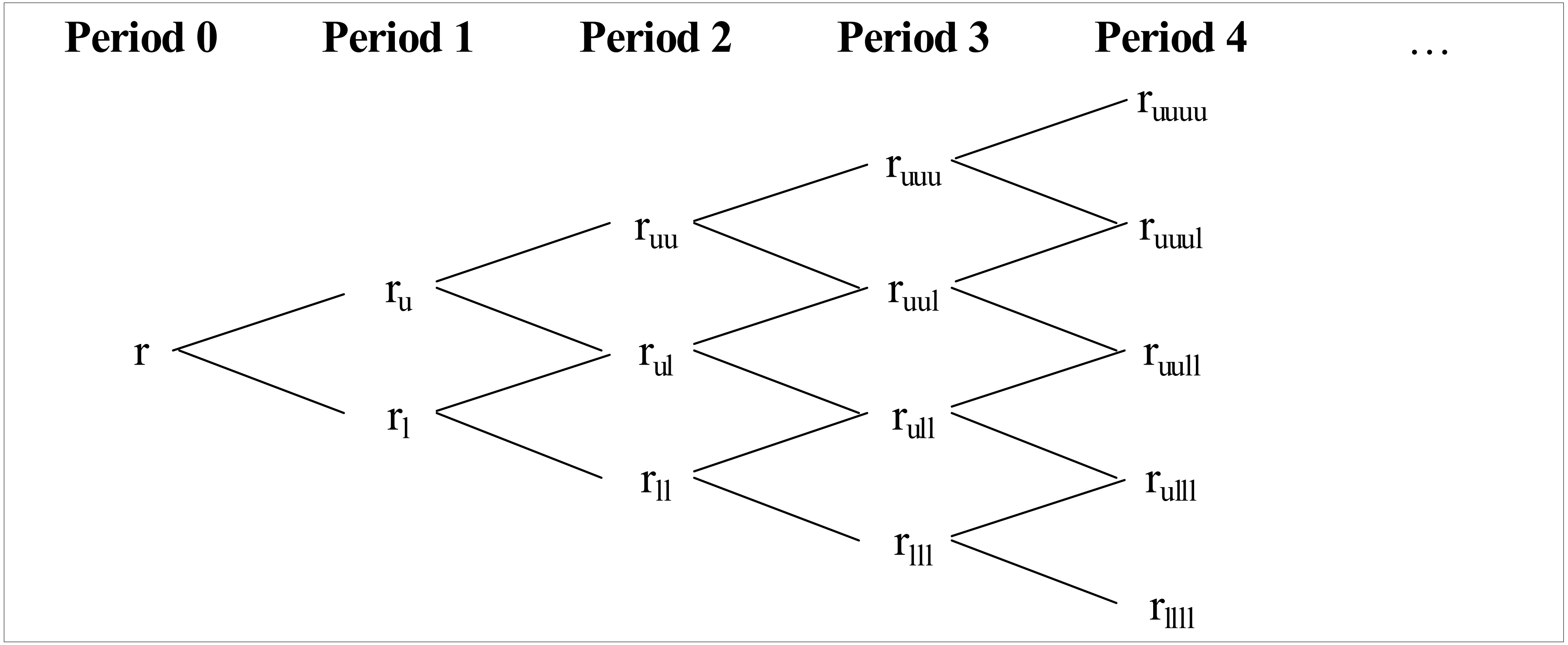
Source: Hull (2017)

2.2.1 – Interest Rate Trees

- General binomial model
  - Given the current level of short-term rate r, the next-period short rate can take only two possible values: an upper value r_u and a lower value r_l, with equal probability 0.5
  - In period 2, the short-term interest rate can take on four possible values: r_{uu}, r_{ul}, r_{lu}, r_{ll}
  - More generally, in period *n*, the short-term interest rate can take on 2^n values => very time-consuming and computationally inefficient

- Recombining trees
  - Means that an upward-downward sequence leads to the same result as a downward-upward sequence (regardless being binomial or trinomial trees)
  - For example, $r_{ul} = r_{lu}$
  - Only (n+1) different values at period n

INTEREST RATE TREE - Recombining



INTEREST RATE TREE – analytical

- We may write down the binomial process as

$$\Delta r_t \equiv r_{t+1} - r_t = \sigma\varepsilon_t$$

  where ε_t are independent variables taking on values (+1,-1) with probability (1/2,1/2)

- Problem: rates can take on negative values with positive probability

- Fix that problem by working with logs

$$\Delta \ln r_t \equiv \ln r_{t+1} - \ln r_t = \sigma\varepsilon_t$$

$$\Rightarrow r_{t+1} = r_t \times \exp(\sigma\varepsilon_t) = r_t \times \begin{pmatrix} u = \exp(\sigma) \\ d = \exp(-\sigma) \end{pmatrix}$$

  with probability (1/2,1/2)

- More general models (that can also be written on log rates, for any time increment and any mean):

$$\Delta r_t \equiv r_{t+1} - r_t = \sigma \varepsilon_t$$

$$\Downarrow$$

$$\Delta r_t \equiv r_{t+\Delta t} - r_t = \mu\left(t, \Delta t, r_t\right) + \sigma\left(t, \Delta t, r_t\right)\varepsilon_t$$

- Specific case – assuming that the drift and the variance are proportional to the time increment:

$$\Delta r_t \equiv r_{t+\Delta t} - r_t = \mu \Delta t + \sigma \sqrt{\Delta t}\, \varepsilon_t$$

- Continuous-time limit (Merton (1973)):

$$dr_t \equiv r_{t+dt} - r_t = \mu dt + \sigma dW_t$$

INTEREST RATE TREE – calibration

- Calibration of the model is performed so as to make model consistent with the current term structure.
- We have at date 0 (working in logs):

$$\Delta \ln r_0 \equiv \ln r_{\Delta t} - \ln r_0 = \mu\Delta t + \sigma\varepsilon_0\sqrt{\Delta t}$$

(From $\Delta r_t \equiv r_{t+\Delta t} - r_t = \mu\Delta t + \sigma\sqrt{\Delta t}\varepsilon_t$)

- As the uncertainty source is the path of the interest rate (up or down), the difference between interest rates in $t+\Delta t$ will be originated by the random factor (the deterministic factor will be the same if the interest rate increases or decreases):

$$\Rightarrow \ln r_u - \ln r_l = 2\sigma\sqrt{\Delta t} \quad \text{or} \quad r_u = r_l \exp\left(2\sigma\sqrt{\Delta t}\right)$$

  - If we take as given an estimate for σ and the current yield curve y_t, we iteratively find the values r_u, r_l, r_{uu}, r_{ul}, r_{lu}, r_{ll}, etc., consistent with the input data.

EXAMPLE

- Consider a 2 period tree (t=0 and t=1 => $\Delta t = 1$)
- The price 1 year from now of a 2-year Treasury bond (at the par value, i.e. coupon rate = yield) can take 2 values:

- P_u - associated with r_u (price with the interest rate increasing)
- P_l - associated with r_l (price with the interest rate decreasing):

Coupon rate, which is equal to the yield for a bond at par.

$$P_u = \frac{100 + y_2}{1 + r_u} \text{ and } P_d = \frac{100 + y_2}{1 + r_l}$$

r_u and r_l must be seen as the 2 possible future values in t=0 of the 1-period interest rate in t=1

NPV at *t*=1 of the future cash-flows of the bond - redemption and the last coupon (y_2, the 2-period yield at *t*=0, that corresponds to the coupon rate, as it is assumed that the bond is at par value), as in *t*=1 there is only one remaining period for the bond => the future cash-flows in *t*=1 are the redemption and the last coupon.

The uncertainty in *t*=0 about the bond price in *t*=1 stems from the uncertainty about the interest rate, which may have increased or decreased.

EXAMPLE

- Given that

$$\Rightarrow \ln r_u - \ln r_l = 2\sigma\sqrt{\Delta t} \quad \text{or} \quad r_u = r_l \exp\left(2\sigma\sqrt{\Delta t}\right)$$

- taking expectations at time 0, we find an equation that can be solved for r_u and r_l , replacing r_u by the previous expression (and being Δt = 1) and taking into consideration that
-

$$P_u = \frac{100 + y_2}{1 + r_u} \quad \text{and} \quad P_d = \frac{100 + y_2}{1 + r_l} \quad :$$

Bond price in t=0 is the expected value of the future cash-flows – the coupon in t=1 and the bond price also in t=1, which is the NPV at t=1 of the cash-flows to be paid in t=2.
The bond price in t=0 is also equal to 100, as it is assumed that the bond is at par.
The probability for each bond price in t=1, with r_u or r_l, is ½.

1st year coupon, as in t=0 we have 2 coupons ahead

$\Delta t = 1$

$$100 = \frac{1}{2}\left(\frac{\dfrac{100 + y_2}{1 + r_l \exp\left(2\sigma\right)} + y_2}{1 + y_1} + \frac{\dfrac{100 + y_2}{1 + r_l} + y_2}{1 + y_1} \right)$$

Discounted by y_1 as these are cash-flows that will occur in t=1

2.2.2 – CT Single Factor Models

- General expression for a single-factor continuous-time model (from the continuous time limit - Merton (1973))

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

- The term *W* denotes a Brownian motion - process with independent normally distributed increments: $dW_t = \varepsilon_t \sqrt{dt}$
  - *dW* represents the instantaneous change.
  - It is stochastic (uncertain)
  - It is a stochastic variable with a normal distribution with zero mean and variance *dt*

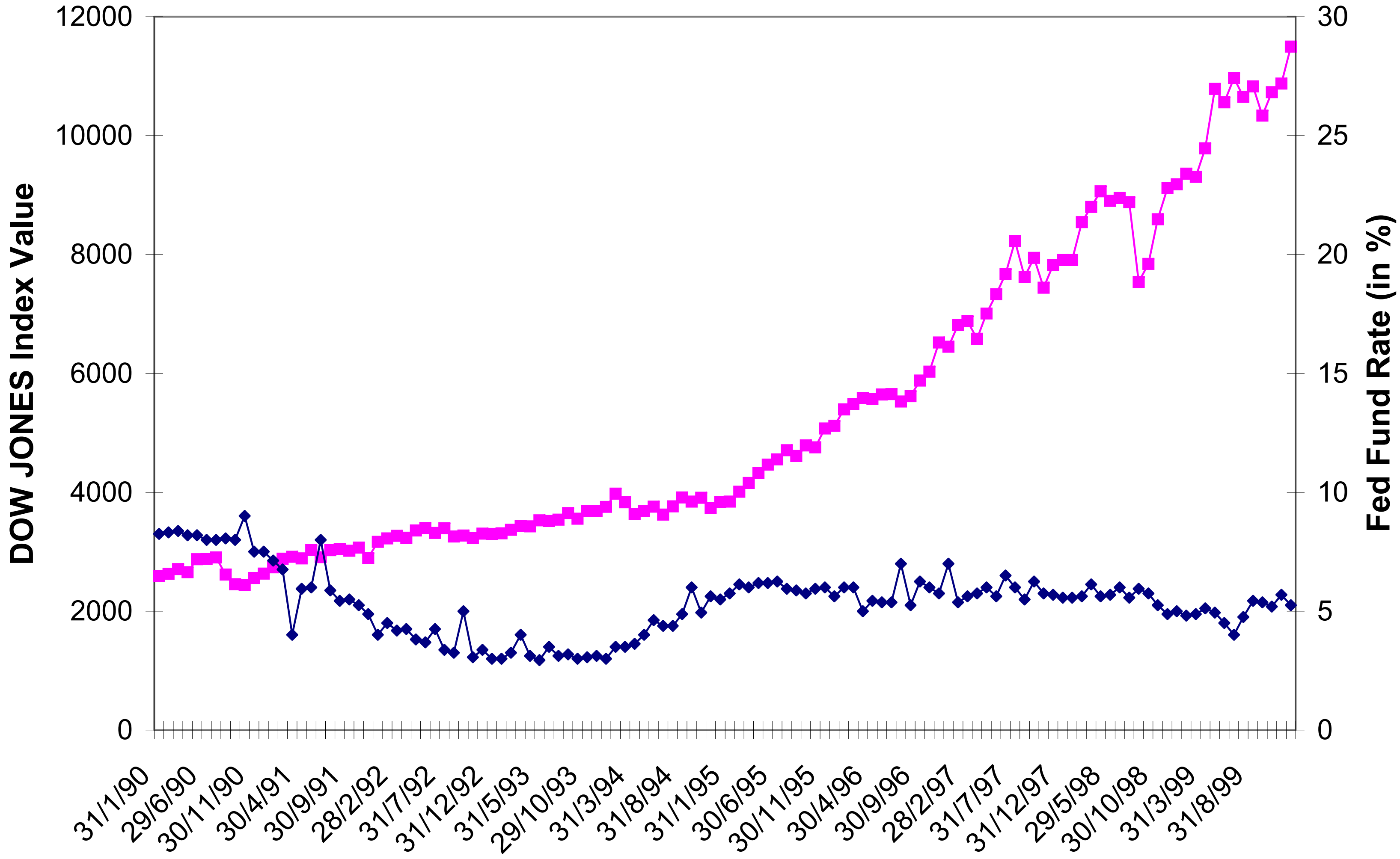
WHAT IS A GOOD MODEL?

- A good model is a model that is consistent with reality

- Stylized facts about the dynamics of the term structure
  - Fact 1: (nominal) interest rates are positive
  - Fact 2: interest rates are mean-reverting
  - Fact 3: interest rates with different maturities are imperfectly correlated
  - Fact 4: the volatility of interest rates evolves (randomly) in time

- A good model should also be
  - Tractable
  - Parsimonious

Empirical Facts 1, 2 and 4



Empirical Fact 3

	1M	3M	6M	1Y	2Y	3Y	4Y	5Y	6Y	7Y	8Y	9Y	10Y
1M	1												
3M	0.992	1											
6M	0.775	0.775	1										
1Y	0.354	0.3	0.637	1									
2Y	0.214	0.165	0.42	0.901	1								
3Y	0.278	0.246	0.484	0.79	0.946	1							
4Y	0.26	0.225	0.444	0.754	0.913	0.983	1						
5Y	0.224	0.179	0.381	0.737	0.879	0.935	0.981	1					
6Y	0.216	0.168	0.352	0.704	0.837	0.892	0.953	0.991	1				
7Y	0.228	0.182	0.35	0.661	0.792	0.859	0.924	0.969	0.991	1			
8Y	0.241	0.199	0.351	0.614	0.745	0.826	0.892	0.936	0.968	0.992	1		
9Y	0.238	0.198	0.339	0.58	0.712	0.798	0.866	0.913	0.95	0.981	0.996	1	
10Y	0.202	0.158	0.296	0.576	0.705	0.779	0.856	0.915	0.952	0.976	0.985	0.99	1

Daily changes in French swap markets in 1998

Most Popular Endogenous short rate CT Models

- Most important types of one-factor interest rate (x_t) models, being the short-term rate the single factor (i.e. endogenous models):

(from $dr_t = \mu(t,r_t)dt + \sigma(t,r_t)dW_t$)

1. **Vasicek (1977):**

$$dx_t = k(\theta - x_t)dt + \sigma dW_t, \quad \alpha = (k, \theta, \sigma).$$

Constant volatility model with mean-reversion to θ (as it is na Ornstein-Uhlenback process

2. **Cox-Ingersoll-Ross (CIR, 1985):**

$$dx_t = k(\theta - x_t)dt + \sigma\sqrt{x_t}dW_t, \quad \alpha = (k, \theta, \sigma), \quad 2k\theta > \sigma^2.$$

Stochastic volatility model with mean-reversion to θ

Most Popular Endogenous short rate CT Models

- Most important types of one-factor interest rate (x_t) models, being the short-term rate the single factor (i.e. endogenous models):

(from $dr_t = \mu(t, r_t)dt + \sigma(t, r_t)dW_t$)

- 1. Vasicek (1977):

Vasicek, O., 1977, "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

$$dr = \alpha(\gamma - r)dt + \rho dz$$

  - The spot rate follows an Ornstein-Uhlenbeck process => contrary to the random walk (Wiener process), which diverge to infinite values in the long-run, this process converges to the long-term mean γ.

VASICEK MODEL

- Therefore, the Vasicek model has some peculiarities that make it attractive:

- Gaussian disturbances;
- Constant volatility, making the model more tractable;
- Mean reverting – expected value of the short rate tends to a constant value γ with velocity given by α.

- Drawbacks:

- Rates can assume negative values with positive probability.
- Gaussian distributions for the rates are not compatible with the market implied distributions.

COX-INGERSOLL AND ROSS MODEL

- 2. CIR (1985):

Cox, Ingersoll, and Ross. 1985, "A Theory of the Term Structure of Interest Rates", *Econometrica*, Vol 53, March.

$$dr = \kappa(\theta - r)\,dt + \sigma\sqrt{r}\,dz_1$$

- The CIR model is also mean reverting, like Vasicek.
- However, the volatility is not constant, depending on the short-term rate.
- This stochastic volatility brings the model closer to reality.
- However, the model becomes less tractable as it requires the single factor to be positive.

POPULAR EXOGENOUS SHORT RATE CT MODELS

- Exogenous short-rate models are built by suitably modifying the above endogenous models.
- The basic strategy that is used to transform an endogenous model into an exogenous model is the inclusion of time-varying parameters.
- Therefore, interest rates become determined not only by the short-term rates but also by a time-varying drift:

1. **Ho-Lee:**

$$dx_t = \theta(t)\ dt + \sigma\ dW_t.$$

2. **Hull-White (Extended Vasicek):**

$$dx_t = k(\theta(t) - x_t)dt + \sigma dW_t.$$

3. **Hull-White (Extended CIR):**

$$dx_t = k(\theta(t) - x_t)dt + \sigma\sqrt{x_t}\ dW_t\ .$$

2.2.3 – CT MULTI FACTOR MODELS

MOST POPULAR MODELS

1. **Fong and Vasicek (1991) model -** short rate and its volatility (*v*) as two state variables

   H. G. Fong and O. A. Vasicek: Fixed-income volatility management. Journal of Portfolio Management, 41-56, 1991.

$$dr = \alpha(\bar{r} - r)dt + \sqrt{v}dz_1$$

$$dv = \gamma(\bar{v} - v)dt + \xi\sqrt{v}dz_2$$

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

Longstaff, F. A. and E. S. Schwartz, "Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model," *Journal of Finance*, 47, 4 (September 1992): 1259–82.

- Longstaff and Schwartz (1992) use the same two state variables (the short rate and its volatility) , but with a different specification.
- The starting point is a two-factor model, where the drift is governed by the two factors or state variables, while the variance is a function of only one of them:

$$\frac{dQ}{Q} = (\mu X + \theta Y)\, dt + \sigma \sqrt{Y}\, dZ_1$$

- With this specification, it is ensured that the drift and the variance are not perfectly correlated.

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- The dynamics of the state variables are as follows:

$$dX = (a - bX)\,dt + c\sqrt{X}\,dZ_2$$

$$dY = (d - eY)\,dt + f\sqrt{Y}\,dZ_3$$

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- With the rescaling of the state variables to $x = X/c^2$ and $y = Y/f^2$, the dynamics of state variables are as follows:

$$dr = \left(\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha} r - \frac{\xi - \delta}{\beta - \alpha} V\right) dt + \alpha\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}\, dZ_2 + \beta\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}\, dZ_3,$$

$$dV = \left(\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha} r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha} V\right) dt + \alpha^2\sqrt{\frac{\beta r - V}{\alpha(\beta - \alpha)}}\, dZ_2 + \beta^2\sqrt{\frac{V - \alpha r}{\beta(\beta - \alpha)}}\, dZ_3$$

where $\gamma = a/c^2$, $\delta = b$, $\eta = d/f^2$, $\xi = e$, r is the instantaneous riskless rate,

where $\alpha = \mu c^2$ and $\beta = (\theta - \sigma^2) f^2$

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- Relevant features:

(i) All parameters are positive;

(ii) *r* is non-negative, since both state variables follow square root processes;

(iii) *r* has a long-run stationary distribution with mean and variance:

$$E[r] = \frac{\alpha\gamma}{\delta} + \frac{\beta\eta}{\xi} \qquad Var[r] = \frac{\alpha^2\gamma}{2\delta^2} + \frac{\beta^2\eta}{2\xi^2}$$

(iv) Volatility also has a stationary distribution with mean

$$E[V] = \frac{\alpha^2\gamma}{\delta} + \frac{\beta^2\eta}{\xi} \qquad Var[V] = \frac{\alpha^4\gamma}{2\delta^2} + \frac{\beta^4\eta}{2\xi^2}$$

(v) *r* depends on volatility, but volatility also depends on *r;*

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- Closed-form expressions for riskless discount bond prices with τ maturity ($\tau = 0 \Rightarrow F = 1$)

$$F(r, V, \tau) = A^{2\gamma}(\tau) B^{2\eta}(\tau) \exp(\kappa\tau + C(\tau) r + D(\tau) V),$$

where

$$A(\tau) = \frac{2\phi}{(\delta + \phi)(\exp(\phi\tau) - 1) + 2\phi},$$

$$B(\tau) = \frac{2\psi}{(\nu + \psi)(\exp(\psi\tau) - 1) + 2\psi},$$

$$C(\tau) = \frac{\alpha\phi(\exp(\psi\tau) - 1) B(\tau) - \beta\psi(\exp(\phi\tau) - 1) A(\tau)}{\phi\psi(\beta - \alpha)}$$

$$D(\tau) = \frac{\psi(\exp(\phi\tau) - 1) A(\tau) - \phi(\exp(\psi\tau) - 1) B(\tau)}{\phi\psi(\beta - \alpha)},$$

and

$$\nu = \xi + \lambda,$$

$$\phi = \sqrt{2\alpha + \delta^2},$$

$$\psi = \sqrt{2\beta + \nu^2},$$

$$\kappa = \gamma(\delta + \phi) + \eta(\nu + \psi).$$

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- YTM of riskless discount bonds with τ maturity:

$$Y_\tau = -(\kappa\tau + 2\gamma \ln A(\tau) + 2\eta \ln B(\tau) + C(\tau)r + D(\tau)V)/\tau$$

↓

- For a given τ maturity, the yield is a linear function of *r* and *V*.

2.2.3 – CT MULTI FACTOR MODELS

MOST POPULAR MODELS

2. Longstaff and Schwartz (1992) model

- It can be shown that:

$\tau \rightarrow 0 \Rightarrow Y_t \rightarrow r$

$\tau \rightarrow \infty \Rightarrow Y_t$ tends to a constant $\gamma(\phi - \delta) + \eta(\psi - \nu)$

↓

- The current values of *r* and *V* become less relevant for very distant cash-flows

↓

- The current term structure is irrelevant for the determination of very long interest rates.

2.2.3 – CT Multi Factor Models

Most Popular Models

2. Longstaff and Schwartz (1992) model

- This model offers a much larger variety of shapes than single factor models, with one inflexion point for the slope and the convexity.
- Instantaneous expected return for a discount bond:

$$r + \lambda \frac{(\exp(\psi\tau) - 1)B(\tau)}{\psi(\beta - \alpha)}(\alpha r - V)$$

- Subtracting r from the previous result, one obtains the risk premium.
- For a given τ maturity, the term premium is a linear function of *r* and *V*, depending on λ (market price of risk):
  - λ <0 => term premium > 0.
  - λ =0 => term premium = 0 => Expectations theory holds.
- For small τ, the term premium is an increasing function of *r*.

3. Balduzzi et al. (1996) models

Balduzzi, P., S. R. Das, S. Foresi, and R. Sundaran, 1996, "A Simple Approach to Three-Factor Affine Term Structure Models," *The Journal of Fixed Income*, 6, 14–31.

- Balduzzi et al. (1996) suggest the use of a three-factor model by adding the mean of the short-term rate to a 2-factor model.

$$dr = \mu_r(r, \theta, t)dt + \sigma_r(r, V, t)dz$$

$$d\theta = \mu_\theta(\theta, t)dt + \sigma_\theta(\theta, t)dw$$

$$dV = \mu_V(V, t)dt + \sigma_V(V, t)dy$$

$$dr = \kappa(\theta - r)dt + \sqrt{V}\,dz$$

$$d\theta = \alpha(\beta - \theta)dt + \eta\,dw$$

$$dV = a(b - V)dt + \phi\sqrt{V}\,dy$$

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) $P_t = E_t[M_{t+1} \dots M_{t+n}]$

- 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:

(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:

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

- Applying the already used statistical result

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

to (6) it is obtained

(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:

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

- In equation (10) , $\beta_{i_{t+1},M_{t+1}}$ is the coefficient of a regression of i_{t+1} on M_{t+1}.
- Therefore, it measures the correlation between the asset's return and the stochastic discount factor (sdf) or the 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:

(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) illustrate a basic result in finance theory:</u> the excess return of any asset over the risk-free asset depends on the covariance 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 $\left[-\frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]},\frac{\sigma_{M_{t+1}}\sigma_{i_{t+1}}}{E_t[M_{t+1}]}\right]$.

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

CAPM

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

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

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

- (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.

- **CCAPM:** an asset will pay a higher return or is riskier when the covariance 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.

- Equation (1) in logs:

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

- 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:

(14) $$p_t = \underbrace{E_t\left[m_{t+1} + p_{t+1}\right]}_{\mu} + 0.5 \cdot \underbrace{Var_t\left[m_{t+1} + p_{t+1}\right]}_{\sigma^2}$$

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

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

- **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:

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

$V(Z_t)$ - 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.

ε_t - 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 (15)):

(16)

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

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

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

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 (17), we have the factors as follows:

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

- Asset prices are also log-linear functions of the factors.

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

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 factor 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 = 0 => (from (19)) $A_0 = B_0 = 0$

- According to (15) and (19), the 1st term on the RHS of (14) is in (20):

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

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

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

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

- Using the factor definition in (17) $z_{t+1} = (I-\Phi)\theta + \Phi z_t + V(z_t)^{1/2}\varepsilon_{t+1}$ we get from (20):

(21)

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

- 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 => (21) may be simplified to:

(22)

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

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

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

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

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

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

- Evidencing the independent terms and the terms in z_t ,

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

- From

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

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

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

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

- Putting in evidence the independent terms and the terms in z_t, from (25) one obtains:

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

- Considering that the continuously compounded yield is

- Comparing the coefficients on the RHS of

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

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

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

the recursive restrictions in (27) and (28) are obtained:

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

- Considering that the continuously compounded yield is:

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

- From (29) and

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

the yield curve is defined as:

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

- From equations (27), (28) and (30)

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

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

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

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):

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

- Correspondingly, using the definition of the factors in

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

- and solving backwards, one gets:

(32)
$$E_t(z_{i,t+n}) = (1-\phi_i)\theta_i + \phi_i z_{i,t+n-1} = (1-\phi_i)\theta_i + \phi_i\left[(1-\phi_i)\theta_i + \phi_i z_{i,t+n-2}\right]$$
$$= \ldots = \sum_{j=1}^{n}\left[\phi_i^{j-1}(1-\phi_i)\theta_i\right] + \phi_i^n z_{i,t}$$

- Given that the expression in the sum corresponds to the sum of the first *n*-terms of a geometric progression with rate *f* 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:

(33)
$$E_t(z_{i,t+n}) = \left[(1-\phi_i)\theta_i\right]\frac{1-\phi_i^n}{1-\phi_i} + \phi_i^n z_{i,t} = \theta_i(1-\phi_i^n) + \phi_i^n z_{i,t}$$

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

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

- and plug

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

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

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

$$= \xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] E_t\left(z_{t+n}\right)$$

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

- From

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

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

one gets the variance of interest rates:

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

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

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

- From

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

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

one gets the instantaneous or one-period forward curve:

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

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

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

and the short-term rate in

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

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

(38)

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

- 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 (38) can be simplified as in (39):

(38)
$$\begin{aligned}
\Lambda_{n,t} &= E_t p_{n,t+1} - p_{n+1,t} - y_{1,t} \\
&= E_t(-A_n - B_n{}^T z_{t+1}) + (A_{n+1} + B_{n+1}{}^T z_t) - \left(\xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t\right) \\
&= -A_n - B_n{}^T[(I - \Phi)\theta + \Phi z_t] + (A_{n+1} + B_{n+1}{}^T z_t) - \left[\xi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i + \left(\gamma^T - \frac{1}{2}\sum_{i=1}^{k}\lambda_i^2 \beta_i\right) z_t\right] \\
&= \left\{A_{n+1} - [A_n + \xi + B_n{}^T (I - \Phi)\theta] - \frac{1}{2}\sum_{i=1}^{k}\lambda_i^2 \alpha_i\right\} + \left[B_{n+1}{}^T - \left(\gamma^T + B_n{}^T \Phi - \frac{1}{2}\sum_{i=1}^{k}\lambda_i^2 \beta_i\right)\right] z_t
\end{aligned}$$

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

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

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

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

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

(41) $$E_t p_{n,t+1} - p_{n+1,t} = \left\{p_{n+1,t} - E_t\left[m_{t+1}\right] - 0.5 \cdot Var_t\left[m_{t+1} + p_{n,t+1}\right]\right\} - p_{n+1,t}$$
$$= -E_t\left[m_{t+1}\right] - \frac{Var_t\left(m_{t+1}\right) + Var_t\left(p_{n,t+1}\right) + 2Cov\left(m_{t+1}, p_{n,t+1}\right)}{2}$$

- Given that the $Cov\left(m_{t+1}, p_{n,t+1}\right) = Cov\left(m_{t+1}, i_{n,t+1}\right)$, as $p_{n,t+1}$ is the only stochastic component in the rate of return, the previous equation is equal to:

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

- According to

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

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

- Solving in order to $E_t[p_{t+1}]$, having in mind that p_0=0, we get the price of the short-term bond:

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

- From

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

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

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

(38)
$$\begin{aligned}
\Lambda_{n,t} &= E_t p_{n,t+1} - p_{n+1,t} - y_{1,t} \\
&= E_t\left(-A_n - B_n{}^T z_{t+1}\right) + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left(\xi - \frac{1}{2}\lambda^T \alpha \lambda + \left[\gamma^T - \frac{1}{2}\lambda^T \beta^T \lambda\right] z_t\right) \\
&= -A_n - B_n{}^T\left[(I-\Phi)\theta + \Phi z_t\right] + \left(A_{n+1} + B_{n+1}{}^T z_t\right) - \left[\xi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i + \left(\gamma^T - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right) z_t\right] \\
&= \left\{A_{n+1} - \left[A_n + \xi + B_n{}^T (I-\Phi)\theta\right] - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \alpha_i\right\} + \left[B_{n+1}{}^T - \left(\gamma^T + B_n{}^T \Phi - \frac{1}{2}\sum_{i=1}^{k} \lambda_i^2 \beta_i\right)\right] z_t
\end{aligned}$$

the term premium will be equal to

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

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

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

we get

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

the covariance in

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

(46) $-B_n^T COV_t\left(z_{t+1}, m_{t+1}\right)$

- Consequently, equation (44) for **the term premium** becomes equivalente to:

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

- From

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

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

the term premium in

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

may be written as:

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

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.

↓

- Vasicel model can have the following DK characterisation:

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 diferences 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 and 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 in harder to do in iterative econometric processes.

- **Shortcomings of the forward rates under constant volatility:**

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

as β_i=0, the forward rate may be written as:

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

- As the last term of the RHS of

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

is zero when the volatility is constant, each factor loading in a multifactor Vasicek model corresponds to:

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

- The one-period forward rate may thus be written

From

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

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

- 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 limiting 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.

- **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 (15)

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

Factors - first-order autoregressive processes with zero mean (corresponds to considering the differences between the "true" factors and their means): From (17)

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

Bond prices: From (19)

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

Yield curve: From (30)

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

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

Factor loadings: From (27), (28) and (52)

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

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

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

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

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

Short-term interest rate:

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

From (31)

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

Given the common normalisation $p_{ot} = 0, \quad A_0 = B_{10} = B_{20} = 0,$

- 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:

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

From (30)

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

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

Volatility curve:

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

From (35)

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

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:

(61)

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

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

From (30)

$$\Lambda_{n,t} = E_t p_{n,t+1} - p_{n+1,t} - y_{1,t}$$

$$= -\sum_{i=1}^{k}\left[\lambda_i B_{i,n} + \frac{B_{i,n}^{\;2}}{2}\right]\alpha_i - \sum_{i=1}^{k}\left(\lambda_i B_{i,n} + \frac{B_{i,n}^{\;2}}{2}\right)\beta_i^T z_t$$

- 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.

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

- Maximum likelihood procedure – 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

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

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

where $y_{1,t}, \ldots, y_{l,t}$ are the l zero-coupon yields at time t with maturities $j = 1, \ldots, u$ periods and $w_{1,t}, \ldots, w_{l,t}$ 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

(63)
$$\underset{(k\times1)}{Z_t} = \underset{(k\times1)}{C} + \underset{(k\times k)}{F} \cdot \underset{(k\times1)}{Z_{t-1}} + G \underset{(k\times1)}{\upsilon_t}$$

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

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 $\upsilon_t \sim N(0,Q)$
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:

(64)
$$
\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ Var_t(y_{1,t+1}) \\ \vdots \\ Var_t(y_{l,t+1}) \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} \\ \vdots & \vdots \\ b_{1,l} & b_{2,l} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \end{bmatrix}
$$

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 second 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, as follows.

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

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

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

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

- From the 2 previous equations:

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

- As stated before, we have $z_{i,t+1} = \varphi_i z_{it} + \sigma_i \varepsilon_{i,t+1}$ → $\rho = \varphi_2$, $\rho u_{t+1} = \sigma_2 \varepsilon_{2,t+1}$

$$z_{2,t} = E_t(\pi_{t+1} - \overline{\pi}) = \rho(\pi_t - \overline{\pi}) = \rho\tilde{\pi}_t$$ → $$\tilde{\pi}_t = \frac{1}{\varphi_2} z_{2,t}$$

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

$$(68)\quad \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ Var_t(y_{1,t+1}) \\ \vdots \\ Var_t(y_{l,t+1}) \\ \tilde{\pi}_t \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \\ 0 \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} \\ \vdots & \vdots \\ b_{1,l} & b_{2,l} \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & b_{\pi} \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \\ \upsilon_{\pi} \end{bmatrix}$$

- **Major drawback:** it implies the 2nd factor to explain simultaneously the inflation as well as 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 not allow a simple 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:

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

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:

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

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

- 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:

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

(72) $$-m^*_{t+1} = \delta^* + \frac{\lambda_1^{*2}}{2}\sigma_1^{*2} + z^*_{1t} + \lambda^*_1\sigma^*_1\varepsilon^*_{t+1} + \frac{\lambda_2^2}{2}\sigma_2^2 + z_{2t} + \lambda_2\sigma_2\varepsilon_{t+1}$$

- Remaining equations:

(73) $-p^*_{n,t} = A^*_n + B^*_{1,n} z^*_{1t} + B_{2,n} z_{2t}$

(74) $z^*_{1,t+1} = \varphi^*_1 z^*_{1t} + \sigma^*_1 \varepsilon^*_{1,t+1}$

(75) $y^*_{n,t} = \frac{1}{n}\left(A^*_n + B^*_{1,n} z^*_{1t} + B_{2,n} z_{2t}\right)$

- 2-country model with (common) inflation:

(76)
$$\begin{bmatrix} y_{1,t} \\ \vdots \\ y_{l,t} \\ y^*_{1,t} \\ \vdots \\ y^*_{l,t} \\ Var_t\left(y_{1,t+1}\right) \\ \vdots \\ Var_t\left(y_{l,t+1}\right) \\ Var_t\left(y^*_{1,t+1}\right) \\ \vdots \\ Var_t\left(y^*_{l,t+1}\right) \\ \widetilde{\pi}_t \end{bmatrix} = \begin{bmatrix} a_{1,t} \\ \vdots \\ a_{l,t} \\ a^*_{1,t} \\ \vdots \\ a^*_{l,t} \\ a_{l+1,t} \\ \vdots \\ a_{2l} \\ a^*_{l+1,t} \\ \vdots \\ a^*_{2l} \\ 0 \end{bmatrix} + \begin{bmatrix} b_{1,1} & b_{2,1} & 0 \\ \vdots & \vdots & \vdots \\ b_{1,l} & b_{2,l} & 0 \\ 0 & b_{2,1} & b^*_{1,1} \\ \vdots & \vdots & \vdots \\ 0 & b_{2,l} & b^*_{1,l} \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \\ 0 & b_\pi & 0 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ z^*_{1,t} \end{bmatrix} + \begin{bmatrix} \upsilon_{1,t} \\ \vdots \\ \upsilon_{l,t} \\ \upsilon^*_{1,t} \\ \vdots \\ \upsilon^*_{l,t} \\ \upsilon_{l+1,t} \\ \vdots \\ \upsilon_{2l,t} \\ \upsilon^*_{l+1,t} \\ \vdots \\ \upsilon^*_{2l,t} \\ \upsilon_\pi \end{bmatrix}$$

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

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

- Using $\bar{Z}_o$ and P_0 and following (17), the optimal estimator for Z_1 will be given by:

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

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

(78)

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

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

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

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

- 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 forecast of y_t conditional on X_t and ζ_{t-1} directly from

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

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

- Therefore, from (62) and (79), we have the following expression for the forecasting error:

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

- From (80), the conditional variance-covariance matrix of the estimation error of the observation vector will be:

(81)

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

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

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

$$f(Y_t|I_{t-1}) = (2\pi)^{-1/2}\left|HP_{t|t-1}H' + R\right|^{-1/2} \cdot$$

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

- 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:

(83)
$$E\{[Y_t - E(Y_t|X_t,\zeta_{t-1})][Z_t - E(Z_t|X_t,\zeta_{t-1})]'\} = E\{[H(Z_t - \hat{Z}_{t|t-1}) + w_t][Z_t - \hat{Z}_{t|t-1}]'\}$$
$$= HE[(Z_t - \hat{Z}_{t|t-1})(Z_t - \hat{Z}_{t|t-1})']$$
$$= HP_{t|t-1}$$

- Therefore, using (79), (81) and (83), the conditional distribution of the vector (Y_t, Z_t) is:

(84)
$$\begin{bmatrix} Y_t|X_t,\zeta_{t-1} \\ Z_t|X_t,\zeta_{t-1} \end{bmatrix} \sim N\left(\begin{bmatrix} AX_t + H_{t|t-1} \\ \hat{Z}_{t|t-1} \end{bmatrix}, \begin{bmatrix} HP_{t|t-1}H' + R & HP_{t|t-1} \\ P_{t|t-1}H' & P_{t|t-1} \end{bmatrix} \right)$$

- Consequently, following (84), 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:

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

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

- Following this update, a new estimate for these estimates can be obtained, generalizing (77) and (78):

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

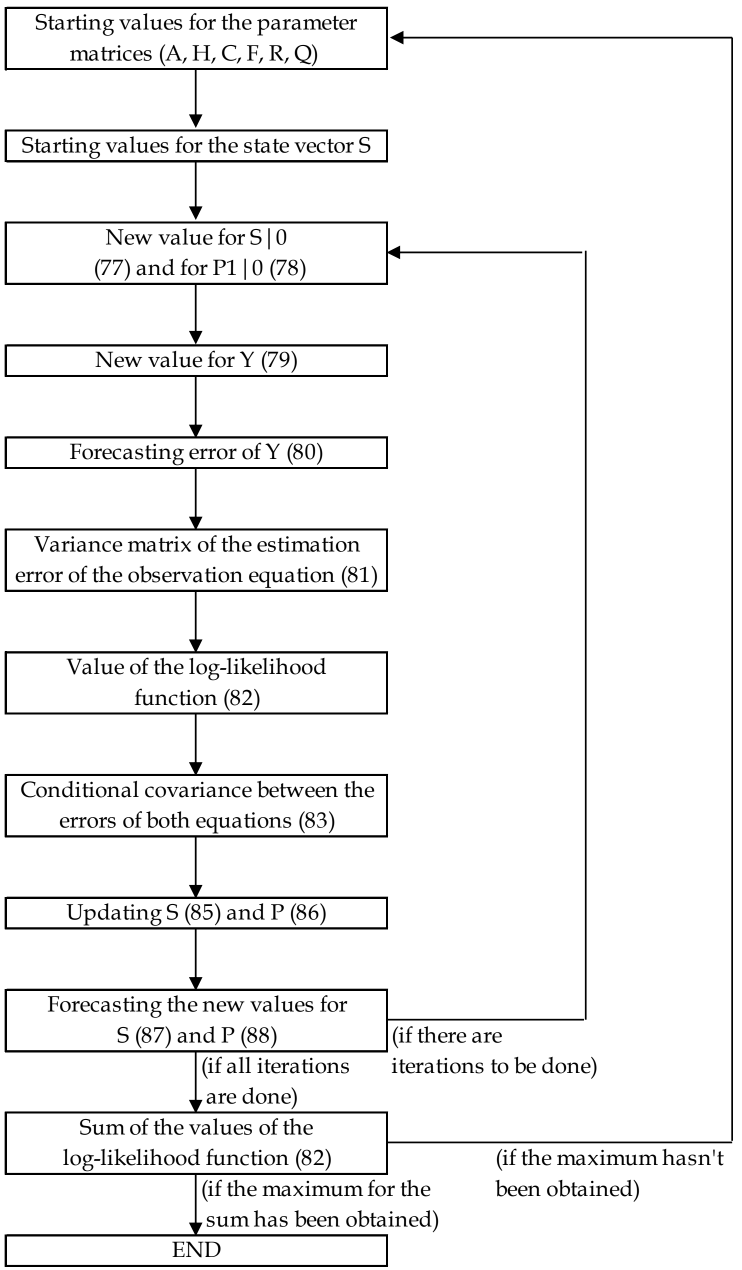
(88)

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

→ **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 err Y_t of

- Concluding, the Kalman Filter may be applied after specifying starting values for $\hat{Z}_{1|0}$ and $P_{1|0}$ using equations (79), (81), (85), and (86) and iterating on equations (87) and (88).


Starting values for the parameter matrices (A, H, C, F, R, Q)
Starting values for the state vector S
New value for S|0 (77) and for P1|0 (78)
New value for Y (79)
Forecasting error of Y (80)
Variance matrix of the estimation error of the observation equation (81)
Value of the log-likelihood function (82)
Conditional covariance between the errors of both equations (83)
Updating S (85) and P (86)
Forecasting the new values for S (87) and P (88)
(if there are iterations to be done)
(if all iterations are done)
Sum of the values of the log-likelihood function (82)
(if the maximum hasn't been obtained)
(if the maximum for the sum has been obtained)
END

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