Fixed Income and Credit Risk

Lecture 2

Professor	Assistant	Program
Fulvio Pegoraro	Roberto Marfè	MSc. Finance

Fall Semester 2012

UNIL | Université de Lausanne

Fixed Income and Credit Risk

Lecture 2 - Part I

Duration and Convexity of Bond Prices

Outline of Lecture 2 - Part I

2.1.1 **Introduction**

2.1.2 **Duration**

2.1.3 **Properties of Coupon Bond Duration**

2.1.4 **Duration of a Portfolio**

2.1.5 Dollar Duration

2.1.6 Convexity

2.1.7 Only one risk factor in the "Duration-Convexity setting"

2.1.8 Is there really a Free Lunch in the Duration model ?

2.1.1 Introduction

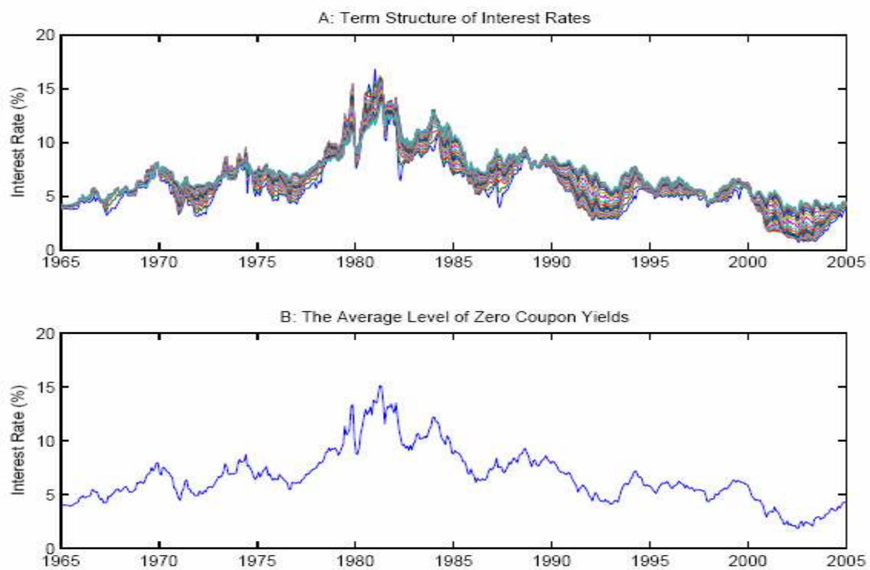
- Interest rates change substantially over time, and their variation poses large risks to financial institutions, portfolio managers, corporations, government and households.

- This chapter discusses the basics of interest rate risk management.

- In particular, we discuss how to measure risk for fixed income instruments, by introducing the notion of duration and convexity.

From Veronesi (2010, Chapter 3)

Figure 3.1 Zero Coupon Bond Yields and the Level of Interest Rates: 1965 - 2005



Data Source: CRSP.

- □ **Example: The Savings and Loan Debacle** in the 1980s is a standard example of what can go wrong when interest rates shift.

- □ A Savings and Loan ($S\&L$, say) is a kind of bank that earns a large part of its revenues from the difference between the *long-term mortgages* it provides to home owners and the short-term deposit rate it offers to depositors.

- □ When interest rates increased at the end of the 1970s, $S\&L$ were receiving their fixed coupons from mortgages contracted in the past, when rates were low, BUT suddenly they had to pay interest on deposits at the new higher deposit rates.

- Because depositors could choose were to put their money, banks were forced to offer high deposit rates, otherwise depositors would withdraw their deposits and invest in other securities, such as Treasuries.

- A withdraw of funds is the worst nightmare for a bank, as depositor's money is not in the bank any longer: it has been loaned to others (to firms).

- The spread between the rate earned on assets and the (higher) rate paid on liabilities quickly put many $S\&L$ out of business.

☐ The example above calls for:

(a) a systematic methodology to assess the riskiness of a bond portfolio to movements in interest rates;

(b) and a methodology to effectively manage such risk.

☐ Let us tackle the former problem thanks to the concept of **Duration**.

2.1.2 Duration

☐ **Definition:** *The* **Duration** *of a coupon bond (a security) with price CB is the (negative of the) percent sensitivity of the price CB to a small parallel shift in the level of interest rates. That is, let $Y(t,T)$ be the term structure of interest rates at time t.*

- *Consider a uniform shift of size dY across rates that brings the yield curve to $Y^*(t,T) = Y(t,T) + dY$ thus inducing $CB^* = CB + dCB$.*

- *The duration (approximately) measures the magnitude of dCB induced by dY.*

- ☐ It is a measure of the sensitivity of the coupon bond price to a change (a shift!) in interest rates.

- ☐ We use the yield to maturity $Y^{CB}(t,T) = Y$ of the bond as a proxy of the whole term structure of interest rates. Indeed, this YTM can be seen as an *average* of the spot rates discounting the risk-less cash flows.

- ☐ If the term structure of interest rates is *flat*, then the YTM is the term structure.

- ☐ How the duration of a coupon bond is calculated ? There are two main definitions: the **Macaulay Duration** and the **Modified Duration**.

□ The coupon bond price (assuming annual payments) at date $t = 0$ is given by:

$$CB(0,T) = \sum_{i=1}^{T} (C_i) \times (1+Y)^{-i},$$

and it is a **non-linear function** of Y.

□ Differentiating $CB(0,T)$ with respect to Y gives:

$$\frac{dCB(0,T)}{dY} = -\frac{1}{1+Y} \sum_{i=1}^{T} (i \times C_i) \times (1+Y)^{-i}.$$

□ We multiply both sides of the equation by $dY/CB(0,T)$ to get:

$$\frac{dCB(0,T)}{CB(0,T)} = -\frac{1}{1+Y} \times \frac{\sum_{i=1}^{T} (i \times C_i) \times (1+Y)^{-i}}{\sum_{i=1}^{T} (C_i) \times (1+Y)^{-i}} dY.$$

- We define the *Duration D* as:

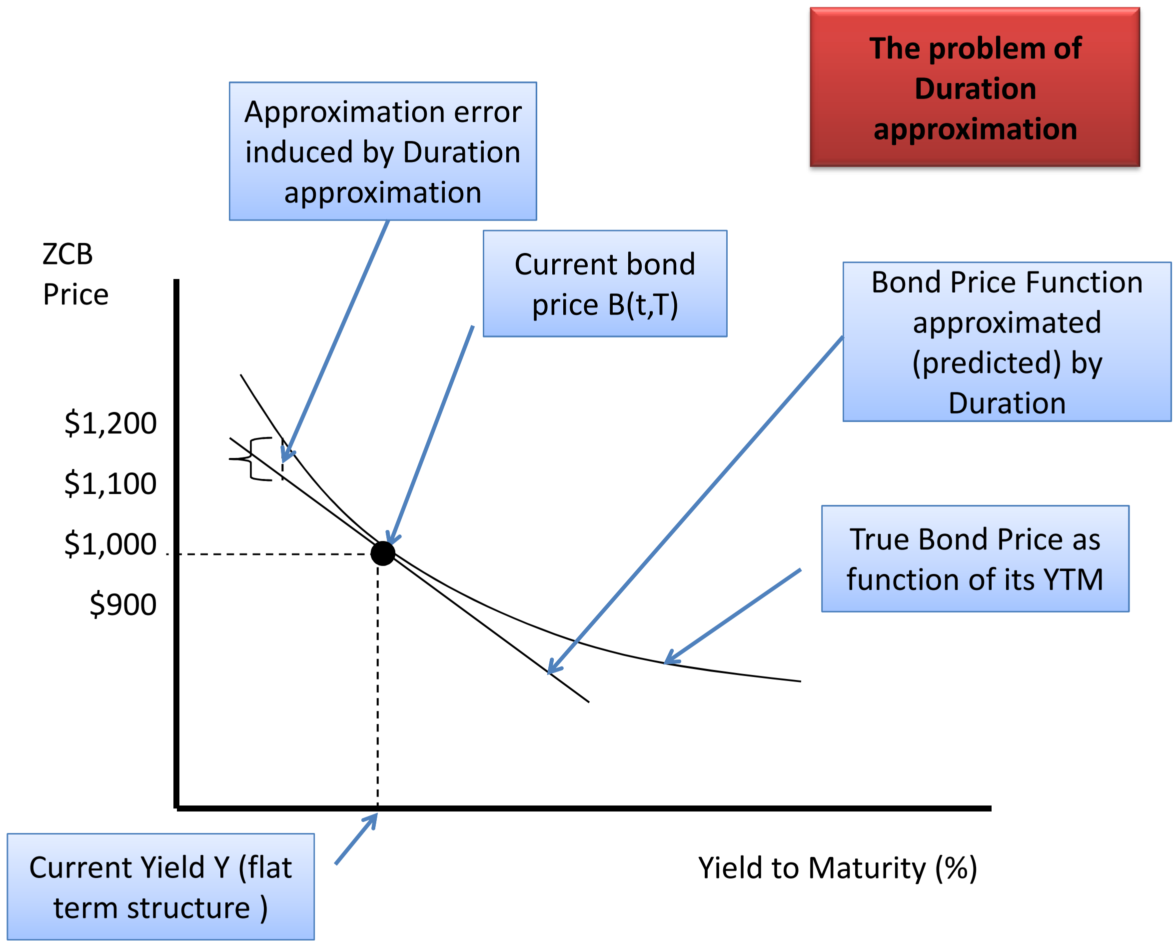
$$D = \frac{\sum_{i=1}^{T} (i \times C_i) \times (1+Y)^{-i}}{\sum_{i=1}^{T} (C_i) \times (1+Y)^{-i}} = \sum_{i=1}^{T} i \times \frac{C_i \times (1+Y)^{-i}}{CB(0,T)},$$

- We can now write $dCB(0,T)/CB(0,T)$ as:

$$\frac{dCB(0,T)}{CB(0,T)} = -\frac{D}{1+Y} dY \;\;\Rightarrow\;\; D = -\frac{dCB(0,T)}{CB(0,T)} \times \frac{(1+Y)}{dY}.$$

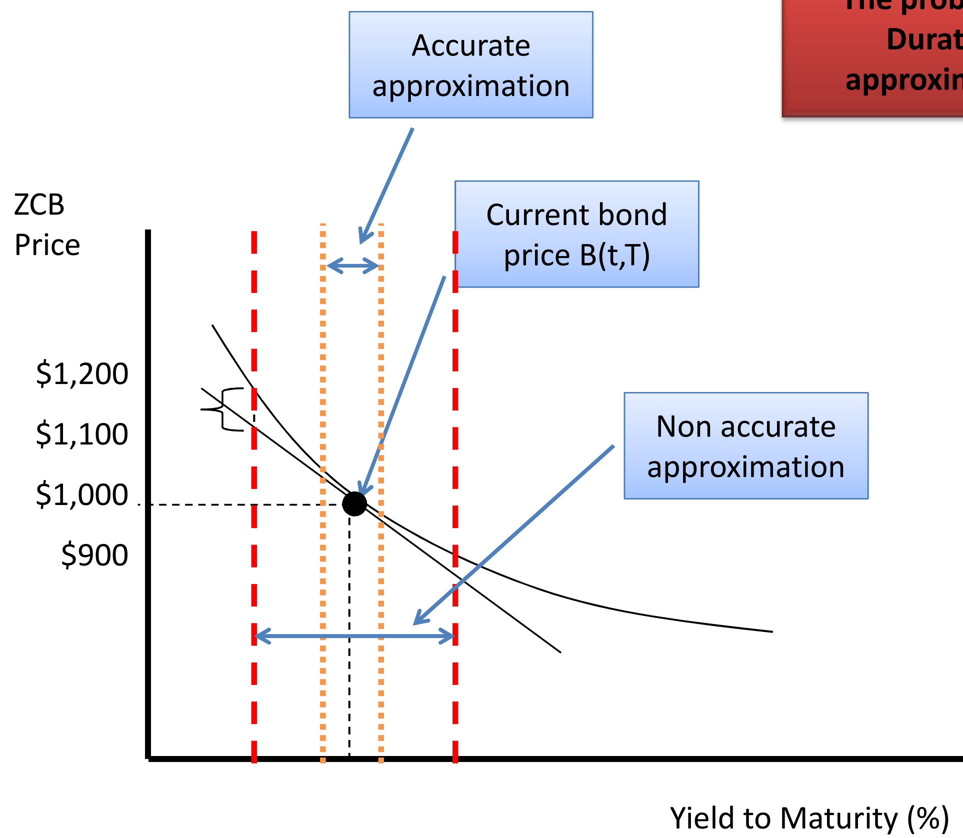
- D is called the *Macaulay duration*, and it is the weighted average of coupon dates (expressed in years) until the maturity of the bond, with the weights being the present values of the cash flows divided by the bond price.

□ It acts as a measure of **first-order sensitivity** of the bond price with respect to changes in the YTM (or **parallel shift** of the flat term structure).



The problem of Duration approximation
Approximation error induced by Duration approximation
Current bond price B(t,T)
Bond Price Function approximated (predicted) by Duration
True Bond Price as function of its YTM
ZCB Price
$1,200
$1,100
$1,000
$900
Current Yield Y (flat term structure)
Yield to Maturity (%)

The problem of Duration approximation



- Thus, Duration can be used to approximate/predict bond price changes, given a (very small!) change in the YTM.

- Let us consider at date $t = 0$ a coupon bond with a constant coupon rate c, a face value of C_T and time to maturity T. Let us denote with $CB(0,T)$ the bond price, $CB^c(0,T)$ the present value of coupon payments, and $CB^{C_T}(0,T)$ the present value of the principal payment. Then:

$$CB(0,T) = CB^c(0,T) + CB^{C_T}(0,T)\,,$$

$$\text{with } CB^c(0,T) = \sum_{i=1}^{T}(c \times C_T) \times (1+Y)^{-i} = (c \times C_T)\frac{1-(1+Y)^{-T}}{Y}\,,$$

$$\text{and } CB^{C_T}(0,T) = C_T(1+Y)^{-T}\,.$$

□ From previous relationships we can write:

$$D = -\frac{(1+Y)}{CB(0,T)}\frac{dCB(0,T)}{dY} = -\frac{(1+Y)}{CB(0,T)} \times \left(\frac{dCB^c(0,T)}{dY} + \frac{dCB^{C_T}(0,T)}{dY}\right) ,$$

□ and

$$\frac{dCB^c(0,T)}{dY} = \frac{-CB^c(0,T)}{Y} + \frac{(Tc)\, CB^{C_T}(0,T)}{Y(1+Y)} ,$$

$$\frac{dCB^{C_T}(0,T)}{dY} = \frac{-C_T T}{(1+Y)^{T+1}} = \frac{-CB^{C_T}(0,T)T}{1+Y} .$$

□ We thus obtain the following closed-form formula for D:

$$D = \frac{CB^c(0,T)}{CB(0,T)}\left(1+\frac{1}{Y}\right) + \frac{CB^{C_T}(0,T)}{CB(0,T)}\left(1-\frac{c}{Y}\right)T .$$

Special cases of this formula include:

a) *Zero-coupon bonds :* $c = 0$ and thus $CB^c(0,T) = 0$ and $CB^{C_T}(0,T) = CB(0,T)$ with $C_T = 1$. This means that $D = T$: the duration of a ZCB is equal to its residual maturity.

b) *Perpetuities :* there is no final repayment ($CB^{C_T}(0,T) = 0$ and $CB(0,T) = CB^c(0,T)$) in this case; we thus obtain $D = (1+Y)/Y$.

c) *Par bonds :* by definition, a par bond is a coupon bond such that $c = Y$ and thus $CB(0,T) = C_T$. Then:

$$D = (1 - (1+Y)^{-T}) \times \left(1 + \frac{1}{Y}\right) .$$

□ Another well known measure of duration, the **Modified Duration**, is given by:

$$D_{mod} = \frac{D}{1+Y}, \text{ so that}$$

$$\frac{dCB(0,T)}{CB(0,T)} = -D_{mod} \times dY \quad \Rightarrow \quad D_{mod} = -\frac{1}{dY} \times \frac{dCB(0,T)}{CB(0,T)}.$$

- **Example 1:** A \$100 million bond has modified duration equal 10, $D_{mod} = 10$.

→ This implies that one basis point increase in the level of interest rates $dY = .01\%$ generates a swing in the bond value of:

$$dCB = -10 \times \$100\,\texttt{million} \times \frac{0.01}{100} = -\$100,000\,.$$

↪ That is, the investor stands to lose 100,000 for every basis point increase in the (flat) term structure.

- **Example 2:** An investor has \$100 million invested in 5-year ZCBs, thus $D_{mod} = 5$ (years).

→ This implies that one basis point increase in the level of interest rates $dY = .01\%$ generates price reduction of:

$$dB = -5 \times \$100\,\texttt{million} \times \frac{0.01}{100} = -\$50,000\,.$$

↪ That is, the investor stands to lose $50,000$ for every basis point increase in the (flat) term structure.

2.1.3 Properties of Coupon Bond Duration

- ☐ For a given time to maturity and YTM, the duration decreases as the coupon rate increases:

  - the higher the coupon rate, the larger are the cash flows in the near future compared to long-term future. Cash flows that arrive sooner rather than later are less sensitive to changes in interest rates.

  - Thus an increase in c_i implies lower sensitivity to changes in the discount rate $(1+Y)^{-i}$ (or $B(t,T)$).

- For a given time to maturity and coupon rate, duration decreases as the YTM increases.
  - a higher YTM implies that short-term cash flows have a relatively higher weight in the value of the bond, and thus a lower sensitivity to changes in YTM.
- For a given coupon rate and YTM, the duration increases as the maturity increases.

2.1.4 Duration of a Portfolio

□ Consider a portfolio of $M = 2$ securities: the portfolio is made of N_1 units of security 1, and N_2 units of security 2. Let P_1 and P_2 be the prices of these two securities. The value of the portfolio is then $\Pi = N_1 \times P_1 + N_2 \times P_2$.

→ The duration of these two assets is

$$D_{mod,i} = -\frac{1}{P_i} \times \frac{dP_i}{dY}, \quad i \in \{1, 2\}.$$

→ The duration of the portfolio is:

$$D_{mod,\Pi} = -\frac{1}{\Pi} \times \frac{d\Pi}{dY} = -\frac{1}{\Pi} \times \frac{d(N_1 \times P_1 + N_2 \times P_2)}{dY} = -\frac{1}{\Pi}\left[N_1 \times \frac{dP_1}{dY} + N_2 \times \frac{dP_2}{dY}\right]$$

☐ The duration of the portfolio can thus be written :

$$
\begin{aligned}
D_{mod,\Pi} &= \frac{1}{\Pi}\left[N_1 \times P_1 \times \left(-\frac{1}{P_1}\frac{dP_1}{dY}\right) + N_2 \times P_2 \times \left(-\frac{1}{P_2}\frac{dP_2}{dY}\right)\right] \\
&= \pi_1\, D_{mod,1} + \pi_2\, D_{mod,2}\,,
\end{aligned}
$$

with $\pi_i = \dfrac{N_i \times P_i}{\Pi}$, $i \in \{1,2\}$.

☐ The duration of a portfolio of M securities is therefore:

$$
D_{mod,\Pi} = \sum_{i=1}^{M} \pi_i\, D_{mod,i}\,.
$$

→ the duration of the portfolio is a weighted average of the durations of the bonds in the portfolio, if all the bonds have the same YTM.

□ **Example 3:** A bond portfolio manager has \$100 million invested in 5-year ZCBs and \$200 million invested in 10-year ZCBs. What is the impact of a one basis point parallel shift of the term structure on the value of the portfolio?

□ We can answer this question by computing the duration of the portfolio: The 5-year and 10-year ZCBs have duration of 5 and 10, respectively. The total portfolio value is \$300 million.

→ The duration of the portfolio is : $D_{mod,\Pi} = \frac{100}{300} 5 + \frac{200}{300} 10 = 8.3$. Therefore, a one basis point increase in interest rates generates a portfolio loss of:

$$\$300\,\texttt{million} \times 8.3 \times 0.01\% = \$249,000.$$

2.1.5 Dollar Duration

- The previous definitions of Durations implicitly assume that the today's price of the asset is strictly positive: $CB > 0$, $\Pi > 0$. However, in many cases the assets or the portfolio we are interested in have a value exactly equal to zero. In that case we resort to the **Dollar Duration**.

- **Definition:** *The* **Dollar Duration** $D^{\$}$ *of a security* P *and portfolio* Π *are*

$$D_P^{\$} = -\frac{dP}{dY}, \quad D_{\Pi}^{\$} = \sum_{i=1}^{M} N_i \, D_i^{\$}$$

- **Definition:** *The* **price value of a basis point**, *denoted PV01 or PVBP, of a security* P *is defined by* $PV01 = D_P^{\$} \times dY$.

☐ The concept of Duration we have just presented is based on the following assumptions: i) YTM or interest rates variations are **infinitesimal**, ii) the term structure of interest rates is **flat**, and iii) shift in the term structure are **parallel**:

→ implicitly, there is **only one particular risk factor**: the one producing the infinitesimal parallel shift in the flat term structure.

☐ Starting from the next slide we relax assumption i) and discuss the notion of *Convexity* of a bond.

☐ Then we show why assumptions ii)-iii) can (seem to) generate risk-less arbitrage.

2.1.6 Convexity

☐ Let us consider now a Taylor expansion of the coupon bond price considered as function of the yield to maturity, namely $CB(0,T) = CB(Y)$:

$$CB(Y) = CB(Y_0) + \frac{dCB(Y)}{dY}|_{Y=Y_0}(Y - Y_0) + \frac{1}{2}\frac{d^2CB(Y)}{dY^2}|_{Y=Y_0}(Y - Y_0)^2 + \ldots$$

where Y_0 and Y denotes the YTM before and after the variation.

☐ Let us limit ourselves to the first two terms of the expansion to get:

$$\Delta CB(Y) = \frac{dCB(Y)}{dY}|_{Y=Y_0}\Delta Y + \frac{1}{2}\frac{d^2CB(Y)}{dY^2}|_{Y=Y_0}(\Delta Y)^2\,,$$

where $\Delta CB(Y) = CB(Y) - CB(Y_0)$ and $\Delta Y = Y - Y_0$ (**not infinitesimal!**).

☐ If we divide the LHS and RHS of that relation by $CB(Y)$ we obtain an expression of the relative price change as a function of a (modified) duration term and a **convexity** term:

$$\frac{\Delta CB(Y)}{CB(Y)} = -D_{mod}\Delta Y + \frac{\kappa}{2}(\Delta Y)^2,$$

$$\text{where } D_{mod} = -\frac{dCB(Y)}{dY}\frac{1}{CB(Y)},$$

$$\text{and } \kappa = \frac{1}{CB(Y)}\frac{d^2CB(Y)}{dY^2} \text{ is the } \textit{convexity}.$$

☐ For bond with fixed, risk-less cash flows, we can differentiate $CB(0,T) = \sum_{i=1}^{T}(C_i)\times(1+Y)^{-i}$ twice with respect to Y to get:

$$\frac{d^2CB(Y)}{dY^2} = \frac{2C_1}{(1+Y)^3} + \frac{6C_2}{(1+Y)^4} + \ldots + \frac{T(T+1)C_T}{(1+Y)^{T+2}}.$$

- Convexity in this case is:

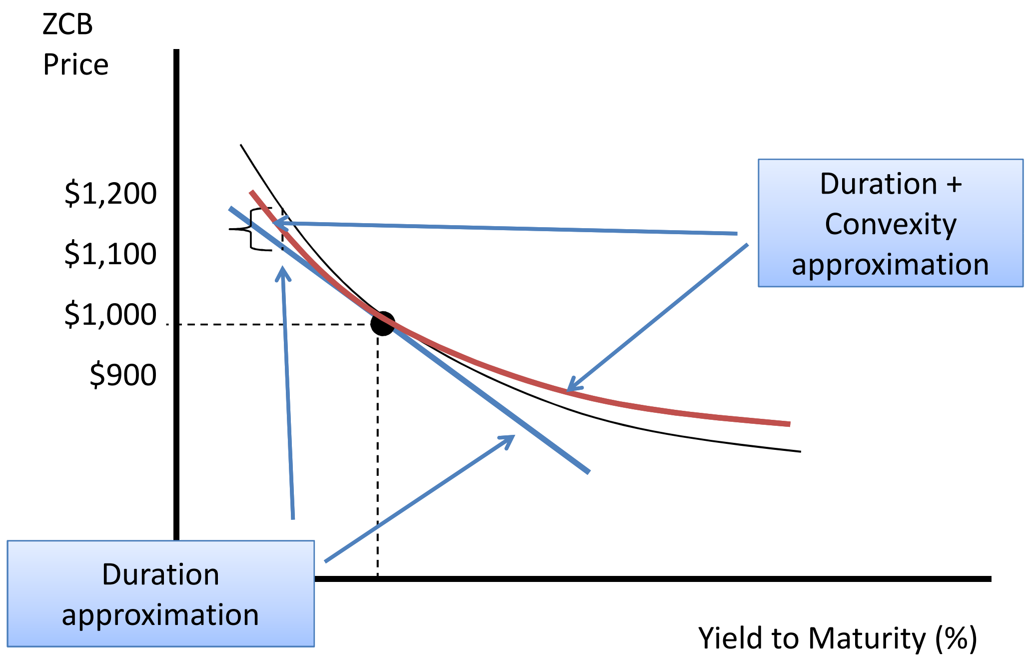
$$\kappa = \frac{\displaystyle\sum_{i=1}^{T} \frac{(i \times (1+i) \times C_i)}{(1+Y)^{i+2}}}{\displaystyle\sum_{i=1}^{T} \frac{C_i}{(1+Y)^i}} = \sum_{i=1}^{T} \frac{(i \times (1+i) \times C_i)}{(1+Y)^{i+2}} / CB(0,T)$$

- Duration is a linear approximation to the sensitivity of the bond price to changes in the YTM (parallel shift of the flat term structure). Convexity provides a second-order approximation of that sensitivity.

- Convexity, like duration, generally increases with maturity and decreases with the coupon rate and the YTM.

- The convexity of the ZCB is $\kappa = \dfrac{T\,(T+1)}{(1+Y)^2}$.

Duration and Convexity approximation



□ In the case of continuously compounded YTM R, the convexity of the ZCB is $\kappa = T^2$.

□ The **convexity of the portfolio** is:

$$\kappa_\Pi(Y) = \frac{d^2\Pi(Y)/dY^2}{\Pi(Y)} = \sum_{m=1}^{M} \pi_m \kappa_m(Y)$$

where $\kappa_m(Y)$ denotes the convexity of the m-th bond in the portfolio.

→ the convexity of the portfolio is a weighted average of the convexities of the bonds in the portfolio (if all the bonds have the same YTM).

☐ **Example 4:** A corporation has \$100 million (par) of a 10-year coupon bond that pays a 5% semi-annual coupon. Assume that the term structure of the interest rates is flat at $Y = 4.5\%$. The price of the bond is $CB = \$103.58$, implying a position of \$103.50 million, a (modified) duration of $D = 8.03$ and convexity $\kappa = 73.87$.

☐ If Y moves from 4.5% to 5.5%, the new price decline to \$95.63 with an associated *exact* loss of $dCB/CB = -7.67\%$.

☐ With a duration-base approximation, we have a loss of:

$$\frac{dCB}{CB} \approx -D \times 0.01 = -0.0803 \approx -8\% \,.$$

$\rightarrow$ Adding a convexity term instead entails a more precise approximate loss of:

$$\frac{dCB}{CB} \approx -D \times 0.01 + \frac{1}{2}\,\kappa\,(0.01)^2 = -0.07662 \approx -7.66\%\,.$$

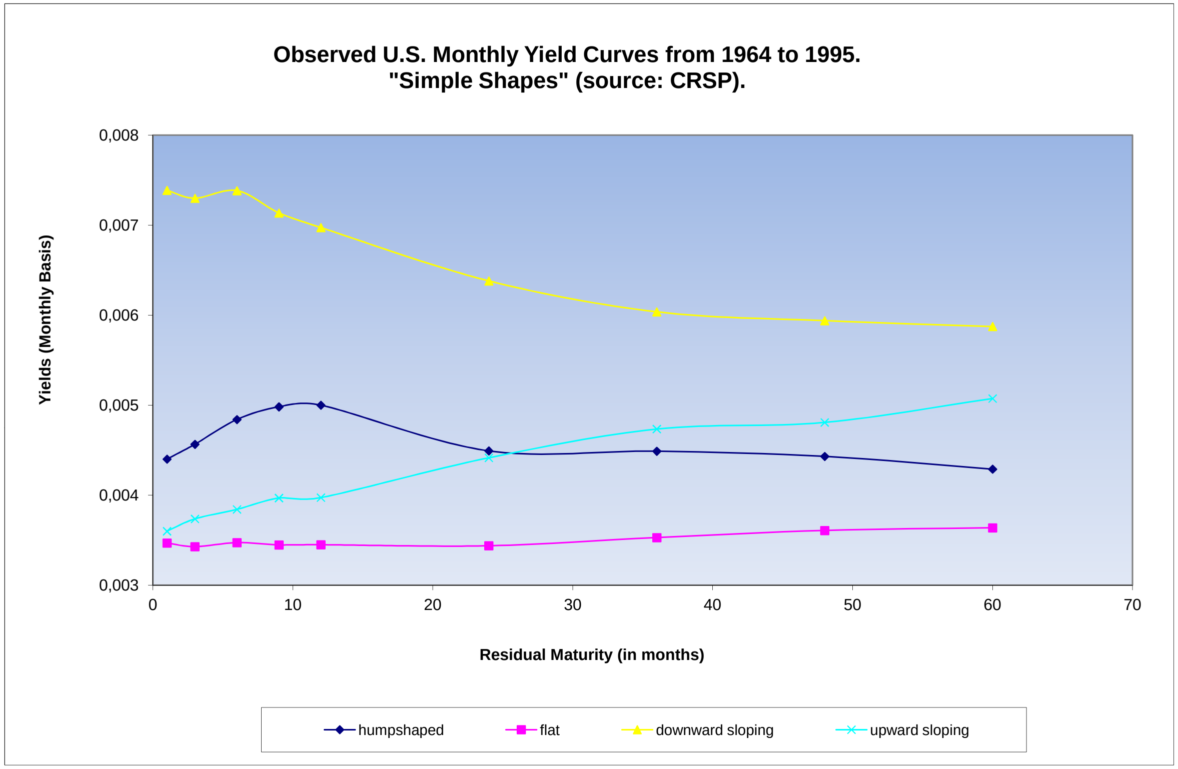
☐ In other words, the *Duration and Convexity* measure of $\frac{dCB}{CB}$ is more precise than the *Duration only* measure.

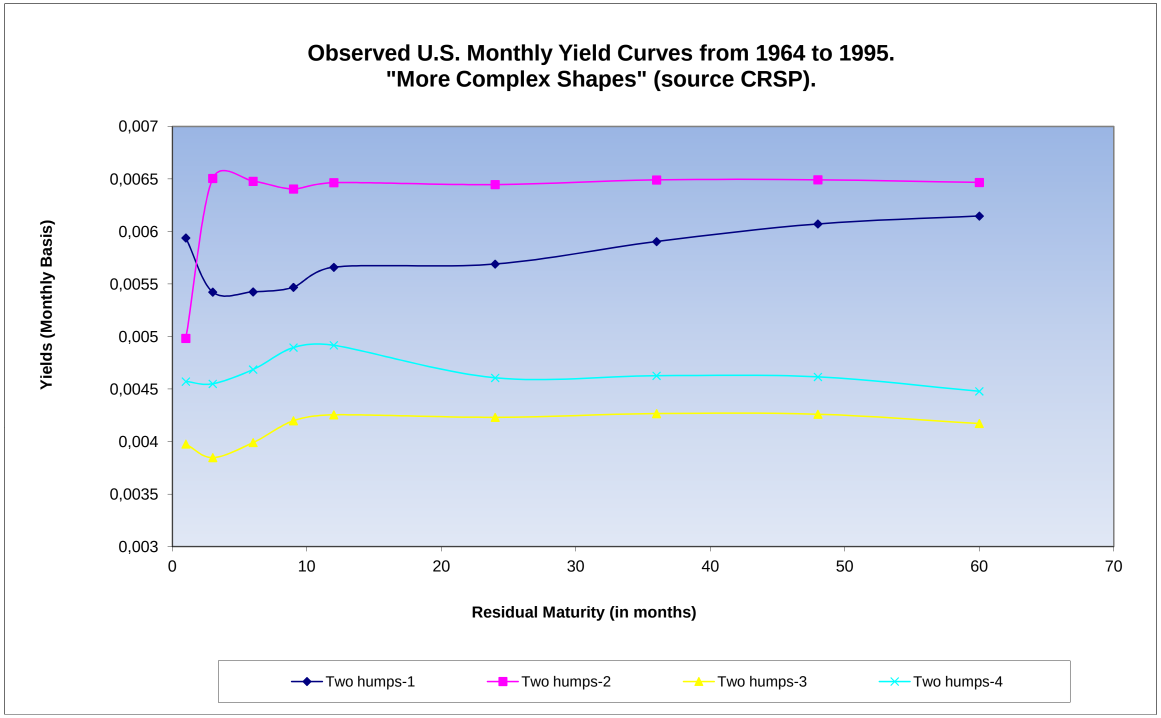
2.1.7 Only one risk factor in the "Duration-Convexity setting"

☐ The concepts of Duration and Convexity, presented in the previous slides, assumed parallel shift of a flat term structure.

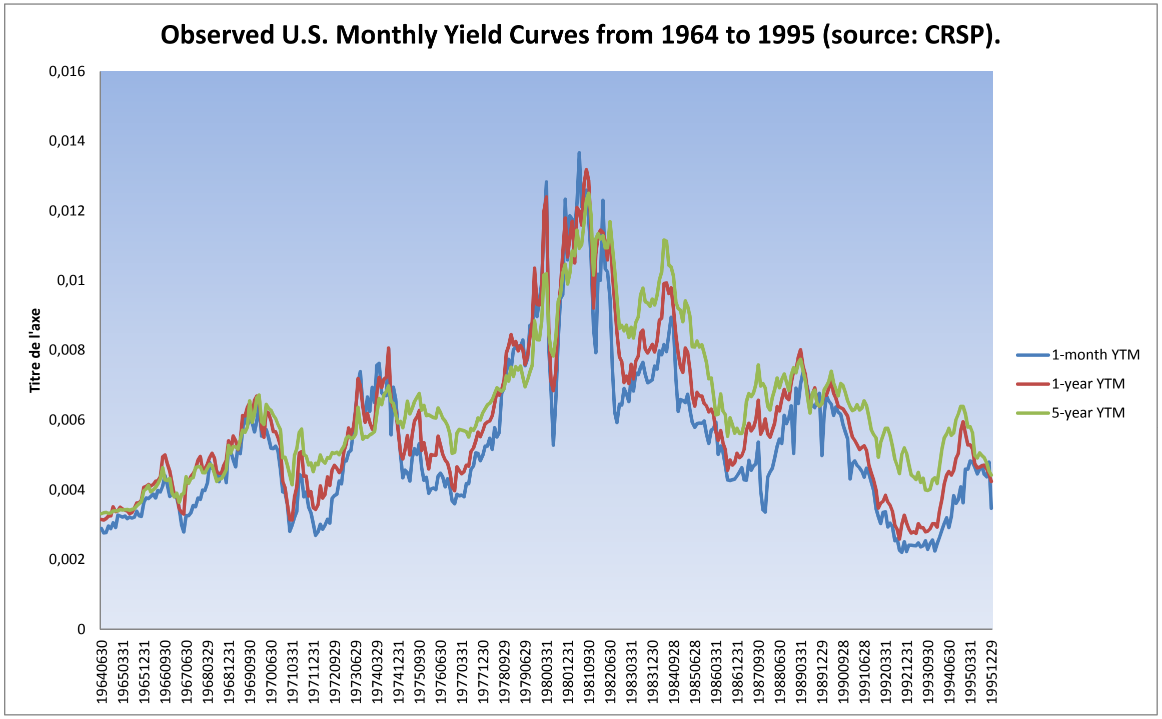
☐ This means that in the Duration-Convexity setting there is only one "FACTOR" determining bond price variations over time and for any residual maturity. In other words, there is only one source of risk. It is the YTM Y which is identified with the flat yield curve.

☐ In reality, the yield curve is not flat: see the two following graphs.


Observed U.S. Monthly Yield Curves from 1964 to 1995.
"Simple Shapes" (source: CRSP).
Yields (Monthly Basis)
0,008
0,007
0,006
0,005
0,004
0,003
0
10
20
30
40
50
60
70
Residual Maturity (in months)
humpshaped
flat
downward sloping
upward sloping


Observed U.S. Monthly Yield Curves from 1964 to 1995.
"More Complex Shapes" (source CRSP).
0,007
0,0065
0,006
0,0055
0,005
0,0045
0,004
0,0035
0,003
0
10
20
30
40
50
60
70
Yields (Monthly Basis)
Residual Maturity (in months)
Two humps-1
Two humps-2
Two humps-3
Two humps-4

☐ It does not move in parallel fashion (distance is not the same).

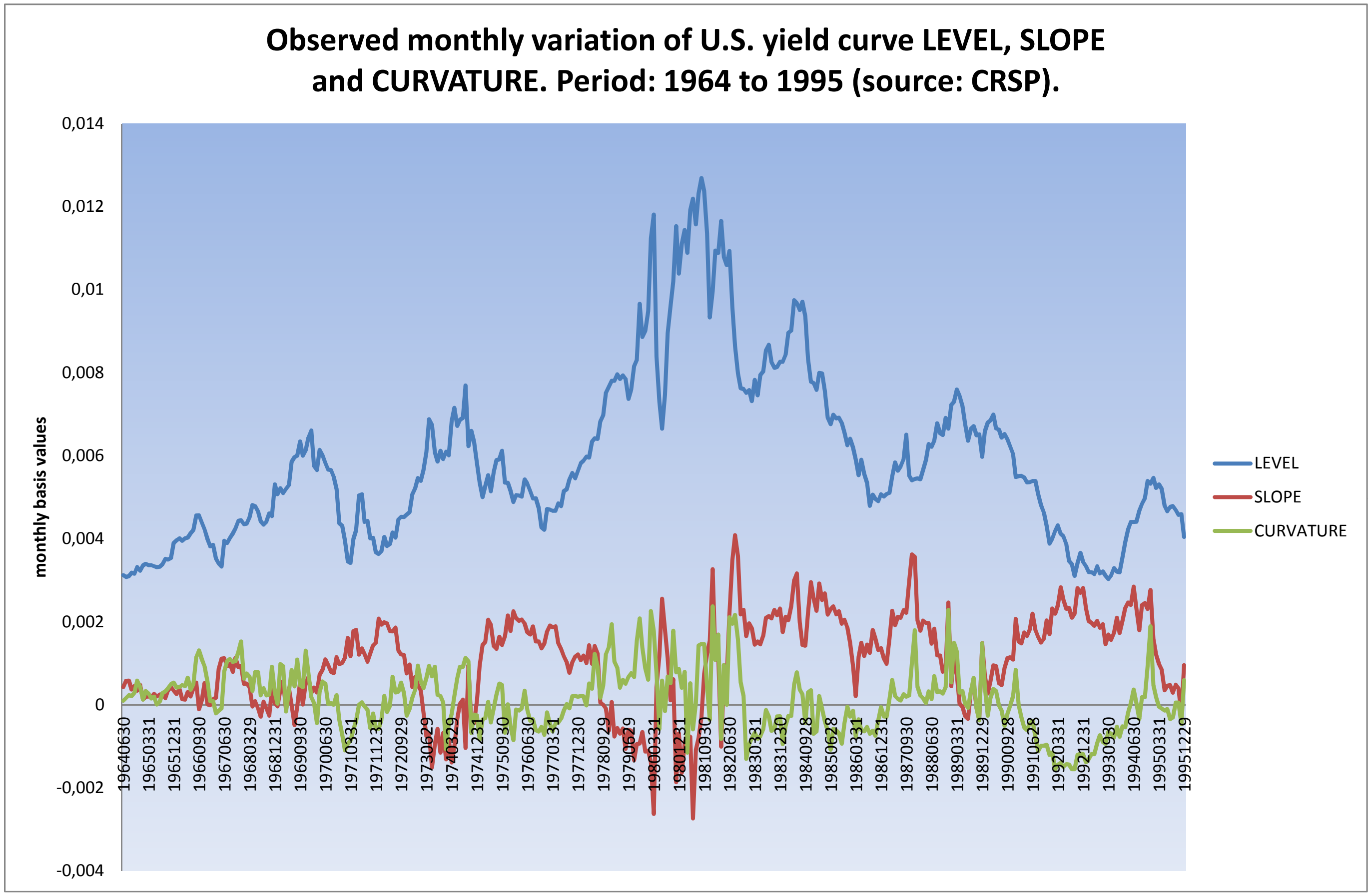


- The yield curve is characterized by time variations in its average level, slope and curvature.

- The single factor Y can be identified with the average LEVEL of interest rates.

- The time series of the Term Spread $= R(t, t+5y) - R(t, t+1m)$ can be seen as a measure of the variation over time of the yield curve SLOPE.

- The time series of the Butterfly Spread $= -R(t, t+1m) + 2\,R(t, t+1y) - R(t, t+5y)$ can be seen as a measure of the variation over time of the yield curve CURVATURE.



Observed monthly variation of U.S. yield curve LEVEL, SLOPE and CURVATURE. Period: 1964 to 1995 (source: CRSP).
monthly basis values
0,014
0,012
0,01
0,008
0,006
0,004
0,002
0
-0,002
-0,004
LEVEL
SLOPE
CURVATURE

$\Rightarrow$ We need for more FACTORS organized by sophisticated interest rate models able to:

- account for (to explain!) a **non-flat term structure**
- moving over time and maturities in a **realistic way** (i.e. close to the data : time varying level, slope and curvature)
- and compatible with the **no-arbitrage principle** (i.e. a set of "fair" prices).

2.1.8 Is there really a Free Lunch in the Duration model ?

☐ Are the parallel shifts of the term structure (implicit in the Duration setting) *acceptable*, in the sense that they do not allow for "free lunch" ?

☐ If such free lunch was possible, then these assumptions should be revised in order to better explain the reality of bond markets : there are very few arbitrage opportunities!

☐ An **arbitrage opportunity** is a feasible trading strategy involving two or more securities with either of the following characteristics :

a) it does not cost anything at initiation, and it generates a sure positive profit by certain date in the future;

b) it generates a positive profit at initiation, and it has a sure non-negative payoff by a certain date in the future.

☐ The **no-arbitrage** condition requires that no arbitrage opportunities exist.

☐ see Veronesi (2010).

□ The presumption of a free lunch in the "Duration setting" comes from the following fact:

→ two portfolios having the *same value (=price)* and the *same duration* will generally have *different convexities*.

↪ in a world where movement in flat term structure are parallel, a strategy where one buy the high-convexity portfolio and sells the low-convexity one seems to be an arbitrage opportunity.

The Example:

- Let us consider at date $t = 0$ a portfolio A of a 5-year zero-coupon bond (with price $B(0,5)$), and a portfolio B invested in 1-year and 10-year zero-coupon bonds (with price $B(0,1)$ and $B(0,10)$, respectively).

- The relative weights on 1-year and 10-year ZBCs are chosen such that portfolios A and B have the *same value* and the *same modified duration*. Portfolio B is called "barbell portfolio".

- Let us denote by q_1 and q_{10} the number of 1-year and 10-year ZCBs in portfolio B, and by $D_{mod,i}$ the modified duration of an i-year ZCB.

☐ We have the following linear system:

$$B(0,5) = B(0,1)q_1 + B(0,10)q_{10} \quad \text{(equal value constraint)}\,,$$

$$\Longrightarrow \quad B(0,5)D_{mod,5} = (B(0,1)D_{mod,1})q_1 + (B(0,10)D_{mod,10})q_{10}$$

$$\text{(duration-matching constraint)}\,.$$

☐ Let us assume that all bonds have a face value of 100 and interest rates are 5% (flat term structure), then:

$$B(0,1) = 100(1.05)^{-1} = 95.238\,,\ B(0,5) = 100(1.05)^{-5} = 78.353\,,$$

$$B(0,10) = 100(1.05)^{-10} = 61.391\,,\ D_{mod,1} = \tfrac{1}{1.05} = 0.952 \text{ year}\,,$$

$$D_{mod,5} = \tfrac{5}{1.05} = 4.762 \text{ years}\,,\ D_{mod,10} = \tfrac{10}{1.05} = 9.524 \text{ years}\,.$$

☐ The system simplify to:

$$78.353 = 95.238q_1 + 61.391q_{10}\,,$$

$$373.108 = 90.703q_1 + 584.676q_{10} \quad \Rightarrow \quad q_1 = 0.457 \text{ and } q_{10} = 0.576.$$

- Which of portfolio A or portfolio B is more convex? From the ZCB convexity formula we obtain : $\kappa_1 = 1.81$, $\kappa_5 = 27.21$ and $\kappa_{10} = 99.77$. The convexity of the portfolio B is therefore:

$$\kappa_B = \frac{B(0,1)q_1\kappa_1 + B(0,10)q_{10}\kappa_{10}}{B(0,1)q_1 + B(0,10)q_{10}} \approx 45$$

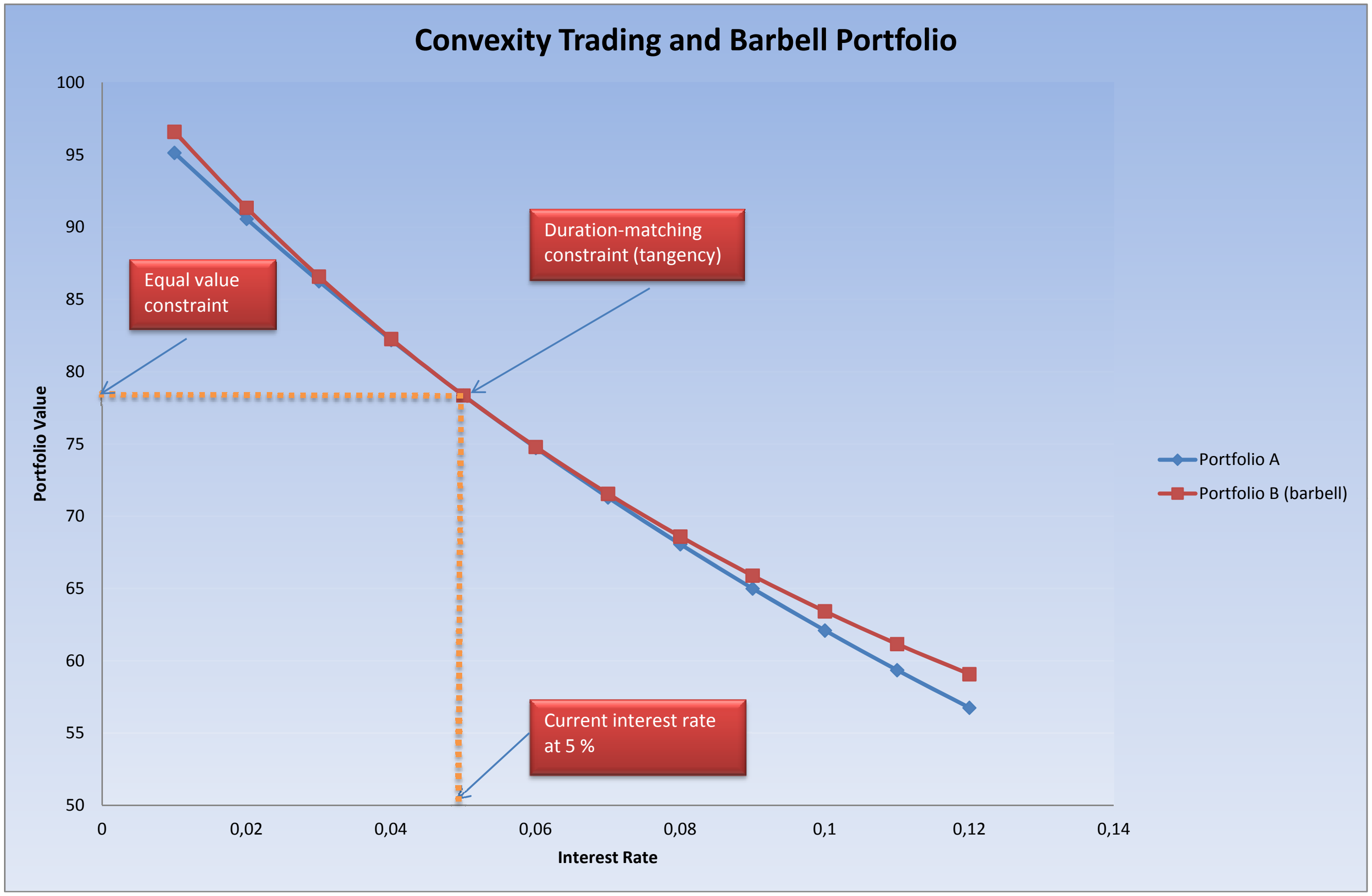
while the convexity of portfolio A is $\kappa_A = \kappa_5 = 27.21 \Rightarrow \kappa_B > \kappa_A$.

- The free lunch (in this simple setting) is obtained selling portfolio A (intermediate maturity) and buying portfolio B (at date $t = 0$). To illustrate this point, let us imagine that the term structure has a upward shift (to 6%) or a downward shift (to 4%) (like a binomial distribution).

☐ Starting from the rate $Y = 5\%$ we have $\Pi_A(5\%) = \Pi_B(5\%) = 78.353$. We buy portfolio B (with large convexity) and we sell A (smaller convexity) and thus we have zero net profit at $t = 0$.

☐ If, at $t = 1$, we move to the **up scenario**, $\Pi_B(6\%) = 74.79 > \Pi_A(6\%) = 74.73$, and if we move to the **down scenario** we have $\Pi_B(4\%) = 82.27 > \Pi_A(4\%) = 82.19$.

$\hookrightarrow$ At $t = 1$, portfolio B is always more valuable than portfolio A, and therefore I earn the difference once I close the position. I have a positive return exploiting the convexity.


Convexity Trading and Barbell Portfolio
Duration-matching constraint (tangency)
Equal value constraint
Current interest rate at 5 %
Portfolio Value
Interest Rate
100
95
90
85
80
75
70
65
60
55
50
0
0,02
0,04
0,06
0,08
0,1
0,12
0,14
Portfolio A
Portfolio B (barbell)

⇒ Free Lunch !...apparently !

☐ When we move from date $t = 0$ to date $t = 1$ (where we find up/down scenario), we assume that the 3 ZCBs maintain the same residual maturity.

☐ We have proposed a "very (very!) simple" example interested to highlight the limit of the Duration model, as we have done with the level-slope-curvature analysis of the observed U.S. yield curves.

☐ In reality, the convexity trading strategy does not represent an arbitrage opportunity. Why ? Because, in this analysis we do not take into account the *time dimension* that naturally affects bond prices over time.

☐ For instance, a ZCB price can increases over time simply because times passes (and it approaches the maturity date), even if interest rates do not move.

☐ What happens ? The gain in value from higher convexity is compensated by a lower gain due to the passage of time (in dynamic investment strategy this relation is known as the Theta-Gamma relation).

Fixed Income and Credit Risk

Lecture 2 - Part II

Extracting Yield Curves from Bond Prices

Outline of Lecture 2 - Part II

2.2.4 Exponential-Polynomial Families

2.2.4.1 The Exponential-Polynomial class for forward rates

2.2.4.2 The Nelson-Siegel (1987) family

2.2.4.3 The Svensson (1994) family

2.2.4.4 The Gurkaynak, Sack and Wright (2007) data base on nominal yields

2.2.4.5 The Gurkaynak, Sack and Wright (2010) TIPS yield curve data base

2.2.5 **The Principal Component Analysis of the Yield Curve**

2.2.5.1 Principal Component Analysis

2.2.5.2 PCA of the yield curve

2.2.5.3 PCA, Factors and yield curve information

2.2.1 Introduction

☐ We have seen during **Lecture 1** how important is the **Discount Function** $B(t, T)$:

$i)$ The associated (continuously compounded) ZCB **Yield Curve** $R(t, T) = -\frac{\ln B(t,T)}{T-t}$ provides a clear picture of how date-t spot interest rates changes as a function of the maturity date T.

$ii)$ From $B(t,T)$ or $R(t,T)$, we can determine the **forward rate curve** $R(t,\tau,T)$ or the **par yield curve** $\rho(t,T)$:

$\hookrightarrow$ the first one provides a date-t information of future spot rates.

$\hookrightarrow$ the second one gives the date-t market interest for a bullet bond.

$iii)$ We have seen that a coupon bond can be interpreted as a portfolio of ZCBs : C_1 units of ZCBs maturing at T_1, C_2 units of ZCBs maturing at T_2, ... etc. :

$$CB(t,T) = \sum_{i=1}^{n} 1_{\{t<T_i\}} C_i B(t,T_i).$$

- ☐ In most markets only a few ZCBs are traded, so that information about the discount function must be inferred from market prices of coupon bonds.

- ☐ The purpose of this lecture is to present methods to extract or estimate the term structure of interest rates from prices of coupon bonds at a given date.

- ☐ In Section 2.2.2 I will present the so-called **bootstrapping technique**, based on the construction of ZCB prices by means of certain portfolios of coupon bonds.

  $\hookrightarrow$ this is possible only in bond markets with sufficiently many coupon bonds with regular payment dates and maturities.

□ In Sections 2.2.3 and 2.2.4 I will present two alternative techniques based on the assumption that the discount function $B(t,T)$ is specified as a function of unknown parameters (and the time-to-maturity $T-t$): $B(t,T) = f(t,T,\theta)$:

$a)$ the vector of unknown parameters θ is estimated in order to obtain the best possible fit (for a given estimation criterion) of observed bond prices by theoretical ones.

$b)$ $f(t,T,\theta)$ is typically a polynomial, or an exponential function of the time-to-maturity $\tau = T - t$ (for a given t) or some combination.

c) this is consistent with the idea that $B(t,T)$ and $R(t,T)$ are continuous and smooth function of τ. Indeed, if $R(t,T_1) >> R(t,T_1+\varepsilon)$:

$\Rightarrow$ Bond owners would probably shifts from low-yield to high-yield bonds, and bond issuers would shift to the low-yield maturity. These changes in the supply and demand will determine a reduction of the gap driving $R(t,T_1)$ close to $R(t,T_1+\varepsilon)$.

2.2.2 Bootstrapping technique

- ☐ Several bond markets issue and trade a very small number of zero-coupon bonds. Usually, such zero-coupon bonds have a very short maturity.

- ☐ If we want to determine market zero-coupon yields for longer maturities, we have to extract information from the prices of traded coupon bonds. In some (not all!) markets we have the possibility to construct some longer-term zero-coupon bonds by forming portfolios of traded coupon bonds.

- ☐ Market prices of these ”synthetical” zero-coupon bonds and the associated zero-coupon yields can then be derived. Let us start from a simple example.

Example - Consider a bond market at date t where two bullet bonds are traded. A ZCB expiring in one year and selling for $CB(t, t+1) = 97$ and a coupon bond, with an annual coupon rate of 5%, expiring in two years and selling for $CB(t, t+2) = 95$. Both have annual payments and face value of 100.

– from the traded ZCB : $97 = 100 \times B(t, t+1) \Rightarrow B(t, t+1) = 0.97$ and thus

$R(t, t+1) = -\ln B(t, t+1) = 0.03046$.

– from the traded CB :

$$\begin{aligned} 95 &= 5 \times B(t, t+1) + 105 \times B(t, t+2) \\ &= 5 \times 0.97 + 105 \times B(t, t+2) \;\Rightarrow\; B(t, t+2) = 0.8586\,, \end{aligned}$$

and thus $R(t, t+2) = -\frac{1}{2} \ln B(t, t+2) = 0.07623$.

☐ From the last relation we observe that:

$$B(t, t+2) = \frac{1}{105} CB(t, t+2) - \frac{5}{105} B(t, t+1)\,,$$

we can construct a ZCB with residual maturity of 2 years as a portfolio of $1/105$ units of $CB(t, t+2)$ (sell) and $-5/105$ units of the 1-year ZCB (buy).

☐ This simple example considers only 2 bonds associated to the maturities $t+1$ and $t+2$. This setting can easily be generalized to more periods and more assets.

□ Let us assume to have, at date t, M coupon bonds with:

- regularly increasing maturities $\{t+1, t+2, \ldots, t+M\}$, respectively,
- coupon payments at each period
- and occurring at the same dates.

□ In that case, we can construct recursively (as in the example!) the market discount factors $B(t, t+1), B(t, t+2), \ldots, B(t, t+M)$ and the associated yield to maturities $R(t, t+1), R(t, t+2), \ldots, R(t, t+M)$.

□ We are able to construct a finite set of points of the yield curve, and up to $t+M$. This methodology is called **bootstrapping** or **yield curve stripping**.

- The yield curve stripping also applies to the case where the maturity dates of the M bonds are not all different and regularly increasing as in the previous case.

- Let us consider M bonds having at most M different payment dates. Let us denote the payment of the coupon bond $i \in \{1, \ldots, M\}$ at time $t + j$, with $j \in \{1, \ldots, M\}$, by $C_{i,j}$. We can organize all these cash flows in a **square** matrix $\mathbf{C} = (C_{i,j})$.

- Clearly, some of the elements $C_{i,j}$ may be zero (e.g., if the maturity date is before $t + M$).

- Let us organize the coupon bond and ZCB prices in the vectors $\mathbf{P} = [CB(t, t+1), \ldots, CB(t, t+M)]'$ and $\mathbf{B} = [B(t, t+1), \ldots, B(t, t+M)]'$.

- From the relation between coupon bond and ZCB prices we can write:

$$\mathbf{P} = \mathbf{C} \times \mathbf{B}$$

- Given this linear system based on the $(M \times M)$ square matrix $\mathbf{C}$, if the bonds payments are such that $\mathbf{C}$ is non-singular (e.g., at any date there is at least a payment), then a unique solution for $\mathbf{B}$ exists: $\mathbf{B} = \mathbf{C}^{-1}\mathbf{P}$.

☐ Observe that the bootstrap method is able to construct ZCB prices, that is, spot rates, for a **finite number** of maturities, while the term structure of interest rates is a curve.

☐ What we do if we need the discount factor for an intermediate maturity (e.g., 2 years and 3 months)?

$\rightarrow$ **We can interpolate !** Let us imagine to know the interest rates $R(t, T_i)$ and $R(t, T_{i+1})$ at the maturities T_i and T_{i+1} (respectively) and let us assume that we are interested to find the yield with maturity $T_j \in (T_i, T_{i+1})$.

☐ By linear interpolation we have:

$$R(t,T_j) = \frac{(T_j - T_i) \times R(t,T_{i+1}) + (T_{i+1} - T_j) \times R(t,T_i)}{T_{i+1} - T_i};$$

we are simply drawing lines between the points $(T_i, R(t,T_i))$ and $(T_{i+1}, R(t,T_{i+1}))$.

☐ The implementation of the bootstrap method is based on several **assumptions/limits** about the **available coupon bonds** we use to extract the market discount factors (and the associated yield to maturities).

□ **Namely :**

$i)$ we cannot construct a yield $R(t, T)$ for $T > M$;

$ii)$ we can extract only a finite number of yields and the use of interpolation generate non smooth (spot and forward) curves;

$iii)$ we need one payment date at each period and identical payment dates or, more generally, the **square matrix C** has to be non-singular.

☐ What happens if $\mathbf{C}$ is not square ? That is, if

number M (say) of coupon bonds $\neq$

number m (say) of payment dates (i.e., maturities) ?

☐ Let us assume that $M > m$ and let us present the following example. At date t we have in the market three bonds:

a) a one-year bullet bond with an annual coupon rate of 10%, face value 100 and price $CB(t, t+1) = 100$;

b) a two-year bullet bond with an annual coupon rate of 5%, face value 100 and price $CB(t, t+2) = 90$;

c) a two-year coupon bond (same payment dates as in b)) paying 58 in one year, 54 in two years, price $SB(t, t+2) = 98$ (it is called *serial bond*).

□ The market discount factors $B(t, t+1)$ and $B(t, t+2)$ (i.e., the prices of the ZCBs with unitary face value) must satisfy the following system of the 3 equations:

$$
\begin{aligned}
100 &= 110B(t, t+1) \\
90 &= 5B(t, t+1) + 105B(t, t+2) \\
98 &= 58B(t, t+1) + 54B(t, t+2)\,.
\end{aligned}
$$

□ Let us denote:

$$
A = \begin{bmatrix} 110 & 0 \\ 5 & 105 \\ 58 & 54 \end{bmatrix}, \quad b = \begin{bmatrix} 100 \\ 90 \\ 98 \end{bmatrix}
$$

☐ We have : $rank(A : b) = rank(A) + 1 \Rightarrow$ no solution exists! It does not exist a system $(B(t,t+1), B(t,t+2))'$ of market discount factors, satisfying the no-arbitrage principle, compatible with existing (at date t) coupon market prices.

☐ **Remember :** for a given $m \times n$ matrix A, and an associated linear system $Ax = b$, we have:

$$\begin{aligned} rank(A:b) &= rank(A) = n \;\Rightarrow\; \text{unique solution}\,; \\ rank(A:b) &= rank(A) < n \;\Rightarrow\; \text{multiple solution}\,; \\ rank(A:b) &= rank(A) + 1 \;\Rightarrow\; \text{no solution}\,. \end{aligned}$$

☐ If we consider only the first two assets:

$$A = \begin{bmatrix} 110 & 0 \\ 5 & 105 \end{bmatrix}, \; b = \begin{bmatrix} 100 \\ 90 \end{bmatrix} \;\Rightarrow\; B(t,t+1) = 0.9091\,, \;\; B(t,t+2) = 0.8139\,.$$

☐ If we consider only the last two assets:

$$A = \begin{bmatrix} 5 & 105 \\ 58 & 54 \end{bmatrix}, \quad b = \begin{bmatrix} 90 \\ 98 \end{bmatrix} \quad \Rightarrow \quad B(t,t+1) = 0.9330\,, \;\; B(t,t+2) = 0.8127\,.$$

☐ If we take the first solution as “correct”, the no-arbitrage price of the *serial bond* should be:

$$58 \times 0.9091 + 54 \times 0.8139 = 96.98 < 98.00\,,$$

☐ The market price at date t is “too expansive” (w.r.t. to the no-arbitrage principle). In other words, the serial bond is mispriced relative to the two coupon bonds: we can exploit this **arbitrage opportunity** selling the **expensive** serial bond and buying a portfolio of the two bullet bonds that replicate the serial one.

☐ What we typically observe in bond markets is that, at a given date t, the number of bond prices M is much smaller than that of payment dates m (say) : $M << m$. It happens in particular for long maturities where, typically, the number of bonds is small.

☐ Thus : $rank(A : b) = rank(A) < m \Rightarrow$ we have multiple solutions.

☐ Many set of discount factors can be compatible both with observed prices and no-arbitrage principle.

☐ Moreover, several entries of **C** are equal to zero, because of bonds with (many) different payment dates.

☐ One can (try to) choose the data set such that cash flows are at the same points in time and the matrix $\mathbf{C}$ is not entirely full of zeros.

☐ Nevertheless, the regression still has big problems!

☐ There are as many parameters as cash flow dates, and there is nothing to smooth the discount curve found from the regression.

☐ Thus, the discount factors of similar maturity can be very different.

☐ An alternative and preferable methodology would be to estimate a **parametrized smooth** yield curve from the market rates (coupon bond prices).

☐ Indeed, in the next sections we focus on *cubic splines, cubic B-splines* and on the *Exponential-Polynomial class of curves* (Nelson and Siegel (1987) family and on the Svensson (1994) generalization).

2.2.3 Splines (*see Munk (2008) and Filipovic (2009)*)

2.2.3.1 Cubic splines

- □ In this and in the following sections we will consider methods to estimate the **entire discount function** $T \mapsto B(t,T)$ (up to some large T). For ease of exposition we will adopt the notation $B(t,t+h)$ with $h \geq 0$. A similar notation is adopted for yields and forward rates: $R(t,t+h)$ and $R(t,t+\tau,t+\tau+h)$.

- □ We will assume that the discount function be described by a parametric function with parameter values estimated in such a way to get a close match between the observed and theoretical bond prices.

□ The methodology presented in this section is a version of the cubic splines approach introduced by McCulloch (1971) and then modified by McCulloch (1975) and Litzenberger and Rolfo (1984).

□ Let us imagine to observe, at date t, M bonds with maturity dates $T_1 \leq T_2 \leq \ldots \leq T_M$. Now, we divide the time-to-maturity axis into subintervals defined by the *knot points* $0 = h_0 < h_1 < \ldots < h_k = T_M - t$.

□ A spline approximation of the discount function $B(t, t+h)$ is based on an expression like:

$$B(t, t+h) = \sum_{j=0}^{k-1} G_j(h) \mathbb{I}_j(h)$$

☐ The $G_j(h)$'s are basis functions, and the $\mathbb{I}_j(h)$ are step functions:

$$\mathbb{I}_j(h) = \begin{cases} 1 \text{ if } h \geq h_j \\ 0 \text{ otherwise}. \end{cases}$$

☐ In other words, we have:

$$\begin{aligned} B(t, t+h) &= G_0(h) \text{ for } h \in [h_0, h_1), \\ &= G_0(h) + G_1(h) \text{ for } h \in [h_1, h_2), \\ &= G_0(h) + G_1(h) + \ldots + G_j(h) \text{ for } h \in [h_j, h_{j+1}), \\ & \text{etc.} \end{aligned}$$

☐ We require the $G_j(h)$'s functions to be continuous and twice differentiable and ensuring a smooth transition in the knot points h_j.

☐ A **polynomial spline** is a spline with polynomials as basis functions. Let us consider a **cubic spline** where:

$$G_j(h) = \alpha_j + \beta_j(h - h_j) + \gamma_j(h - h_j)^2 + \delta_j(h - h_j)^3\,,$$

where $\alpha_j\,, \beta_j\,, \gamma_j$ and δ_j are constants.

☐ For $h \in [0, h_1)$, we have:

$$B(t, t + h) = \alpha_0 + \beta_0 h + \gamma_0 h^2 + \delta_0 h^3\,,$$

since $B(t,t) = 1 \;\Rightarrow\; \alpha_0 = 1$.

☐ For $h \in [h_1, h_2)$, we have:

$$\begin{aligned} B(t, t + h) = {} & (1 + \beta_0 h + \gamma_0 h^2 + \delta_0 h^3) \\ & + [\alpha_1 + \beta_1(h - h_1) + \gamma_1(h - h_1)^2 + \delta_1(h - h_1)^3]\,. \end{aligned}$$

☐ To guarantee a smooth transition from the first to the second subinterval, that is at the knot $h = h_1$, we impose:

$$\begin{aligned}
\lim_{h \to h_1^-} B(t, t+h) &= \lim_{h \to h_1^+} B(t, t+h) = B(t, t+h_1) \\
\lim_{h \to h_1^-} B'(t, t+h) &= \lim_{h \to h_1^+} B'(t, t+h) < \infty \\
\lim_{h \to h_1^-} B''(t, t+h) &= \lim_{h \to h_1^+} B''(t, t+h) < \infty \,.
\end{aligned}$$

☐ The first condition ensures the continuity at the point $h = h_1$. The second condition guarantee that $B(t, t+h)$ has no kink at h_1 and the second one impose a further degree of smoothness around h_1.

☐ These three conditions imply (exercise!) : $\alpha_1 = 0$, $\beta_1 = 0$ and $\gamma_1 = 0$. Iterating for the other knots, we similarly find $\alpha_j = \beta_j = \gamma_j = 0$ for all $j = 1, \ldots, k-1$.

☐ The cubic spline is therefore reduced to:

$$B(t, t+h) = (1 + \beta_0 h + \gamma_0 h^2 + \delta_0 h^3) + \sum_{j=1}^{k-1} \delta_j (h - h_j)^3 \mathbb{I}_j(h) \,.$$

☐ Let $t_1, t_2, \ldots, t_m$ denote the time distance between the date t and each of the payment dates in the set of all payment dates of the bonds we observe. Let $C_{i,j}$ denote the payment of bond i in t_j periods.

☐ From the no-arbitrage relation linking coupon and zero-coupon bond prices we have:

$$CB_i(t, t+h_i) = \sum_{j=1}^{m} C_{i,j} B(t, t+t_j) \,,$$

where $CB_i(t, t+h_i) =$ observed market price,

and $B(t, t+t_j) =$ unknown discount factors.

□ Since not all ZCBs $B(t, t+t_j)$ involved in the relation are traded in the market, we will allow for a deviation $\varepsilon_t^{(i)}$ so that:

$$CB_i(t, t+h_i) = \sum_{j=1}^{m} C_{i,j} B(t, t+t_j) + \varepsilon_t^{(i)} \,.$$

□ We assume that $\varepsilon_t^{(i)} \sim IIN(0, \sigma^2)$ for all i, and that $\varepsilon_t^{(i)} \perp \varepsilon_{t+k}^{(j)}$ for all $i \neq j$ and $k \in \mathbb{R}$.

□ We write, for any $i \in \{1, \ldots, M\}$, the following regression:

$$CB_i(t, t+h_i) = \sum_{j=1}^{m} C_{i,j} \left\{ (1 + \beta_0 t_j + \gamma_0 t_j^2 + \delta_0 t_j^3) + \sum_{\ell=1}^{k-1} \delta_l (t_j - h_\ell)^3 \mathbb{I}_\ell(t_j) \right\} + \varepsilon_t^{(i)} \,.$$

- Given the observed prices and payment schemes of the M bonds, the $k+2$ parameters $\theta = (\beta_0, \gamma_0, \delta_0, \delta_1, \ldots, \delta_{k-1})'$ can be estimated by Ordinary Least Squares (OLS).

- The estimated market discount function is therefore given by:

$$B(t, t+h) = (1 + \widehat{\beta}_0 h + \widehat{\gamma}_0 h^2 + \widehat{\delta}_0 h^3) + \sum_{j=1}^{k-1} \widehat{\delta}_j (h - h_j)^3 \mathbb{I}_j(h)\,,$$

where $\widehat{\theta} = (\widehat{\beta}_0, \widehat{\gamma}_0, \widehat{\delta}_0, \ldots, \widehat{\delta}_{k-1})'$ denotes the OLS estimates.

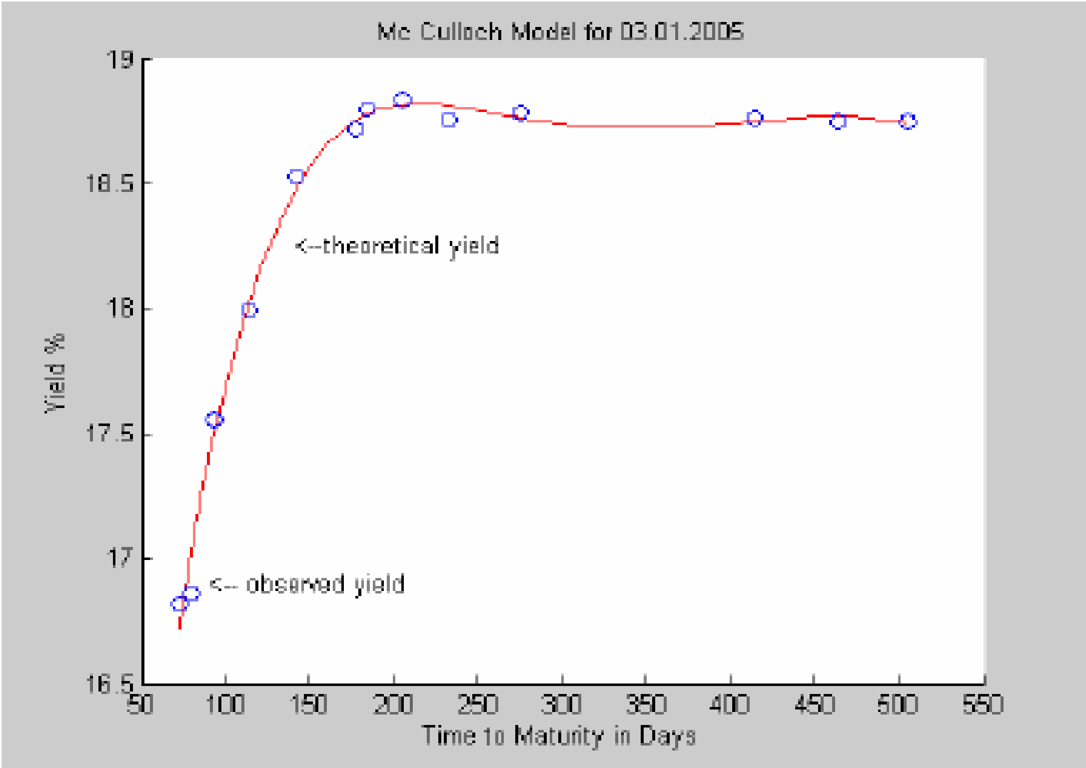
- How the number of subintervals k and the knot points h_j are chosen ?

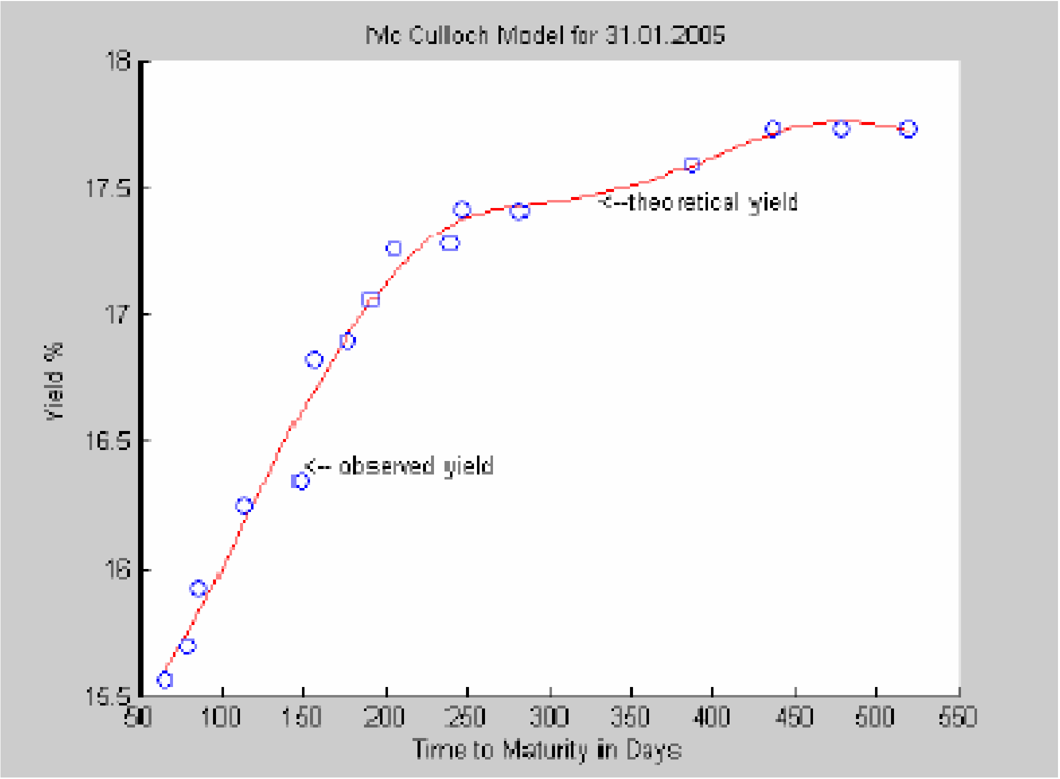
□ Following McCulloch (1971, 1975), let us assume that k be the nearest integer to $\sqrt{M}$ and the knot points be defined by:

$$h_j = T_{\tau_j} + \omega_j (T_{\tau_{j+1}} - T_{\tau_j}) \,, \text{ where } \tau_j = \left\lfloor j \times \tfrac{M}{k} \right\rfloor \,, \; \omega_j = j \times \tfrac{M}{k} - \tau_j \,,$$

and in particular $h_k = T_M$.

□ Alternatively, the knot points can be placed at (for instance) 1 year, 5 years and 10 years to maturity, so that the intervals broadly fit the short-term, intermediate-term and long-term segments of the bond market.


Mc Culloch Model for 03.01.2005
Yield %
Time to Maturity in Days
<--theoretical yield
<-- observed yield
19
18.5
18
17.5
17
16.5
50
100
150
200
250
300
350
400
450
500
550



Mc Culloch Model for 31.01.2005
18
17.5
17
16.5
16
15.5
50
100
150
200
250
300
350
400
450
500
550
Yield %
Time to Maturity in Days
<--theoretical yield
<-- observed yield

2.2.3.2 Cubic B-splines

- In the previous section we have seen that the cubic spline is represented by a collection of nodes $h_j \in \{h_1, \ldots, h_k\}$, and by interval specific coefficients $\beta_0, \gamma_0, \delta_0, \delta_1, \ldots, \delta_{k-1}$. The space of cubic splines with nodes on a prescribed grid comprise a finite vector space.

- A much-more-widely used form of splines, forming a basis for cubic splines, are the cubic B-spline. This methodology require the additional specification of three knots below h_0 and three knots above h_k, giving us :

$$h_{-3} < h_{-2} < h_{-1} < h_0 < h_1 < \ldots < h_k < h_{k+1} < h_{k+2} < h_{k+3}\,.$$

☐ The $k+3$ cubic B-splines are:

$$\psi_j(h) = \sum_{a=j}^{j+4} \left(\prod_{b=j, b\neq a}^{j+4} \frac{1}{h_b - h_a} \right) (h - h_a)^3 \mathbb{I}_a(h)\,, \quad j \in \{-3, \ldots, k-1\}\,.$$

☐ The B-spline $\psi_j(h) \geq 0$ in the interval $[h_j, h_{j+4}]$ and zero outside.

☐ Using the above specified B-splines, we can write the ZCB price as:

$$B(t, t+h) = z_1 \psi_1(h) + \ldots + z_s \psi_s(h)\,,$$

$$s = k+3\,,$$

☐ and assuming m ZCB prices $B(t, t+h_i)$ with $i \in \{1, \ldots, m\}$:

$$b(z) = \begin{pmatrix} B(t, t+h_1) \\ \vdots \\ B(t, t+h_m) \end{pmatrix} = \begin{pmatrix} \psi_1(h_1) & \ldots & \psi_s(h_1) \\ \vdots & \ddots & \vdots \\ \psi_1(h_m) & \ldots & \psi_s(h_m) \end{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ z_s \end{pmatrix} =: \Psi z\,,$$

☐ we can write the following linear optimization problem in terms of observed coupon bond prices:

$$\min_{z \in \mathbb{R}^s} ||\mathbf{P} - \mathbf{C} \times \Psi z||^2$$

☐ If the $(M \times s)$ matrix $\mathbf{A} = \mathbf{C} \times \Psi$ has full rank, the unique unconstrained solution is:

$$z^* = (\mathbf{A'A})^{-1}\mathbf{A}'\mathbf{P}\,.$$

☐ Nevertheless, we have the constraint:

$$B(t,t) = z_1\psi_1(0) + \ldots + z_s\psi_s(0) = 1\,.$$

□ *Example [see Chapter 3 in Filipovic (2009)] :* Let us take market prices for UK gilts, 04/09/96

	Coupon(%)	Next coupon	Maturity date	Dirty price
Bond 1	10.00	15/11/96	15/11/96	103.82
Bond 2	9.75	19/01/97	19/01/98	106.04
Bond 3	12.25	26/09/96	26/03/99	118.44
Bond 4	9.00	03/03/97	03/03/00	106.28
Bond 5	7.00	06/11/96	06/11/01	101.15
Bond 6	9.75	27/02/97	27/08/02	111.06
Bond 7	8.50	07/12/96	07/12/05	106.24
Bond 8	7.75	08/03/97	08/09/06	98.49
Bond 9	9.00	13/10/96	13/10/08	110.87

☐ The actual date t is the 04/09/96. The coupon payments are semiannual. The day-count convention is actual/365.

☐ In terms of our notation : we have $M = 9$ bonds, and the payment dates are $T_1 = 26/09/96$, $T_2 = 13/10/96$, $T_3 = 06/11/97$, ...

☐ Bond 1 has 1 payment date, Bond 2 has 3 payment dates, Bond 3 has 6 payment dates .... In particular, we have:

$$m = 1 + 3 + 6 + 7 + 11 + 12 + 19 + 20 + 25 = 104\,.$$

☐ The cash flow matrix $\mathbf{C} = (C)_{ij}$ is (9×104).

☐ Let us assume to select the following knot points:

$$\{-20, -5, -2, 0, 1, 6, 8, 11, 15, 20, 25, 30\}$$

thus, $h_{-3} = -20$, $h_{-2} = -5$, ..., $h_0 = 0$, ..., $h_8 = 30$, and therefore we are going to use $s = k + 3 = 8$ B-splines for the estimation.

☐ The estimation with the 8 B-splines leads to:

$$\min_{z \in \mathbb{R}^8} ||\mathbf{P} - \mathbf{C} \times \Psi z||^2 = ||\mathbf{P} - \mathbf{C} \times \Psi z^*||^2 = 0.23$$

☐ Now, if we use only 5 B-splines with the 9 knot points:

$$\{-10, -5, -2, 0, 4, 15, 20, 25, 30\}$$

☐ the estimation leads to:

$$\min_{z \in \mathbb{R}^5} ||\mathbf{P} - \mathbf{C} \times \Psi z||^2 = ||\mathbf{P} - \mathbf{C} \times \Psi z^*||^2 = 0.39$$

☐ There is a trade-off between the regularity (degree smoothness) of the estimated curve and its ability to fit the data.

Limits :

- □ The discount function estimated *via* cubic splines usually has a realistic shape (over the maturities and over time) for maturities less that the longest one used in the data set.

- □ Even if there is nothing in the approach explicitly imposing to the discount function to be positive and decreasing as far as the residual maturity increases, this will almost always be the case.

- □ Nevertheless, as $h \to +\infty$, the cubic spline $B(t, t + h) \to \infty$.

- Estimating the discount function $B(t, t+h)$ leads to **unstable** and **irregular** yield and forward curves.

- The problems are typically observed over short and long maturities of the curve.

- In particular : $i)$ the zero-coupon rates $R(t, t+h)$ will often increase or decrease significantly for maturities $h \to T_M$; $ii)$ the derived forward rate curve will typically be quite *non-smooth* near the knot points, and the curve shape tends to be quite sensitive to the observed bond prices and the location of the knot points.

☐ From the previous example, and the above mentioned limits of the cubic spline and cubic B-spline methodology to extract the discount function (and, thus, the yield and forward rate curve), we observe that we **need methodologies**

a) able to extract smooth yield and forward curves for any maturity

b) that do not fluctuate "too much" (unrealistically)

c) and realistically flatten toward the long end.

$\rightarrow$ it seems advisable to directly estimate the yield or forward curve is such a way to satisfy a), b) and c).

$\Rightarrow$ *Exponential-Polynomial Families.*

2.2.4 Exponential-Polynomial Families (*see Cairns (2004) and Filipovic (2009)*)

☐ The purpose of this section is to introduce the parametric curve families which are used by most central banks to construct the term structure of interest rates.

☐ The main feature is to capture (to fit) as much as possible the structure of market interest rates with a small number of parameters.

☐ These parametric families are typically defined for the instantaneous forward rate $f(t, \tau)$.

□ The **instantaneous forward rate** is defined as:

$$f(t,\tau) = \mathrm{lim}_{T\to\tau} R(t,\tau,T)\,,$$

and the function $\tau \mapsto f(t,\tau)$ is called the **term structure of instantaneous forward rates** or the **instantaneous forward rate curve**.

□ From:

$$f(t,\tau) = -\frac{\partial \ln B(t,\tau)}{\partial\tau} = -\frac{\partial B(t,\tau)/\partial\tau}{B(t,\tau)}\,,$$

→ we get:

$$B(t,\tau) = \exp\left[- \textstyle\int_t^\tau f(t,u)du\right] \quad \Rightarrow R(t,\tau) = \frac{1}{\tau - t}\int_t^\tau f(t,u)du\,.$$

2.2.4.1 The Exponential-Polynomial class for forward rates

□ This class of forward-rate curves [see Björk and Christensen (1999)] is defined as a combination of polynomials and exponentials with different rates of decay.

□ This class of curves is defined as follows:

$$G(x) = p_0(x) + \sum_{i=1}^{K} p_i(x)\,\mathrm{exp}[-\alpha_i x]\,,$$

$$\text{with } \alpha_i \in R_+ \;\; \forall i \in \{0, \ldots, K\}\,.$$

and where $p_i(x)$ is any polynomial with $deg(p_i) \leq n_i \; \forall i \in \{0, \ldots, K\}$.

□ Writing the polynomial $p_i(x)$ as:

$$p_i(x) = \sum_{j=0}^{n_i} z_{i,j} x^j \,, \ \ \forall i \in \{0, \dots, K\}$$

we observe that $p_i(x)$ is determined by its $(n_i + 1)$-dimensional vector of coefficients $z_i = (z_{i,0}, \dots, z_{i,n_i})'$.

□ The entire exponential family is therefore specified by the mapping $G(x) = G(z, \alpha, x)$ with $\alpha = (\alpha_1, \dots, \alpha_K)' \in R_+^K$ and $z = (z_1, \dots, z_K)$.

□ Let us now identify, for any date t, the class of instantaneous forward rate curves $\tau \mapsto f(t, \tau)$ with the class of curves $G(\tau - t)$ introduced above.

☐ For ease of notation, let $\tau = t + h$ with $h > 0$. Then we have:

$$f(t, t+h) = p_{0,t}(h) + \sum_{i=1}^{K} p_{i,t}(h) \exp[-\alpha_{i,t} h] \,,$$

$$\text{where } \; p_{i,t}(h) = \sum_{j=0}^{n_i} z_{i,j,t} h^j \,.$$

☐ Well known special classes are: **`Nelson and Siegel (1987)`**

$$f(t, t+h) = z_{0,0,t} + [z_{1,0,t} + z_{1,1,t} h] \exp[-\alpha_{1,t} h] \,.$$

☐ **`Svensson (1994)`**

$$f(t, t+h) = z_{0,0,t} + [z_{1,0,t} + z_{1,1,t} h] \exp[-\alpha_{1,t} h] + z_{2,1,t} h \exp[-\alpha_{2,t} h] \,.$$

☐ `Wiseman (1994)`

$$f(t, t+h) = z_{0,0,t} + \sum_{i=1}^{k} z_{i,0,t} \,\mathsf{exp}[-\alpha_{i,t} h]\,.$$

☐ `Cairns (1998), the restricted exponential model`

$$f(t, t+h) = z_{0,0,t} + \sum_{i=1}^{k} z_{i,0,t} \,\mathsf{exp}[-\alpha_{i} h]\,.$$

☐ In these models, the "dependence on t" of parameters $z_{i,j,t}$ and $\alpha_{i,t}$ simply indicates that they are re-estimated at each date t (using date t bond market prices) in order to build the date t forward curve.

2.2.4.2 The Nelson and Siegel (1987) family

□ We have seen that Nelson and Siegel (1987) specify the forward rate curve in the following way:

$$f(t, t+h) = z_{0,0,t} + [z_{1,0,t} + z_{1,1,t}h] \, \exp[-\alpha_{1,t}h] \, .$$

□ Let us now take their original parametrization:

$$f(t, t+h) = \beta_0 + \beta_1 \, e^{(-h\,/\,\theta)} + \beta_2 \, \frac{h}{\theta} \, e^{(-h\,/\,\theta)} \, , \;\; \theta > 0 \, .$$

where $(\beta_0, \beta_1, \beta_2, \theta)'$ are parameters (applying to all residual maturities h) to be estimated at each date t.

□ Interpretation of the parameters:

$$\theta = \text{scale (or location) parameter}\,, \;\; \beta_0 = \text{level parameter}\,,$$

$$\beta_1 = \text{slope parameter}\,, \;\; \beta_2 = \text{curvature parameter}\,,$$

$i)$ The role of $B_0(h) = \beta_0$ $(\beta_0 \neq 0)$: mostly determine **long-term forward rates**

indeed, if $h \to +\infty$ (and $\theta > 0$) then $f(t, t+h) \to f(t, t+\infty) = \beta_0\,,$

which is a long-term forward rate

$ii.1)$ The role of $B_1(h) = \beta_1\, e^{(-h\,/\,\theta)}$ $(\beta_1 \neq 0)$:

if $h \to 0$ then $f(t, t+h) \to r(t) = \beta_0 + \beta_1$ =the risk-free rate

$\Rightarrow$ $r(t) - f(t, t+\infty) = \beta_1\,, \Rightarrow$ it drives the spread between

the spot rate and the long-term forward rate.

$ii.2)$ The term $B_1(h) = \beta_1\, e^{(-h\,/\,\theta)}$ mostly affects **short-term forward rates** :

$$B_1(h)|_{h=0} = \beta_1\,, \;\text{ and }\; \lim_{h\to+\infty} B_1(h) = 0\,.$$

$iii.1)$ The role of $B_2(h) = \beta_2 \dfrac{h}{\theta} e^{(-h\,/\,\theta)}$ $(\beta_2 \neq 0)$:

$$B_2(h)|_{h=0} = 0\,, \text{ and } \lim_{h\to+\infty} B_2(h) = 0\,;$$

$$\frac{dB_2(h)}{dh} = \frac{1}{\theta} e^{-h/\theta} \left[\beta_2 \left(1 - \frac{h}{\theta}\right)\right] \;\; \forall h > 0\,; \;\; \frac{dB_2(h)}{dh} = 0 \;\; \Leftrightarrow \;\; h = \theta\,,$$

$$\frac{dB_2(h)}{dh} > 0 \text{ iff } \beta_2 > 0 \text{ and } h < \theta\,; \text{or } \beta_2 < 0 \text{ and } h > \theta\,.$$

$iii.2)$ The term $B_2(h)$ mostly affects, compared to $B_1(h)$, the **medium-term forward rates**.

$iv)$ The first role of θ : it is indicated as a **scale** parameter, given that it measures the rate at which short-term and medium term components decay to zero.

$v.1$) The second role of θ : for fixed (β_1, β_2), it determines the **location** of the (admitted) "hump" in the in the forward rate curve (if it exists). Indeed:

$$\frac{df(t,t+h)}{dh} = \frac{1}{\theta}e^{-h/\theta}\left[\beta_2\left(1-\frac{h}{\theta}\right)-\beta_1\right]\ ;\ \ \frac{df(t,t+h)}{dh} = 0 \ \Leftrightarrow\ h = \theta\frac{\beta_2-\beta_1}{\beta_2}\,;$$

$v.2$) There exists $h^* > 0$ such that $\dfrac{df(t,t+h^*)}{dh} = 0$ if and only if:

$$\beta_2 > 0 \text{ and } \beta_2 > \beta_1\,;\ \text{or } \beta_2 < 0 \text{ and } \beta_2 < \beta_1\,;$$

$v.3$) For given (β_1, β_2) such that exists h^* making $\dfrac{df(t,t+h^*)}{dh} = 0$, this residual maturity h^* increases as far as θ increases. This is the reason why we also indicate θ as a location parameter for the "hump" (if it exists). If we assume $\theta = h$ $(B_2(h)' = 0)$, we have $\frac{d}{dh}f(t,t+h) = 0$ iff $\beta_1 = 0$ (that is, $r(t) = f(t,t+\infty)$).

☐ From the above interpretation, we observe that the term structure is determined by THREE FACTORS :

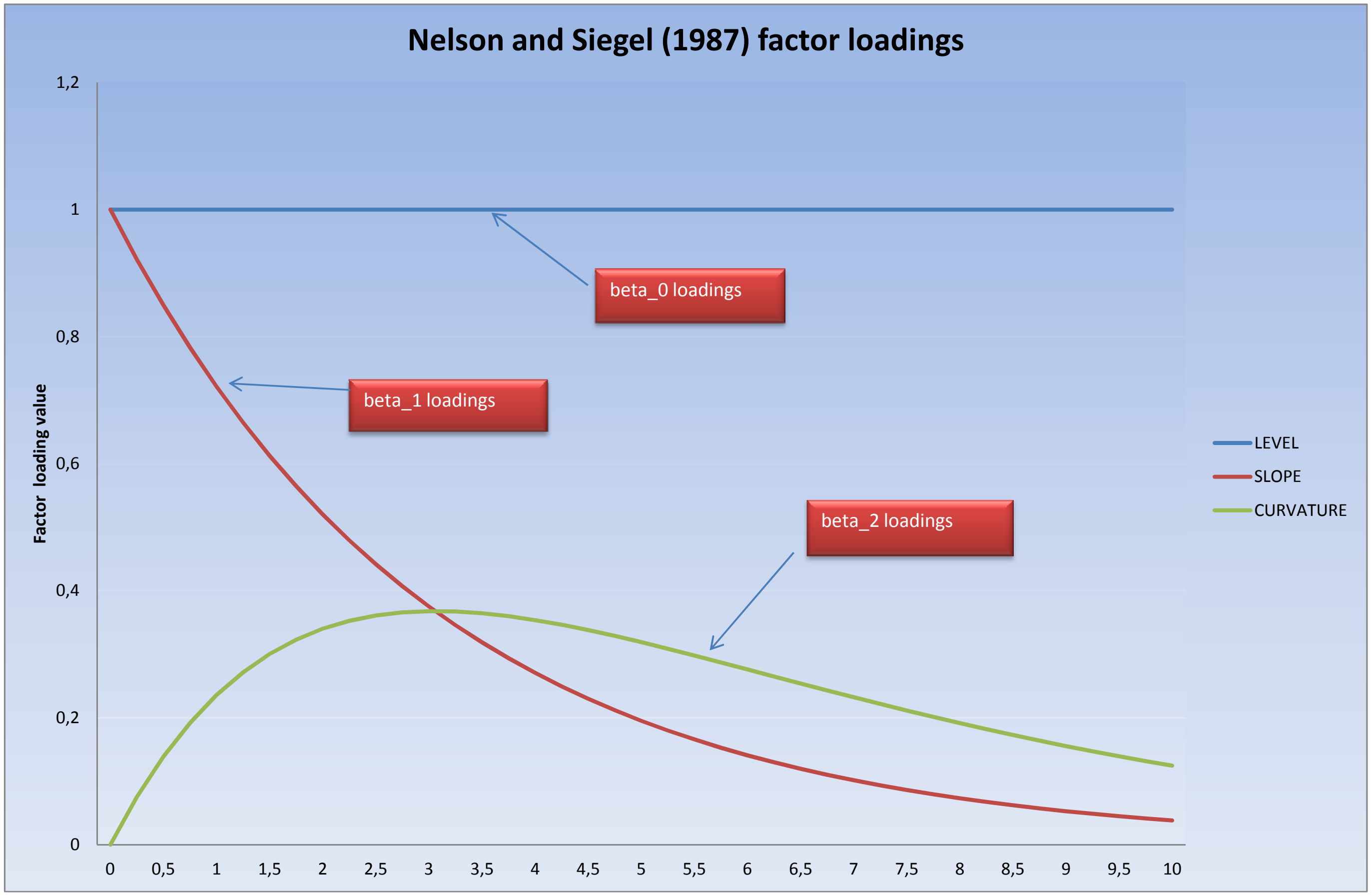
$$B_0(h) = \beta_0 \text{ can be seen as a } \textit{Level Factor},$$

$$B_1(h) = \beta_1\, e^{(-h\,/\,\theta)} \text{ as a } \textit{Slope Factor},$$

$$B_2(h) = \beta_2\, \frac{h}{\theta}\, e^{(-h\,/\,\theta)} \text{ as a } \textit{Curvature Factor}, .$$

☐ The Nelson and Siegel (1987) yield curve $R(t, t+h)$ is given by:

$$\begin{aligned} R(t, t+h) &= \frac{1}{h}\int_0^h f(t, t+u)du\,, \\ &= \beta_0 + (\beta_1 + \beta_2)\frac{1 - e^{(-h\,/\,\theta)}}{h\,/\,\theta} - \beta_2\, e^{(-h\,/\,\theta)}\,, \\ &= a + b\frac{1 - e^{(-h\,/\,\theta)}}{h\,/\,\theta} + c\, e^{(-h\,/\,\theta)}\,. \end{aligned}$$



Nelson and Siegel (1987) factor loadings
Factor loading value
1,2
1
0,8
0,6
0,4
0,2
0
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
5
5,5
6
6,5
7
7,5
8
8,5
9
9,5
10
beta_0 loadings
beta_1 loadings
beta_2 loadings
LEVEL
SLOPE
CURVATURE

□ We can also write:

$$R(t,t+h) \;=\; \beta_0 + \beta_1 \frac{1 - e^{(-h\,/\,\theta)}}{h\,/\,\theta} + \beta_2 \frac{1 - \left(1 + \frac{h}{\theta}\right) e^{(-h\,/\,\theta)}}{h\,/\,\theta}$$

□ The discount function $B(t,t+h)$ is:

$$B(t,t+h) \;=\; \exp\left[-a\,h - b\,\theta(1 - e^{(-h\,/\,\theta)}) - c\,h\,e^{(-h\,/\,\theta)}\right] .$$

□ We write, for any $i \in \{1,\ldots,M\}$, the following regression:

$$CB_i(t,t+h_i) = \sum_{j=1}^{m} C_{i,j} \left\{\exp\left[-a\,h_i - b\,\theta(1 - e^{(-h_i\,/\,\theta)}) - c\,h_i\,e^{(-h_i\,/\,\theta)}\right]\right\} + \varepsilon_t^{(i)} .$$

$\rightarrow$ and parameters are estimated by non-linear regression techniques

[see Gallant (1987)].

2.2.4.3 The Svensson (1994) family

☐ Svensson (1994) generalizes the Nelson-Siegel curve adding a third term. Precisely, he assumes:

$$f(t, t+h) = \beta_0 + \beta_1\, e^{(-h\,/\,\theta_1)} + \beta_2\, \frac{h}{\theta_1}\, e^{(-h\,/\,\theta_1)} + \beta_3\, \frac{h}{\theta_2}\, e^{(-h\,/\,\theta_2)}\,.$$

☐ As for Nelson and Siegel (1987), we have:

$$f(t, t) = r(t) = \beta_0 + \beta_1\,, \;\; f(t, t+\infty) = \beta_0$$

☐ but, the presence of $B_3(h) = \beta_3\, \frac{h}{\theta_2}\, e^{(-h\,/\,\theta_2)}$ allows for two “humps” in the forward curve thanks to θ_1 and θ_2.

□ This is important, given that we frequently observe two humps:

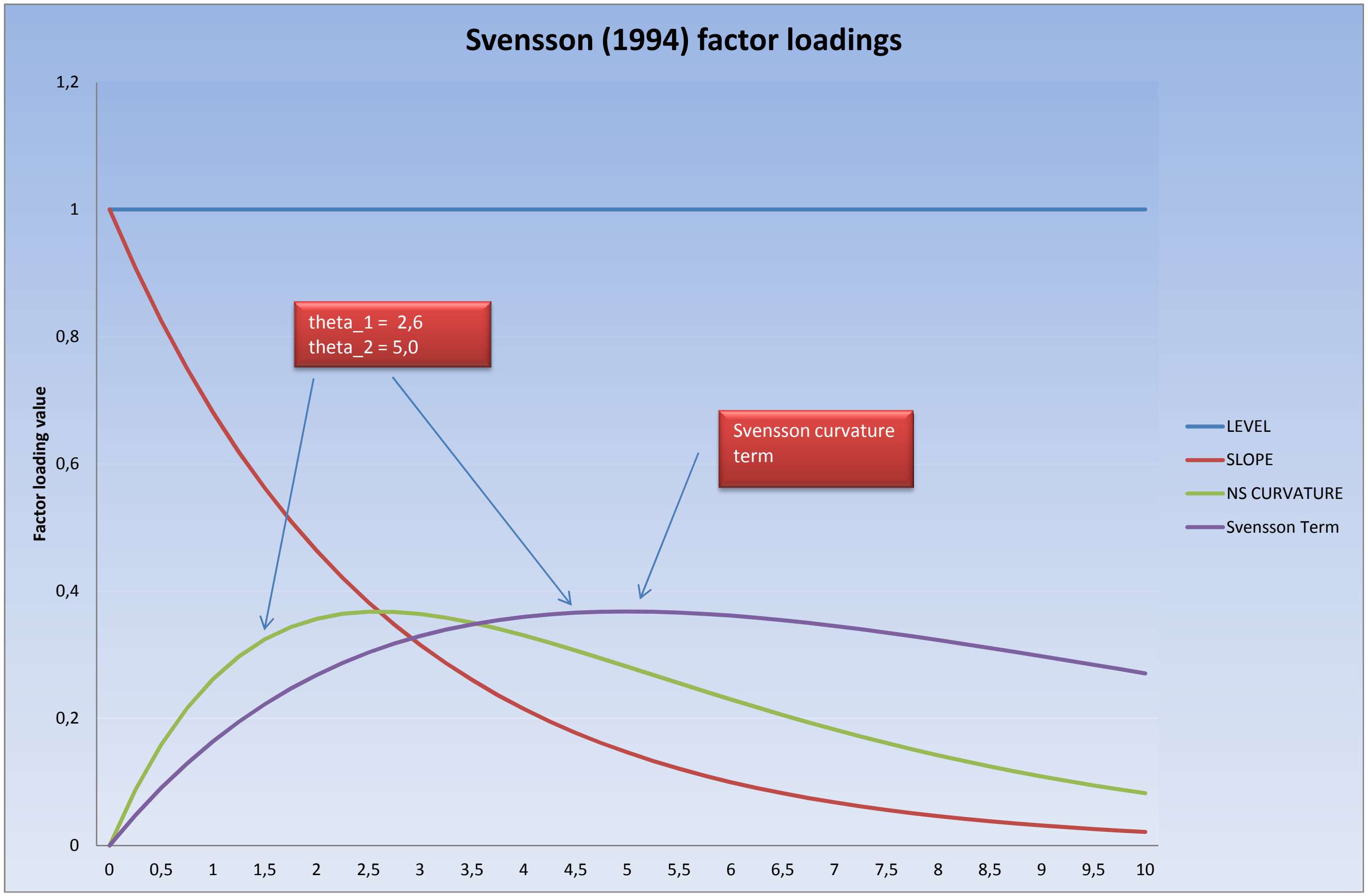
the first at short maturities associated to monetary policy expectations,

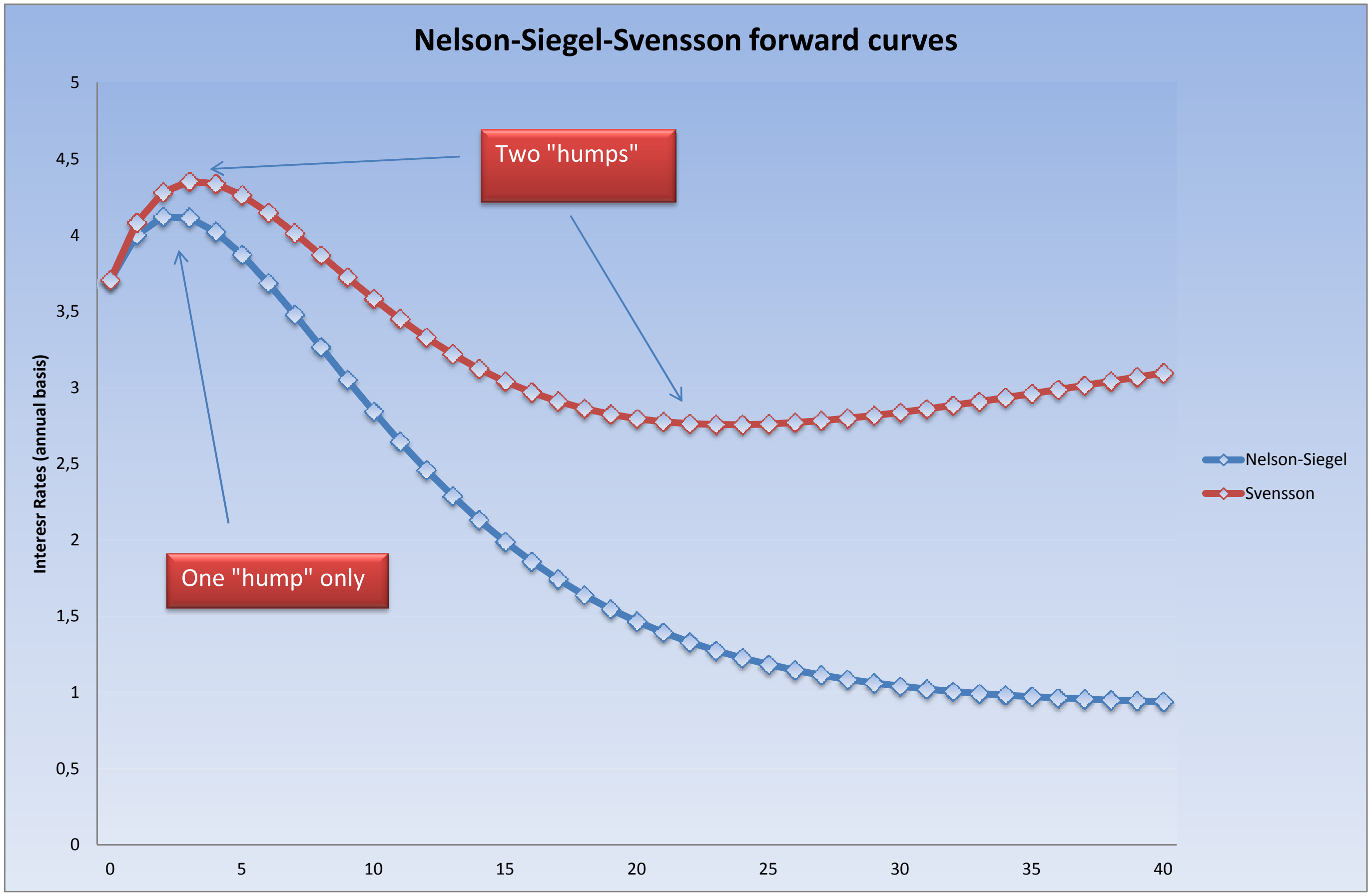
the second one at long maturities to catch convexity effects.

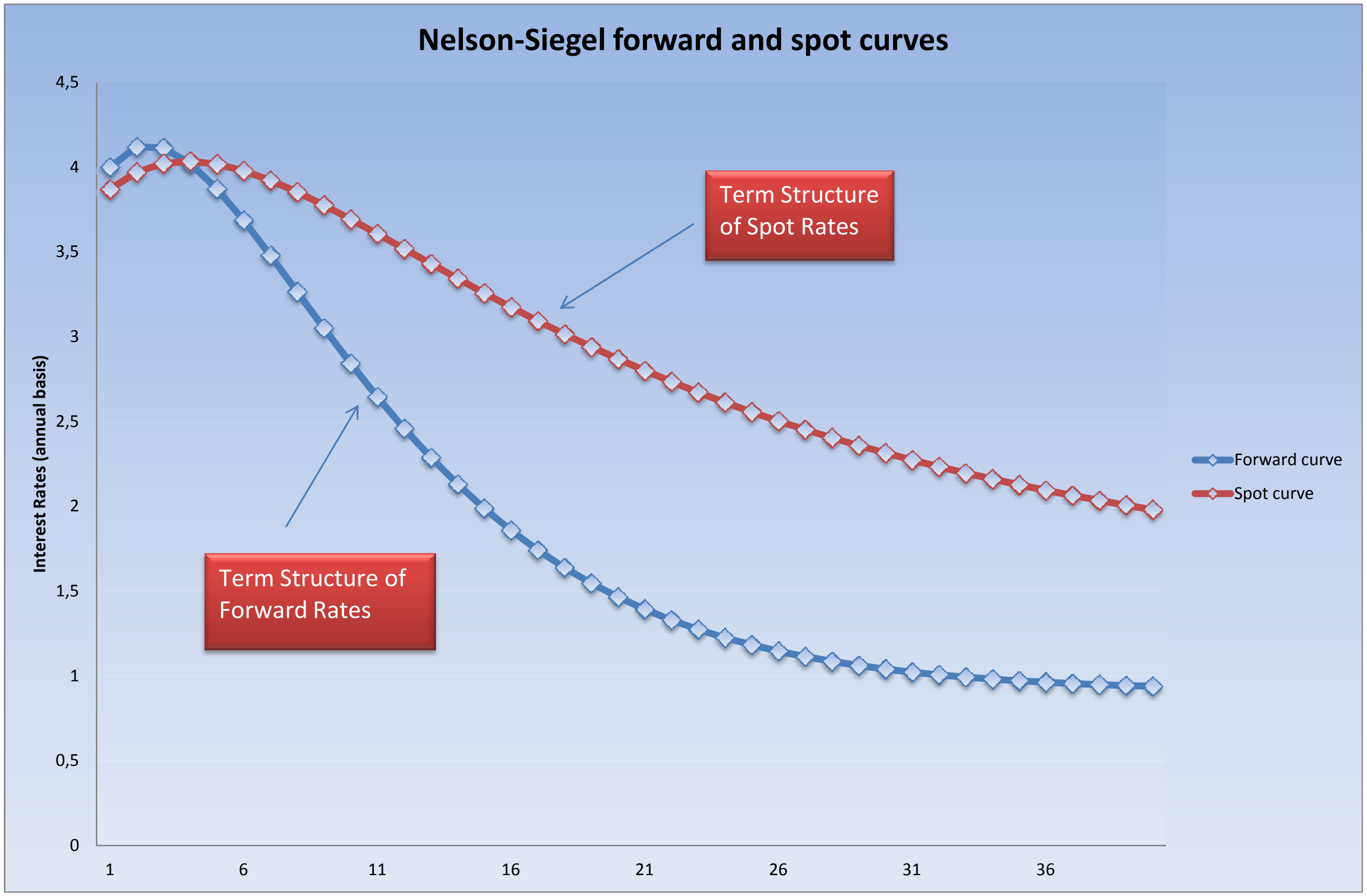
□ The new parameters β_3 and θ_2 provide additional flexibility, compared with the Nelson-Siegel curve, and in particular over short and medium maturities.

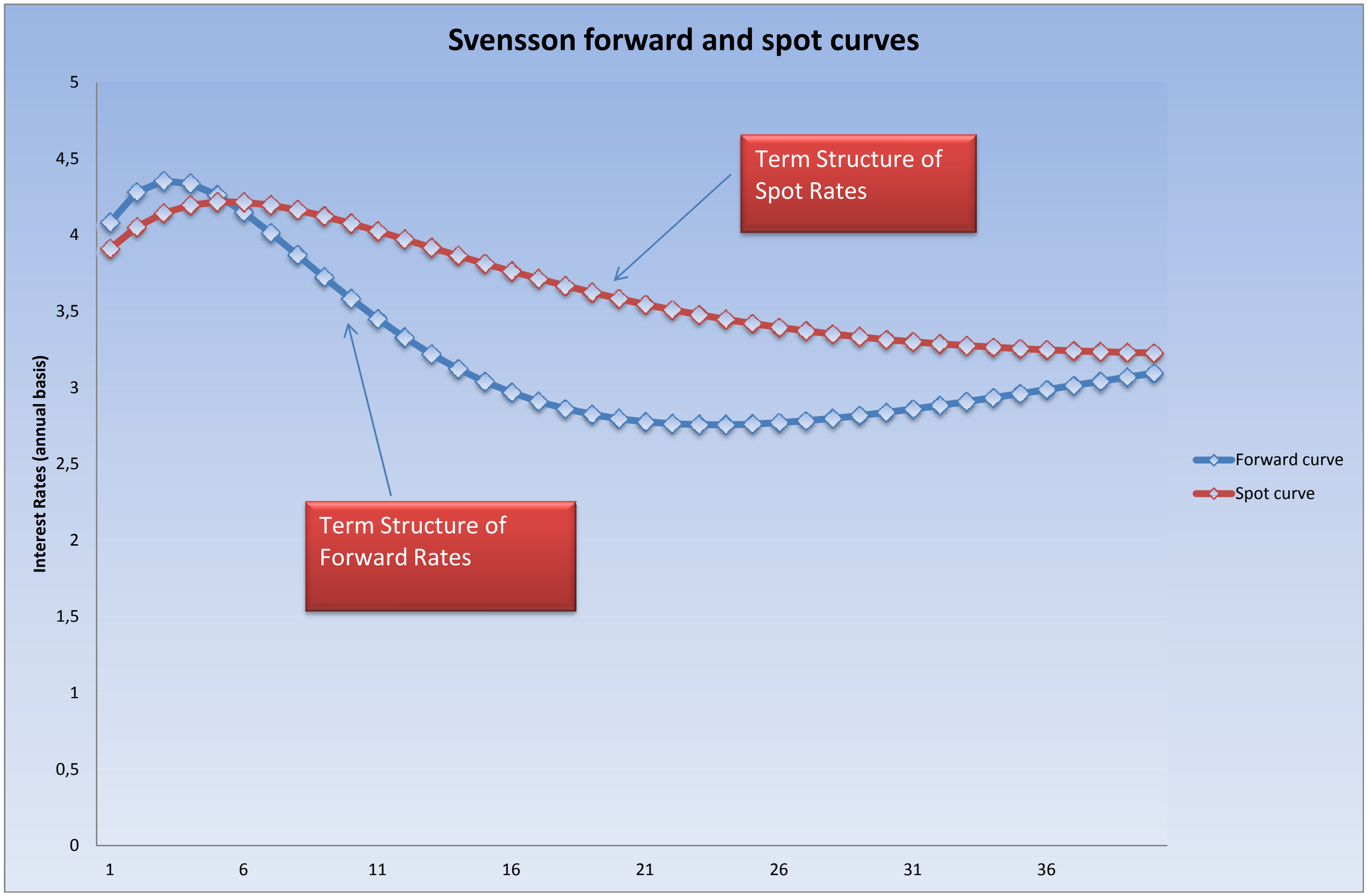
□ The Svensson (1994) yield curve $R(t, t+h)$ is given by:

$$R(t, t+h) = \beta_0 + \beta_1 \frac{1 - e^{(-h/\theta_1)}}{h/\theta_1} + \beta_2 \frac{1 - \left(1 + \frac{h}{\theta_1}\right) e^{(-h/\theta_1)}}{h/\theta_1} + \beta_3 \frac{1 - \left(1 + \frac{h}{\theta_2}\right) e^{(-h/\theta_2)}}{h/\theta_2}.$$

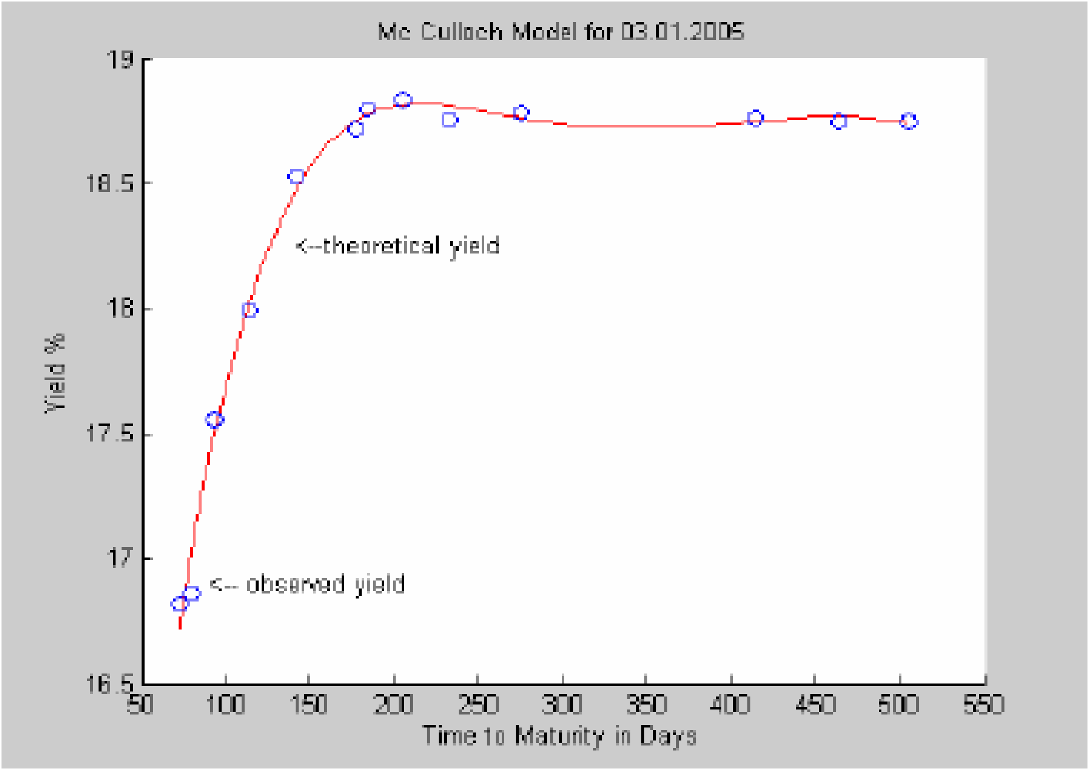

Svensson (1994) factor loadings
Factor loading value
1,2
1
0,8
0,6
0,4
0,2
0
0
0,5
1
1,5
2
2,5
3
3,5
4
4,5
5
5,5
6
6,5
7
7,5
8
8,5
9
9,5
10
theta_1 = 2,6
theta_2 = 5,0
Svensson curvature term
LEVEL
SLOPE
NS CURVATURE
Svensson Term

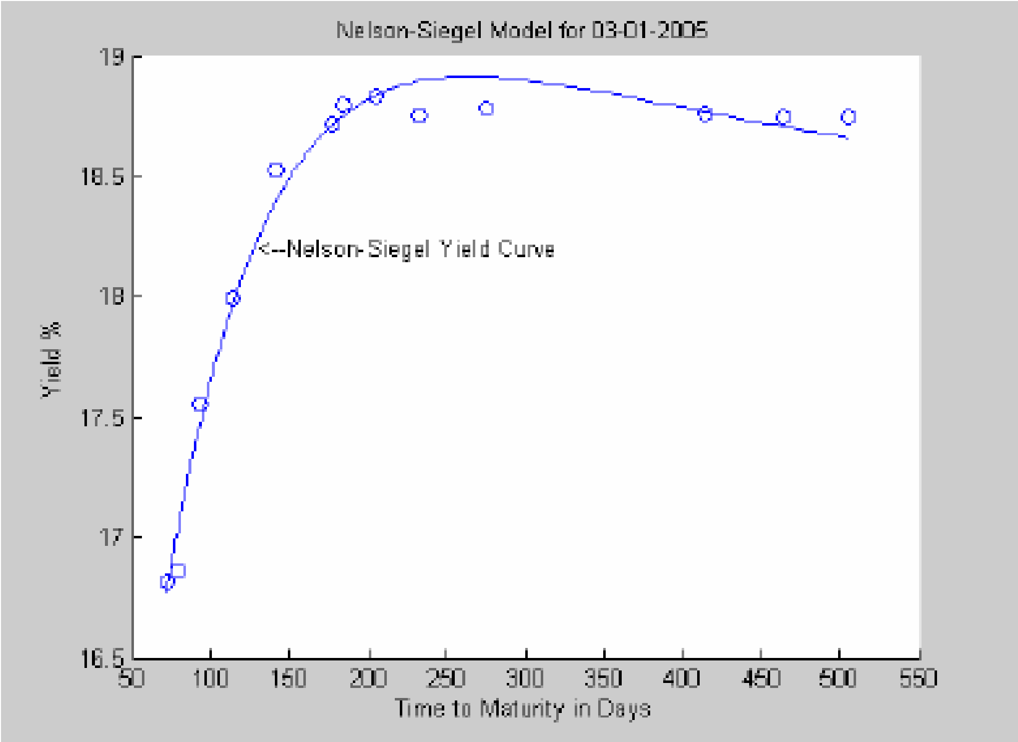

Nelson-Siegel-Svensson forward curves
Interesr Rates (annual basis)
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
0
5
10
15
20
25
30
35
40
Two "humps"
One "hump" only
Nelson-Siegel
Svensson

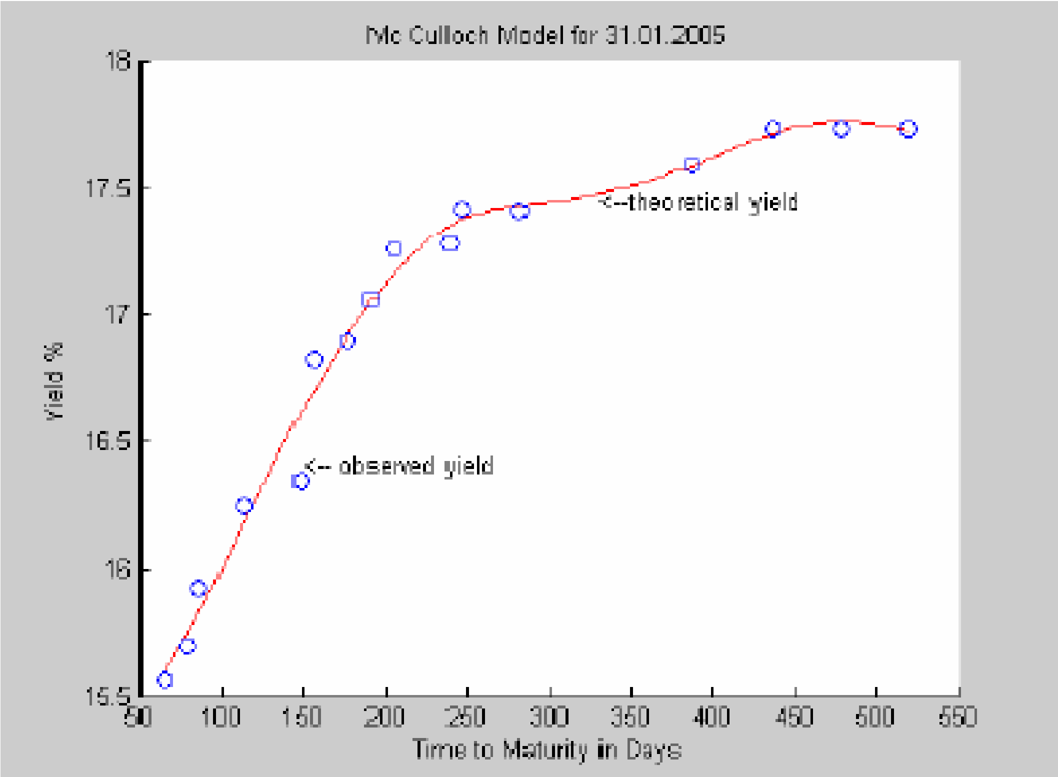

Nelson-Siegel forward and spot curves
Interest Rates (annual basis)
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
1
6
11
16
21
26
31
36
Term Structure of Spot Rates
Term Structure of Forward Rates
Forward curve
Spot curve



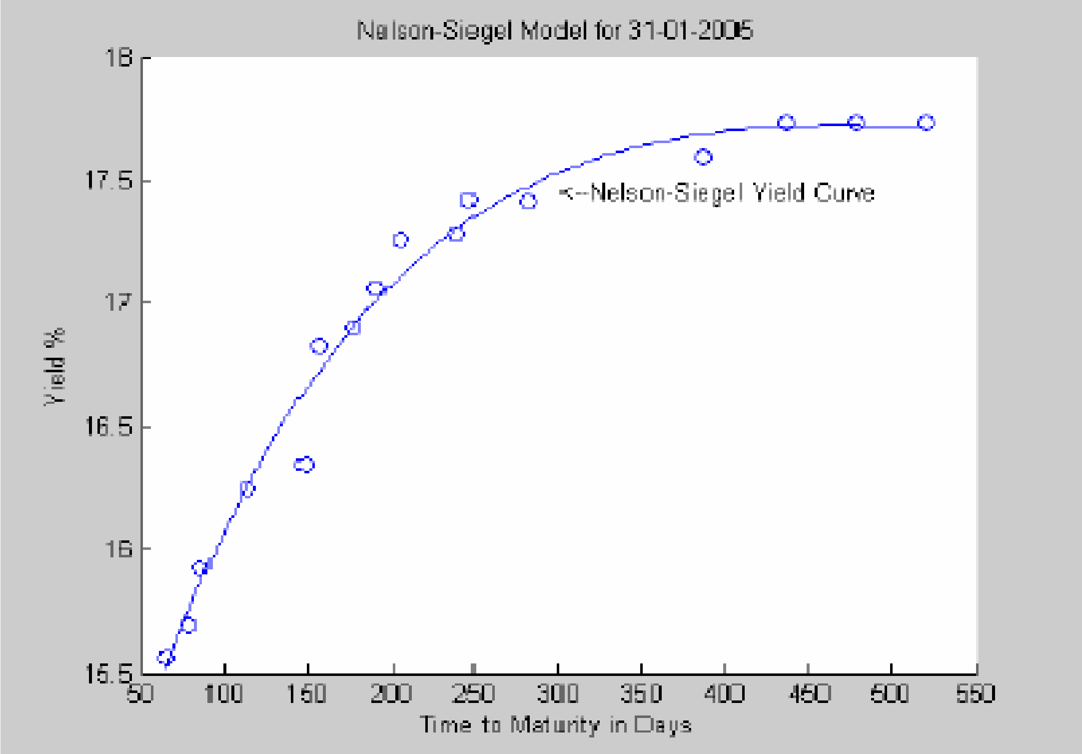
Svensson forward and spot curves
Interest Rates (annual basis)
5
4,5
4
3,5
3
2,5
2
1,5
1
0,5
0
1
6
11
16
21
26
31
36
Term Structure of Spot Rates
Term Structure of Forward Rates
Forward curve
Spot curve


Mc Culloch Model for 03.01.2005
Yield %
Time to Maturity in Days
<--theoretical yield
<-- observed yield
19
18.5
18
17.5
17
16.5
50
100
150
200
250
300
350
400
450
500
550


Nelson-Siegel Model for 03-01-2005
19
18.5
18
17.5
17
16.5
Yield %
<--Nelson-Siegel Yield Curve
50
100
150
200
250
300
350
400
450
500
550
Time to Maturity in Days



Mc Culloch Model for 31.01.2005
Yield %
Time to Maturity in Days
<--theoretical yield
<-- observed yield
18
17.5
17
16.5
16
15.5
50
100
150
200
250
300
350
400
450
500
550


Nelson-Siegel Model for 31-01-2005
<--Nelson-Siegel Yield Curve
Yield %
Time to Maturity in Days
18
17.5
17
16.5
16
15.5
50
100
150
200
250
300
350
400
450
500
550

2.2.4.4 The Gurkaynak, Sack and Wright (2007) data base on nominal yields

□ Among the several U.S. yield curve data bases characterizing the empirical studies in the fixed income literature [Fama and Bliss (1987), McCulloch and Kwon (1993)], recently Gurkaynak, Sack and Wright (2007) have proposed daily U.S. yield curve estimates.

□ The sample period is from 1961 to the present : it is updated regularly and is available (for free!) at

`http://www.federalreserve.gov/Pubs/feds/2006/200628/200628abs.html`

`http://www.federalreserve.gov/econresdata/researchdata.htm`

- The methodology they use to estimate (daily!) the forward curve $f(t, t+h)$ is using the Svensson (1994) family, that is a generalization of the Nelson and Siegel (1987) family.

- The estimation is based on observations of only T-notes and T-bonds prices (maturities, at the issuing date, ranging from 2 years to 30 years).

□ At any date t in the sample, they thus exclude:

- all T-bills : their market is disconnected (segmented) from that of T-notes and T-bonds.

- all T-notes and T-bonds with residual maturity less than 3 months : they are highly affected by lack of liquidity of short-term investors.

- the two most recently issued T-notes and T-bonds: they are called "on-the-run" and "first off-the-run". They are excluded because of their greater liquidity.

☐ This paper tells us (among others) that we have to select a "coherent/homogeneous" set of bond prices in order to build a reliable term structure of interest rates. The bonds have to:

- denominated in the same currency;
- of the same credit quality (default-free in our case!)
- reflect the same level of liquidity
- not be affected by option like features (callable bonds)

☐ Ideally, the bonds should ideally differ in terms of their maturity and coupons.

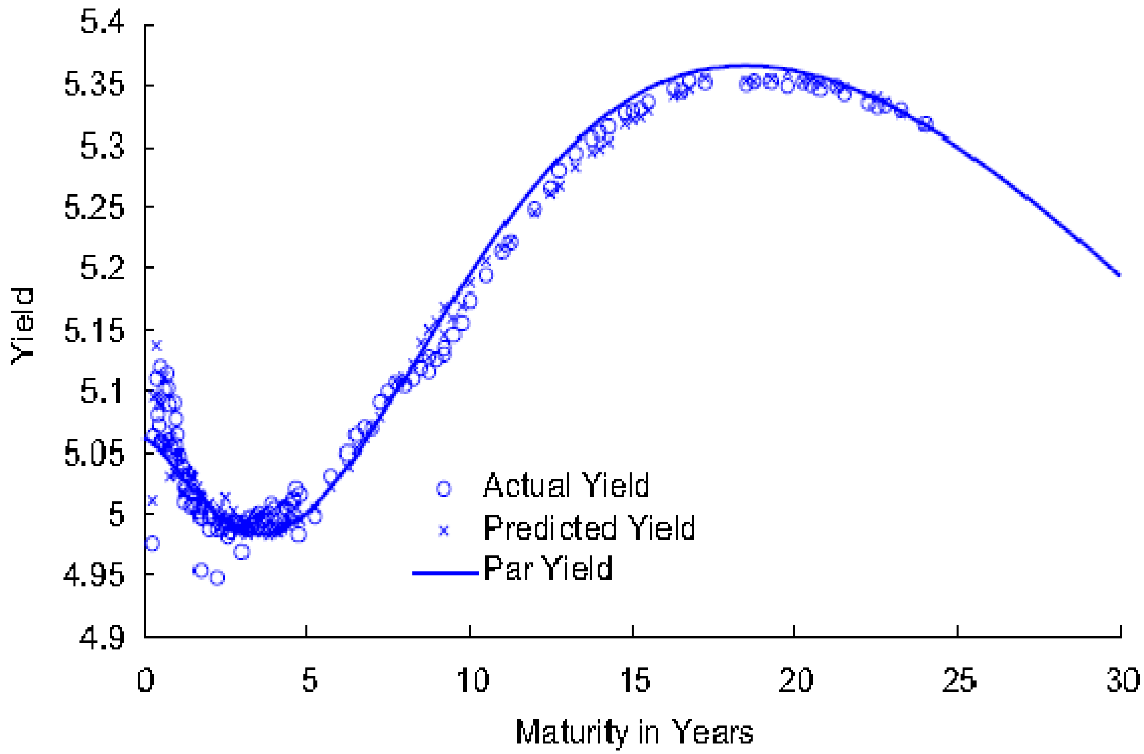
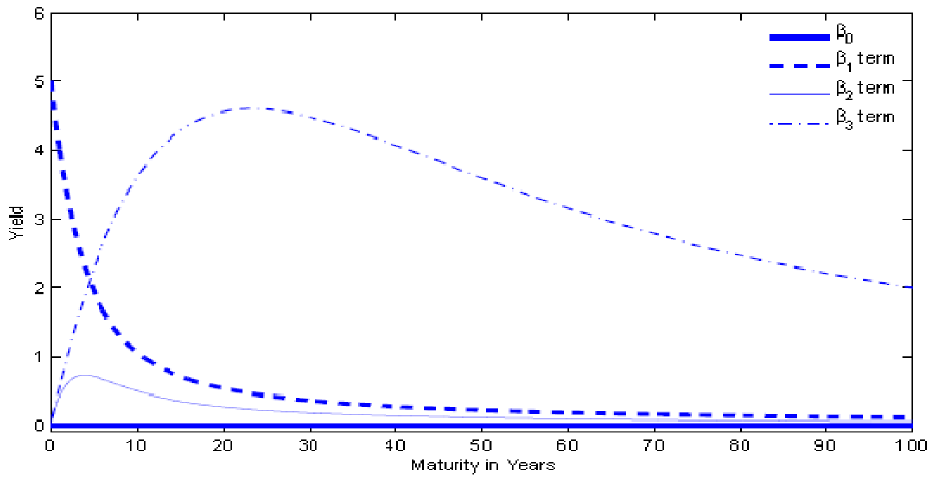


Fig. 1. Par-yield curve on May 9, 2006.

Figure 5: Decomposition of the Yield Curve on May 3, 2006



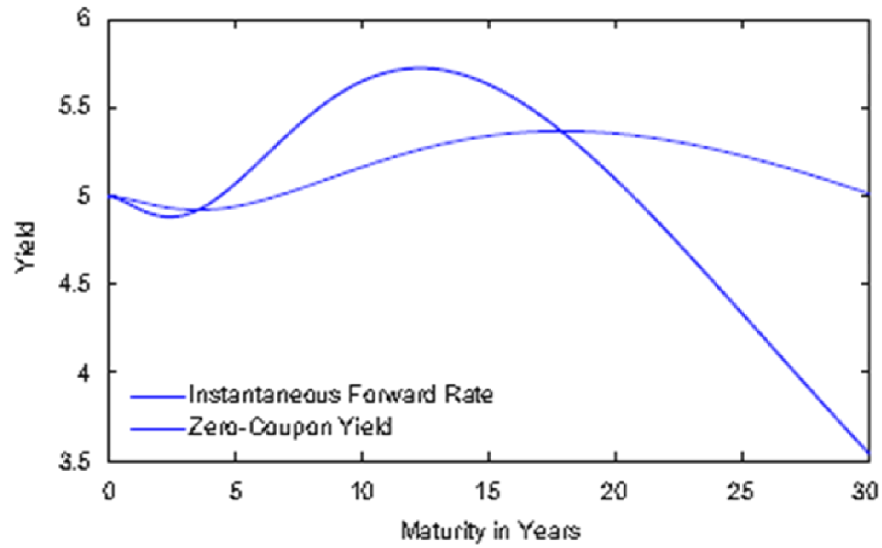


Fig. 3. Zero-coupon yield curve and forward rates on May 9, 2006.

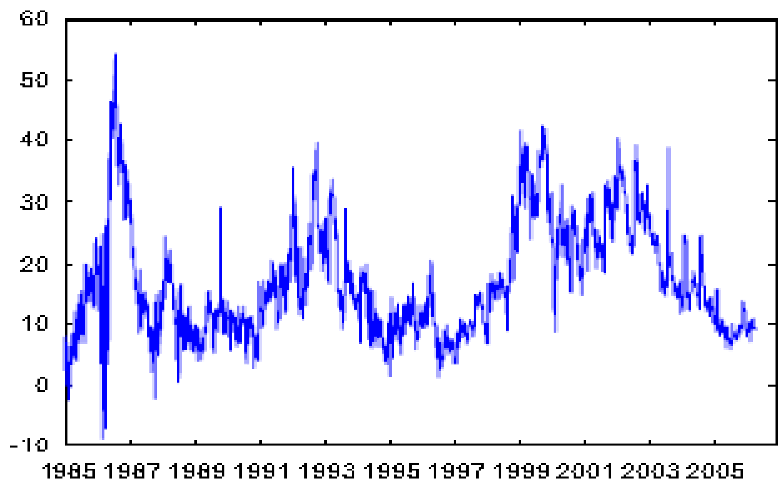


Fig. 5. Premium for the on-the-run 10-year Treasury note.

- The Nelson-Siegel-Svensson methodology is widely used in practice (central banks, in particular)

- The **main feature** is that they are **parsimonious**:

  - small number of parameters (compared to splines);

  - the curves we fit (forward curve) and then we derive (spot curve) are smooth over all maturities and over time.

- Obviously the **main limit** is the **lack of flexibility**: these methods cannot replicate all the shapes we cross in reality.

2.2.4.5 The Gurkaynak, Sack and Wright (2010) TIPS yield curve data base

- They estimate the Nelson-Siegel-Svensson yield curve on TIPS from the start of 1999 to the present.

- It is updated regularly and is available (again for free!) at

  `http://www.federalreserve.gov/econresdata/researchdata.htm`

- As mentioned in Lecture 1, comparison (difference) with the corresponding nominal yield curve allows measures of inflation compensation (or break-even inflation rates) to be computed.

Figure 2: Par TIPS Yield Curve on June 1, 2005

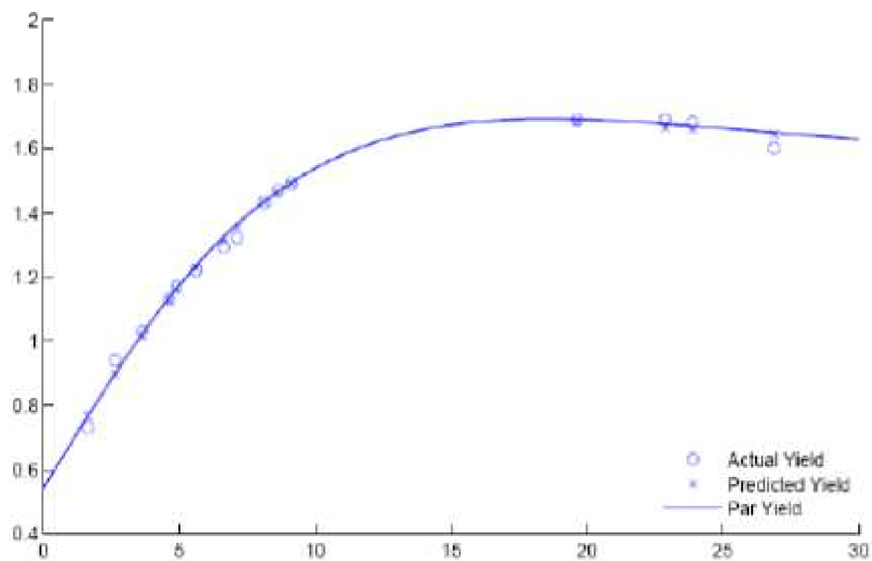


Figure 4: Zero-Coupon and Forward Rates on June 1, 2005

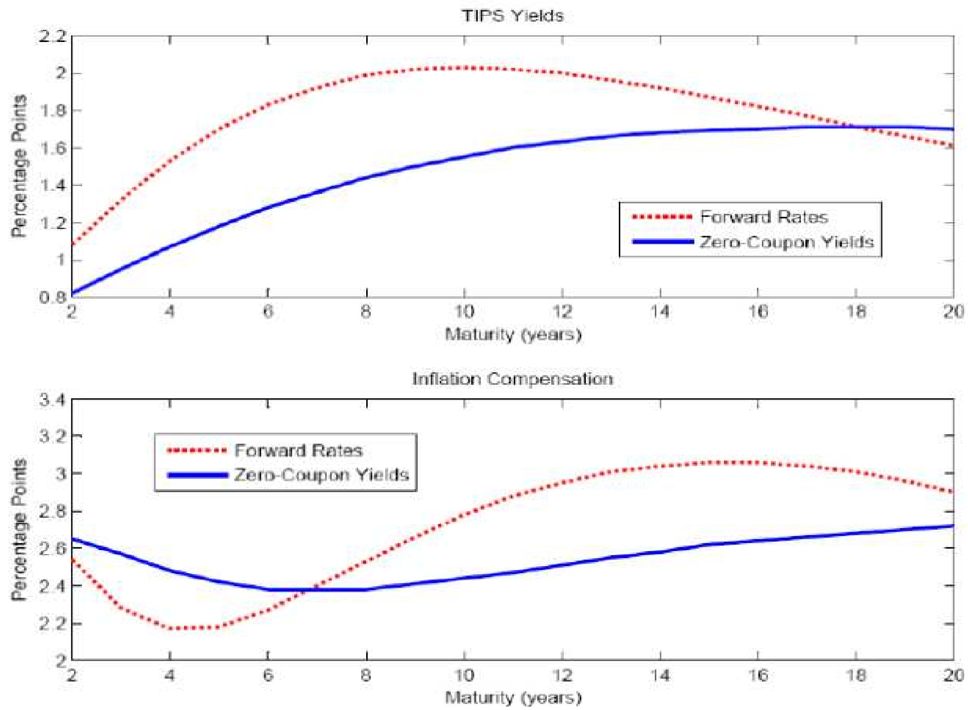


Figure 5: Zero Coupon Yields: TIPS and Nominal

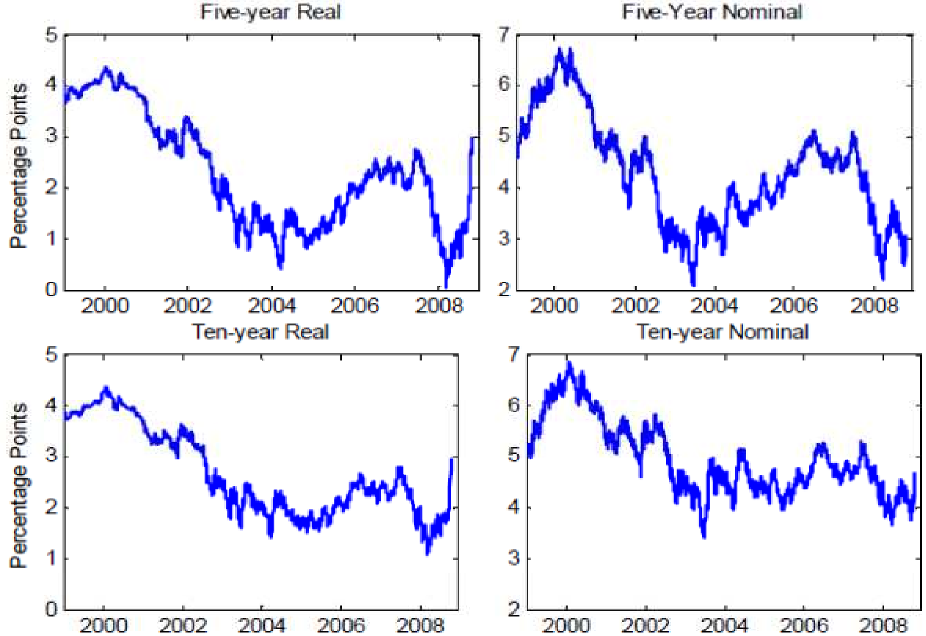
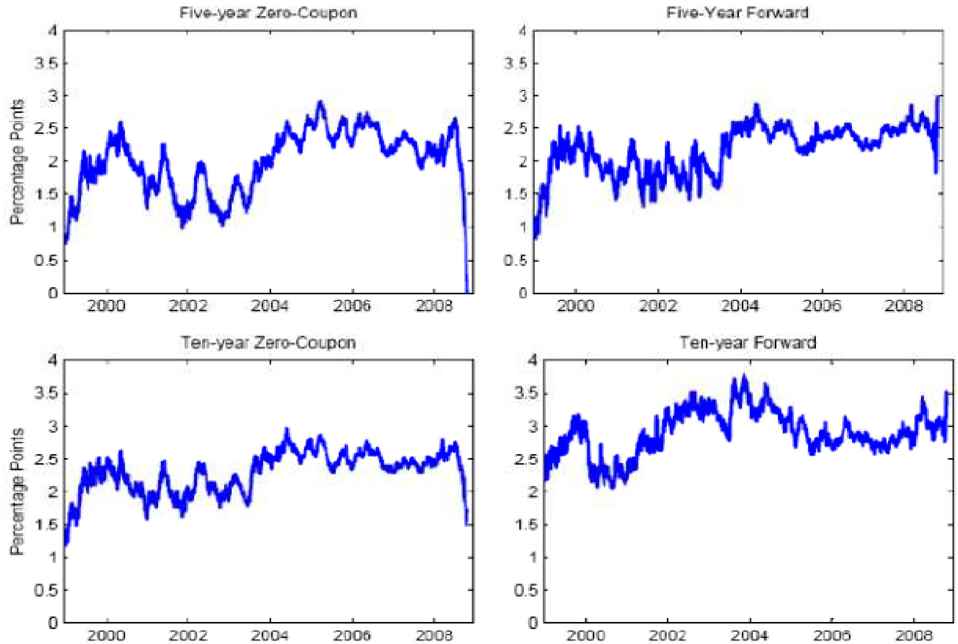


Figure 6: Inflation Compensation



2.2.5 The Principal Component Analysis of the Yield Curve

- When we have presented the Nelson and Siegel (1987) forward rate curve, we have seen that is was possible to identify **THREE FACTORS** determining the possible shapes:

$$B_0(h) = \beta_0 \text{ can be seen as a } \textit{Level Factor},$$

$$B_1(h) = \beta_1 \, e^{(-h\,/\,\theta)} \text{ as a } \textit{Slope Factor},$$

$$B_2(h) = \beta_2 \, \frac{h}{\theta} \, e^{(-h\,/\,\theta)} \text{ as a } \textit{Curvature Factor}, .$$

- Are these **THREE DRIVING FORCES** of the term structure of interest rates model dependent or model independent ?

□ In other words :

– are these three factors simply arbitrarily assumed, from the beginning, by the model [the Nelson and Siegel (1987) model in our case]

– or are they really present in (suggested by) the interest rates data ?

□ The answer to that question is extremely important, given that it leads to understand how many factors a term structure model should incorporate in order to be close to the observations (i.e. well specified).

□ The well known methodology followed by the term structure literature is the **Principal Component Analysis (PCA)**.

2.2.5.1 Principal Component Analysis

- The Principal Component Analysis (PCA) is a well known dimension reduction technique in multivariate analysis.

- It is used to extract the factors explaining most of the variability of the multivariate phenomenon we observe : a p-dimensional random process $X_t = (X_{1,t}, \ldots, X_{p,t})'$ observed over time.

- PCA produces a lower dimensional description of $(X_t)_{t=1,2,\ldots}$ searching for linear combinations of X_t with the largest variances [see Härdle and Simar (2003) : *Applied Multivariate Statistical Analysis*, Springer].

☐ The key mathematical principles behind the PCA are the **Spectral Decomposition theorems** of linear algebra.

The Jordan Decomposition Theorem (JD) : *Each symmetric* $(p \times p)$ *matrix* $\mathcal{A}$ *can be written as:*

$$\mathcal{A} = \Gamma \Lambda \Gamma' = \sum_{j=1}^{p} \lambda_j \gamma_j \gamma_j' \,, \text{ where } \Lambda = diag(\lambda_1, \dots, \lambda_p) \text{ contains the}$$

p *eigenvalues of the matrix* $\mathcal{A}$*, and where* $\Gamma = (\gamma_1, \dots, \gamma_p)$ *is an orthogonal matrix* ($\Gamma^{-1} = \Gamma'$*, i.e.* $\gamma_i' \gamma_j = 0 \; \forall i \neq j$ *with* $||\gamma_j|| = 1$) *where the* j^{th} *column is the* j^{th} *eigenvector* γ_j *of* $\mathcal{A}$

Useful application : if $\mathcal{A} = \Gamma \Lambda \Gamma'$ then, for any $\alpha \in \mathbb{R}$, we have $\mathcal{A}^{\alpha} = \Gamma \Lambda^{\alpha} \Gamma'$ where $\Lambda^{\alpha} = diag(\lambda_1^{\alpha}, \dots, \lambda_p^{\alpha})$.

Singular Value Decomposition Theorem (SVD) : Each $(n \times p)$ matrix $\mathcal{A}$, with $rank(\mathcal{A}) = r$, can be decomposed as:

$\mathcal{A} = \Gamma \Lambda \Delta'$, where Γ is $(n \times r)$ and Δ is $(p \times r)$.

Both Γ and Δ are column orthonormal, i.e. $\Gamma'\Gamma = \Delta'\Delta = I_r$.

$\Lambda = diag(\lambda_1^{1/2}, \ldots, \lambda_r^{1/2})$, $\lambda_j > 0$. The values $\lambda_1, \ldots, \lambda_r$ are the non-zero eigenvalues of the matrices $\mathcal{A}\mathcal{A}'$ and $\mathcal{A}'\mathcal{A}$.

Γ and Δ consist of the corresponding r eigenvectors of these matrices.

- **Principal Components of a Random Vector :** let us consider a $\mathbb{R}^p$-valued square-integrable random vector $X = (x_1, \ldots, x_p)'$ with mean vector $\mu = \mathbb{E}[X]$ and variance-covariance matrix $\Sigma = \mathbb{V}[X]$.

□ Since Σ is symmetric and positive semi-definite, then from JD we have that $\Sigma = \Gamma\Lambda\Gamma'$ with $\lambda_i \geq 0 \ \forall i \in \{1, \ldots, p\}$.

□ The **principal components transform** of X is defined as:

$Y = \Gamma'(X - \mu)$, which can be seen as a re-centering and rotation of X;

and each $Y_i = \gamma_i'(X - \mu)$ is the projection of $X - \mu$ onto

the i^{th} eigenvector γ_i of Γ. $Y_i = i^{th}$ *principal component of* X.

γ_i is also called the i^{th} vector of *loadings* of X.

- □ We thus obtain the decomposition $X = \mu + \Gamma Y = \mu + \sum_{j=1}^{p} \gamma_j Y_j$, and we have that $\mathbb{E}[Y] = 0$ and $\mathbb{V}[Y] = \Gamma'\Sigma\Gamma = \Gamma'(\Gamma\Lambda\Gamma')\Gamma = \Lambda$.

- □ This means that the principal components of X are uncorrelated and have variances $\mathbb{V}[Y_j] = \lambda_j \ \forall j$, which can be ordered from the largest to the smallest: $\mathbb{V}[Y_1] = \lambda_1 \geq \ldots \geq \mathbb{V}[Y_p] = \lambda_p \geq 0$.

- □ It can be shown (exercise!) that the 1^{st} principal component, Y_1, has maximal variance among all standardized linear combinations of X. That is:

$$\mathbb{V}[\gamma_1' X] = \max_{\{\delta:||\delta||=1\}} \{\mathbb{V}[\delta' X]\}$$

- For $j \in \{2, \ldots, p\}$, the j^{th} principal component Y_j can be shown to have maximal variance among all such linear combinations that are orthogonal to the first $(j-1)$ linear combinations.

- We also have that:

$$\sum_{j=1}^{p} \mathbb{V}[X_j] = trace(\boldsymbol{\Sigma}) = \sum_{j=1}^{p} \lambda_j = \sum_{j=1}^{p} \mathbb{V}[Y_j]$$

- The quantity:

$$\textstyle\sum_{j=1}^{k} \lambda_j / \sum_{j=1}^{p} \lambda_j$$

represents the amount of variability in X explained by the first k principal components $Y_1, \ldots, Y_k$.

- **Sample Principal Components :** Now let assume to observe our p-dimensional vector over time. Let us therefore denote with $x_t = (x_{1,t}, \dots, x_{p,t})'$ the date t observation (realization) of the random vector $X_t = (X_{1,t}, \dots, X_{p,t})'$, and let us assume to have T observations.

☐ We organize the T observations of each component $x_{i,t}$ in the following ($T \times p$) matrix: $\mathbf{X} = [x'_1, \dots, x'_T]$ where each rows $x'_j = (x_{1,j}, \dots, x_{p,j})'$ is a sample realization (observation) of the random vector X_t at date $t = j$.

☐ Let us assume that X_t is independently and identically distributed (i.i.d.) with $\mathbb{E}[X_t] = \mu$ and $\mathbb{V}[X_t] = \mathbf{\Sigma}$ $\forall t$.

□ We consider the empirical $(T \times T)$ covariance matrix:

$$\widehat{\Sigma}_{ij} = \widehat{\mathsf{Cov}}[X_{i,t}, X_{j,t}] = \frac{1}{T}\sum_{t=1}^{T}(x_{i,t} - \widehat{\mu}_i)(x_{j,t} - \widehat{\mu}_j)$$

where $\widehat{\mu} = (\widehat{\mu}_1, \ldots, \widehat{\mu}_p)' = \frac{1}{T}\sum_{t=1}^{T} x_t$ denotes the empirical mean.

□ Given that $\widehat{\Sigma}$ is positive semi-definite, the above PCA applies :

$\widehat{\Sigma} = \widehat{\Gamma}\widehat{\Lambda}\widehat{\Gamma}' = \sum_{j=1}^{p} \widehat{\lambda}_j \widehat{\gamma}_j \widehat{\gamma}_j'$, where $\widehat{\Lambda} = diag(\widehat{\lambda}_1, \ldots, \widehat{\lambda}_p)$ and where

$\widehat{\Gamma} = (\widehat{\gamma}_1, \ldots, \widehat{\gamma}_p)$ is an orthogonal matrix ($\widehat{\Gamma}^{-1} = \widehat{\Gamma}'$, i.e. $\widehat{\gamma}_i'\widehat{\gamma}_j = 0 \ \forall i \neq j$)

with j^{th} column the j^{th} eigenvector $\widehat{\gamma}_j$ of $\widehat{\Sigma}$.

- The first empirical principal component is the T-dimensional vector $\mathbf{y_1}$ given by:

$$\mathbf{y_1} = (\mathbf{X} - 1_T \widehat{\mu}')\widehat{\gamma}_1 \,.$$

- We thus can write the p empirical principal components in the following compact form:

$$\mathbf{Y} = (\mathbf{X} - 1_T \widehat{\mu}')\widehat{\Gamma}\,, \text{ where } \mathbf{Y} = (\mathbf{y_1}, \ldots, \mathbf{y_p}) \text{ is a } (T \times p) \text{ matrix}\,.$$

$$\text{Cov}[y_i, y_j] = \tfrac{1}{T}\textstyle\sum_{t=1}^{T} y_{i,t} y_{j,t} = \widehat{\lambda}_i \text{ iif } i = j\,, (= 0 \text{ otherwise})\,.$$

- Remember that the PCA is sensitive to scale changes: if we multiply one variable by a scalar we obtain different eigenvalues and eigenvectors. Thus the PCA should be applied to data having the same scale in each variable. Otherwise, we have to use the Normalized Principal Component Analysis (NPCA).

2.2.5.2 PCA of the yield curve

- ☐ Now, let us assume that our $(T \times p)$ data matrix $\mathbf{X} = [x'_1, \ldots, x'_T]$ contains at each row $x'_t = (x_{1,t}, \ldots, x_{p,t})'$ the sample realization (observation) of the following random vector containing, at date t, the increments of spot rate curve that is:

$$x_{i,t} = R(t, t + \tau_i)$$

with $\tau_i \in \{\tau_1, \ldots, \tau_p\}$ set of maturities.

- ☐ The first difference is also (frequently) taken (as a standard practice) given that interest rates are not i.i.d.. They are highly serial dependent, i.e. Cov$[R(t, t + \tau_i), R(t + k, t + \tau_i)] >> 0$ for $k \neq 0$ (and quite close to one!). Nevertheless, results do not change a lot.

☐ Now, let us apply the PCA to the US yields, observed quarterly from 1964:Q1 to 2009:Q4 [from the Gurkaynak, Sack and Write data base], with maturities 3 months, 6 months, 9 months, 1 year, 5 years and 10 years. We find:

PC	Explained Variance (%)	Cumulative Variance (%)	Eigenvalue
1	0.9592	0.9592	5.7554
2	0.0376	0.9968	0.2255
3	0.0025	0.9993	0.0148
4	0.0006	0.9999	0.0038
5	0.0001	1.0000	0.0005
6	0.0000	1.0000	5.02E-06

☐ The first 3 eigenvector associated to the first 3 PC are:

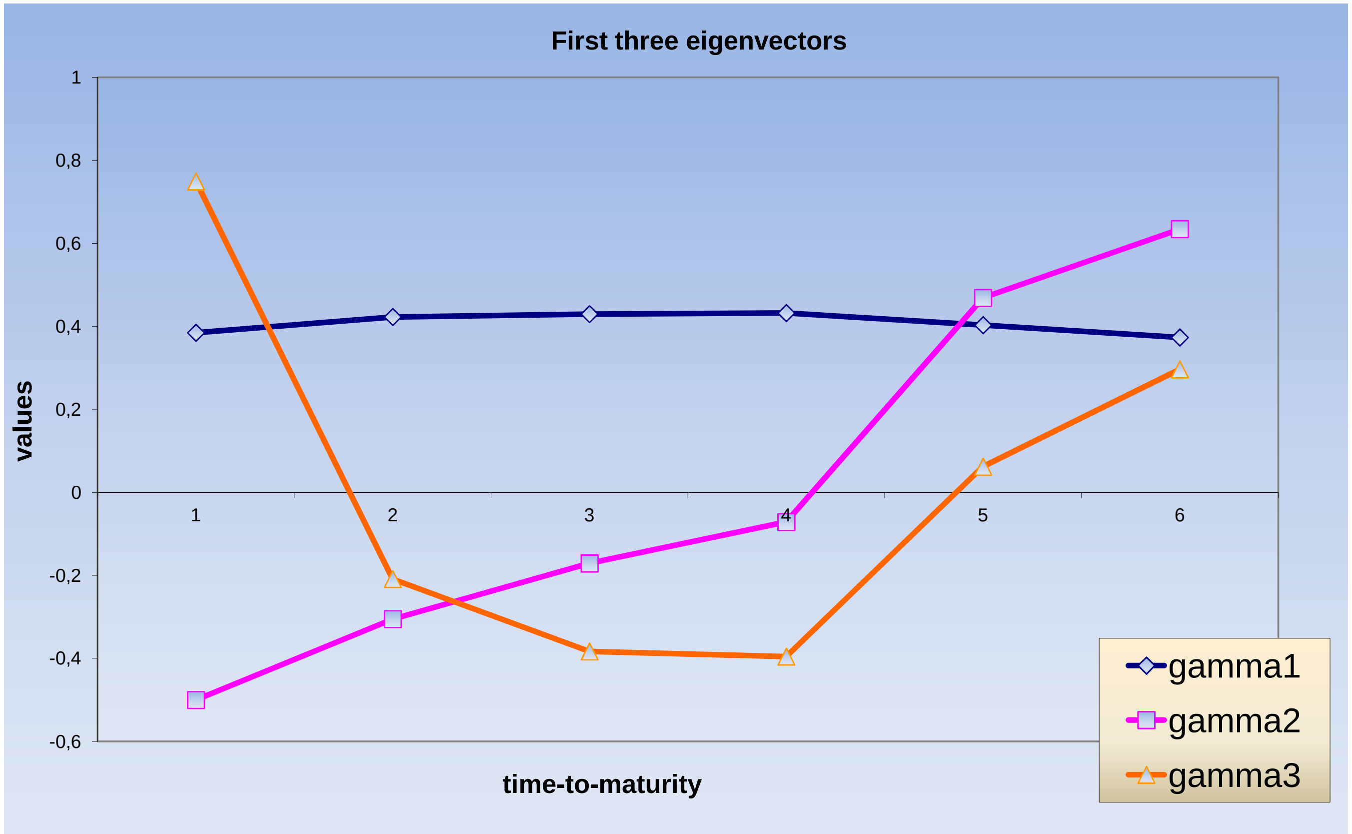
$R(t, t+\tau_i)$	$\widehat{\gamma}_1$	$\widehat{\gamma}_2$	$\widehat{\gamma}_3$
3m	0.4067	-0.4178	0.7569
6m	0.4124	-0.3011	-0.1058
9m	0.4141	-0.2207	-0.3442
1y	0.4151	-0.1505	-0.4350
5y	0.4074	0.4341	-0.1419
10y	0.3934	0.6892	0.2968

☐ The last two tables suggest that we can approximate the time variability of $R_t = [R(t, t+3m), \ldots, R(t, t+10y)]'$ in the following way:

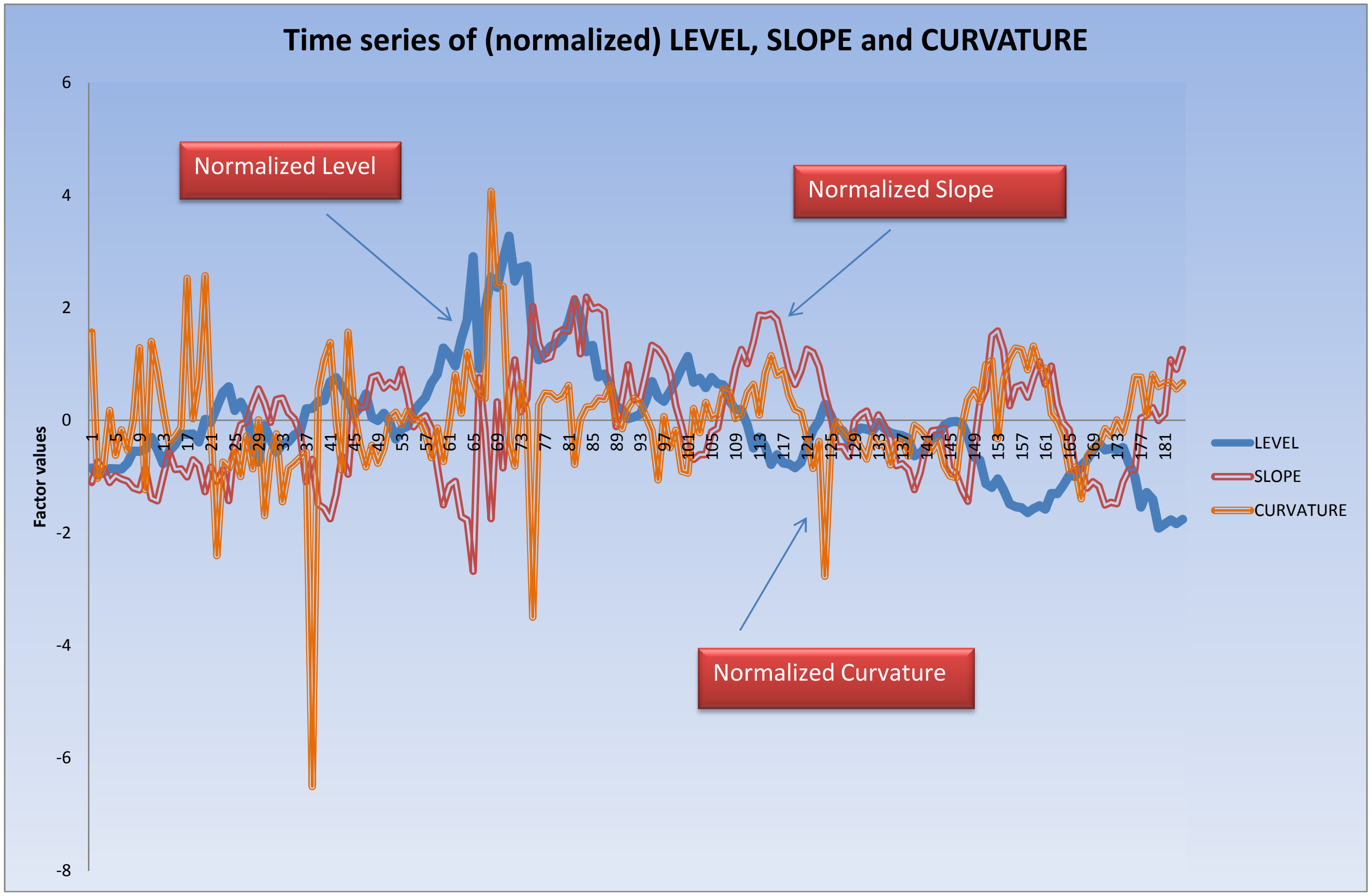
$$R_t = \widehat{\mu} + \widehat{\gamma}_1 \mathbf{y_{1,t}} + \widehat{\gamma}_2 \mathbf{y_{2,t}} + \widehat{\gamma}_3 \mathbf{y_{3,t}}$$

☐ Same conclusions are obtained for yield variations (ΔR_t) [see Piazzesi (2003)].

☐ If we plot these 3 eigenvectors and the time series of the three *PCs* we find :



First three eigenvectors
values
1
0,8
0,6
0,4
0,2
0
-0,2
-0,4
-0,6
1
2
3
4
5
6
time-to-maturity
gamma1
gamma2
gamma3



Time series of (normalized) LEVEL, SLOPE and CURVATURE
Normalized Level
Normalized Slope
Normalized Curvature
Factor values
LEVEL
SLOPE
CURVATURE

□ This means that, for fixed $\mathbf{y_{2,t}}$ and $\mathbf{y_{3,t}}$, if we assume a variation of $\mathbf{y_{1,t}}$ (the first PC), then the effect is transferred on the yield curve by $\widehat{\gamma}_1$ which is almost constant over maturities. Thus, the entire yield curve has a parallel shift. For this reason the first PC is interpreted as a LEVEL FACTOR.

□ Following the same reasoning, the second and third PC are interpreted as SLOPE and CURVATURE FACTORS, respectively [see Litterman and Scheinkman (1991)].

□ We find the same interpretations as in Nelson and Siegel (1987).

2.2.5.3 PCA, Factors and yield curve information

- PCA is adapted to Gaussian i.i.d. (*independent identically distributed*) processes and, anyway, the role of factors is judged on the basis of the explained variance and not in terms of the contribution to fit the observations.

- Well known empirical studies have highlighted that at least two or three factors (maybe four or five) are required by a yield curve model to match the dynamics and the shapes of the term structure, and this is regardless the sample period and the kind of used data [see Dai and Singleton (2000, 2002, 2003), Duffee (2002), Cheridito, Filipovic and Kimmel (2002), Duarte (2004)].

☐ Now, let us take a look to *skewness, kurtosis* and $ACF(k)$ of the data base used in the PCA.

☐ **Remember:**

$i)$ for a Gaussian random variable *skewness = 0, kurtosis = 3*,

$ii)$ the Autocorrelation at lag $k > 0$, denoted $ACF(k)$, for the data-t yield $R(t,h)$ (with residual maturity h) is given by $ACF(k) = Corr(R(t+k,h), R(t,h))$ (for $k = 0$, $ACF(0) = 1$);

$iii)$ if a process is i.i.d, then $ACF(k) = 0$ for any $k > 0$.

Yields	1-Q	2-Q	3-Q	4-Q	20-Q	40-Q
Mean	0.015	0.015	0.015	0.015	0.017	0.018
Std. Dev.	0.007	0.007	0.007	0.007	0.007	0.006
Skewness	0.84	0.74	0.69	0.67	0.77	0.94
Kurtosis	4.41	4.08	3.95	3.90	3.58	3.60
ACF(1)	0.91	0.92	0.93	0.93	0.95	0.96
ACF(4)	0.74	0.76	0.77	0.77	0.83	0.85
ACF(8)	0.48	0.52	0.54	0.55	0.68	0.73
ACF(12)	0.35	0.38	0.41	0.43	0.60	0.64
ACF(16)	0.31	0.34	0.36	0.37	0.50	0.53

☐ So, yields are not really i.i.d. and Gaussian-distributed.

☐ This result is the same regardless the sample period, the number of residual maturities and the methodology followed to construct the database.

☐ Observe that, in the case of yields, ACF is decreasing in k (for any fixed h) and increasing in h (for any fixed k).

- Cochrane and Piazzesi (2005, AER) show that a particular combination of forward rates successfully forecasts excess bond returns.

- This linear combination of forward rates is not completely linked to the first three Principal Components (LEVEL, SLOPE and CURVATURE).

- Thus, it seems that there is information useful to predict yield variations (i.e., bond returns), over and above the one contained in the three PCs.