Model of The Term Structure

Carlo A Favero

February 2016

Contents

1 Asset Pricing with Time-Varying Expected Returns **1**

2 Bond Returns: Yields-to-Maturity, Duration and Holding Period Returns **3**

2.1 Zero-Coupon Bonds 3

2.2 A simple model of the term structure 4

2.3 Forward Rates 5

2.3.1 Instantaneous forward rates 6

3 Financial Factor Models of the Term Structure: **6**

4 A general state-space representation **8**

5 Assessing the importance of the risk premium **8**

5.1 Single-equation evidence on the ET 8

5.2 The VAR based evidence 12

6 No-Arbitrage Affine Factor Models **14**

6.1 Risk Premium 16

6.2 Extensions 17

7 Appendix: Coupon Bonds **18**

7.1 Zero-Coupon Equivalent 18

7.2 Log-linearization 19

1. Asset Pricing with Time-Varying Expected Returns

Different approaches have been used in finance to model time-varying expected returns, they are all understood within the context of a basic model that stems from the assumption of the absence of "arbitrage opportunities" (i.e. by the impossibility of making profits without taking risk). Consider a situation in which in each period k state of nature can occur and each state has a probability $\pi(k)$, in the absence of arbitrage opportunities the price of an asset i at time t can be written as follows:

$$P_{i,t} = \sum_{s=1}^{k} \pi_{t+1}(s) m_{t+1}(s) X_{i,t+1}(s)$$

where $m_{t+1}(s)$ is the discounting weight attributed to future pay-offs, which (as the probability π) is independent from the asset i, $X_{i,t+1}(s)$ are the payoffs of the assets (in case of stocks we have $X_{i,t+1} = P_{t+1} + D_{t+1}$, in case of zero coupon bonds, $X_{i,t+1} = P_{t+1}$), and therefore returns on assets are defined as $1 + R_{s,t+1} = \frac{X_{i,t+1}}{P_{i,t}}$.For the safe asset, whose payoffs do not depend on the state of nature, we have:

$$\begin{aligned} P_{s,t} &= X_{i,t+1} \sum_{s=1}^{k} \pi_{t+1}(s) m_{t+1}(s) \\ 1 + R_{s,t+1} &= \frac{1}{\sum_{j=1}^{m} \pi_{t+1}(s) m_{t+1}(s)} \end{aligned}$$

In general, we can write:

$$\begin{aligned} P_{i,t} &= E_t (m_{t+1} X_{i,t+1}) \\ 1 + R_{s,t+1} &= \frac{1}{E_t (m_{t+1})} \end{aligned}$$

consider now a risky asset :

$$\begin{aligned} E_t (m_{t+1} (1 + R_{i,t+1})) &= 1 \\ Cov (m_{t+1} R_{i,t+1}) &= 1 - E_t (m_{t+1}) E_t (1 + R_{i,t+1}) \\ E_t (1 + R_{i,t+1}) &= -\frac{Cov (m_{t+1} R_{i,t+1})}{E_t (m_{t+1})} + (1 + R_{s,t+1}) \end{aligned}$$

Turning now to excess returns we can write:

$$E_t (R_{i,t+1} - R_{s,t+1}) = -(1 + R_{s,t+1}) \, cov (m_{t+1} R_{i,t+1})$$

Assets whose returns are low when the stochastic discount factor is high (i.e. when agents values payoffs more) require an higher risk premium, i.e. an higher excess return on the risk-free rate. Turning to predictability at different horizon, if you consider the case in which t is defined by taking two points in time very close to each other the safe interest rate will be approximately zero and m will not vary too much across states. The constant expected return model (with expected returns equal to zero) is compatile with the no-arbitrage approach at high-frequency. However, consider now the case of low frequency, when t is defined by taking two very distant points in time; in this case safe interest rate will be diffferent from zero and m will vary sizeably across different states. The constant expected return model is not a good approximation at long-horizons. Predictability is not a symptom of market malfunction but rather the consequence of a fair compensation for risk taking, then it should reflect attitudes toward risk and variation in market risk over time. Different theories on the relationship between risk and asset prices should then be assessed on the basis of their ability of explaining the predictability that emerges from the data.

Also, different theories or return predictability can be interpreted as different thoeries of the determination of m. On the one hand we have theories of m based on rational investor behaviour, on the other hand we have alternative approaches based on psycological models of investor behaviour.

2. Bond Returns: Yields-to-Maturity, Duration and Holding Period Returns

Analyzing bonds is crucial to understand how information on expected inflation can be extracted bythe term structure. We distinguish between two type of bonds: those paying a coupon each given period and those that do noy pay a coupon but just reimburse the entire capital upon maturity (zero-coupon bonds). We shall work mainly with zero-coupon bonds but all our results can be extended to coupon bonds.

Cash-flows from different type of bonds:

	$t+1$	$t+2$	$t+3$	...	T
general	CF_{t+1}	CF_{t+2}	CF_{t+3}	...	CF_T
coupon bond	C	C	C	...	$1+C$
1-period zero	1	0	0	...	0
2-period zero	0	1	0	...	0
...				...	
(T–t) -period zero	0	0	0	...	1

2.1. *Zero-Coupon Bonds*

Define the relationship between price and yield to maturity of a zero-coupon bond as follows:

$$P_{t,T} = \frac{1}{(1+Y_{t,T})^{T-t}}, \tag{1}$$

where $P_{t,T}$ is the price at time t of a bond maturing at time T, and $Y_{t,T}$ is yield to maturity. Taking logs of the left and the right-hand sides of the expression for $P_{t,T}$, and defining the continuously compounded *yield*, $y_{t,T}$, as $\log(1+Y_{t,T})$, we have the following relationship:

$$p_{t,T} = -\left(T-t\right) y_{t,T}, \tag{2}$$

which clearly illustrates that the elasticity of the yield to maturity to the price of a zero-coupon bond is the maturity of the security. In other words the duration of the bond equals maturity as no coupons are paid.

Table 2.:Price and YTM of zero-coupon bonds							
Maturity	1	2	3	5	7	10	20
$P_{t,T}$	0.9524	0.9070	0.8638	0.7835	0.7106	0.6139	0.3769
$Y_{t,T}$	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
$p_{t,T}$	−0.0487	−0.0976	−0.1464	−0.2439	−0.3416	−0.4879	−0.9757
$y_{t,T}$	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488	0.0488

The one-period uncertain holding-period return on a bond maturing at time T, $r_{t,t+1}^{T}$, is then defined as follows:

$$r_{t,t+1}^{T} \equiv p_{t+1,T} - p_{t,T} = -\left(T-t-1\right) y_{t+1,T} + \left(T-t\right) y_{t,T} \tag{3}$$
$$= y_{t,T} - \left(T-t-1\right)\left(y_{t+1,T} - y_{t,T}\right),$$
$$= \left(T-t\right) y_{t,T} - \left(T-t-1\right) y_{t+1,T}, \tag{4}$$

which means that yields and returns differ by the a scaled measure of the change between the yield at time $t+1$, $y_{t+1,T}$, and the yield at time t, $y_{t,T}$. Think of a situation in which the one-year YTM stands at 4.1 per cent while the 30-year YTM stands at 7 per cent. If the YTM of the thirty year bonds goes up to 7.1 per cent in the following period, then the period returns from the two bonds is the same.

2.2. *A simple model of the term structure*

Apply the no arbitrage condition to a one-period bond (the safe asset) and a T-period bond:

$$E_t\left(r_{t,t+1}^{T} - r_{t,t+1}^{1}\right) = E_t\left(r_{t,t+1}^{T} - y_{t,t+1}\right) = \phi_{t,t+1}^{T}$$
$$E_t\left(r_{t,t+1}^{T}\right) = y_{t,t+1} + \phi_{t,t+1}^{T}$$

Solving forward the difference equation $p_{t,T} = p_{t+1,T} - r_{t,t+1}^{T}$, we have :

$$y_{t,T} = \frac{1}{(T-t)}\sum_{i=0}^{n-1} E_t\left(r_{t+i,t+i+1}^{T}\right)$$
$$= \frac{1}{(T-t)}\sum_{i=0}^{n-1} E_t\left(y_{t+i,t+i+1} + \phi_{t+i,t+i+1}^{T}\right)$$

Which makes clear that YTM on long bond depends on future expected returns on one-period bond and the average risk premium over the maturity of the long bond. The simplest model of the term structure, the expectations theory, posits that the average risk premium is constant.

2.3. *Forward Rates*

Forward rates are returns on an investment at time t, made in the future at time t' with maturity at time T. The return on this strategy is equivalent to the return on a strategy that buys at time t zero coupon with maturity T and sells at time t the same amount of bonds with maturity t'.

The price of the investment strategy is $\left(-\left(T-t\right)y_{t,T}+\left(t'-t\right)y_{t,t'}\right)$and using the usual formula that links prices to returns we have :

$$f_{t,t',T}=\frac{\left(T-t\right)y_{t,T}-\left(t'-t\right)y_{t,t'}}{T-t'} \tag{5}$$

Applying the general formula to specific maturities we have :

$$f_{t,t+1,t+2}=2y_{t,t+2}-y_{t,t+1} \tag{6}$$

$$f_{t,t+2,t+3}=3y_{t,t+3}-2y_{t,t+2} \tag{7}$$

$$f_{t,t+3,t+4}=4y_{t,t+4}-3y_{t,t+3} \tag{8}$$

$$f_{t,t+n-1,t+n}=ny_{t,t+n}-(n-1)y_{t,t+n-1} \tag{9}$$

Using all these equations we have:

$$\begin{aligned} y_{t,t+n} &= \frac{1}{n}\left(y_{t,t+1}+f_{t,t+1,t+2}+f_{t,t+2,t+3}+...+f_{t,t+n-1,t+n}\right) \qquad (10) \\ y_{t,t+n} &= \frac{1}{n}\sum_{i=0}^{n-1}E_t\left(y_{t+i,t+i+1}+\phi^T_{t+i,t+i+1}\right) \end{aligned}$$

therefore each spot rate can be interpreted as an average of future one period forward. Putting together this evidence with the model for the term structure we have:

$$f_{t,t+i,t+i+1}=E_t\left(y_{t+i,t+i+1}+\phi^T_{t+i,t+i+1}\right)$$

Think of using forward rates to assess the impact of monetary policy. Let us analyze a potential movement of spot and forward rates around a shift in the central bank target rate.

Before CB intervention

1-year spot and forward rates					
maturity	i=1	i=2	i=3	i=4	i=5
$y_{t,t+i}$	0.05	0.05	0.05	0.05	0.05
$f_{t,t+i,t+i+1}$	0.05	0.05	0.05	0.05	

After CB intervention :

1-year spot and forward rates					
maturità	i=1	i=2	i=3	i=4	i=5
$y_{t,t+i}$	0.06	0.06	0.05	0.045	0.04
$f_{t,t+i,t+i+1}$	0.06	0.03	0.03	0.02	

Please remember that the interpretion of future forward as expected rates requires some assumption on the risk premium.

2.3.1. **Instantaneous forward rates**

Define the instanteous froward as the forwad rate on the contract with infinitesimal maturity:

$$f_{t,t'} = \lim_{T \to t'} f_{t,t',T} \tag{11}$$

given the sequence of forward rates you can define forward rate at any settlement date as follows :

$$f_{t,t',T} = \frac{\int_{\tau=t'}^{T} f_{\tau t} d\tau}{(T-t')}$$

As a consequence the relationship between spot and forward rate is written as:

$$y_{t,T} = \frac{\int_{\tau=t}^{T} f_{\tau t} d\tau}{(T-t)}$$

and therefore

$$f_{t,T} = y_{t,T} + (T-t)\frac{\partial y_{t,T}}{\partial T} \tag{12}$$

so instantaneous forward rates and spot rates coincide at the very short and very long-end of the term structure, forward rates are above spot rates when the yield curve slopes positevely and forward rates are below spot rates when the yield curve slopes negatively.

3. **Financial Factor Models of the Term Structure:**

The empirical financial literature has put the concept just illustrated at work to build model of the term structure that can be used to interpolate nontraded maturity and to forecast yield at each maturity.

Gurkanyak et al. estimate the following interpolant at each point in time, by non-linear least squares, on the cross-section of yields:

$$\begin{aligned} y_{t,t+k} &= L_t + SL_t \frac{1-\exp\left(-\frac{k}{\tau_1}\right)}{\frac{k}{\tau_1}} + C_t^1 \left(\frac{1-\exp\left(-\frac{k}{\tau_1}\right)}{\frac{k}{\tau_1}} - \exp\left(-\frac{k}{\tau_1}\right)\right) + \\ &+ C_t^2 \left(\frac{1-\exp\left(-\frac{k}{\tau_2}\right)}{\frac{k}{\tau_2}} - \exp\left(-\frac{k}{\tau_2}\right)\right) \end{aligned} \tag{13}$$

which is an extension originally proposed by Svensson(1994) on the original parameterization adopted by Nelson and Siegel (1987) that sets $C_t^2 = 0$.

Forward rates are easily derived as

$$f_{tk} = L_t + SL_t \exp\left(-\frac{k}{\tau_1}\right) + C_t^1 \frac{k}{\tau_1} \exp\left(-\frac{k}{\tau_1}\right) + C_t^2 \frac{k}{\tau_2} \exp\left(-\frac{k}{\tau_2}\right) \tag{14}$$

When maturity k goes to zero forward ans spot rates coincide at $L_t + SL_t$, and when maturity goes to infinite forward and spot coincide at L_t. Terms in C_t^1 and C_t^2 describes two humps starting at zero at different starting points and ending at zero. As discussed in Diebold and Li (2002) the above interpolant is very flexible and capable of accommodating several stylized facts on the term structure and its dynamics. In particular, L_t, SL_t, C_t^1, C_t^2, which are estimated as parameters in a cross-section of yields, can be interpreted as latent factors. L_t has a loading that does not decay to zero in the limit, while the loading on all the other parameters do so, therefore this parameter can be interpreted as the long-term factor, the level of the term-structure. The loading on SL_t is a function that starts at 1 and decays monotonically towards zero; it may be viewed a short-term factor, the slope of the term structure. In fact, $r_t^{rf} = L_t + SL_t$ is the limit when k goes to zero of the spot and the forward interpolant. We naturally interpret r_t^{rf} as the risk-free rate. Obviously SL_t, the slope of the yield curve, is nothing else than the minus the spread in Campbell-Shiller. C_t are medium term factor, in the sense that their loading start at zero, increase and then decay to zero (at different speed). Such factors capture the curvature of the yield curve. In fact, Diebold and Li show for example that C_t^1 tracks very well the difference between the sum of the shortest and the longest yield and twice the yield at a mid range (2-year maturity). The repeated estimation of loadings using a cross-section of eleven US yields at different maturities allows to construct a time-series for four factors.

4. **A general state-space representation**

To generalize the NS approach we can put the dynamics of the term structure in a state-space framework. Yields with different maturities are collected in a vector $y_t = [y_{t,t+1} \;, y_{t,t+2} \;, \ldots, y_{t,t+k}]'$. Equation (15) is the measurement equation, in which different yields $y_{t,t+n}$ are assumed to be determined by a set of state variables, collected in the vector X_t:

$$y_{t,t+n} = \frac{-1}{n}\left(A_n + B_n' X_t\right) + \varepsilon_{t,t+n} \qquad \varepsilon_t \sim i.i.d.N(0, \sigma^2 I) \tag{15}$$

$$X_t = \mu + \Phi X_{t-1} + v_t \qquad v_t \sim i.i.d.N(0, \Omega) \tag{16}$$

In the case of original NS we have

$$B_n' = \left[-n \;, -\left(\frac{1-e^{-\lambda n}}{\lambda}\right) \;, -\left(\frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n}\right)\right] \text{ and } A_n = 0$$

Note that factors can be first extracted in the cross-sectional dimension and then their time-series behaviour can be studied in a VAR or the model can be estimated simultaneously, using an unobservable component method such as the Kalman filter.

5. **Assessing the importance of the risk premium**

We consider two approaches to assessing the importance of the risk premium: single equation based-evidence and VAR based evidence.

5.1. *Single-equation evidence on the ET*

The traditional evidence is mainly based on a single-equation, limited information, approach.

To assess the importance of RP, three different implications of the Expectations theory can be brought to the data:

$$y_{t,T} - (T-t-1)\, E_t\left(y_{t+1,T} - y_{t,T}\right) = y_{t,t+1} + \phi_{t,T}. \tag{17}$$

$$f_{t,t+i,t+i+1} = E_t\left(y_{t+i,t+i+1} + \phi^T_{t+i,t+i+1}\right) \tag{18}$$

$$y_{t,t+n} - y_{t,t+1} = \sum_{i=1}^{n-1}\left(1-\frac{1}{n}\right) E_t \Delta y_{t+i,t+i+1} + \frac{1}{n}\sum_{i=1}^{n-1}\phi^T_{t+i,t+i+1} \tag{19}$$

(a) Estimate the following model :

$$y_{t+1,T} - y_{t,T} = \beta_0 + \beta_1 \frac{1}{T-t-1}\left(y_{t,T} - y_{t,t+1}\right) + u_{t+1}$$

to test $\beta_0 = 0, \beta_1 = 1$.

(b) Estimate the following model :

$$
\begin{aligned}
S^*_{t,n} &= \beta_0 + \beta_1 S_{t,n} + u_t \\
S_{t,n} &= y_{t,t+n} - y_{t,t+1} \\
S^*_{t,n} &= \sum_{i=1}^{n-1} \left(1 - \frac{1}{n}\right) \Delta y_{t+i,t+i+1}
\end{aligned}
$$

to test $\beta_0 = 0, \beta_1 = 1$.

(c) Estimate the following model :

$$
(y_{t+i,t+i+1} - y_{t,t+1}) = \beta_0 + \beta_1 (f_{t,t+i,t+i+1} - y_{t,t+1}) + u_{t+i+1}
$$

to test $\beta_0 = 0, \beta_1 = 1$.

(d) Estimate the following model :

$$
y_{t,T} - (T - t - 1)(y_{t+1,T} - y_{t,T}) - y_{t,t+1} = \beta_0 + \beta_1 (f_{t,t+i,t+i+1} - y_{t,t+1}) + u_{t+i+1}
$$

to test $\beta_1 = 0$.

The empirical evidence shows that:

- i) high yields spreads fare poorly in predicting increases in long rates(see Campbell, 1995)
- ii) the change in yields does not move one-to-one with the forward spot spread (see Fama and Bliss,1986)
- iii) period excess returns on long-term bond are predictable using the information in the forward-spot spread (see Cochrane,1999,Cochrane-Piazzesi 2005)

Table 10.3. Regression coefficients $\hat{\beta}_n$ and $\hat{\gamma}_n$.

Dependent variable	Long bond maturity (n)						
	2	3	6	12	24	48	120
Long-yield changes (10.2.16)	0.003 (0.191)	−0.145 (0.282)	−0.835 (0.442)	−1.435 (0.599)	−1.448 (1.004)	−2.262 (1.458)	−4.226 (2.076)
Short-rate changes (10.2.18)	0.502 (0.096)	0.467 (0.148)	0.320 (0.146)	0.272 (0.208)	0.363 (0.223)	0.442 (0.384)	1.402 (0.147)

Long bond maturities are measured in months. The first row reports the estimated regression coefficient $\hat{\beta}_n$ from (10.2.16), with an asymptotic standard error (in parentheses) calculated to allow for heteroskedasticity in the manner described in the Appendix. The second row reports the the estimated regression coefficient $\hat{\gamma}_n$ from (10.2.18), with an asymptotic standard error calculated in the same manner, allowing also for residual autocorrelation. The expectations hypothesis of the term structure implies that both $\hat{\beta}_n$ and $\hat{\gamma}_n$ should equal one. The underlying data are monthly zero-coupon bond yields over the period 1952:1 to 1991:2, from McCulloch and Kwon (1993).

TABLE 2—FAMA-BLISS EXCESS RETURN REGRESSIONS

Maturity n	β	Small T	R^2	$\chi^2(1)$	p-val	EH p-val
2	0.99	(0.33)	0.16	18.4	⟨0.00⟩	⟨0.01⟩
3	1.35	(0.41)	0.17	19.2	⟨0.00⟩	⟨0.01⟩
4	1.61	(0.48)	0.18	16.4	⟨0.00⟩	⟨0.01⟩
5	1.27	(0.64)	0.09	5.7	⟨0.02⟩	⟨0.13⟩

Notes: The regressions are $rx_{t+1}^{(n)} = \alpha + \beta(f_t^{(n)} - y_t^{(1)}) + \varepsilon_{t+1}^{(n)}$. Standard errors are in parentheses "( )", probability values in angled brackets "⟨ ⟩". The 5-percent and 1-percent critical values for a $\chi^2(1)$ are 3.8 and 6.6.

5.2. *The VAR based evidence*

The VAR based evidence is due to Campbell-Shiller(1987) and considers the following version of the ET

$$S_{t,T} = S^*_{t,T} = \sum_{j=1}^{T-t-1} \gamma^j E[\Delta r_{t+j} \mid I_t] \tag{20}$$

(20) shows that a necessary condition for the ET to hold puts constraints on the long-run dynamics of the spread. In fact, the spread should be stationary being a weighted sum of stationary variables. Obviously, stationarity of the spread implies that, if yields are non-stationary, they should be cointegrated with a cointegrating vector (1,-1). However, the necessary and sufficient conditions for the validity of the ET impose restrictions both on the long-run and the short run dynamics.

Assuming[1] that $R_{t,T}$ and r_t are cointegrated with a cointegrating vector (1,-1), CS construct a bivariate stationary VAR in the first difference of the short-term rate and the spread :

$$\begin{aligned} \Delta r_t &= a(L)\Delta r_{t-1} + b(L)S_{t-1} + u_{1t} \\ S_t &= c(L)\Delta r_{t-1} + d(L)S_{t-1} + u_{2t} \end{aligned} \tag{21}$$

Stack the VAR as:

$$\begin{bmatrix} \Delta r_t \\ . \\ . \\ \Delta r_{t-p+1} \\ S_t \\ . \\ . \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & . & . & a_p & b_1 & . & . & b_p \\ 1 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & 1 & 0 & 0 & . & . & 0 \\ c_1 & . & . & c_p & d_1 & . & . & d_p \\ 0 & . & . & 0 & 1 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta r_{t-1} \\ . \\ . \\ \Delta r_{t-p} \\ S_{t-1} \\ . \\ . \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ . \\ . \\ 0 \\ u_{2t} \\ . \\ . \\ 0 \end{bmatrix} \tag{22}$$

This can be written more succinctly as:

$$z_t = Az_{t-1} + v_t \tag{23}$$

The ET null puts a set of restrictions which can be written as :

$$g' z_t = \sum_{j=1}^{T-1} \gamma^j h' A^{j\prime} z_t \tag{24}$$

[1] In fact, the evidence for the restricted cointegrating vector which constitutes a necessary condition for the ET to hold is not found to be particularly strong in the original CS work.

where $g\prime$ and $h\prime$ are selector vectors for S and Δr correspondingly (i.e. row vectors with 2p elements, all of which are zero except for the p+1st element of $g\prime$ and the first element of $h\prime$ which are unity). Since the above expression has to hold for general z_t, and, for large T, the sum converges under the null of the validity of the ET, it must be the case that:

$$g' = h'\gamma A(I - \gamma A)^{-1} \tag{25}$$

which implies:

$$g'(I - \gamma A) = h'\gamma A \tag{26}$$

and we have the following constraints on the individual coefficients of VAR(21):

$$\{c_i = -a_i, \forall i\}, \{d_1 = -b_1 + 1/\gamma\}, \{d_i = -b_i, \forall i \neq 1\} \tag{27}$$

The above restrictions are testable with a Wald test. By doing so using US data between the fifties and the eighties Campbell and Shiller (1987) rejected the null of the ET. However, when CS construct a theoretical spread $S^*_{t,T}$, by imposing the (rejected) ET restrictions on the VAR they find that, despite the statistical rejection of the ET, $S^*_{t,T}$ and $S_{t,T}$ are strongly correlated.

6. **No-Arbitrage Affine Factor Models**

The standard response in finance to the empirical rejection of the Expectations Theory has been modelling the term structure based on the assumption that there are no riskless arbitrage opportunities among bonds of various maturities. The standard model is based on three components: a transition equation for the state vector relevant for pricing bonds, made traditionally of latent factors, an equation which defines the process for the risk-free one-period rate and a relation which associates the risk premium with shocks to the state vector, defined as a linear function of the state of the economy. In such structure, the price of a j-period nominal bond is a linear function of the factors. Unobservable factors and coefficients in the bond pricing functions are jointly estimated by maximum likelihood methods (see, for example, Chen and Scott(1993)). This type of models usually provides a very good within sample fit of different yields but do not perform well in forecasting. Duffee(2002) shows that the forecasts produce by no-arbitrage models with latent factors do not outperform the random walk model.

First building block is the dynamics of the factors determining risk premium:

$$X_t = \mu + \Phi X_{t-1} + \Sigma \epsilon_t$$

interpret this as a companion form representation.

Second is a specification for the one-period rate r_t which is assumed to be a linear function of the factors:

$$r_t = \delta_0 + \delta_1' X_t$$

The third is a pricing kernel. The assumption of no-arbitrage guarantees the existence of a risk-neutral measure Q such that the price of any asset V_t that does not pay any dividends at time t+1 satisfies the following relation:

$$V_t = E_t^Q \left(\exp\left(-r_t\right) V_{t+1}\right)$$

the Radon-Nikodym derivative (which converts the risk neutral measure to the data-generating measure) is denoted by ξ_{t+1}. So for any random variable Z_{t+1} we have

$$E_t^Q \left(Z_{t+1}\right) = E_t \left(\xi_{t+1} Z_{t+1}\right) / \xi_t$$

The assumption of no-arbitrage allows us to price any nominal bond in the economy.

Assume that ξ_{t+1} follows the log-normal process:

$$\xi_{t+1} = \xi_t \exp\left(-\frac{1}{2}\lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1}\right)$$

where λ_t are the time-varying market prices of risk associated with the sources of uncertainty ϵ_t. Parameterize λ_t as an affine process:

$$\lambda_t = \lambda_0 + \lambda_1 X_t$$

define the pricing kernel m_{t+1} as

$$m_{t+1} = \exp\left(-r_t\right) \xi_{t+1}/\xi_t$$

substituting from the processes for the short-rate and ξ_{t+1} we have:

$$m_{t+1} = \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t' \lambda_t - \lambda_t' \varepsilon_{t+1}\right)$$

Now the total one-period gross return of any nominal asset satisfies:

$$E_t\left(m_{t+1} R_{t+1}\right) = 1$$

If p_t^n represents the price of an n-period zero coupon bond, then we can use this equation to compute recursively bond prices as:

$$p_t^{n+1} = E_t\left(m_{t+1} p_{t+1}^n\right)$$

Now guess that the log of bond prices are linear functions of the state variable we have:

$$p_t^n = \exp\left(A_n + B_n' X_t\right)$$

This guess is easily verified for the one-period bond, in which case we have:

$$\begin{aligned} p_t^1 &= E_t\left(m_{t+1}\right) = \exp\left(-r_t\right) \\ &= \exp\left(\delta_0 + \delta_1' X_t\right) \end{aligned}$$

But it also applies to n-period bonds, in which case we have:

$$
\begin{aligned}
p_t^{n+1} &= E_t\left(m_{t+1}p_{t+1}^n\right) \\
&= E_t \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\varepsilon_{t+1} + A_n + B_n' X_{t+1}\right) \\
&= \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t'\lambda_t + A_n\right) E_t\left[\exp\left(-\lambda_t'\varepsilon_{t+1} + B_n' X_{t+1}\right)\right] \\
&= \exp\left(-\delta_0 - \delta_1' X_t - \frac{1}{2}\lambda_t'\lambda_t + A_n\right) E_t\left[\exp\left(-\lambda_t'\varepsilon_{t+1} + B_n'\left(\mu + \Phi X + \Sigma\epsilon_{t+1}\right)\right)\right] \\
&= \exp\left(-\delta_0 + \left(B_n'\Phi - \delta_1'\right) X_t - \frac{1}{2}\lambda_t'\lambda_t + A_n + B_n'\mu\right) \\
&\quad E_t\left[\exp\left(-\lambda_t'\varepsilon_{t+1} + B_n'\Sigma\epsilon_{t+1}\right)\right] \\
&= \exp\left(-\delta_0 + \left(B_n'\Phi - \delta_1'\right) X_t + A_n + B_n'\left(\mu - \Sigma\lambda_0\right) + \frac{1}{2}B_n'\Sigma\Sigma' B_n - B_n'\Sigma\lambda_1 X_t\right)
\end{aligned}
$$

here the last step uses log-normality and the fact that $\lambda_t'\lambda_t = \lambda_t' var\left(\varepsilon_{t+1}\right)\lambda_t$.By matching coefficients we now have:

$$
\begin{aligned}
A_{n+1} &= -\delta_0 + A_n + B_n'\left(\mu - \Sigma\lambda_0\right) + \frac{1}{2}B_n'\Sigma\Sigma' B_n \\
B_{n+1}' &= B_n'\left(\Phi - \Sigma\lambda_1\right) - \delta_1'
\end{aligned}
$$

To sum up, we can characterize a traditional Affine TS model as follows:

$$
y_{t,t+n} = \frac{-1}{n}\left(A_n + B_n' X_t\right) + \varepsilon_{t,t+n} \qquad \varepsilon_t \sim i.i.d.N(0, \sigma^2 I) \tag{28}
$$

$$
X_t = \mu + \Phi X_{t-1} + v_t \qquad v_t \sim i.i.d.N(0, \Omega) \tag{29}
$$

$$
b_{n+1} = \tfrac{1}{(n+1)}\left[\sum_{i=0}^{n}\left(\Phi' - \lambda_1'\Omega\right)^i\right] b_1
$$

$$
a_{n+1} = a_1 - \tfrac{1}{(n+1)}\sum_{i=0}^{n} B^{(i)}, \text{ where } B^{(i)} = B_i'\left(\mu - \Omega\lambda_0\right) + \tfrac{1}{2}B_i'\Omega B_i.
$$

6.1. *Risk Premium*

Given the knowledge of the model parameters the risk premium can be derived naturally:

$$
RP_t^n = y_{t,t+n} - \frac{1}{n}E_t\sum_{j=0}^{n-1} r_{t+j}
$$

$$
\begin{aligned}
E_t r_{t+j} &= \delta_0 + \delta_1' E_t X_{t+j} \\
&= \delta_0 + \delta_1'\left(\bar{\mu} + \Phi^j\left(X_t - \bar{\mu}\right)\right)
\end{aligned}
$$

where

$$\begin{aligned} X_t &= \mu + \Phi X_{t-1} + \Sigma\epsilon_t \\ \bar{\mu} &= (I - \Phi)^{-1}\mu \end{aligned}$$

and, in absence of measurement error, we have:

$$\begin{aligned} y_{t,t+n} &= \frac{-1}{n}\left(A_n + B_n' X_t\right) \\ &= \bar{A}_n + \bar{B}_n' X_t \end{aligned}$$

$$\begin{aligned} RP_t^n &= \bar{A}_n - \sum \frac{\delta_0}{n} - \delta_1' \left(I - \frac{1}{n}\sum_{j=0}^{n-1} \Phi^j \right) \bar{\mu} + \\ &+ \left(\bar{B}_n' - \delta_1' \frac{1}{n}\sum_{j=0}^{n-1} \Phi^j \right) X_t \end{aligned}$$

but $B'_{n+1} = B'_n \left(\Phi - \Sigma\lambda_1\right) - \delta'_1$,so when $\Sigma = 0$ or $\lambda_1 = 0$ the term multiplying X_t in the last expression vanishes and the risk premium becomes constant.

6.2. *Extensions*

Recently the no-arbitrage approach has been extended to include some observable macroeconomic factors in the state vector and to explicit allow for a Taylor-rule type of specification for the risk-free one period rate. Ang and Piazzesi(2002) and Ang, Piazzesi and Wei(2003) show that the forecasting performance of a VAR improves when no-arbitrage restrictions are imposed and that augmenting non-observable factors models with observable macroeconomic factors clearly improves the forecasting performance. Hordahl et al.(2003) and Rudebusch and Wu(2003) use a small scale macro model to interpret and parameterize the state vector; forecasting performance is improved and models have also some success in accounting for the empirical failure of the Expectations Theory.

This extension of the small information set using macroeconomic variable can be further expanded by moving to large-information set. In this case rather than including in the state vector some specific macroeconomic variables, common factors can be extracted from a large panel of macroeconomic variables using static principal components, as suggested by Stock and Watson (2002).

7. Appendix: Coupon Bonds

The relationship between price and yield to maturity of a constant coupon (C) bond is given by:

$$P_{t,T}^{c} = \frac{C}{\left(1+Y_{t,T}^{c}\right)} + \frac{C}{\left(1+Y_{t,T}^{c}\right)^{2}} + ... + \frac{1+C}{\left(1+Y_{t,T}\right)^{T-t}}.$$

To apply the framework we ahve considered so far to Coupon Bonds therea are two alternatives.

1) Find the zero-coupon equivalent
2) Consider log-linearization

7.1. *Zero-Coupon Equivalent*

Define the discount function at time t of a zero-coupon paying one unit at $t+m$:

$$P_{t,t+m}^{ZC} = \exp\left(-my_{t,t+m}\right) \tag{30}$$

$$= \frac{1}{\left(1+Y_{t,t+m}\right)^{m}} = D_{t,t+m} \tag{31}$$

Consider now a coupon-bond:

$$P_{t,t+m} = \sum_{k=1}^{m} cD_{t,t+k} + D_{t,t+m} \tag{32}$$

Coupon bonds are nothing else than a bundle of zero coupons. For coupon bonds youa have available yield to maturity computed, given the price, as :

$$P_{t,t+m} = \sum_{k=1}^{m} c\exp\left(-k\bar{y}_{t,t+m}\right) + D_{t,t+m}$$

The tow formula can be used to build a zero coupon equivalent curve, given that you have available at least one zero coupon for the shortest maturity. the methodology is recursive and it is called "bootstrapping" :

$$Y_{t,t+m} = \left(\frac{1+\bar{Y}_{t,t+m}}{1-\sum_{k=1}^{m-1}\frac{\bar{Y}_{t,t+m}}{\left(1+Y_{t,t+k}\right)^{k}}}\right)^{1/m} - 1$$

The following table illustrates an application

YTM and zero coupon equivalent

YTM	$\bar{Y}_{t,t+1}$	$\bar{Y}_{t,t+2}$	$\bar{Y}_{t,t+3}$	$\bar{Y}_{t,t+4}$	$\bar{Y}_{t,t+5}$	$\bar{Y}_{t,t+6}$	$\bar{Y}_{t,t+7}$	$\bar{Y}_{t,t+8}$	$\bar{Y}_{t,t+9}$
	4.69	4.64	4.72	4.82	4.92	5.01	5.10	5.17	5.23
zero-coupon equivalent	$Y_{t,t+1}$	$Y_{t,t+2}$	$Y_{t,t+3}$	$Y_{t,t+4}$	$Y_{t,t+5}$	$Y_{t,t+6}$	$Y_{t,t+7}$	$Y_{t,t+8}$	$Y_{t,t+9}$
	4.69	4.64	4.72	4.83	4.94	5.04	5.14	5.22	5.29

Given the one-year zero coupon the zero coupon equivalent forthe two year is generated as follows :

$$\frac{\bar{Y}_{t,t+2}}{1+\bar{Y}_{t,t+2}}+\frac{1+\bar{Y}_{t,t+2}}{\left(1+\bar{Y}_{t,t+2}\right)^{2}}=\frac{Y_{t,t+1}}{1+\bar{Y}_{t,t+2}}+\frac{1+\bar{Y}_{t,t+2}}{\left(1+Y_{t,t+2}\right)^{2}}$$

from which we have:

$$Y_{t,t+2}=\left(\frac{1+\bar{Y}_{t,t+2}}{1-\frac{\bar{Y}_{t,t+2}}{1+Y_{t,t+1}}}\right)^{1/2}-1$$

Having obtained $Y_{t,t+2}$,the equation :

$$\frac{\bar{Y}_{t,t+3}}{1+\bar{Y}_{t,t+3}}+\frac{\bar{Y}_{t,t+3}}{\left(1+\bar{Y}_{t,t+3}\right)^{2}}+\frac{1+\bar{Y}_{t,t+3}}{\left(1+\bar{Y}_{t,t+3}\right)^{3}}=\frac{\bar{Y}_{t,t+3}}{1+\bar{Y}_{t,t+2}}+\frac{1+\bar{Y}_{t,t+3}}{\left(1+Y_{t,t+2}\right)^{2}}+\frac{1+\bar{Y}_{t,t+3}}{\left(1+Y_{t,t+3}\right)^{3}}$$

can be solved for $Y_{t,t+3}$:

$$Y_{t,t+3}=\left(\frac{1+\bar{Y}_{t,t+3}}{1-\sum_{k=1}^{2}\frac{\bar{Y}_{t,t+3}}{\left(1+Y_{t,t+k}\right)^{k}}}\right)^{1/3}-1$$

by iteration the full term structure is then derived

7.2. *Log-linearization*

When the bond is selling at par, the yield to maturity is equal to the coupon rate. To measure the length of time that a bondholder has invested money for we need to introduce the concept of duration:

$$D_{t,T}^{c}=\frac{\frac{C}{\left(1+Y_{t,T}^{c}\right)}+2\frac{C}{\left(1+Y_{t,T}^{c}\right)^{2}}+...+(T-t)\frac{1+C}{\left(1+Y_{t,T}\right)^{T-t}}}{P_{t,T}^{c}}$$
$$=\frac{C\sum_{i=1}^{T-t}\frac{i}{\left(1+Y_{t,T}^{c}\right)^{i}}+\frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}}}{P_{t,T}^{c}}.$$

Note that when a bond is floating at par we have:

$$\begin{aligned}
D^c_{t,T} &= Y^c_{t,T} \sum_{i=1}^{T-t} \frac{i}{\left(1+Y^c_{t,T}\right)^i} + \frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}} \\
&= Y^c_{t,T} \frac{\left((T-t)\frac{1}{1+Y^c_{t,T}} - (T-t) - 1\right)\frac{1}{\left(1+Y^c_{t,T}\right)^{T-t+1}} + \frac{1}{1+Y^c_{t,T}}}{\left(1-\frac{1}{1+Y^c_{t,T}}\right)^2} + \frac{(T-t)}{\left(1+Y_{t,T}\right)^{T-t}} \\
&= \frac{1-\left(1+Y^c_{t,T}\right)^{-(T-t)}}{1-\left(1+Y^c_{t,T}\right)^{-1}},
\end{aligned}$$

because when $|x| < 1$,

$$\sum_{k=0}^{n} k x^k = \frac{(nx - n - 1)\, x^{n+1} + x}{(1-x)^2}.$$

Duration can be used to find approximate linear relationships between log-coupon yields and holding period returns. Applying the log-linearization of one-period returns to a coupon bond we have:

$$\begin{aligned}
p_{c,t,T} - c &= -r^c_{t+1} + k + \rho\left(p_{c,t+1,T} - c\right) \\
r^c_{t+1} &= k + \rho p_{c,t+1,T} + (1-\rho)\, c - p_{c,t,T}.
\end{aligned}$$

When the bond is selling at par, $\rho = (1+C)^{-1} = \left(1+Y^c_{t,T}\right)^{-1}$. Solving this expression forward to maturity delivers:

$$p_{c,t,T} = \sum_{i=0}^{T-t-1} \rho^i \left(k + (1-\rho)\, c - r^c_{t+1+i}\right).$$

The log yield to maturity $y^c_{t,T}$ satisfies an expression with the same structure:

$$\begin{aligned}
p_{c,t,T} &= \sum_{i=0}^{T-t-1} \rho^i \left(k + (1-\rho)\, c - y^c_{t,T}\right) = \frac{1-\rho^{T-t-1}}{1-\rho}\left(k + (1-\rho)\, c - y^c_{t,T}\right) \\
&= D^c_{t,T}\left(k + (1-\rho)\, c - y^c_{t,T}\right).
\end{aligned}$$

By substituting this expression back in the equation for linearized returns we have the expression

$$r^c_{t+1} = D^c_{t,T} y^c_{t,T} - \left(D^c_{t,T} - 1\right) y^c_{t+1,T},$$

that illustrates the link between continuously compounded returns and duration.

A simple model of the term structure of coupon bonds Consider the relation between the return on a riskless one period short-term bill, r_t, and a long term bond bearing a coupon C, the one-period return on the long-term bond $H_{t,T}$ is a non-linear function of the log yield to maturity $R_{t,T}$. Shiller (1979) proposes the *linearization* illustrated in the previous section which takes duration as constant and considers the following approximation in the neighborhood $y_{t,T} = y_{t+1,T} = \bar{y} = C$:

$$\begin{aligned}
H_{t,T} &\simeq D_T y_{t,T} - (D_T - 1)\, y_{t+1,T} \\
D_T &= \frac{1 - \gamma^{T-t-1}}{1-\gamma} = \frac{1}{1-\gamma_T} \\
\gamma_T &= \left\{1 + \bar{y}\left[1 - 1/(1+\bar{y})^{T-t-1}\right]^{-1}\right\}^{-1} \\
\lim_{T \longrightarrow \infty} \gamma_T &= \gamma = 1/(1+\bar{y})
\end{aligned}$$

solving this expression forward :

$$y_{t,T} = \sum_{j=0}^{T-t-1} \gamma^j (1-\gamma) H_{t+j,T} + \gamma^{T-t} y_{T-1,T}$$

In this case, by equating one-period risk-adjusted returns, we have:

$$E\left[\frac{y_{t,T} - \gamma y_{t+1,T}}{1-\gamma} \mid I_t\right] = r_t + \phi_{t,T} \tag{33}$$

From the above expression, by recursive substitution, under the terminal condition that at maturity the price equals the principal,we obtain:

$$y_{t,T} = y^*_{t,T} + E[\Phi_T \mid I_t] = \frac{1-\gamma}{1-\gamma^{T-t}} \sum_{j=0}^{T-t-1} \gamma^j E[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \tag{34}$$

where the constant $\Phi_{t,T}$ is the term premium over the whole life of the bond:

$$\Phi_{t,T} = \frac{1-\gamma}{1-\gamma^{T-t}} \sum_{j=0}^{T-t-1} \gamma^j \phi_{t+j,T}$$

For long-bonds, when $T - t$ is very large, we have :

$$y_{t,T} = y^*_{t,T} + E[\Phi_T \mid I_t] = (1-\gamma) \sum_{j=0}^{T-t-1} \gamma^j E[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t]$$

Subtracting the risk-free rate from both sides of this equation we have:

$$\begin{aligned}
S_{t,T} = y_{t,T} - r_t &= \sum_{j=1}^{T-1} \gamma^j E[\Delta r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \\
&= S^*_{t,T} + E[\Phi_T \mid I_t]
\end{aligned}$$