Monetary Economics 2
F.C. Bagliano - September 2017

Notes on:

R. King and A. Kurmann "Expectations and the term structure of interest rates: evidence and interpretation", *Federal Reserve Bank of Richmond Economic Quarterly*, Fall 2002

1 Main points

The paper provides a thorough assessment of the empirical relevance of the (rational) *expectations theory* of the term structure of interest rates using monthly US data for the 1-month ("short") and the 20-year ("long") interest rates over the 1951-2001 period. The analysis exploits two sets of empirical implications of the expectations theory, concerning:

1. the existence of **cointegration** between short and long interest rates and therefore the presence of a **common stochastic trend** driving the two rates. This implication is formalized and tested by means of a (cointegrated) *vector error-correction* $(VECM)$ system.

2. the existence of a set of **cross-equation restrictions** that the theory (assuming rational expectations) imposes on an (unrestricted) VAR model describing the dynamics of short and long rates and their difference (i.e. the term spread).

2 Stylized facts

1. **Short** (R) and **long** (R^L) **rates**: large and persistent *variations* over time with *comovements*

2. **Spread** $(S \equiv R^L - R)$: much more *stable* over time

$\Rightarrow$ if the rates are non-stationary, then there might be evidence of cointegration (implying a stationary term spread), with a common stochastic trend determining the comovement feature; simple tests: ADF tests on interest rate series to test for non-stationarity and ADF test on spread to test for cointegration with $(1\ -1)$ cointegrating vector imposed.

3. Strong evidence of **predictability** in the *spread* behaviour (highly serially correlated), much less in the *interest rate changes* behaviour.

3 Theory

Basic theoretical relationships between long and short interest rates derived from expectation theory.

Yield to maturity of a N-period coupon bond at t, R_t^L, defined from price of the bond P_t^L:

$$P_t^L = \sum_{j=1}^{N} \frac{C_{t+j}}{(1+R_t^L)^j}$$

For a **consol** ($N \to \infty$) the yield to maturity is defined as:

$$P_t^L = \sum_{j=1}^{\infty} \frac{C}{(1+R_t^L)^j} = \frac{C}{R_t^L}$$

and the **holding period yield** (from t to $t+1$) is:

$$H_{t+1} = \frac{P_{t+1}^L + C - P_t^L}{P_t^L} = \frac{\frac{1}{R_{t+1}^L} + 1}{\frac{1}{R_t^L}} - 1 = \frac{R_t^L}{R_{t+1}^L} + R_t^L - 1$$

Using a first-order Taylor expansion around $R_t^L = R_{t+1}^L = R^L$ (average long rate over sample):

$$\begin{aligned} \frac{R_t^L}{R_{t+1}^L} &\simeq 1 - \left. \frac{R_t^L}{\left(R_{t+1}^L\right)^2} \right|_{R_t^L = R_{t+1}^L} \cdot (R_{t+1}^L - R_t^L) = 1 - \frac{1}{R^L}\,(R_{t+1}^L - R_t^L) \\ &= 1 + \theta\,(R_t^L - R_{t+1}^L) \qquad \text{with} \;\; \theta = \frac{1}{R^L} \end{aligned}$$

from which

$$H_{t+1} = \theta\,(R_t^L - R_{t+1}^L) + R_t^L$$

Defining the discount factor $\beta \equiv \frac{1}{1+R^L}$:

$$H_{t+1} = \frac{1}{1-\beta}\,R_t^L - \frac{\beta}{1-\beta}\,R_{t+1}^L$$

The **expectations theory** states that investors equate the (expected) holding period yield on long bonds $E_t H_{t+1}$ to the short-term interest rate R_t plus a (potentially time-varying) excess holding period return (risk premium) k_t:

$$E_t H_{t+1} = R_t + k_t$$

yielding

$$R_t^L = \beta\, E_t R_{t+1}^L + (1-\beta)\,(R_t + k_t) \tag{1}$$

Solving forward (imposing convergence):[1]

$$\begin{aligned}
R_t^L &= (1-\beta)\sum_{j=0}^{\infty}\beta^j\left(E_tR_{t+j}+E_tk_{t+j}\right)\\
&= \underbrace{(1-\beta)\sum_{j=0}^{\infty}\beta^j\,E_tR_{t+j}}_{\substack{\text{weighted average of}\\ \text{future expected short rates}}} + \underbrace{(1-\beta)\sum_{j=0}^{\infty}\beta^j\,E_tk_{t+j}}_{\substack{K_t\,=\,\text{weighted average of}\\ \text{future expected risk premia}}}
\end{aligned}$$

Basic implications (assuming constant expected risk premium, $K_t = K$):

1. *permanent and temporary changes in short rates* have different effects on long rates.
   As an example, if the short rate follows an $AR(1)$ process

   $$R_t = \rho R_{t-1} + e_t^R \quad \Rightarrow \quad E_tR_{t+j} = \rho^j R_t$$

   then

   $$\begin{aligned}
   R_t^L &= (1-\beta)\sum_{j=0}^{\infty}\beta^j\left(\rho^jR_t\right) + \underbrace{(1-\beta)\sum_{j=0}^{\infty}\beta^j\,E_tk_{t+j}}_{K}\\
   &= \frac{1-\beta}{1-\beta\rho}R_t + K
   \end{aligned}$$

---

[1] Substituting forward for one period we get:

$$\begin{aligned}
R_t^L &= \beta\,\underbrace{\left[\beta\,E_tR_{t+2}^L + (1-\beta)\left(E_tR_{t+1}+E_tk_{t+1}\right)\right]}_{E_t\,R_{t+1}^L} + (1-\beta)\left(R_t+k_t\right)\\
&\equiv (1-\beta)\left(R_t+k_t\right) + (1-\beta)\,\beta\left(E_tR_{t+1}+E_tk_{t+1}\right) + \beta^2\,E_tR_{t+2}^L
\end{aligned}$$

and after forward substitution for J periods:

$$\begin{aligned}
R_t^L &= (1-\beta)\left(R_t+k_t\right) + (1-\beta)\,\beta\left(E_tR_{t+1}+E_tk_{t+1}\right) + \ldots +\\
&\quad +(1-\beta)\,\beta^J\left(E_tR_{t+J}+E_tk_{t+J}\right) + \beta^{J+1}E_tR_{t+J+1}^L\\
&= (1-\beta)\sum_{j=0}^{J}\beta^J\left(E_tR_{t+J}+E_tk_{t+J}\right) + \beta^{J+1}E_tR_{t+J+1}^L
\end{aligned}$$

If, letting $J\to\infty$, $\lim_{J\to\infty}\beta^{J+1}E_tR_{t+J+1}^L = 0$ (convergence assumption), then we obtain

$$R_t^L = (1-\beta)\sum_{j=0}^{\infty}\beta^j\left(E_tR_{t+j}+E_tk_{t+j}\right)$$

and

$$S_t \equiv R_t^L - R_t = \frac{\beta}{1-\beta\rho}(\rho - 1)R_t + K$$

so that

$$\begin{array}{llllll} \text{if} & \rho & \rightarrow & 1 & \Rightarrow & R_t^L \rightarrow R_t + K \quad \text{and} \quad S_t \rightarrow K \\ \text{if} & \rho & \rightarrow & 0 & \Rightarrow & R_t^L \rightarrow (1-\beta)\, R_t + K \quad \text{and} \quad S_t = -\beta\, R_t + K \end{array}$$

2. the *spread* is an indicator of future changes in short rates:

$$\begin{aligned} R_t^L - R_t &= (1-\beta) \sum_{j=0}^{\infty} \beta^j \left(E_t R_{t+j} - R_t \right) + K \\ &= \sum_{j=1}^{\infty} \beta^j \, E_t \, \Delta R_{t+j} + K \end{aligned}$$

3. testable "*efficient market*" implication: the spread should forecast future changes in the long rate[2]

$$R_t^L - R_{t-1}^L = \frac{1-\beta}{\beta} \left(R_{t-1}^L - R_{t-1} \right) - \frac{1-\beta}{\beta} K + \xi_t$$

   where $\xi_t = R_t^L - E_{t-1} R_t^L$ is the expectational error (that should be orthogonal to any variable in the agents' information set at time $t-1$).

[2]Lagging (1) by one period we get

$$R_{t-1}^L = \beta \, E_{t-1} R_t^L + (1-\beta)\left(R_{t-1} + K \right)$$

and using $R_t^L = E_{t-1} R_t^L + \xi_t$ to substitute for $E_{t-1} R_t^L$ and rearranging:

$$R_t^L = \frac{1}{\beta} R_{t-1}^L - \frac{1-\beta}{\beta} R_{t-1} - \frac{1-\beta}{\beta} K + \xi_t$$

Finally, subtracting R_{t-1}^L from both sides we obtain the above expression for $R_t^L - R_{t-1}^L$.

4 Cointegration and common trends

Implications of the expectation theory for the time-series properties of interest rates and their differential:

$$S_t \equiv R_t^L - R_t = \sum_{j=1}^{\infty} \beta^j \, E_t \underbrace{\Delta R_{t+j}}_{I(0)} + \underbrace{(1-\beta) \sum_{j=0}^{\infty} \beta^j \, E_t k_{t+j}}_{I(0)}$$

if changes in the short rate are stationary and expected risk premia are stationary, then the term spread is stationary:

$$S_t \equiv R_t^L - R_t \sim I(0)$$

The spread can then be tested for stationarity (using unit root tests). More rigorously, it is possible to test for *cointegration* between R^L and R in a **bivariate cointegrated VAR system**.

4.1 Testing for cointegration

Start from a bivariate VAR in the levels of the two interest rates (ignoring constant terms):

$$\begin{pmatrix} R_t \\ R_t^L \end{pmatrix} = \begin{pmatrix} a^*(L) & b^*(L) \\ c^*(L) & d^*(L) \end{pmatrix} \begin{pmatrix} R_{t-1} \\ R_{t-1}^L \end{pmatrix} + \begin{pmatrix} e_t^R \\ e_t^L \end{pmatrix}$$

reparameterized in first differences with the term in levels capturing the long-run cointegration properties of the series:

$$\begin{pmatrix} \Delta R_t \\ \Delta R_t^L \end{pmatrix} = \begin{pmatrix} a(L) & b(L) \\ c(L) & d(L) \end{pmatrix} \begin{pmatrix} \Delta R_{t-1} \\ \Delta R_{t-1}^L \end{pmatrix} + \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} \begin{pmatrix} R_{t-1} \\ R_{t-1}^L \end{pmatrix} + \begin{pmatrix} e_t^R \\ e_t^L \end{pmatrix}$$

Imposing cointegration between R^L and R with cointegrating vector $(1, -1)$ we obtain the *vector error-correction* $(VECM)$ representation

$$\begin{pmatrix} \Delta R_t \\ \Delta R_t^L \end{pmatrix} = \begin{pmatrix} a(L) & b(L) \\ c(L) & d(L) \end{pmatrix} \begin{pmatrix} \Delta R_{t-1} \\ \Delta R_{t-1}^L \end{pmatrix} + \begin{pmatrix} \alpha_R \\ \alpha_L \end{pmatrix} (R_{t-1}^L - R_{t-1}) + \begin{pmatrix} e_t^R \\ e_t^L \end{pmatrix}$$

Under cointegration the term in levels is important in determining the (error-correction) interest rate dynamics and the $VECM$ should fit better the data than a simple VAR in first differences (with no term in levels). A formal test of the better performance of the $VECM$ (and therefore of cointegration between R^L and R) may be carried out using a *likelihood ratio test* (comparing the values of the log-likelihood of the $VECM$ and the VAR).

4.2 Deriving the common stochastic trend

If the two rates are cointegrated, then they share a *common stochastic trend*. According to the popular characterization of the stochastic trend given by Beveridge and Nelson (*Journal of Monetary Economics* 1981),[3] for a non stationary interest rate R_t, the stochastic trend $\bar{R}_t$ can be defined as the long-run forecast, i.e.

$$\bar{R}_t = \lim_{h\to\infty} E_t\, R_{t+h}$$

or, in terms of expected future changes (since $R_{t+h} = R_{t-1} + \Delta R_t + \Delta R_{t+1} + ... + \Delta R_{t+h}$):

$$\bar{R}_t = R_{t-1} + \sum_{j=0}^{\infty} E_t\, \Delta R_{t+j}$$

Expected future (long and short) interest rate changes may be obtained by means of the $VECM$ system using the following procedure, applied to the case of one lag only in the first difference terms, that is

$$\begin{aligned} \Delta R_t &= a\, \Delta R_{t-1} + b\, \Delta R^L_{t-1} + \alpha_R S_{t-1} + e^R_t \\ \Delta R^L_t &= c\, \Delta R_{t-1} + d\, \Delta R^L_{t-1} + \alpha_L S_{t-1} + e^L_t \end{aligned}$$

This $VECM$ can be expressed as a $VAR(1)$ system for the three stationary variables ΔR_t, ΔR^L_t, and S_t, after deriving S_t from

$$\begin{aligned} \Delta S_t &\equiv \Delta R^L_t - \Delta R_t \\ &= (c-a)\, \Delta R_{t-1} + (d-b)\, \Delta R^L_{t-1} + (\alpha_L - \alpha_R)\, S_{t-1} + (e^L_t - e^R_t) \\ \Rightarrow\ S_t &= (c-a)\, \Delta R_{t-1} + (d-b)\, \Delta R^L_{t-1} + (1 + \alpha_L - \alpha_R)\, S_{t-1} + (e^L_t - e^R_t) \end{aligned}$$

Augmenting the original VAR with S_t we obtain:

$$\begin{pmatrix} \Delta R_t \\ \Delta R^L_t \\ S_t \end{pmatrix} = \begin{pmatrix} a & b & \alpha_R \\ c & d & \alpha_L \\ c-a & d-b & 1+\alpha_L-\alpha_R \end{pmatrix} \begin{pmatrix} \Delta R_{t-1} \\ \Delta R^L_{t-1} \\ S_{t-1} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} e^R_t \\ e^L_t \end{pmatrix}$$

or more compactly, with $\mathbf{x}'_t = \begin{pmatrix} \Delta R_t & \Delta R^L_t & S_t \end{pmatrix}$ and $\mathbf{e}'_t = (e^R_t \;\; e^L_t)$:

$$\mathbf{x}_t = \mathbf{M}\, \mathbf{x}_{t-1} + \mathbf{G}\, \mathbf{e}_t$$

[3]Beveridge and Nelson showed that any (univariate) non-stationary series modelled as an $ARIMA(p, 1, q)$ process can be decomposed into a permanent component (given by the stochastic trend and represented by a random walk, possibly with drift) and a transitory (stationary) component. Stock and Watson (*Journal of the American Statistical Asssociation* 1988) extended this decomposition to a multivariate cointegrated framework.

Expected future values of $\mathbf{x}$ are then computed as

$$E_t \mathbf{x}_{t+j} \equiv E(\mathbf{x}_{t+j} \mid \mathbf{x}_t) = \mathbf{M}^j \, \mathbf{x}_t$$

To select the expected value of the first two elements of vector $\mathbf{x}$, corresponding to ΔR and ΔR^L, we make use of the two "selection" vectors $\mathbf{h}'_R = (1\ 0\ 0)$ and $\mathbf{h}'_L = (0\ 1\ 0)$, obtaining

$$\begin{aligned} E_t \Delta R_{t+j} &\equiv E(\Delta R_{t+j} \mid \mathbf{x}_t) = \mathbf{h}'_R \, \mathbf{M}^j \, \mathbf{x}_t \\ E_t \Delta R^L_{t+j} &\equiv E(\Delta R^L_{t+j} \mid \mathbf{x}_t) = \mathbf{h}'_L \, \mathbf{M}^j \, \mathbf{x}_t \end{aligned}$$

The stochastic trends of the short and long interest rates can then be computed as

$$\begin{aligned} \bar{R}_t &= R_{t-1} + \mathbf{h}'_R \, [\mathbf{I} - \mathbf{M}]^{-1} \, \mathbf{x}_t \\ \bar{R}^L_t &= R^L_{t-1} + \mathbf{h}'_L \, [\mathbf{I} - \mathbf{M}]^{-1} \, \mathbf{x}_t \end{aligned}$$

where use has been made of (for the short-rate and similarly for the long-rate):

$$\begin{aligned} \sum_{j=0}^{\infty} E_t \, \Delta R_{t+j} &= \sum_{j=0}^{\infty} \mathbf{h}'_R \, \mathbf{M}^j \, \mathbf{x}_t \\ &= \mathbf{h}'_R \left(\sum_{j=0}^{\infty} \mathbf{M}^j \right) \mathbf{x}_t = \mathbf{h}'_R \, [\mathbf{I} - \mathbf{M}]^{-1} \, \mathbf{x}_t \end{aligned}$$

Finally, given cointegration between R^L and R, the spread S is stationary and its long-run forecast $\bar{S}$ is simply related to the stochastic trend in the long and short rates as follows:

$$\bar{S}_t = \bar{R}^L_t - \bar{R}_t = \lim_{h \to \infty} E_t \, R^L_{t+h} - \lim_{h \to \infty} E_t \, R_{t+h} = \lim_{h \to \infty} E_t \, S_{t+h} = K$$

where K denotes the unconditional mean of the spread (capturing the term premium).

The behaviour of the short-term and long-term interest rates can then be decomposed into a *permanent* and a *transitory* components, the former being the stochastic trend ($\bar{R}_t$ and $\bar{R}^L_t$), and the latter simply derived as $R_t - \bar{R}_t$ and $R^L_t - \bar{R}^L_t$. This decomposition allows to relate the fluctuations of the spread around its unconditional mean ($K = \bar{S}$) to the *temporary* components of the long and short rates (since permanent shifts given by the common stochastic trend move both interest rates by the same amount with no effect on S), exploiting the identity:

$$\begin{aligned} S_t - \bar{S} &\equiv \left(R^L_t - R_t\right) - (\bar{R}^L_t - \bar{R}_t) \\ &= (R^L_t - \bar{R}^L_t) - (R_t - \bar{R}_t) \end{aligned}$$

from which we get

$$var(S_t) = var(R_t^L - \bar{R}_t^L) + var(R_t - \bar{R}_t) - 2\,cov\left[(R_t^L - \bar{R}_t^L), (R_t^L - \bar{R}_t^L)\right]$$

Note: Panel C of Table 5 in the paper reads (correlation coefficients in bold below the diagonal):

Long-Short Spread S		
Total	Temporary R^L	Temporary R
1.93	0.56	-1.37
0.96	0.18	-0.38
-0.99	**-0.91**	0.98

5 Rational expectations tests

The expectations theory of the term structure of interest rates is an example of a more general class of theories that make use of the assumption of rational expectations. Such rational expectations models have the common feature of imposing restrictions across the equations describing the time-series dynamics of the variables involved ("cross-equation restrictions").

To see this point in the context of the term structure of interest rates, let us consider a $VAR(4)$ model for the change in the short rate and the term spread:

$$\begin{aligned} \Delta R_t &= \sum_{i=1}^{4} a_i\, \Delta R_{t-i} + \sum_{i=1}^{4} b_i\, S_{t-i} + e_t^{\Delta R} \\ S_t &= \sum_{i=1}^{4} c_i\, \Delta R_{t-i} + \sum_{i=1}^{4} d_i\, S_{t-i} + e_t^{S} \end{aligned}$$

Under the assumption of a constant term premium (dropped for simplicity), the expectations theory implies that the spread is

$$S_t = \sum_{j=1}^{\infty} \beta^j\, E_t\, \Delta R_{t+j}$$

which imposes a set of testable cross-equation restrictions on the VAR model for ΔR and S.

5.1 Deriving the restrictions on the VAR

Under the expectations hypothesis, the current spread reflects future changes in the short rate expected on the basis of the full information set available to agents in the market, $\mathbf{\Omega}_t$. In order to generate expected values of future ΔR, the econometrician uses an information set, $\boldsymbol{\omega}_t$, which is (necessarily) a subset of the market's set: $\boldsymbol{\omega}_t \subset \mathbf{\Omega}_t$. The *law of iterated expectations* is used to prove that the above formula for the spread is unchanged even if the information set used to form expectations is reduced.

In fact, given $\boldsymbol{\omega}_t \subset \mathbf{\Omega}_t$ we have

$$E\left[E\left(\Delta R_{t+j} \mid \mathbf{\Omega}_t\right) \mid \boldsymbol{\omega}_t\right] = E\left(\Delta R_{t+j} \mid \boldsymbol{\omega}_t\right)$$

Applying this property to the term spread formula (conditioning both sides of the equation on the limited econometrician's information set $\boldsymbol{\omega}_t$) we get, using the fact that $S_t \subset \boldsymbol{\omega}_t$:

$$\begin{aligned} E\left(S_t \mid \boldsymbol{\omega}_t\right) &= \sum_{j=1}^{\infty} \beta^j\, E\left[E\left(\Delta R_{t+j} \mid \mathbf{\Omega}_t\right) \mid \boldsymbol{\omega}_t\right] \\ \Rightarrow \quad S_t &= \sum_{j=1}^{\infty} \beta^j\, E\left(\Delta R_{t+j} \mid \boldsymbol{\omega}_t\right) \end{aligned}$$

Therefore, according to the expectations theory, the current spread is a discounted sum of future changes of the short rate expected on the basis of the limited information set available to the econometrician $\boldsymbol{\omega}_t$. In the context of the VAR model above, $\boldsymbol{\omega}_t$ contains the current and lagged values of ΔR_t and S_t. The expectations of future ΔR_{t+j} can then be constructed using the so-called "companion form" of the $VAR(4)$ model, written as a $VAR(1)$:

$$\begin{pmatrix} \Delta R_t \\ \Delta R_{t-1} \\ \Delta R_{t-2} \\ \Delta R_{t-3} \\ S_t \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \end{pmatrix} = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & b_1 & b_2 & b_3 & b_4 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ c_1 & c_2 & c_3 & c_4 & d_1 & d_2 & d_3 & d_4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \Delta R_{t-1} \\ \Delta R_{t-2} \\ \Delta R_{t-3} \\ \Delta R_{t-4} \\ S_{t-1} \\ S_{t-2} \\ S_{t-3} \\ S_{t-4} \end{pmatrix} + \begin{pmatrix} e_t^{\Delta R} \\ 0 \\ 0 \\ 0 \\ e_t^S \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

or more compactly:

$$\boldsymbol{\omega}_t = \mathbf{M}\, \boldsymbol{\omega}_{t-1} + \mathbf{e}_t$$

From this VAR, expectations are easily formed as:

$$E\left(\boldsymbol{\omega}_{t+j} \mid \boldsymbol{\omega}_t\right) = \mathbf{M}^j\, \boldsymbol{\omega}_t$$

Using the "selection" vector $\mathbf{h}'_{\Delta R} = (1 \;\; 0 \;\; ... \;\; 0)$ expected future changes in the short rate are derived as

$$E\left(\Delta R_{t+j} \mid \boldsymbol{\omega}_t\right) = \mathbf{h}'_{\Delta R}\, \mathbf{M}^j\, \boldsymbol{\omega}_t$$

Using these expectations in the formula for the spread implied by the theory and using the "selection" vector $\mathbf{h}_S$ (with 1 in the position corresponding to S_t and 0 elsewhere) to express S_t as $\mathbf{h}'_S\; \boldsymbol{\omega}_t$,we get:

$$\begin{aligned} S_t = \mathbf{h}'_S\; \boldsymbol{\omega}_t &= \sum_{j=1}^{\infty} \beta^j\, \mathbf{h}'_{\Delta R}\, \mathbf{M}^j\, \boldsymbol{\omega}_t = \mathbf{h}'_{\Delta R} \left(\sum_{j=1}^{\infty} \beta^j\, \mathbf{M}^j \right) \boldsymbol{\omega}_t \\ &= \mathbf{h}'_{\Delta R}\, \beta\, \mathbf{M}\, \left[\mathbf{I} - \beta\, \mathbf{M}\right]^{-1}\, \boldsymbol{\omega}_t \end{aligned}$$

$$\Rightarrow \qquad \mathbf{h}'_S = \mathbf{h}'_{\Delta R}\, \beta\, \mathbf{M}\, \left[\mathbf{I} - \beta\, \mathbf{M}\right]^{-1}$$

This is a set of 8 (non-linear) cross-equation restrictions on the 16 coefficients of the bivariate VAR (elements of the rows of $\mathbf{M}$ corresponding to ΔR_t and ΔS_t) imposed by the expectations theory.[4] To see the restrictions more clearly we can make them linear by multiplying both sides of the equation by $[\mathbf{I} - \beta\, \mathbf{M}]$, getting

$$\mathbf{h}'_S\, \left[\mathbf{I} - \beta\, \mathbf{M}\right] = \mathbf{h}'_{\Delta R}\, \beta\, \mathbf{M}$$

which corresponds to the following set of 8 restrictions on the VAR parameters:

$$\begin{aligned} a_i &= -c_i \quad \text{for } i = 1, 2, 3, 4 \\ b_1 &= \frac{1}{\beta} - d_1 \\ b_i &= -d_i \quad \text{for } i = 2, 3, 4 \end{aligned}$$

This set of restrictions can be tested after estimating (only) the unrestricted VAR by means of a *Wald test*, or by means of a *likelihood ratio test* after estimating both the unrestricted and the restricted forms of the VAR (in both cases a value for $1/\beta \equiv 1 + R^L$ must be assumed; in the paper a value of 1.0056, corresponding to a sample average monthly long rate of 0.56%, yelding an average annual rate of 6.67%).

[4] More generally, given a bivariate VAR in first differences with p lags, the total number of parameters to be estimated is $4\,p$ and the number of cross-equations restrictions imposed by the (rational) expectations theory of the term structure is $2\,p$.

6 Expectations and the spread

Very often, empirical tests of the cross-equations restrictions imposed by the (pure) expectations theory on the bivariate dynamics of the short rate and the spread yield strong rejections. Even the allowance of some time-variation in the term premium (instead of assuming that it is constant over time as under the pure expectations theory) does not rescue the theory in empirical tests. Nonetheless, the expectations theory is important to understand the behaviour of interest rates and to explain fluctuations in the spread.

One way to assess the merits of the theory is, after relaxing the assumption of a time-invariant term spread, to construct the component of the spread that can be explained by expectations of future changes in short rates and evaluate its correlation with the spread itself; in the process, an estimate of the term premium is derived as the difference between the spread and the expectations component.

6.1 Decomposing the spread

Introducing a time-varying term premium, the expectations theory predicts that the spread is the sum of the following two components:

$$S_t = \underbrace{\sum_{j=1}^{\infty} \beta^j \, E_t \, \Delta R_{t+j}}_{\substack{\text{expectations} \\ \text{component}}} + \underbrace{(1-\beta) \sum_{j=0}^{\infty} \beta^j \, E_t k_{t+j}}_{\substack{\text{term premium} \\ \text{component}}}$$

where expectations are formed relative the the full market's information set $\mathbf{\Omega}_t$. Conditioning onto the (smaller) econometrician's information set $\boldsymbol{\omega}_t \subset \mathbf{\Omega}_t$, it is possible to construct a measure of the expectations component (so-called "theoretical spread") and (subtracting this estimate from the actual spread) of the term premium component. Using the current (i.e. dated t) information set $\boldsymbol{\omega}_t$ and the bivariate VAR model of the previous section, the *expectations component* is

$$\sum_{j=1}^{\infty} \beta^j \, E\left(\Delta R_{t+j} \mid \boldsymbol{\omega}_t\right) = \mathbf{h}'_{\Delta R} \, \beta \, \mathbf{M} \, \left[\mathbf{I} - \beta \, \mathbf{M}\right]^{-1} \, \boldsymbol{\omega}_t$$

and the *term premium* component is derived as

$$(1-\beta) \sum_{j=0}^{\infty} \beta^j \, E\left(k_{t+j} \mid \boldsymbol{\omega}_t\right) = S_t - \mathbf{h}'_{\Delta R} \, \beta \, \mathbf{M} \, \left[\mathbf{I} - \beta \, \mathbf{M}\right]^{-1} \, \boldsymbol{\omega}_t$$

In the data the expectations component ("theoretical spread") is very much correlated with the actual spread (with a correlation coefficient of 0.99), pointing to some validity of the expectations theory of the term structure.

Moreover, the variance of the spread can be decomposed (in a different but related way as that presented in section 4.2) as:

$$var(S_t) = var\left(\sum_{j=1}^{\infty} \beta^j \, E_t \, \Delta R_{t+j}\right) + var\left((1-\beta)\sum_{j=0}^{\infty} \beta^j \, E_t k_{t+j}\right) + 2cov(.,.)$$

In the data, although the variance of the term-premium component is not very large, the covariance term is positive and relatively high: periods of high expectations component occur when the term premium is also high, so that the variability of the spread is significantly higher than that of the expectations component only.

7 Conclusions

Several conclusions can be drawn from the results of the paper:

1. the levels of the short-term and long-term interest rates (non-stationary series) are driven by a **common stochastic trend**; the term spread is stationary (cointegrating relation)

2. changes in the **long rate** (ΔR^L) are closely correlated with changes in the stochastic trend: R^L is a **good indicator of the** (common) **stochastic trend** in interest rates. Changes in the short rate (ΔR) are much less correlated with changes in the stochastic trend, being driven instead by important temporary fluctuations

3. the **spread** (not affected by the common stochastic trend in interest rates) is a **good indicator of changes in the temporary component of short-term rates**

4. the **rational expectations** (cross-equation) **restrictions** imposed by the expectations theory of the term structure are **rejected** by formal statistical tests **but the theory provides nevertheless an useful approximation for practical purposes**.