Level, Slope, Curvature: Characterising the Yield Curve in a Cointegrated VAR Model

Julia V. Giese

Nuffield College, University of Oxford

Abstract:

Empirical evidence on the expectations hypothesis of the term structure is in-conclusive and its validity widely debated. Using a cointegrated VAR model of US treasury yields, this paper extends a common approach to test the theory. If, as we find, spreads between two yields are non-stationary, the expectations hypothesis fails. However, we present evidence that *differences* between two spreads are stationary. This suggests that the curvature of the yield curve may be a more meaningful indicator of expected future interest rates than the slope. Furthermore, we characterise level and slope by deriving the common trends inherent in the cointegrated VAR, and establish feedback patterns between them and the macroeconomy.

1   Introduction

Investigating relations between yields of different maturities was one of the first applications of cointegration analysis (see, for example, Engle and Granger (1987), or Campbell and Shiller (1987)). The initial bivariate approach was extended to the multivariate case by Hall, Anderson, and Granger (1992) among others. The study concentrated on a set of short-term maturities and found one common trend. This was believed to corroborate the expectations theory of the term structure, which says that a longer-term bond rate is just the average of expected one-period rates for the duration of the bond plus some constant term premium. According to this hypothesis, the spreads between different maturities make up the cointegrating vectors. Given only one common trend, it was concluded that the term premia therefore must be mean-reverting if not constant.

Shea (1992) examined a broader set of yields, including long-term maturities up to twenty-five years. His results support the findings of Hall, Anderson, and Granger (1992) for the short end of the yield curve but reject stationarity of the spreads between longer-term maturities. Building on Shea (1992), other researchers find up to three common trends when including yields of longer maturities. Zhang (1993) demonstrates this for US data while Carstensen (2003) looks at German data. They argue that their findings suggest that term premia are in fact non-stationary and that additional common trends have interpretations familiar from factor models of the yield curve.

This paper seeks to extend past understanding of driving forces behind the yield curve by employing a cointegrated vector auto-regression (CVAR) model on monthly data of US treasury zero-coupon yields over the period 1987 to 2000.[1] We show that there is strong evidence for two common trends, implying that not all independent spreads can be stationary.

However, weighted differences between pairs of spreads are found to be stationary, and hence two term premia cointegrate. This suggests that while investors' preferences with respect to a certain maturity vary over time without reverting back to a mean, their relative preferences between two maturities are stationary. A conclusion from this finding is that we should look at the curvature of the yield curve (approximated by the weighted difference between two spreads) if we are interested in the interest rate expectations embodied in the term structure. It enables policymakers to deduce whether the rate of change in interest rates is expected to diminish or increase in the long run compared with the medium run. The finding may also be interesting for traders trading on mean-reversion properties of the yield curve.

Our analysis of the yield levels' non-stationary common trends through the Granger-Johansen representation, introduced in Engle and Granger (1987) and extended by Johansen (e.g. Johansen (1996)), confirms the results. It is, as far as we know, a novel application to the term structure of interest rates: both Zhang (1993) and Carstensen (2003) arrive at their conclusions using factor representations. We

---

[1]The data was kindly provided by Diebold and Li and uses the unsmoothed Fama-Bliss methodology to construct the zero-coupon series. See Bliss (1997), and Fama and Bliss (1987), for details. To my knowledge, this particular series has not been updated to include more recent years. We use it nonetheless because the way it was constructed is particularly suited for the purposes at hand.

find that one common trend acts on the level of the yield curve and the other on the slope, giving them interpretations of a level and slope factor.

Hence, this paper shows empirically that the common trend analysis is related to common factor models, often used in financial economics to model the yield curve. This literature finds up to three factors to be sufficient to explain the yield curve's shape, often level, slope and curvature.[2] Rather than relying on assessments of explanatory power, cointegration theory provides powerful and thoroughly understood methods for doing inference on the number of cointegrating relations and thus directly on the number of common trends. In addition, the analysis of the Granger-Johansen representation allows us to characterise the driving forces behind each common trend, and to link them with macroeconomic variables.

The paper is structured as follows: In Section 2 we develop the theoretical model based on Hall, Anderson, and Granger (1992). Section 3 introduces the CVAR and presents the cointegration analysis. Section 4 concludes. The computations were made using *CATS* and *PcGive*.[3]

2 Theoretical Framework

Define b_t^m as the yield at time t of a zero-coupon bond with maturity m, $m = 1, 2, 3, \ldots$. Similarly, let the forward rate at time t of period j be f_t^j, giving the linearised no-arbitrage condition[4]

$$b_t^m = \frac{1}{m} \sum_{j=1}^{m} f_t^j. \tag{1}$$

The term premium, l_t^j, is defined as

$$f_t^j = E_t \left(b_{t+j-1}^1 \right) + l_t^j, \tag{2}$$

and is paid for the release from risk in contracting for a future debt immediately. Combining (1) and (2) and setting $\frac{1}{m} \sum_{j=1}^{m} l_t^j = L_t^m$, we get

$$b_t^m = \frac{1}{m} \sum_{j=1}^{m} \left[E_t \left(b_{t+j-1}^1 \right) + l_t^j \right] = \frac{1}{m} \sum_{j=1}^{m} E_t \left(b_{t+j-1}^1 \right) + L_t^m. \tag{3}$$

Equation (3) can be interpreted in terms of the expectations hypothesis. Its pure version would suggest that the term premium, L_t^m, is zero, allowing the yield to maturity only to be determined by expectations of future short-term yields. A constant

---

[2]See among others Litterman and Scheinkman (1991), Knez, Litterman, and Scheinkman (1994), Nelson and Siegel (1987), as well as Duffee (2002), for discussions of different factor models. The former two only place structure on the factors and not the loadings, e.g. using principal components. Nelson and Siegel (1987) introduce a latent factor model where the factors are unobserved but the loadings represent level slope and curvature. The latter uses an affine latent factor model, imposing a no-arbitrage condition.

[3]See Dennis (2006) on *CATS* and Doornik and Hendry (2007) on *PcGive*.

[4]The relationship in (1) is an approximation derived from taking logs of $b_t^m = \left[(1 + f_t^1)(1 + f_t^2) \ldots (1 + f_t^m) \right]^{\frac{1}{m}} - 1$.

term premium is consistent with a less strict interpretation, while a stationary term premium is more flexible still. We will concentrate exclusively on the latter, most generous, version of the hypothesis and find evidence in the data to reject even it.

Rearranging Equation (3) gives a convenient representation in terms of the spread between yields of different maturities which is tested empirically in this paper, i.e.

$$
\begin{aligned}
b_t^m - b_t^1 &= (\frac{1}{m} - 1)b_t^1 + \frac{1}{m} \sum_{j=1}^{m-1} E_t\left(b_{t+j}^1\right) + L_t^m \\
&= \frac{1}{m} E_t(\Delta b_{t+1}^1 + \Delta b_{t+2}^1 + \Delta b_{t+1}^1 + \Delta b_{t+3}^1 + \Delta b_{t+2}^1 + \Delta b_{t+1}^1 + ... \\
&\quad + \Delta b_{t+m-1}^1 + \Delta b_{t+m-2}^1 + \Delta b_{t+m-3}^1 + ...) + L_t^m \\
&= \frac{1}{m} \sum_{j=1}^{m-1} (m - j) E_t\left(\Delta b_{t+j}^1\right) + L_t^m.
\end{aligned} \tag{4}
$$

Since bond yields are well approximated by processes integrated of at most order one (I(1)), their differences are integrated of order zero (I(0)) and the first term on the right hand side of Equation (4) is stationary. If the term premium was stationary, we would expect the spreads to be I(0) because a process determined by two stationary processes is itself stationary. On the other hand, non-stationary spreads found in the data would imply a non-stationary term premium.

Extending the framework to weighted differences between spreads, Equation (5) shows that if we find the spreads to be pairwise cointegrating, the weighted differences between the term premia of differing maturities have to be stationary:[5]

$$
\left(b_t^m - b_t^n\right) - c\left(b_t^n - b_t^1\right) =
$$

$$
\frac{1}{m} \sum_{j=1}^{m-1} (m - j) E_t\left(\Delta b_{t+j}^1\right) - \frac{1+c}{n} \sum_{j=1}^{n-1} (n - j) E_t\left(\Delta b_{t+j}^1\right) + L_t^m - (1 + c)L_t^n, \tag{5}
$$

where c is a constant weight.

A deviation in a mean-reverting process is informative because we can judge the observation against its long-run equilibrium while we do not have a point of reference for a non-stationary process. In the case of non-stationary spreads driven by a non-stationary term premium, the stationary part of the process is indeterminable. However, given cointegrated term premia, a deviation from the typical curvature may well reflect changes in future interest rate expectations, rather than changes in preferences. In this case, the yield curve's curvature may be used as an indication of expectations on the future path of interest rates.

---

[5]A curve going through three points A, B and C with $A \geq B \geq C$, can be described by a quadratic whose curvature, the second derivative, is given by $-\frac{1}{2}((A - B) - (B - C))$ if the distance between A and B, $\overline{AB}$, is the same as that between B and C, $\overline{BC}$. However, when distances between points are not equal, as is the case for maturities considered here, the curvature is the weighted expression $\frac{1}{M+N}(N(A - B) - M(B - C))$ where M is $\overline{AB}$ and N is $\overline{BC}$, i.e. a weighted rather than exact difference between spreads should be characteristic of the curvature of the yield curve.

In the following analysis we show how a CVAR can be used to test the theoretical models in terms of stationarity and thus assess the expectation theory of the term structure in its conventional notation and the extension discussed above.

3 A CVAR Model

Our model consists of monthly end-of-period yields for US treasury zero-coupon bonds of five different maturities, namely for the one-month, three-month, eighteen-month, four-year and ten-year bonds. The choice of variables reflects the structure of the yield curve with very short-term as well as medium- and long-term maturities. Zhang (1993) includes 19 yields of different maturities in the initial analysis. Given the nature of VAR models, however, a smaller set with still many parameters should suffice here. Future work should examine the robustness of the results with respect to the dimension of the system and the choice of maturities included in the analysis.

3.1 Properties of the Data

The subsequent analysis will focus on the time span of Alan Greenspan's chairmanship of the Fed from August 1987 to the end of our sample in December 2000. Samples that reach further back exhibit problems of non-constancy in the parameters when analysed in a linear CVAR framework. This is in line with Baba, Hendry, and Starr (1992) who find in their study of US money demand that even risk-adjusted spreads may change when yields reach previously unknown levels as they did under the "new operating procedures" in the late 1970s and early 1980s. Giese (2006) extends the present analysis to a sample beginning in January 1970, by fitting a Markov-switching model that allows for regime changes. Results from that model very closely resemble the ones presented here.

A graphical analysis of Figure 1 suggests the zero-coupon series of the one-month, eighteen-month and ten-year yields differ not so much in levels – although the ten-year yield is typically higher than yields of short-term maturities – but more in their differences which show that the long-term yields are less variable. According to the theory presented above, longer-term yields can at least partly be explained by the average expected spot interest rates of all periods to maturity and thus contain information on the shorter end of the yield curve. This aggregation implies that long maturities are less affected by temporary shocks and that the plots of the long-term yields look smoother than those of short-term maturities. The bond yields also appear to be highly persistent, well approximated by I(1) processes.

3.2 Cointegration Analysis

We begin the empirical analysis with a formal definition of the statistical concept to be used, i.e. the CVAR or vector equilibrium correction model (VECM(k)):

$$\Delta x_t = (\Pi, \mu_0) \begin{pmatrix} x_{t-1} \\ 1 \end{pmatrix} + \Gamma_1 \Delta x_{t-1} + ... + \Gamma_{k-1} \Delta x_{t-k+1} + \phi D_t + \varepsilon_t, \quad \varepsilon_t \sim iidN_p(0, \Omega),$$
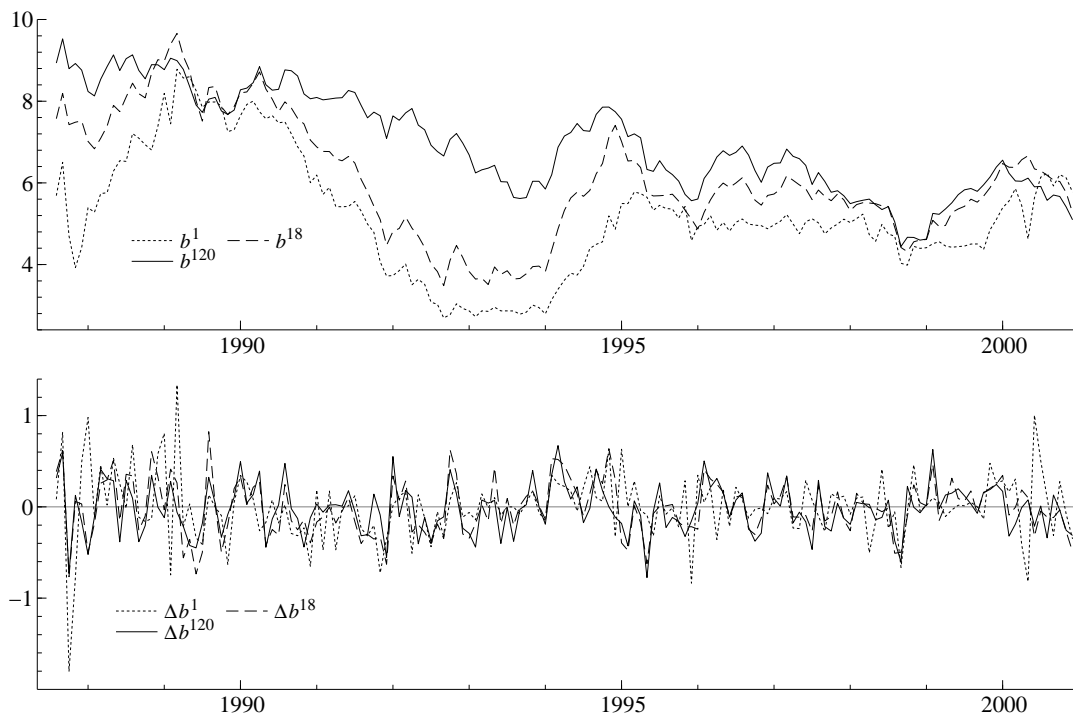
Figure 1: Levels and differences of yields for three maturities, 1987:08-2000:12

with p endogenous variables $x_t = (b_t^1, b_t^3, b_t^{18}, b_t^{48}, b_t^{120})'$ where b_t^m represents the yield at time t of a zero-coupon bond with m months to maturity. The constant μ_0 is restricted to the cointegrating space and D_t is a vector of three impulse dummy variables to allow for one-off shocks unexplained by the variables in the model and not reconcilable with the assumption of normality in the residuals. The dummies take the value one in the months February 1989, December 1990 and May 2000.[6] Since yield levels are non-stationary, Π is of reduced rank r, and we can write $(\Pi, \mu_0) = \alpha(\beta', \beta_0)$.

3.2.1   Links to the Theoretical Model

The theoretical framework presented in Section 2 is tested by determining r which defines the number of stationary cointegrating relations $\beta' x_t$ and non-stationary common trends $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$, and by testing explicit hypotheses on parameters in these relations. In our case where $p = 5$, rank $r = 4$ is equivalent to four cointegrating relations and $p-r = 1$ common trend. This is implied by the expectations hypothesis in its weakest form: according to Equation (4), a stationary term premium can only hold if the four independent spreads between five yields are stationary and thus form cointegrating relations. The studies cited in the Introduction test for this case.

   If $r = 3$, however, there would be two common trends which implies that all

---

[6]Given the length of the sample, we include dummies for standardised residuals which are greater than 3.5. This value is found by determining the critical value corresponding to the probablity of having an outlier in a particular sample size.

Table 1: LR tests for lag determination and restricted short-run structure

Lag deletions	Test statistic			p-value	
$k = 5 \to k = 4$	$\chi^2(25) = 17.022$			[0.881]	
$k = 5 \to k = 3$	$\chi^2(50) = 85.363$			[0.001]	
Γ-restrictions	Δb^1	Δb^3	Δb^{18}	Δb^{48}	Δb^{120}
Γ_1	1	1	0	0	0
Γ_2	0	0	1	1	0
Γ_3	1	1	1	1	1

four spreads cannot be stationary, and the expectations hypothesis even loosely defined fails. Nevertheless, interesting conclusions can still be drawn, for example, the weighted differences between two spreads may be stationary, as formulated in Equation (5). The analysis is thus extended to examine the stationarity of the yield curve's derivatives. While the cointegrating relations can be used to assess the stationarity of the level, slope and curvature, so can the common trends: We find that they represent the non-stationary derivatives of the yield curve. Thus, in the case of three common trends ($r = 2$), even the curvature could be non-stationary.

3.2.2   Specification of the Statistical Model

Before determining the rank r of Π, we choose the lag length k of the model. Nielsen (2006) shows that the location of the characteristic roots is of no consequence when testing for the lag order. The choice of lag length k reflects the persistence of short-run effects, and is determined by a likelihood ratio (LR) test that compares two hypotheses on different lag lengths according to residual auto-correlation. The test statistic is given by

$$\mathrm{LR}(\mathcal{H}_k|\mathcal{H}_{k+i}) = -2\ln Q(\mathcal{H}_k/\mathcal{H}_{k+i}) = T(\ln |\hat{\Omega}_k| - \ln |\hat{\Omega}_{k+i}|), \tag{6}$$

where $\mathcal{H}_k$ is the null hypothesis that there are k lags, while $\mathcal{H}_{k+i}$ is the alternative hypothesis that $k+i$ lags are needed. $\hat{\Omega}_k$ is the estimated variance-covariance matrix of the residuals. The results obtained for our model are presented in Table 1. They give strong evidence for $k = 4$, which means that three Γ_i matrices need to be estimated in the VECM.

Since $k = 4$ involves many parameters, we set insignificant columns in the Γ_is to zero. The resulting short-run structure is also shown in Table 1, where an entry of 1 stands for a fully estimated column, and an entry of 0 for a column with all entries restricted to zero. These restrictions are rejected in an LR test with a p-value of 0.001 ($\chi^2(30) = 59.144$), but not at the 1-percent level in an F-test (p-value: 0.012, F(30,518) = 1.705). In addition, the Schwartz, Hannan-Quinn and Akaike information criteria support the restricted model, and hence we continue with the restrictions on Γ_i imposed.

Furthermore, we test the model for mis-specification, in particular whether the residuals are consistent with the errors behaving according to $\varepsilon_t \sim iid\ N_p(0, \Omega)$. A

Table 2: Vector mis-specification tests

Lag	Test statistic	p-value
LM tests for no auto-correlation:		
1	$\chi^2(25) = 24.444$	[0.494]
	F(25,484) = 0.977	[0.497]
2	$\chi^2(50) = 79.368$	[0.005]
	F(50,573) = 1.483	[0.020]
3	$\chi^2(75) = 107.75$	[0.008]
	F(75,578) = 1.329	[0.041]
4	$\chi^2(100) = 133.57$	[0.014]
	F(100,565)= 1.234	[0.075]
Test for normality:		
	$\chi^2(10) = 24.514$	[0.006]
LM tests for no ARCH-effects:		
1	$\chi^2(225) = 261.746$	[0.047]
2	$\chi^2(450) = 474.927$	[0.201]
3	$\chi^2(675) = 700.646$	[0.240]
4	$\chi^2(900) = 989.998$	[0.019]

discussion of parameter constancy is left until after the system has been identified. Table 2 shows that there may be a problem with auto-correlation: the null hypothesis is rejected for lags 1-2 and 1-3 in the χ^2 but not the F form. However, none of the single-equation tests (not reported) is rejected, suggesting that the problem is not serious. The vector test for normality, also shown in Table 2, suggests that there may be a minor problem with normality, but for individual yields normality is not rejected (test results not shown here). Also, cointegration results have been found to be quite robust to moderate degrees of excess kurtosis (see Gonzalo (1994)). ARCH effects can be rejected, broadly supporting the null hypothesis of homoskedasticity.

3.2.3 Determination of the Cointegration Rank

The trace test seeks to determine which eigenvalues correspond to stationary and which to non-stationary relations. A small eigenvalue indicates a unit root and thus at least a very persistent and possibly non-stationary process. In Table 3 we report the test results, where starred trace statistics and p-values are corrected by the Bartlett factor for small sample size and λ_{r+1} denotes the smallest eigenvalue of rank $r + 1$.

Our economic prior of $r = 4$ is not rejected, implying that there is at least one eigenvalue – 0.025 – that is not statistically different from zero. However, the next smallest eigenvalue has a magnitude of only 0.047, which the test also finds to be not statistically different from zero, supporting $r = 3$. To investigate further whether the data include one or two non-stationary trends, we use information from other indicators.

Table 3: Trace test of the cointegration rank

$p-r$	r	λ_{r+1}	**trace**	**trace***	**5% cv**	*p*-value	*p*-value*
5	0	0.433	184.577	163.903	76.813	0.000	0.000
4	1	0.306	95.453	85.634	53.945	0.000	0.000
3	2	0.156	38.150	32.911	35.070	0.022	0.086
2	3	0.047	11.597	10.224	20.164	0.494	0.624
1	4	0.025	4.015	3.805	9.142	0.422	0.454

Table 4: Modulus of the five largest roots

$r=5$	$r=4$	$r=3$	$r=2$	$r=1$
0.970	1.0	1.0	1.0	1.0
0.939	0.939	1.0	1.0	1.0
0.939	0.939	0.878	1.0	1.0
0.735	0.734	0.735	0.879	1.0
0.735	0.734	0.735	0.682	0.731

The roots z of the characteristic polynomial $\Pi(z) = I_p - \Pi_1 z - \Pi_2 z^2 - ... - \Pi_k z^k$ are shown in Table 4. Three roots are close to one, suggesting $r = 2$. However, imposing $r = 3$ changes the roots and we observe that the 3rd root is now more moderate. With monthly data a root of 0.88 may well imply slow adjustment rather than non-stationarity.

Furthermore, we may graphically assess the stationarity of the cointegrating relations. The relations $\hat{\beta}'_3 x_t$ and $\hat{\beta}'_4 x_t$, shown in Figure 2, support $r = 3$ since $\hat{\beta}'_4 x_t$ appears non-stationary and $\hat{\beta}'_3 x_t$ stationary.

Finally, coefficients in columns 4 and 5 of the unrestricted $\hat{\alpha}$ (*t*-values in brackets) are largely insignificant, implying that we would not learn anything about adjustment if we included more than three cointegrating relations, while coefficients in the first three columns contain information on adjustment (*t*-values ¿ 2.6 in bold face):

$$\hat{\alpha}^u = \begin{pmatrix} \underset{[7.283]}{\mathbf{0.116}} & \underset{[4.608]}{\mathbf{0.074}} & \underset{[-1.873]}{-0.030} & \underset{[-0.978]}{-0.016} & \underset{[-0.164]}{-0.003} \\ \underset{[-0.617]}{-0.008} & \underset{[6.500]}{\mathbf{0.082}} & \underset{[-0.576]}{-0.007} & \underset{[-1.690]}{-0.021} & \underset{[-0.165]}{-0.002} \\ \underset{[-2.728]}{\mathbf{-0.053}} & \underset{[1.989]}{0.039} & \underset{[-2.652]}{\mathbf{-0.052}} & \underset{[-2.106]}{-0.041} & \underset{[-0.520]}{-0.010} \\ \underset{[-2.618]}{\mathbf{-0.056}} & \underset{[1.766]}{0.038} & \underset{[-2.797]}{\mathbf{-0.060}} & \underset{[-1.676]}{-0.036} & \underset{[-1.040]}{-0.022} \\ \underset{[-1.109]}{-0.022} & \underset{[0.495]}{0.010} & \underset{[-1.428]}{-0.028} & \underset{[-1.766]}{-0.034} & \underset{[-1.446]}{-0.028} \end{pmatrix}.$$

Overall, the results suggest $r = 3$, or $p - r = 2$ common trends. Not all four spreads can form cointegrating relations and the term premium is non-stationary.
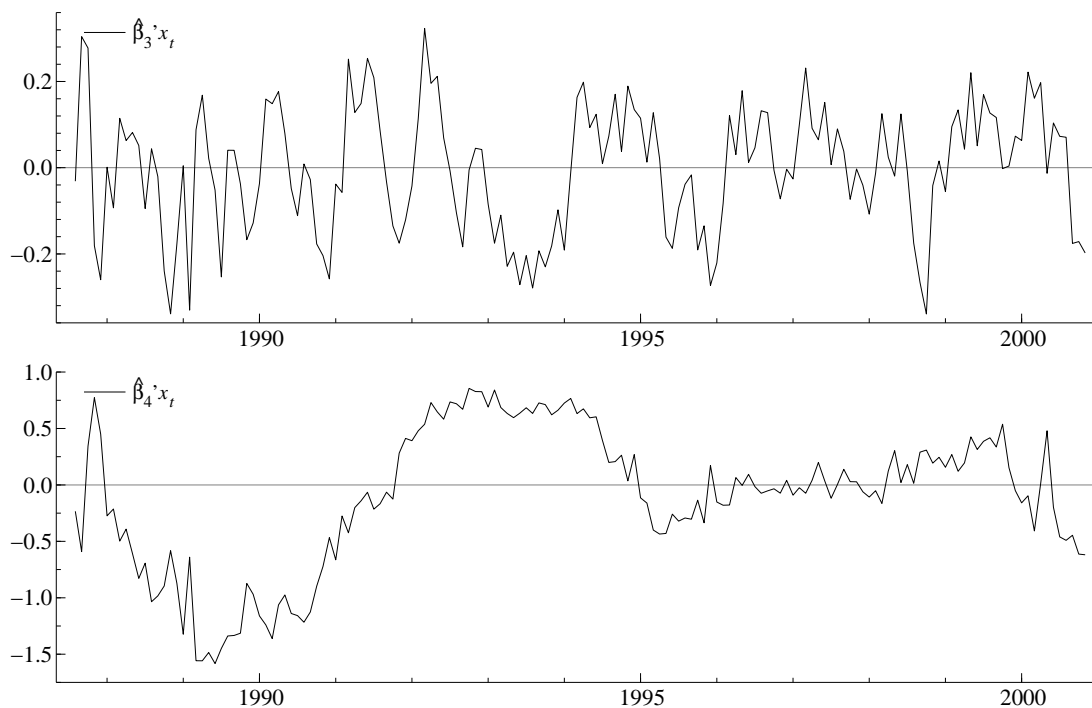
Figure 2: Plots of the estimated $\hat{\beta}'_3 x_t$ and $\hat{\beta}'_4 x_t$

3.2.4   Identification of the Long-Run Structure

In the unrestricted model, α and β' are not uniquely identified but their product Π is. Given $r = 3$, we firstly test some non-identifying hypotheses for the structures of β and α, and secondly impose over-identifying restrictions on β.

Table 5 shows that the spreads including a constant are far from stationary except for the spread between the two short-term yields.[7] The restrictions in each case are formulated as

$$\mathcal{H}^c(r) : \beta^c = (H_1 \varphi_1, \psi),$$

where ψ and φ_1 are $(p+1) \times (r-1)$ and 2×1 matrices, respectively, of unrestricted estimates, and H_1 a $(p+1) \times 2$ known matrix of the form (for the short-term spread)

$$H' = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The restrictions are tested for each spread using an LR test, and the results are supported by Figure 3.

If a variable is not equilibrium correcting, it is "weakly exogenous with respect to β". The hypothesis of a zero row in α is given by

$$\mathcal{H}^c_\alpha(r) : \alpha = H\alpha^c \text{ or equivalently } \mathcal{H}^c_\alpha(r) : R'\alpha = 0 \tag{7}$$

---

[7]Not only are spreads with respect to the one-month yield not stationary, but also spreads between any other combination of yields with different maturities (not shown). Moreover, non-exact spreads, i.e. where (1, -1) is not imposed but coefficients are estimated, are also found to be non-stationary implying that they share more than one common stochastic trend.
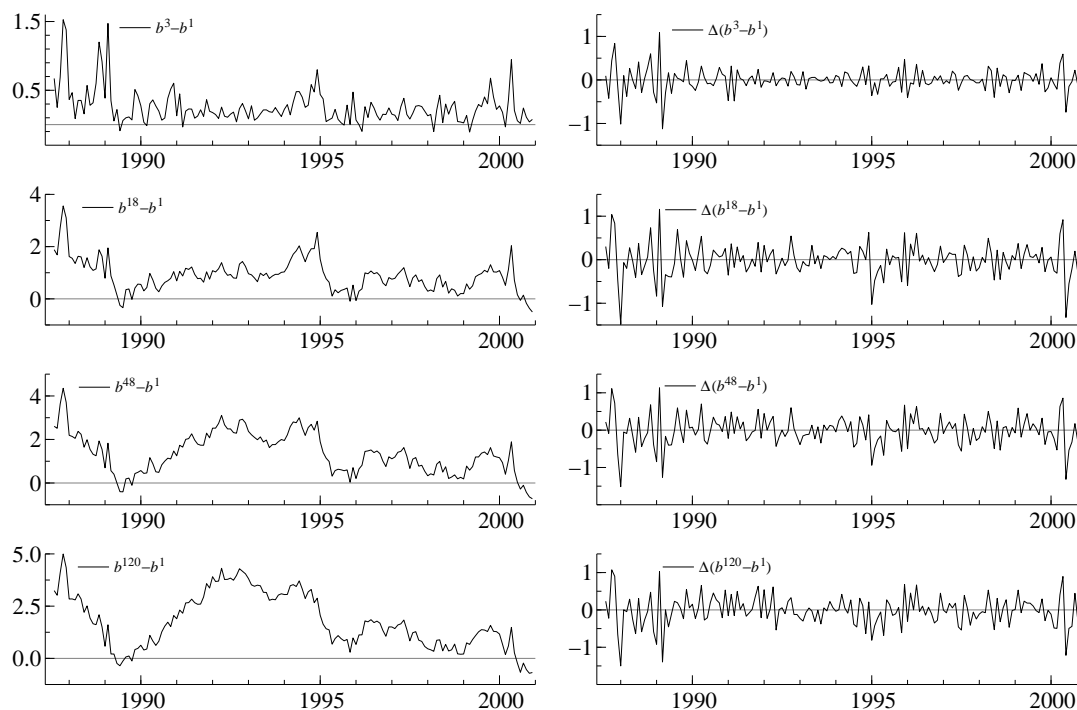
Figure 3: Levels and differences of yield spreads, 1987:08-2000:12

Table 5: Hypothesis tests on β and α for $r = 3$, [p-value]

Tests on β	$b_t^3 - b_t^1$	$b_t^{18} - b_t^1$	$b_t^{48} - b_t^1$	$b_t^{120} - b_t^1$	
Stationary spreads, $\chi^2(4)$	4.791 [0.310]	20.983 [0.000]	21.668 [0.000]	21.971 [0.000]	
Tests on α	b_t^1	b_t^3	b_t^{18}	b_t^{48}	b_t^{120}
Weak exogeneity, $\chi^2(r)$	56.375 [0.000]	32.694 [0.000]	14.346 [0.003]	14.156 [0.003]	2.888 [0.409]
Unit vector, $\chi^2(2)$	0.324 [0.851]	0.747 [0.688]	20.353 [0.000]	13.009 [0.002]	9.271 [0.010]

where H is a $p \times s$ matrix, α^c a $s \times r$ matrix of non-zero α-coefficients ($s = p-$number of restrictions $= 4$), and $R = H_\perp$. The test results in Table 5 suggest that only the ten-year yield can be regarded as non-equilibrium correcting. If we set a row in α to zero, this is translated to a unit vector in $\alpha'_\perp$ as $\alpha'_\perp \alpha = 0$, implying that the cumulated residuals from a weakly exogenous variable form a common trend on their own.

Imposing a unit vector on α is equivalent to setting the corresponding entry in $\alpha'_\perp$ to zero, i.e. a variable that has a unit vector in α is purely adjusting and shocks to such a variable have no permanent effect on any of the variables in the system. The hypothesis can be expressed as follows

$$\mathcal{H}_c^\alpha : \alpha^c = (\mathrm{a}, \tau) \Longleftrightarrow \alpha_\perp = H\alpha_\perp^c \tag{8}$$

where a is a $p \times 1$ vector and τ a $p \times (r - 1)$ matrix. Table 5 shows that the null hypothesis is not rejected for the one- and three-month yields.

Since α and β are jointly determined, we cannot impose all restrictions discussed above at once. In identifying β, we make use of the stationary spread between the short-term maturities and the weighted differences between spreads as discussed theoretically in Section 2, in addition to the weak exogeneity of the ten-year yield in α. The LR test statistic for these restrictions on β and α is 8.309 and follows a $\chi^2(8)$ (p-value: 0.404). The normalised long-run relations are given by (t-values in brackets):

$$\hat{\beta}_1' x_t \;=\; (b_t^3 - b_t^1) \underset{[-10.056]}{-0.214}; \tag{9}$$

$$\hat{\beta}_2' x_t \;=\; \underset{[38.975]}{0.759}(b_t^{18} - b_t^1) \underset{[-12.376]}{-0.241}(b_t^{120} - b_t^{18}) \underset{[-7.808]}{-0.381}; \tag{10}$$

$$\hat{\beta}_3' x_t \;=\; \underset{[-43.137]}{0.467}(b_t^{48} - b_t^{18}) \underset{[-49.314]}{-0.533}(b_t^{120} - b_t^{48}). \tag{11}$$

Equation (9) shows that the slope of the yield curve is stationary for short-term maturities, while Equations (10) and (11) indicate that the medium and long ends of the curve are characterised by an approximately stationary curvature. While the weights in (10) are close to those suggested by the discussion in Footnote 5, the same is not true for (11). In practice, however, the curvature of the long end may be best approximated by equal weights on spreads due to problems of discounting time to maturity: perceptions of time may be compressed.

Besides the restrictions on β, the row in α corresponding to the ten-year yield was set to zero (t-values in brackets):

$$\hat{\alpha} = \begin{pmatrix} \underset{[3.117]}{\mathbf{0.498}} & \underset{[1.432]}{0.140} & \underset{[0.053]}{0.012} \\[4pt] \underset{[-3.669]}{\mathbf{-0.469}} & \underset{[5.533]}{\mathbf{0.434}} & \underset{[-3.478]}{\mathbf{-0.609}} \\[4pt] \underset{[-1.418]}{-0.288} & \underset{[1.367]}{0.170} & \underset{[-2.614]}{\mathbf{-0.726}} \\[4pt] \underset{[-1.158]}{-0.256} & \underset{[1.106]}{0.150} & \underset{[-2.350]}{\mathbf{-0.712}} \\[4pt] \underset{[N.A.]}{0} & \underset{[N.A.]}{0} & \underset{[N.A.]}{0} \end{pmatrix}.$$
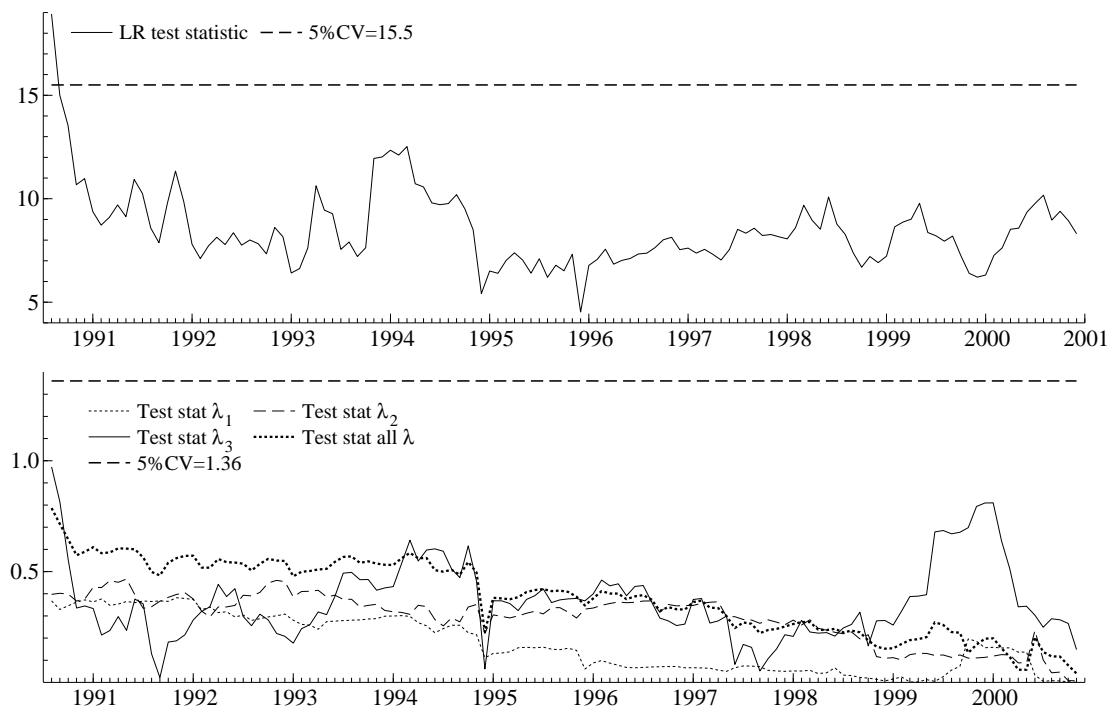
Figure 4: Recursive test statistics of the LR test of over-identifying restrictions and fluctuation test of transformed eigenvalues.

We note that the one-month yield reacts only to the first cointegrating relation. The three-month yield reacts to all three relations, the eighteen-month and four-year yields only to the third, the long-end relation.

Recursive tests suggested by Hansen and Johansen (1999) and discussed in Juselius (2006), Ch. 9, show that coefficients in the restricted CVAR are stable, and that the restrictions are valid over the entire sample. The upper panel in Figure 4 shows the recursively calculated LR test statistic of the over-identifying restrictions, providing evidence for their validity. The recursively computed fluctuation test is given in the lower panel of Figure 4 where we look at the r-largest transformed eigenvalues and their weighted average.[8] At no point is the null hypothesis of recursively estimated eigenvalues being the same as the full sample estimates rejected. Hence, the eigenvalues and corresponding cointegrating relations seem reasonably stable over time.

The finding that the spreads are not stationary by themselves whereas linear combinations of the spreads are, is useful for an improved understanding of how expectations on interest rates are formed and hence for monetary policy. See discussions in Sections 1 and 2. Investors making bets on the yield curve may find it even more profitable to bet on the mean reversion of the weighted differences between spreads – also called butterfly spreads – than on mean reversion of simple spreads. The results here suggest that equilibrium mean reversion is fast for differences be-

---

[8]Transformed eigenvalues are given by $\log(\hat{\lambda}_i) - \log(1 - \hat{\lambda}_i)$.

tween spreads, whereas deviations from spreads are persistent making the timing of bets more difficult.

3.2.5   Identification of Common Trends

The moving average (MA) or Granger-Johansen representation of the CVAR is given by

$$x_t = C \sum_{i=1}^{t} (\varepsilon_i + \phi D_i) + C^*(L)(\varepsilon_t + \phi D_t) + X_0, \tag{12}$$

where $C = \beta_\perp (\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp \equiv \tilde{\beta}_\perp \alpha'_\perp$ with $\Gamma = I - \Gamma_1 - ... - \Gamma_{k-1}$. $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$ defines the $p-r$ non-stationary common trends loaded by coefficients in $\tilde{\beta}_\perp$, while $C^*(L)\varepsilon_t$ denotes the stationary part of the process. X_0 is a function of initial conditions. See Johansen (1996), Ch. 4, for derivations. For calculating the coefficients, we use the model with the above final restrictions imposed on β and α.

Factor models of the yield curve interpret the driving forces of the curve in terms of level, slope and curvature as noted above. We have found two common trends and use the Granger-Johansen representation to interpret them. Equation (13) shows the decomposition of $\hat{C}$ into $\hat{\tilde{\beta}}_\perp \hat{\alpha}'_\perp$ multiplied with the vector of cumulated residuals:

$$\begin{pmatrix} b_t^1 \\ b_t^3 \\ b_t^{18} \\ b_t^{48} \\ b_t^{120} \end{pmatrix} = \begin{pmatrix} 2.008 & -3.956 \\ 2.008 & -3.956 \\ 1.744 & -2.740 \\ 1.302 & -0.697 \\ 0.914 & 1.090 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ -0.020 & 0.065 & -1.000 & 0.963 & 0 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{t} \hat{\varepsilon}_i^1 \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^3 \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^{18} \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^{48} \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^{120} \end{pmatrix} + ...$$

$$\tag{13}$$

For simplicity of notation, we suppress the other terms of Equation (12). The first two variables have identical $\hat{\tilde{\beta}}_\perp$-coefficients because they were restricted to cointegrate. Since the second common trend in (13) has tiny coefficients on the cumulated residuals of the one- and three-month yields, it is useful to reformulate the system as $x_t = (b_t^1, b_t^3, b_t^{18}, b_t^{48} - b_t^{18}, b_t^{120})'$, and test whether $b_t^{48} - b_t^{18}$ is weakly exogenous conditional on the same over-identifying restrictions on β. The LR test statistic for the restrictions on β and α is 8.637 and follows a $\chi^2(11)$ (*p*-value: 0.655) and the additional restrictions seem acceptable. The Granger-Johansen representation simplifies to

$$\begin{pmatrix} b_t^1 \\ b_t^3 \\ b_t^{18} \\ b_t^{48} - b_t^{18} \\ b_t^{120} \end{pmatrix} = \begin{pmatrix} 2.001 & -3.751 \\ 2.001 & -3.751 \\ 1.744 & -2.583 \\ -0.435 & 1.921 \\ 0.921 & 1.013 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{t} \hat{\varepsilon}_i^{120} \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^{48} - \sum_{i=1}^{t} \hat{\varepsilon}_i^{18} \end{pmatrix} + ... \quad ,$$

or equivalently

$$\begin{pmatrix} b_t^1 \\ b_t^3 \\ b_t^{18} \\ b_t^{48} \\ b_t^{120} \end{pmatrix} = \begin{pmatrix} 2.001 & -3.751 \\ 2.001 & -3.751 \\ 1.744 & -2.583 \\ 1.309 & -0.662 \\ 0.921 & 1.013 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{t} \hat{\varepsilon}_i^{120} \\ \sum_{i=1}^{t} \hat{\varepsilon}_i^{48} - \sum_{i=1}^{t} \hat{\varepsilon}_i^{18} \end{pmatrix} + ... \quad .$$

The coefficients in $\hat{\tilde{\beta}}_\perp$ are interpreted as the weights attached to the common trends, or in the language of finance models as the loadings of the factors with respect to each variable. The loadings on the first common trend, the cumulated residuals of the ten-year yield, are in an interval between just under one and two for all variables which implies that it affects all yields similarly, giving it the interpretation of a level factor. This is a plausible result because the ten-year yield contains most information and we should expect shocks to it to influence all yields similarly. The 2nd column of $\hat{\tilde{\beta}}_\perp$ shows coefficients decreasing with maturity, suggesting an interpretation as a slope factor. This interpretation is further strengthened by noting that the second common trend is the spread between the cumulated residuals of the four-year and eighteen-month yields. A positive shock to the spread increases the slope and therefore has a negative effect in particular on the short end of the curve, while a negative shock flattens the yield curve.

In conclusion, our model not only shows that the yield curve is explained by a level and slope factor but gives meaning to them by identifying the cumulated residuals that drive them: the long end of the curve determines the level and the medium sector the slope. The short end does not contain information on yields of other maturities.

3.2.6 Towards a Structural Interpretation

The CVAR residuals used in the determination of the common stochastic trends in (13) are seldom uncorrelated and, therefore, not structurally unique. To attach a structural interpretation to the shocks in the model it is common practise to orthogonalise the residuals. However, Hendry and Mizon (2000) warn against such interpretation because a structural shock should be invariant to changes in the information set. But, unless the model coincides with the data-generating process, residuals are generally not invariant to extensions and omissions. We do not claim that the conditions for a structural interpretation are granted in our case but choose to orthogonalize the residuals as a robustness check of the results in the previous section. In case the conclusions remain reasonably robust we believe this to lend credibility to a causal interpretation of the common trend results.

We use the Granger-Johansen representation but pre-multiply by a rotation matrix B such that the shocks $u_t = B\varepsilon_t$ are uncorrelated:[9]

$$x_t = CB^{-1}\sum_{i=1}^{t}(u_i + B\phi D_i) + C^*(L)B^{-1}(u_t + B\phi D_t) + \tilde{X}_0 \qquad (14)$$

from

$$B\Delta x_t = B\alpha\beta' x_{t-1} + B\Gamma_1\Delta x_{t-1} + B\Gamma_2\Delta x_{t-2} + B\Gamma_3\Delta x_{t-3} + B\phi D_t + B\varepsilon_t, \quad (15)$$

where $B\Omega B' = I$ for orthogonality. The impact matrix $CB^{-1} = \tilde{\beta}_\perp \alpha'_\perp B^{-1}$ has r zero columns due to transitory shocks, and $p - r$ non-zero columns corresponding to

---

[9] B is chosen such that $B = S^{-1}G$, where $G = \begin{pmatrix} \alpha'^{-1} \\ \alpha'_\perp \end{pmatrix}$ and S is found through Choleski decomposition of $G\Omega G'$. Then $Var(u_t) = B\Omega B' = I_p$.

the number of common trends. It reveals how the different variables in the system react to the permanent shocks $\sum_{i=1}^{t} u_{ji}$, $j = 1, 2$. However, due to the rotation of the residuals, the permanent shocks may not have straightforward interpretations.

We have two permanent shocks and to identify the impact matrix need to restrict one entry to zero.[10] We choose the ten-year yield because based on the previous analysis we believe it to be influenced least. The normalised impact matrix is estimated as

$$
\hat{C}\hat{B}^{-1} = \hat{\tilde{\beta}}_{\perp}\hat{\alpha}'_{\perp}\hat{B}^{-1} =
\begin{pmatrix}
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0.992 & 0.761 \\
0 & 0 & 0 & 0.978 & 0.357 \\
0 & 0 & 0 & 0.967 & 0
\end{pmatrix},
$$

where the first three columns reflect impacts from transitory shocks, and the 4th and 5th impacts from permanent shocks (the equal coefficients for the one- and three-month yields are again due to them cointegrating, and to normalisation). The first cumulated permanent shock has nearly identical loadings, indicating a level factor, while the second one is a slope factor. To gain understanding on how to interpret the independent shocks, we examine the rotation matrix given by

$$
\hat{B} =
\begin{pmatrix}
1.000 & -0.008 & -0.653 & 0.828 & -0.587 \\
-0.390 & 1.000 & -0.753 & 0.858 & -0.446 \\
0.061 & -0.576 & 0.083 & 1.000 & -0.816 \\
0.008 & 0.144 & -0.418 & -0.043 & 1.000 \\
0.023 & -0.061 & 1.000 & -0.988 & 0.226
\end{pmatrix},
$$

where the first three rows are due to transitory shocks and rows 4 and 5 are due to permanent shocks. The rotation matrix matches the relations found for the common trends above. The 4th row has small coefficients for all variables except the ten-year yield (and possibly the eighteen-month yield), while the 5th row is essentially the spread between the four-year and eighteen-month yields. Hence, the permanent structural shocks correspond closely to the linear combination of the CVAR residuals that defined the previously estimated common trends, suggesting that they might have been approximately orthogonal from the outset. Therefore, the estimated common stochastic trends might be given a structural interpretation.

3.3 The Common Trends and Macroeconomic Variables

In this section, we use cointegration to determine relationships between the macroeconomy and the common trends from Section 3.2.5 thereby providing a further link to the common factor literature where combining yield factors with macroeconomic variables is a relatively new strand.[11] As demonstrated below this is also possible within the CVAR framework.

---

[10]Since restrictions on $\hat{\alpha}$ are removed when estimating the structural MA, the two formulations of x_t presented in the previous section give identical results up to linear combinations of coefficients. Results presented here are for $x_t = (b_t^1, b_t^3, b_t^{18}, b_t^{48}, b_t^{120})'$.

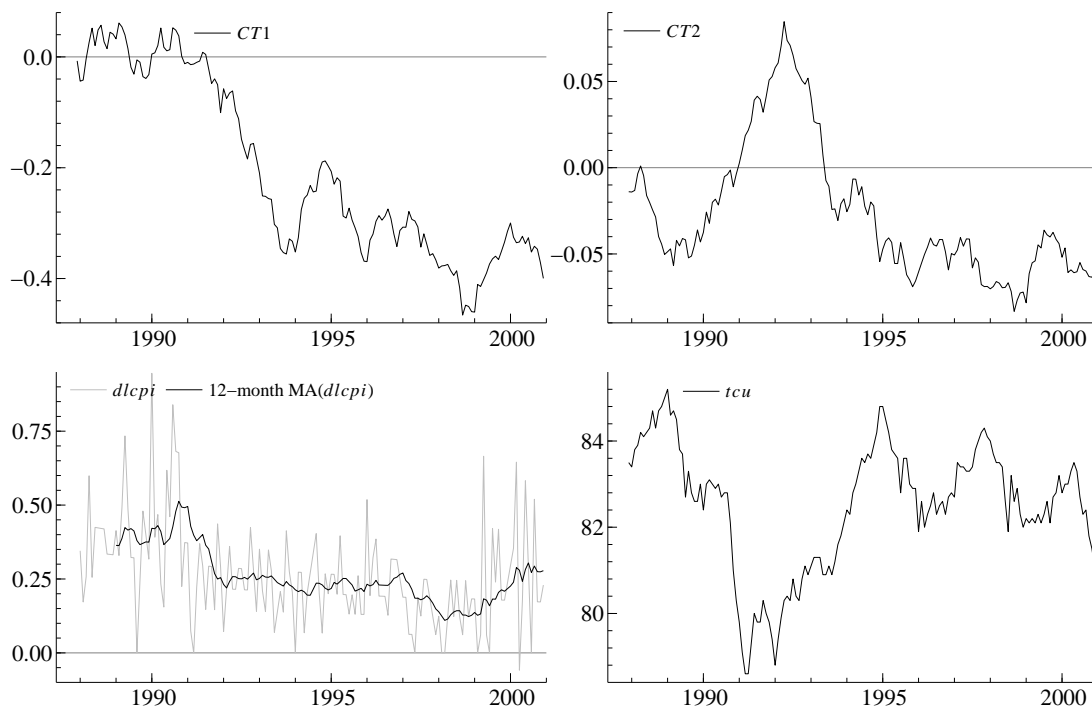[11]See *inter alia* Diebold, Rudebusch, and Aruoba (2006) and Dewachter and Lyrio (2006).

Figure 5: Level plots of the common trends and macroeconomic variables

Variables included in the new CVAR are the two common trends from the yield model ($CT1 = \sum_{i=1}^{t} \hat{\varepsilon}_i^{120}$, the level factor, and $CT2 = \sum_{i=1}^{t} \hat{\varepsilon}_i^{48} - \sum_{i=1}^{t} \hat{\varepsilon}_i^{18}$, the slope factor, both scaled down by 12), the monthly inflation rate based on the log of the consumer price index (*dlcpi*, scaled up by 100), and total capacity utilisation (*tcu*, in percent).[12] The series are shown in Figure 5.

The model is estimated with a lag length of 2 as suggested by LR tests. Mis-specification tests reveal problems only with normality, and three impulse dummies are accordingly included, for January 1990, August 1990 and April 1999. Based on the criteria discussed in Section 3.2.3, the rank is set to $r = 2$. Recursive estimation suggests that parameters are stable over the sample. Together with long-run weak exogeneity of both $CT1$ and *tcu* in α, the following over-identifying restrictions on β are accepted with an LR test statistic of 7.313 ($\chi^2(5)$, *p*-value: 0.198):

$$\hat{\beta}_1' x_t = CT1 - 2.211 dlcpi + 0.768; \qquad (16)$$
$$\hat{\beta}_2' x_t = CT2 - 0.297 dlcpi + 0.030 tcu - 2.396. \qquad (17)$$

The first common trend from the yield model, $CT1$, is positively related with infla-tion in (16), i.e. shocks to the long-term interest rate shifting the yield curve seem associated with inflationary shocks. The second cointegrating relation (17), which involves the second common trend $CT2$, includes inflation and capacity utilisation.

---

[12]The data for all macroeconomic variables was obtained from FRED, the database of the Federal Reserve Bank of St. Louis, and are seasonally adjusted. Alternative monthly measures of economic activity like the number of housing starts and of help-wanted advertising in newspapers gave qualitatively equivalent results, but are not presented here.

A steepening of the slope is associated with an increase in inflation, a decrease in activity or both.

The Granger-Johansen representation of the macro model gives an indication of feedback directions between the yield curve and the macroeconomy. Coefficients of the C-matrix (t-values in brackets) are estimated as

$$\hat{C} = \begin{pmatrix} \mathbf{1.221} & 0.000 & 0.000 & 0.018 \\ \scriptstyle[7.652] & \scriptstyle[0.000] & \scriptstyle[0.000] & \scriptstyle[0.496] \\ -0.039 & 0.000 & 0.000 & \mathbf{-0.041} \\ \scriptstyle[-0.417] & \scriptstyle[0.000] & \scriptstyle[0.000] & \scriptstyle[-1.962] \\ \mathbf{0.552} & 0.000 & 0.000 & 0.008 \\ \scriptstyle[7.652] & \scriptstyle[0.000] & \scriptstyle[0.000] & \scriptstyle[0.496] \\ \mathbf{6.724} & 0.000 & 0.000 & \mathbf{1.430} \\ \scriptstyle[1.941] & \scriptstyle[0.000] & \scriptstyle[0.000] & \scriptstyle[1.860] \end{pmatrix},$$

where the zero columns are due to $CT2$ and inflation being purely adjusting, and the non-zero columns to $CT1$ and capacity utilisation being weakly exogenous. This representation provides evidence that $CT1$ is determined largely by its own cumulated residuals, while $CT2$ is driven by the cumulated residuals of the activity measure. The inflation rate is driven only by the cumulated residuals of $CT1$, while capacity utilisation depends on its own cumulated residuals as well as $CT1$.[13]

Hence feedback in both directions exists between the macroeconomy and the yield curve. The level of the yield curve positively influences the inflation rate and activity measure, but is itself independent of the macroeconomy. This is in contrast to the slope which reacts to the activity measure, but exerts no influence on other variables. Given that *CT1* captures unexplained elements of the long-term yield, it may represent inflation expectations. These may then determine inflation itself by being partly self-fulfilling, and should therefore be useful in forecasting inflation. *CT2* may be interpreted as a term premium following Equation (4). We find the spread to be non-stationary and according to (4) the non-stationary part is the term premium. Since *CT2* is made up of the spread between unexplained parts of the two medium-term yields, this interpretation appears plausible. Related to the macroeconomy, we have that the term premium increases if inflation increases and activity decreases. Its non-stationary component derives from the common trend associated with the activity measure (as shown in the estimated C-matrix), and macro factors may thus prove helpful in forecasting it.

4 Conclusion

The approach introduced in this paper considers the stationarity of the yield curve's derivatives. Past literature has focused on testing the hypothesis of stationary spreads in accordance with the expectation theory but not, in case of rejection, examined the differences between spreads. This extension to the theoretical framework allows us to test the stationarity of weighted differences between spreads in a CVAR of US treasury yields which is accepted. Two term premia of different maturities therefore cointegrate and the curvature of the yield curve may allow a more meaningful assessment of future interest rate expectations than the slope.

---

[13]Where significance is only borderline it becomes stronger when removing the weak exogeneity restrictions.

In addition, the Granger-Johansen representation of the CVAR proved a powerful tool for characterising the non-stationary components of the yield curve, identifying them as level and slope. The cumulated residuals of the ten-year yield make up a common trend associated with a level shift, while the second common trend is the spread between the cumulated residuals of the four-year and eighteen-month yields impacting on the slope of the yield curve. When considering relations between these common trends and macroeconomic variables, we find causation running from the level of the yield curve to inflation, and from an activity measure to the slope of the yield curve.

References

Baba, Y., D. F. Hendry, and R. M. Starr (1992). The demand for M1 in the U.S.A., 1960-1988. *Review of Economic Studies 59*, 25–61.

Bliss, R. R. (1997). Testing term structure estimation methods. *Advances in Futures and Options Research 9*, 197–231.

Campbell, J. Y. and R. J. Shiller (1987). Cointegration and tests of present value models. *Journal of Political Economy 95*, 1062–1088.

Carstensen, K. (2003). Nonstationary term premia and cointegration of the term structure. *Economic Letters 80*, 409–413.

Dennis, J. G. (2006). *CATS in RATS - Cointegration Analysis of Time Series, Version 2*. Evanston, Illinois: Estima.

Dewachter, H. and M. Lyrio (2006). Macro factors and the term structure of interest rates. *Journal of Money, Credit and Banking 38*, 119–140.

Diebold, F. X., G. D. Rudebusch, and S. B. Aruoba (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics 131*, 309–338.

Doornik, J. A. and D. F. Hendry (2007). *Modelling Dynamic Systems Using PcGive 12: Volume II*. London: Timberlake Consultants Press.

Duffee, G. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance 57*, 405–443.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: representation, estimation and testing. *Econometrica 55*, 251–276.

Fama, E. F. and R. R. Bliss (1987). The information in the term structure. *American Economic Review 77*, 680–692.

Giese, J. V. (2006). Characterising the yield curve's derivatives in a regime-changing cointegrated VAR model. Manuscript.

Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics 60*, 203–233.

Hall, A. D., H. M. Anderson, and C. W. Granger (1992). A cointegration analysis of Treasury bill yields. *The Review of Economics and Statistics 74*, 117–126.

Hansen, H. and S. Johansen (1999). Some tests for parameter constancy in the cointegrated VAR. *Econometrics Journal 2*, 306–333.

Hendry, D. F. and G. E. Mizon (2000). Reformulating empirical macro-econometric modelling. *Oxford Review of Economic Policy 16*, 138–159.

Johansen, S. (1996). *Likelihood-based Inference in Cointegrated Vector Auto-regressive Models*. Oxford: Oxford University Press.

Juselius, K. (2006). *The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications*. Oxford: Oxford University Press.

Knez, P. K., R. Litterman, and J. A. Scheinkman (1994). Explorations into factors explaining money market returns. *Journal of Finance 49*, 1861–1882.

Litterman, R. and J. A. Scheinkman (1991). Common factors affecting bond returns. *Journal of Fixed Income 1*, 51–61.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*, 473–489.

Nielsen, B. (2006). Order determination in general vector autoregressions. In H.-C. Ho, C.-K. Ing, and T. Lai (Eds.), *Time Series and Related Topics: In Memory of Ching-Zong Wei*. IMS Lecture Notes and Monograph Series, 52.

Shea, G. S. (1992). Benchmarking the expectations hypothesis of the interest-rate term structure: an analysis of cointegration vectors. *Journal of Business and Economic Statistics 10*, 347–366.

Zhang, H. (1993). Treasury yield curves and cointegration. *Applied Economics 25*, 361–367.

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either rating the article on a scale from 1 (bad) to 5 (excellent) or by posting your comments.

Please go to:

www.economics-ejournal.org/economics/journalarticles/2008-28

The Editor