Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Financial Factors, Macroeconomic Information and the Expectations Theory of the Term Structure of Interest Rates

Andrea Carriero, Carlo A. Favero and Iryna Kaminska

Working Paper n. 253

January 2004

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

Financial Factors, Macroeconomic Information and the Expectations Theory of the Term Structure of Interest Rates

Andrea Carriero, Carlo A. Favero and Iryna Kaminska*

November 6, 2003

Abstract

In this paper we concentrate on the hypothesis that the empirical rejections of the Expectations Theory(ET) of the term structure of interest rates can be caused by improper modelling of expectations. Our starting point is an interesting anomaly found by Campbell-Shiller(1987), when by taking a VAR approach they abandon limited information approach to test the ET, in which realized returns are taken as a proxy for expected returns. We use financial factors and macroeconomic information to construct a test of the theory based on simulating investors' effort to use the model in 'real time' to forecast future monetary policy rates. Our findings suggest that the importance of fluctuations of risk premia in explaining the deviation from the ET is reduced when some forecasting model for short-term rates is adopted and a proper evaluation of uncertainty associated to policy rates forecast is considered.

JEL Classification: E43, E44, E47

Keywords: Expectations Theory, Macroeconomic information in Finance

1 Introduction

The objective of this paper is to provide new evidence on the expectations theory (ET) of the term structure of interest rates.

*Carriero and Kaminska are affiliated with IGIER-Università Bocconi, Favero is affiliated with IGIER and CEPR. Corresponding Author:Carlo Favero, IGIER- Bocconi University, Via Salasco 5, 20124 Milan, Italy, carlo.favero@uni-bocconi.it

How is this possible?

Our starting point is an interesting anomaly which contradicts the bulk of evidence rejecting the expectations model.

The anomaly is reported in the widely cited work by Campbell and Shiller(1987, 1991)(CS), where they implement a bi-variate vector autoregressive (VAR) approach for evaluating present value models. The approach consists in projecting the average of expected future short-term yields onto a subset of the information set used by market participants. Such information set is built by assuming that the first difference of short-term bond yields and the excess holding period returns of long-term bonds on short term bonds are stationary. Under this assumptions, the first difference of the yield on short-term bonds and the yield spreads between long-term and short-term bonds form a bivariate stationary vector-stochastic process. By representing this process as a finite order VAR, a 'theoretical spread', i.e. the spread which would obtain if the expectations theory were true, can be constructed. The equality of the actual spread and the theoretical spread puts a set of restrictions on the coefficients of the estimated VAR. When these restrictions are tested formally using a Wald test, they are rejected. However, despite these negative results, the data produced an anomaly: a very strong correlation between the actual and theoretical spread. Such strong correlation leads Campbell and Shiller to conclude that bivariate analysis suggests that there is an important element of truth to the expectations theory of the term structure.

The work by Campbell and Shiller makes an important exception to the available evidence rejecting the expectations theory. In fact, such evidence is mainly based on a single-equation, limited information, approach. The evidence that high yield spreads fare poorly in predicting increases in long rates(see Campbell, 1995), that the change in yields does not move one-to-one with the forward spot spread(see Fama and Bliss,1987) or that period excess returns on long-term bond are predictable using the information in the forward-spot spread(see Cochrane,2001) is always derived within a single-equation approach which cannot identify if the empirical failure of the model is due to systematic expectations errors, or to shifts in the risk premia. In fact, the tests of the theory are accomplished by the assumption that realized returns are a valid proxy for expected returns. In a recent paper, Elton(1999) clearly asserts that there is ample evidence against the belief that information surprises tend to cancel out over time and hence realized returns cannot be considered as an appropriate proxy for realized returns. Interestingly, Campbell(2001) finds that there is much more truth in the proposition that high yield spreads should forecast long-term increases in short-rates, especially at very short and very long maturities. The failure of the expectations model to predict long rate changes and the (partial) success in the prediction of short rate changes is explained by the role of measurement errors. In fact, in the regression of long rate changes onto the yield spread, changing rational expectations about excess long bond returns act like a measurement error that appears positively in the regressor and negatively in the dependent variable. Conversely, in the regression of short-rate changes onto the yield spread, changing rational expectations about excess long-bond returns act like a measurement error that appears only in the regressor. In the first case a small measurement error can change the sign of the relevant regression coefficient, while in the second case the measurement error biases the coefficient towards zero but cannot affect its sign. These findings on the effects of expectations errors on the tests of the model are confirmed by a number of papers which concentrates on expectations errors by relating them to peso problems or to the very low predictability of short term interest rates. In a famous study Mankiw and Miron, 1986, using data on a three and six month maturity, found evidence in favor of the expectation theory prior to the founding of the Federal Reserve System in 1915. They show that the shift in regime occurred with the founding of the Fed led to a remarkable decrease in the predictability of short-term interest rates. Rudebusch, 1995, and Balduzzi et al., 1997, expand on this evidence by looking at more recent data.

On a different, but clearly related, ground McCallum(1994) is the first to argue that the limited information approach might cause bias in the estimates due to simultaneity. He shows that the anomalous empirical findings based on a single equation evidence can be rationalized with the expectations theory by a recognition of an exogenous term premium plus the assumption that monetary policy involves smoothing of the policy rates together with the responses to the prevailing level of the spread. Interestingly, the bi-variate framework considered by CS matches exactly the scenario used by McCallum to illustrate the simultaneity bias in the single-equation approach. However, McCallum himself notes that a reaction function according to which the Fed reacts to the spread only represents a simplification relative to the actual behaviour of the Fed, which almost certainly responds to recent inflation and output or employment movements, as well as to the spread.

In this paper we concentrate on the hypothesis that the model failure can be caused by expectations errors, we maintain a simultaneous equation approach and we see how far we can get by extending the Campbell-Shiller framework to concentrate on our null of interest.

We consider extensions of the original framework along two dimensions: the testing framework and the information set.

CS test the restrictions imposed by the ET estimating a VAR model over the full available sample and by using only in-sample information. Such procedure cannot simulate the investors' effort to use the model in 'real time' to forecast future monetary policy rates, as the information from the whole sample is used to estimate parameters while investors can use only historically available information to generate (up to n-period ahead) predictions of policy rates. Moreover, the within sample test understate the uncertainty of agents who forecast policy rates by out-of-sample projections. In this paper we use the present value framework to generate real time forecast for future policy rates. At each point in time we estimate, using the historically available information, a model and then we use it to project out-of-sample policy rates up to the nth-period ahead. Given the path of simulated future policy rates, we can construct yield to maturities consistent with the Expectations Theory. Using the historically available information on uncertainty we perform dynamic stochastic simulations and construct confidence bounds around the ET-consistent long-term rates. These bounds reflect explicitly the uncertainty associated with out-of-sample projections. Then we test the ET by checking if the observed long-term rates fluctuate within the bounds. Our procedure allows to evaluate the ET without taking ex-post realized returns as a proxy for ex-ante expected returns, it also allows independent identification of the risk-premia from forecast errors about future policy rates.

We also extend the information set originally considered by CS, which is limited to one long-term interest rate and one short-term policy rate.

Given that the ET links the long-term rate to the path of future policy rates, the natural extension of the original information set is obtained by including arguments in the monetary policy maker reaction function. The reaction function implicit in CS is one in which policy rates respond to long-term interest rates, we extend this including financial factors and macroeconomic information.

Our starting point in the selection of financial factors is the available empirical evidence that a three-factor model is needed to accurately describe the term structure and that the use of term structure related factors is of considerable help in modelling monetary policy rates (see, for example, Ang and Piazzesi(2003)). We derive financial factors by implementing a procedure recently introduced by Diebold and Li(2002), who have shown that accurate term structure forecasts could be produced by estimating autoregressive models for factors corresponding to the level, slope and curvature of the term structure. We estimate factors recursively over the cross-section of yields and we include their time-series in the Vector Autoregressive model. We adopt this approach on the basis of the evidence produced by Diebold and Li on its success for forecasting. Moreover, although this approach does not impose the no-arbitrage condition to derive factors, the relevant parameters are very easily and precisely estimated.

Our starting point in the selection of macroeconomic variables are the typical arguments of a Taylor rule. A growing body of empirical literature has established interest rate rules as a convenient way to model and interpret monetary policy. Interest rate rules, which feature (very) persistent of policy rates responding to central bank's perceptions of (expected) inflation and output gaps (Taylor,1993, Clarida, Gali and Gertler, 1998, 1999, 2000) not only track the data well but are also capable of explaining the high inflation in the seventies in terms of an accommodating behaviour towards inflation in the pre-Volcker era. We are not the first to relate Taylor rules to the term structure. Fuhrer(1996) uses a simple Taylor-rule type reaction function, the expectations model and reduced-form equations for output and inflation to solve for the reaction function coefficients that delivers long-term rates consistent with the expectations theory. He finds that modest and smoothly evolving time-variation in the reaction functions parameters is sufficient to reconcile the expectations model with the long-bond data. Favero(2002) extends Fuhrer framework to derive standard errors for long-term rates consistent with the ET. Our approach of extending the VAR framework is closely related, but very different, from recent work by Roush(2003). Roush considers a VAR model with macro and financial variables to show that the expectations theory of the term structure holds conditional on an exogenous change in monetary policy. The paper adds to the picture the important issue of identification but it does not provide evidence on the impact of the extension of the original CS information set on the outcome of the test for the unconditional validity of cross equation restrictions; moreover, the attention is limited to the within-sample evidence.

The paper is organized as follows. Section 2 illustrates the Present Value Approach and replicates the CS anomaly on our data-set. Section 3 illustrates our testing framework and our extension of the information set. Section 4 discusses our results by interpreting the differences between our results and those generated by testing strategies based on limited-information approaches, Section 5 concludes.

2 The Present Value approach

We adopt respectively the linearized expectations model of Shiller (1979) in the bi-variate framework proposed by CS.

The derivation of this model starts from the assumption that expected one-period holding returns from long-term bonds equal the risk-free short term interest rate plus a term premium:

$$E_t[H_{t,T} \mid I_t] = r_t + \phi_{t,T} \tag{1}$$

where $H_{t,T}$ is the one-period holding return of a bond with maturity T, I_t is the market information set , r_t is the short term interest rate and $\phi_{t,T}$ is a term premium defined over a one-period horizon.

For long term bonds bearing a coupon C, $H_{t,T}$ is a non-linear function of the yield to maturity $R_{t,T}$. We consider the linearization proposed by Shiller(1979) and take the following approximation in the neighborhood $R_{t,T} = R_{t+1,T} = \bar{R} = C$:

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

We then have:

$$E_t\left[\frac{R_{t,T} - \gamma_T R_{t+1,T}}{1-\gamma_T} \mid I_t\right] = r_t + \phi_{t,T} \tag{2}$$

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

$$R_{t,T} = R^*_{t,T} + E[\Phi_T \mid I_t] = \frac{1-\gamma}{1-\gamma^T}\sum_{j=0}^{T-1}\gamma^j E_t[r_{t+j} \mid I_t] + E[\Phi_T \mid I_t] \tag{3}$$

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

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

CS tests the ET[1] by using equation (3) in considering the case of the risk free rate and a very long term bond. In such case, under the null of the ET, we have

[1] In fact CS use de-meaned-variables, that is equivalent to test a weak form of the Expectations Theory, in the sense that de-meaning eliminates a constant risk premium.

$$R_{t,T} = R^*_{t,T} = (1-\gamma)\sum_{j=0}^{T-1}\gamma^j E_t[r_{t+j} \mid I_t] \tag{4}$$

which could be re-written in terms of spread between long and short-term rates,

$$\begin{aligned} S_{t,T} &= S^*_{t,T} = \sum_{j=1}^{T-1}\gamma^j E_t[\Delta r_{t+j} \mid I_t] && (5) \\ S_{t,T} &= R_{t,T} - r_t && (6) \end{aligned}$$

(5) shows that there is necessary condition for the ET to hold which 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.

Having checked 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 :

$$\begin{array}{c} \Delta r_t = a(L)\Delta r_{t-1} + b(L)S_{t-1} + u_{1t} \\ S_t = c(L)\Delta r_{t-1} + d(L)S_{t-1} + u_{2t} \end{array} \tag{7}$$

no-constant is included in the VAR as all series are de-meaned.

Stack the VAR as:

$$\begin{bmatrix} \Delta r_t \\ . \\ . \\ \Delta r_{t-p+1} \\ S_t \\ . \\ . \\ S_{t-p+1} \end{bmatrix} = \begin{bmatrix} a_1 & . & . & a_p & b_1 & . & . & b_2 \\ 1 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & 1 & 0 & 0 & . & . & 0 \\ c_1 & . & . & c_p & d_1 & . & . & d_2 \\ 0 & . & . & 0 & 1 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & . & 0 \\ 0 & . & . & 0 & 0 & . & 1 & 0 \end{bmatrix} \begin{bmatrix} \Delta r_{t-1} \\ . \\ . \\ \Delta r_{t-p} \\ S_{t-1} \\ . \\ . \\ S_{t-p} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ . \\ . \\ 0 \\ u_{2t} \\ . \\ . \\ 0 \end{bmatrix} \tag{8}$$

this can be written more succinctly as:

$$z_t = Az_{t-1} + v_t \tag{9}$$

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

$$g'z_t = \sum_{j=1}^{T-1} \gamma^j h'A'z_t \tag{10}$$

where $g\prime$ and $h\prime$ are row vectors with 2p elements, all of which are zero except for the p+1st element of $g\prime$ and the second 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:

$$g' = h'\gamma A(I - \gamma A)^{-1} \tag{11}$$

which implies:

$$g'(I - \gamma A) = h'\gamma A \tag{12}$$

and we have the following constraints on the individual coefficients of VAR(7):

$$\{c_i = -a_i, \forall i\}, \{d_1 = -b_1 + 1/\gamma_T\}, \{d_i = -b_i, \forall i \neq 1\} \tag{13}$$

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.

We replicated the CS within sample procedure starting form the sample 1981:3-1991:8 and then extending recursively the end point of the sample up to 2002. We have chosen the initial date of the sample to concentrate on an era of homogenous monetary policy, i.e. the Volcker-Greenspan era. Figure 1 and 2 illustrate the anomaly. Figure 1 reports the results of the test for the ET cross-equation restrictions, which shows that the null of is rejected for all the sample end points, after 1996:6 . Figure 2 reports the actual spread, $S_{t,T}$, and the spread, $S'_{t,T}$ obtained by imposing the ET restrictions, even when they are statistically rejected. The two spreads move very closely together, both in periods when the ET restrictions are rejected and in the periods when they are not rejected. This is the fact that lead Campbell and Shiller to conclude that "... deviations from the present value model for bonds are transitory...".

3 A new testing framework with an extended information set

We use a cointegrated VAR framework, in which the original set of variables used by CS is extended by including a vector of variables $\mathbf{X}$. Such vector includes financial factors and macroeconomic variables. At each point in time we estimate, using the historically available information, the following model:

$$\begin{gathered}
\Delta r_t = a_0 + a_1(L)\Delta r_{t-1} + a_2(L)S_{t-1} + a_3(L)\mathbf{X}_{t-1} + u_{1t} \\
S_t = b_0 + b_1(L)\Delta r_{t-1} + b_2(L)S_{t-1} + b_3(L)\mathbf{X}_{t-1} + u_{2t} \\
\mathbf{X}_t = c_0 + c_1(L)\Delta r_{t-1} + c_2(L)S_{t-1} + c_3(L)\mathbf{X}_{t-1} + u_{3t} \\
\begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \end{bmatrix} \backsim N\left[0, \Sigma\right]
\end{gathered}$$

As a starting point, we proceed to test the validity of the ET restrictions on this extended VAR. We then simulate the estimated model forward, to obtain projection for all the relevant policy rates and to construct ET-consistent spread[2] as follows:

$$S'_{t,T} = \sum_{j=1}^{T-1} \gamma^j E_t[\Delta r_{t+j} \mid \Omega_t] \tag{14}$$

In practice we simulate the model forward repeatedly for N draws of its stochastic components. At each repetition, errors are generated for each observation in accordance with the residual uncertainty in the model. We draw residuals from a bivariate normal distribution $N\left(0, \hat{\Sigma}\right)$ where $\hat{\Sigma}$ is the estimated variance-covariance matrix form our VAR . Then ET consistent yields are calculated applying equation (14) to each of the N simulated paths of future expected short-term rates:among these, the 0.5th, 0.05th, and 0.95th quantiles represent respectively the "theoretical" ET-consistent yield and its 95% confidence bounds. The estimation window is then enlarged by one observation and simulation horizon is shifted one period ahead and the same steps are repeated. We implemented this procedure 124 times, estimating recursively the model starting from the sample from 1981:3-1991:8 to conclude with the

[2] For consistency with CS, we simulate the model forward after de-meaning.

sample 1981:3-2001:12 and simulating ahead each time for one-hundred and twenty periods. By doing so we derive the full path of ET-consistent yield and its 95% confidence intervals for the sample 1991:9-2001:12. Point forecasts and their confidence bounds define a region inside which the actual long term rates should lie if the ET holds. Note that the uncertainty associated to our simulated yields is higher than in CS, as it reflects the uncertainty associated with out-of-sample predictions based on the information available in real time rather than that associated to within sample tests. By combining (5) and (14) , we have:

$$\begin{aligned} S_{t,T} &= S'_{t,T} + (S^*_{t,T} - S'_{t,T}) + E[\Phi_T \mid I_t] \qquad (15) \\ S_{t,T} - S'_{t,T} &= \left(\sum_{j=1}^{T-1} \gamma^j E_t[\Delta r_{t+j} \mid I_t] - \sum_{j=1}^{T-1} \gamma^j E_t[\Delta r_{t+j} \mid \Omega_t] \right) + E[\Phi_T \mid I_t] \end{aligned}$$

Equation (15) makes clear that deviation of $S_{t,T}$ from $S'_{t,T}$ can be explained by the effect of the risk premia or by differences between model based forecasts and agents' expectations. Importantly, our choice of recursive estimation of the model to generate projections implicitly allows for a smoothly time varying risk premium. This is a natural extension of the CS framework, in which, variables are demeaned prior to estimation on the full available sample. However, the decomposition of $S_{t,T} - S'_{t,T}$ in the two factors is very difficult, given that neither $\sum_{j=1}^{T-1} \gamma^j E_t[\Delta r_{t+j} \mid I_t]$ nor the risk premium are observable. We shall try and provide some insight by analyzing the performance of our proposed models in forecasting policy rates at different horizons.

4 Financial Factors

To derive financial factors we consider data on zero-coupon equivalent yields for US data measured at the following maturities[3]: 1-month, 2-month, 3-month, 6-month, 9-month, 1-year, 2-year, 3-year, 5-year, 7-year, 10-year. At each point of our time series t we estimate, by non-linear least squares, on the cross-section of eleven yields, the following Nelson-Siegel model:

[3]The data were indly made available by the ECB, and they are posted on Favero's website at the following address: http:/www.igier.uni-bocconi.it/personal/favero
in the section working papers

$$y_{t,t+k} = L_t + S_t \frac{1-\exp\left(-\frac{k}{\tau_1}\right)}{\frac{k}{\tau_1}} + C_t \left(\frac{1-\exp\left(-\frac{k}{\tau_1}\right)}{\frac{k}{\tau_1}} - \exp\left(-\frac{k}{\tau_1}\right) \right) \quad (16)$$

which is implicit in the instantaneous forward curve:

$$f_{tk} = L_t + S_t \exp\left(-\frac{k}{\tau_1}\right) + C_t \frac{k}{\tau_1} \exp\left(-\frac{k}{\tau_1}\right) \quad (17)$$

The parameter τ_1 is kept constant over time[4], as this restriction decreases the volatility of the β parameters, making them more predictable in time. 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, S_t, C_t, 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 S_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, $r_t^{rf} = L_t + S_t$ 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 S_t, the slope of the yield curve, is nothing else than the minus the spread in Campbell-Shiller. C_t is a medium term factor, in the sense that their loading start at zero, increase and then decay to zero (at different speed). Such factor captures the curvature of the yield curve. In fact, Diebold and Li show that it 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 yields at different maturities allows to construct a time-series for our factors. We report in Figure 3 the three factors, while Figure 4 shows the goodness of fit of the Nelson and Siegel interpolation for all yields considered in our sample.

Given these time series, we can estimate the following VAR model with financial factors:

$$\begin{bmatrix} \Delta r_t^{rf} \\ -S_t \\ C_t \end{bmatrix} = A(L) \begin{bmatrix} \Delta r_{t-1}^{rf} \\ -S_{t-1} \\ C_{t-1} \end{bmatrix} + u_t \quad (18)$$

[4]We restrict τ_1 at the value of 0.87, which is the median, over the time series, of the estimated value of τ_1 in a four parameter version of the Nelson-Siegel interpolant.

The above model extends the CS framework by including the curvature of the yield curve. We estimate the VAR recursively[5] and perform both the original within-sample CS test on cross-equation restrictions and our proposed test based on real time out-of sample projections of policy rates. CS note that their specification should constitute a cointegrated VAR if the risk premium is stationary, our extension should not violate this property as the curvature is by construction stationary when the spreads between yields of any two bonds are stationary. We have tested the rank of the long-run matrix in our VAR representation on our full sample using the Johansen(1995) procedure, and found evidence in favour of full rank, in fact we did reject the null of a reduced rank of two at the ninety five confidence level[6]. Figure 5 reports the results of the recursive tests of cross-equation restrictions, while Figure 6 reports the ET-consistent simulated spread between the yield on a ten-year bond and the policy rate along with the spread between the same two yields. The null of the validity of the ET is now never rejected by the within sample procedure for all possible sample splits. The simulation procedure confirms these results by showing that the actual spread is always within the 95 per cent confident intervals produced by forward simulation under the null of the present value model consistent with a weak form of the Expectations Theory.

For comparison we report in Figure 7 the out-of-sample based on the bivariate VAR originally adopted by CS. Interestingly, the simulation procedure in this case does not confirm the results of the within sample procedure. In fact, the simulation procedure never rejects the null, while the within sample tests illustrated in Figure 3 does so virtually for the sample splits considered. Note also that the confidnce intervals associated to the simulated spread are wider than in the case of the model with three financial factors. Overall, the empirical evidence suggests that the original CS work has probably used an incomplete information set and has underestimated uncertainty in the prediction of future policy rates.

It is interesting to consider the time-series behaviour of the difference between actual and simulated series, which depend on risk premia and on the differences between model based forecasts and agents' expectations. There are many observation in which the actual series takes a smaller value than

[5]The lenght of our VAR was set to two on the basis of the Akaike, Schwartz and BIC lag selection criteria.

[6]The trace statistics for the null of at most two cointegrating vectors yieded an observed values of 8.57, while the five per cent critical value is 3.76 (We allowed for a constant restricted to belong to the cointegrating vector)

the simulated one. Given that risk premia cannot be negative, this evidence implies that deviations of simulated from actual variables reflect persistent expectations error for policy rates. Therefore, the approach of proxying ex-ante expected returns with ex-post observed returns, commonly adopted in the limited information, single-equation tests of the ET theory, might seriously underestimate the role of forecasting errors. We shall reconsider this point in section 6 of the paper.

5 Macroeconomic Factors

Interest rate rules, which feature (very) persistent of policy rates responding to central bank's perceptions of (expected) inflation and output gaps (Taylor,1993, Clarida, Gali and Gertler, 1998, 1999, 2000) not only track the data well but are also capable of explaining the high inflation in the seventies in terms of an accommodating behaviour towards inflation in the pre-Volcker era.

The success of Taylor rules points out an obvious potential misspecification of the original Campbell-Shiller framework: the omission of macroeconomic variables to which the monetary policy maker reacts. Interestingly, Fuhrer(1996) uses a simple Taylor-rule type reaction function, the expectations model and reduced-form equations for output and inflation to solve for the reaction function coefficients that delivers long-term rates consistent with the Expectations Theory. He finds that modest and smoothly evolving time-variation in the reaction functions parameters is sufficient to reconcile the expectations model with the long-bond data.

The traditional argument of a Taylor rule are expected inflation and some measure of the output gap. Our framework for simulating policy rates is gear to mimic the decisions of agents in real-time. Orphanides (2001) has shown that data revisions could generate misleading inference. For this reason, as suggested by Evans(2003), we consider as macroeconomic factors variables which are not subject to revision: the CPI inflation and unemployment rate. Our VAR with financial factor and macroeconomic variables takes the following specification:

$$\begin{bmatrix} \Delta r_t^{rf} \\ -S_t \\ C_t \\ \pi_t \\ UN_t \end{bmatrix} = A(L) \begin{bmatrix} \Delta r_{t-1}^{rf} \\ -S_{t-1} \\ C_{t-1} \\ \pi_{t-1} \\ UN_{t-1} \end{bmatrix} + u_t \tag{19}$$

As in the VAR with financial factors our representation is stationary and it allows for the cointegrating relationship which constitute a necessary condition for the ET to hold, being also consistent with the presence of a stationary risk premium[7]. The results of the recursive within sample test and of the simulation based out-of-sample procedure are reported respectively in Figure 8 and Figure 9. Results are very similar to those obtained by the model with three-financial factors only: the present value restrictions are not rejected both by the within sample test and the out-of-sample simulation procedure. An assessment of the incremental benefit of including macroeconomic variables in the model requires an evaluation of the improvement of the forecasting performance for policy rates, to which we devote the next section.

6 Assessing predictability of policy rates

We assess the impact of the enlargement of the information set by considering the predictability of policy rates from the one-month to the 24-month horizon. We consider policy rates predictions based on the following alternative specifications: the random walk model, the CS VAR model, the model augmented with financial factors and the model augmented with macro and financial factors. Consistently with what we have done in the previous sections we derive up to n-step ahead forecasts from the recursive estimation of each proposed model and we then proceed to their evaluation on the basis of a typical efficiency type regression:

$$r_{t+j} = \hat{\beta}_{0j} + \hat{\beta}_{1j} E_t r_{t+j} + \hat{u}_{tj}, \tag{20}$$

where $E_t r_{t+j}$ is the model based prediction for i_{t+j} based on the information set available at time t.

We report in Figure 10 the estimated $\hat{\beta}_{1j}$, actual policy rates at time t with predicted policy rates, based on model simulation run at time $t-j$, with j= 1,...48 months:

Figure 10 clearly shows that using a three factor model for the term structure generates a clear improvement in the forecasting performance with respect to simple univariate and bivariate specifications. The further inclusion of macroeconomic factors does not generate any sizeable improvement. We have also clear evidence of an hump-shape pattern in the performance of the model to forecast future policy rates. Policy rates are predicted very well at

[7] The trace statistics for the null of at most four cointegrating vectors yieded an observed values of 5.22, while the five per cent critical value is 3.76 (We allowed for a constant restricted to belong to the cointegrating vector)

short horizon then the forecasting performance consistently worsens from the three months to the two year horizon, to improve again form the two-year horizon onwards. (Very) Interestingly such pattern matches the evidence often reported in single-equation studies of the implication of the yield spreads for movements in the short rates. When the maturity of the long-bond is three month or less, short rates are generally found to move as predicted by the ET; for maturities between about three months and two years, short rate do not on average react to long-short spreads; and for maturities beyond two years the long-short spread again predicts future short rate movements (see, for example, Roberds and Whiteman(1996)). The match between the predictability "smile" for policy rates and the predictability "smile" for the regression of the change of short rates on the spreads is interesting in that it might be very well the case that the former is caused by the latter when realized ex-post rates are taken as a proxy for ex-ante expected rates in the single equation approach. [8]

The results of our regressions are also comparable with the empirical results recently provided by Rudebusch(2001). Rudebusch runs predictive regressions using expected policy rates implicit in Federal Funds future contracts. Given the availability of future contracts, he considers forecasting horizons up to nine months to obtain results very similar to ours. The very low predictability of policy rates at the six-months and the nine-months horizon is taken as a strong argument supporting the conclusion that monetary policy inertia is an illusion.

Importantly, the fact that the predictive regressions based on model projections and Federal Fund future give very similar results does not contradict the assumption that there is no major discrepancy between the information set used by agents and that implicit in our econometric specification. However, our results are against the conclusion that monetary policy inertia is an illusion. In fact, we predict policy rates using a model which features strong persistence and we still find very little predictability for policy rates at horizons between six-months and one-year. Persistence is only a necessary conditions for predictability of policy rates when they are set according to a rule which react to macroeconomic conditions. In this case stability of the rule, precision in the estimation of parameters, and predictability of financial factors and macroeconomic variables are required along with persistence to generate predictability. Our model-based simulations suggest that these conditions do not occur at

[8] In commenting the predictability smile in single eqaution models Campbell, Lo and McKinley(1997,p.423) state

"... The U-shaped pattern of regression coefficients in Table 10.3 may be explained by reduced forecastability of interest rates movements at horizons around one-year. There may be some short-run forecastability arising from Federal Reserve operating procedures, and some long-run forecastability arising from business-cycle effects on interest rates..."

frequencies between six-months and one-year. Interestingly, predictability improves again for horizons higher than one-year where business-cycle and its effects on interest rates become more predictable.

7 What have we learned? A discussions of our results and their relation to the literature

In this paper we attempted to explain the rejection of the ET model of the term structure and the anomaly observed in CS. We have simulated the real time decision of agents who forecast policy rates by projecting forward a model including financial factors and macro variables to generate long-term rates consistent with the expectations theory along with a confidence interval reflecting the uncertainty associated to out-of-sample forecasting. Our evidence shows that, for different specifications of the information set, the observed long-term yields are contained in the confidence interval generated by our model and hence our results contradicts most of the available evidence on this model. We interpret the difference between our results and those usually reported in the literature as a consequence of the limited information approach adopted in single-equation test of the Expectation Theory and of the within sample approach adopted in the bivariate model of the term structure in CS. Our section on the predictability of policy rates shows how inappropriate is the common practice of using ex-post observed returns as a proxy for ex-ante expected returns and the improvement in forecast generated by adopting an enlarged information set.

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.

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.

No-arbitrage models with observable factors feature a complicated parameterization and cannot accommodate time variation in the parameters of the state vector relevant for pricing bonds. Within this approach, the failure of ET is entirely abscribed to the presence of a time-varying risk premium, which is modelled as a linear function of the state of the economy. There is a lot in common between the latest developments of the no-arbitrage approach and the approach to the term structure proposed in our paper. We share the view on the importance of augmenting the information set with macroeconomic and financial factors to model the yield curve but we concentrate on the role of expectations errors in explaining the failure of the Expectations Theory rather than on the risk premium. The main cost of our approach is that we do not impose no-arbitrage restrictions, while the main advantage is a much more parsimonious (and linear) parameterization, which easily accommodates time-variation in the parameters describing the state vector relevant for pricing bonds. As a consequence, our approach delivers forecasts of the short-rate which clearly outperform the random-walk. Our findings suggests that the importance of fluctuations of risk premia in explaining the deviation from the ET might be reduced when some forecasting model for short-term rates is adopted and a proper evaluation of uncertainty associated to policy rates forecast is considered. We believe that improving the forecasting model for policy rates within a no-arbitrage approach is an important step to assess the relative weight of forecasting errors and risk premia in explaining deviations from the Expectations Theory. This is on our agenda for future research.

References

[1] Ang A. and M.Piazzesi(2003) "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables", Jour-

nal of Monetary Economics, 50, 745-787

[2] Ang A., Piazzesi M. and M.Wei(2003) "What does the yield curve tell us about GDP growth?" paper available from http://www.columbia.edu/~aa610

[3] Balduzzi P., Bertola G., and Foresi S., 1997 "A model of target changes and the term structure of interest rates" Journal of Monetary Economics, 24, 371-399.

[4] Campbell J.,1995, "Some Lessons from the Yield Curve", Journal of Economic Perspectives, 9(3), 129-152.

[5] Campbell J., A. Lo and C.MacKinlay, 1997, *The Econometrics of Financial Markets*, Princeton University Press: Princeton

[6] Campbell,J.,and Shiller,R. "Cointegration and Tests of Present Value Models" J.P.E. 95 (1987) 1062-1088

[7] Chen R.R. and L. Scott, (1993) "Maximum Likelihood estimation for a multi-facor equilibrium model of the term structure of interest rates"Journal of Fixed Income, 3, 14-31.

[8] Clarida R., J. Gali and M. Gertler, 1998 "Monetary policy rules in practice:some international evidence" *European Economic Review*, 42,

[9] Clarida R., J. Gali and M. Gertler,1999 "The science of monetary policy: A new-Keynesian perspective", *Journal of Economic Literature,* XXXVII (4), 1661-1707.

[10] Clarida R., J. Gali and M. Gertler, 2000 "Monetary policy rules and macroeconomic stability: evidence and some theory", *The Quarterly Journal of Economics*, 115, 1, 147-180

[11] Cochrane J.(2001) "Asset Pricing", Princeton University Press

[12] Diebold F.X. and C. Li(2002) "Forecasting the Term Structure of Government Bond Yields", mimeo, University of Pennsylvania

[13] Diebold F.X., G.D. Rudebusch and S.B.Aruoba(2003) "The Macroeconomy and the yield curve", mimeo University of Pennsylvania

[14] Duffee G.R.(2002) "Term Premia and Interest Rate Forecast in Affine Models", Journal of Finance, 57, 1

[15] Elton E.J.(1999) "Expected return, realized return and asset pricing tests", Journal of Finance

[16] Evans C. (2003) "Real-time Taylor rules and the federal funds futures market' Federal Reserve Bank of Chicago, Economic Perspective

[17] Fama E., and R.R. Bliss (1987) "The Information in Long-Maturity Forward Rates" *American Economic Review*, 77, 680-692

[18] Favero C.A.(2002) "Taylor rules and the Term Structure", IGIER Working Paper

[19] Fuhrer J.C., "Monetary Policy Shifts and Long-Term Interest Rates" Quarterly Journal of Economics, 111,4,1183-1209

[20] Hordahl P., O.Tristani and D.Vestin(2003) "A joint econometric model of macroeconomic and term structure dynamics", mimeo ECB

[21] Johansen S.(1995) "Likelihood-based inference in cointegrated vector auto-regressive models", Oxford University Press, Oxford

[22] Mankiw N.G. and J.Miron,1986 "The changing behaviour of the term strucutre of interest rates" ,*The Quarterly Journal of Economics*, 101, 211-221

[23] McCallum(1994) "Monetary Policy and the Term Structure of Interest Rates", NBER Working Paper 4938

[24] Nelson C.R. and A.F. Siegel (1987) "Parsimonious modelling of yield curves", *Journal of Business,* 60, 473-489

[25] Orphanides A. (2000) "The quest for prosperity without inflation", *European Central Bank W.P.*no. 15, March.

[26] Roberds W. and C.H. Whiteman(1996) "Endogenous Term Premia and anomalies in the Term Structure of Interest Rates: Explaining the Predictability Smile", Federal Reserve Bank of Atlanta, WP 96-11

[27] Roush J.(2003) "evidence ncovered: Long-term interest rates, Monetary Policy and the Expectations Theory", mimeo, Board of Governors of the Federal Reserve System

[28] Rudebusch G.D G.D.,1995 "Federal reserve interest rate targeting, rational expectations, and the term structure." Journal of Monetary Economics, 35, 245-274.

[29] Rudebusch G.D.(2002) "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia" Journal of Monetary Economics, 49, 1161-1187

[30] Shiller,R.(1979) "The Volatility of Long Term Interest Rates and Expectations Models of the Term Stucture" J.P.E. 87 1190-1219

[31] Shiller,R.(1981) "Alternative Tests of Rationals Expectations Models: the Case of the Term Structure" J. of Econometrics 16 71-87

[32] Shiller,R.,Campbell,J., and Schoenholtz, K. "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates" Broking Papers in Economic Activity, no.1 1983 173-217

[33] Taylor J.B. (1993) "Discretion versus policy rules in practice", *Carnegie-Rochester Conference Series on Public Policy*, 39, 195-214

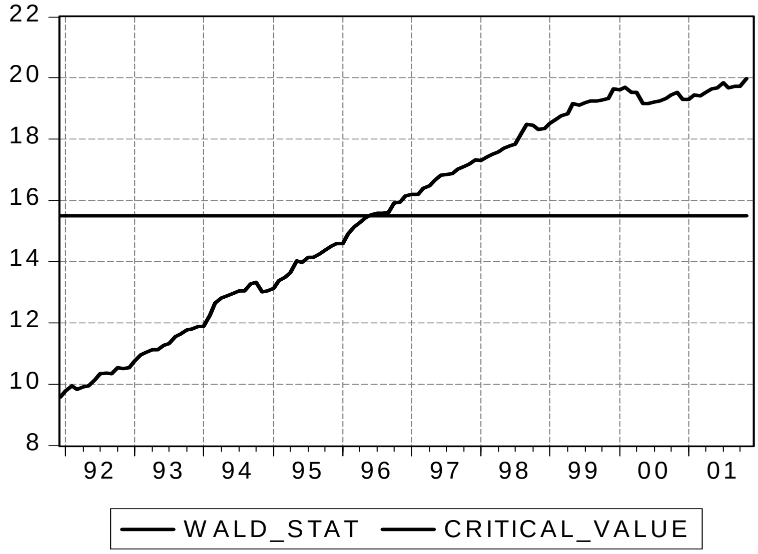


Figure 1: Recursive tests (and five per cent critical value) for the validity of the cross-equation restrictions implied by the Expectations Theory in a VAR with two financial factors (change in policy rates and spread, as in CS)

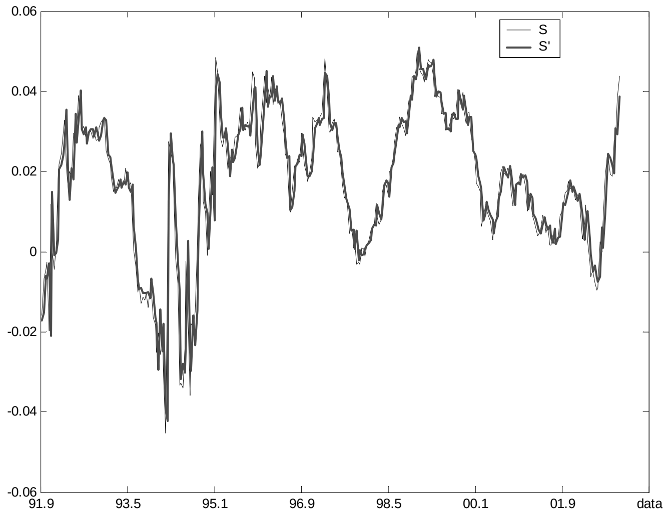


Figure 2: Actual and ET-consistent spreads in the CS model

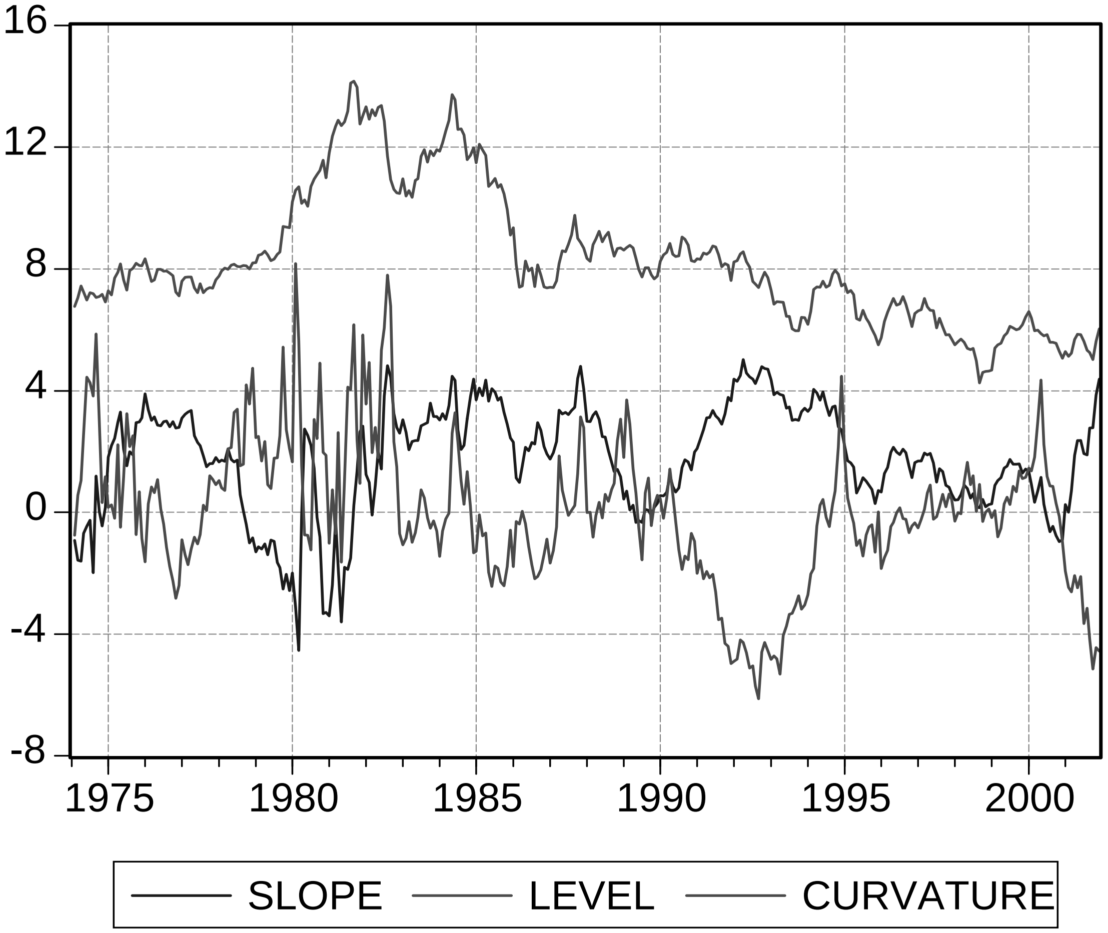
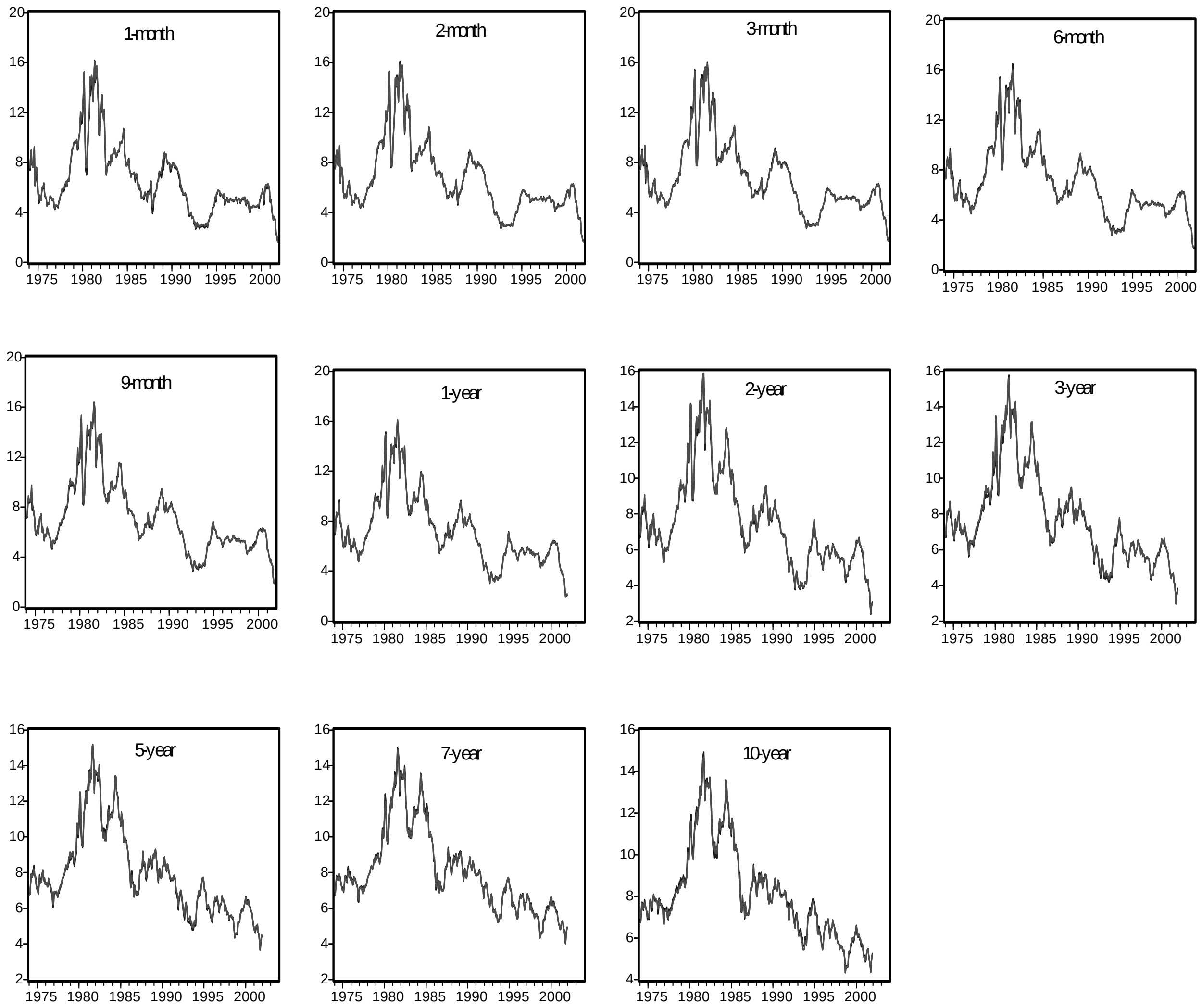


Figure 3: the time series of the three Nelson-Siegel factors for the US yield curve


1-month
2-month
3-month
6-month
9-month
1-year
2-year
3-year
5-year
7-year
10-year

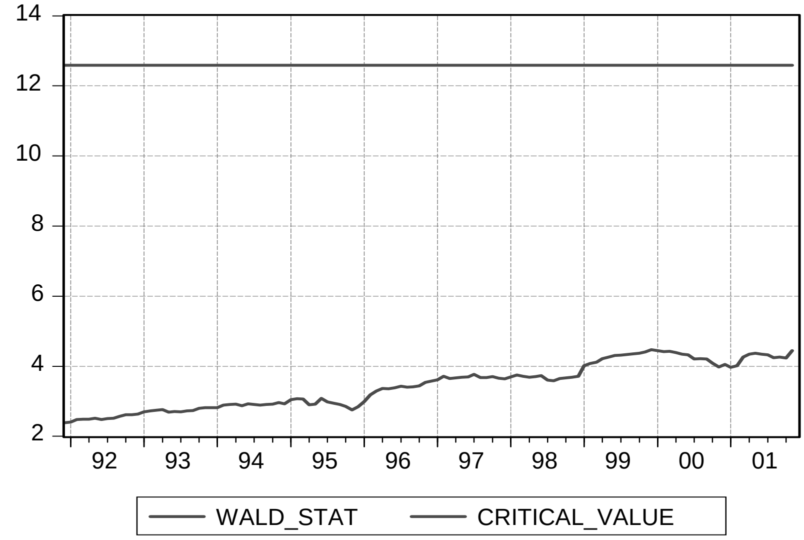
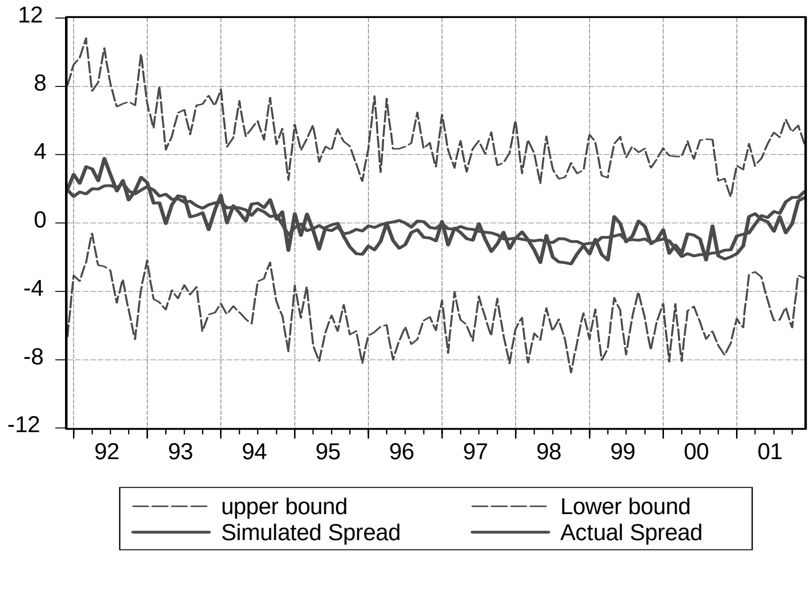


Figure 5: Recursive tests (and five per cent critical value) for the validity of the cross-equation restrictions implied by the Expectations Theory in a VAR with three financial factors (change in policy rates, spread and curvature)



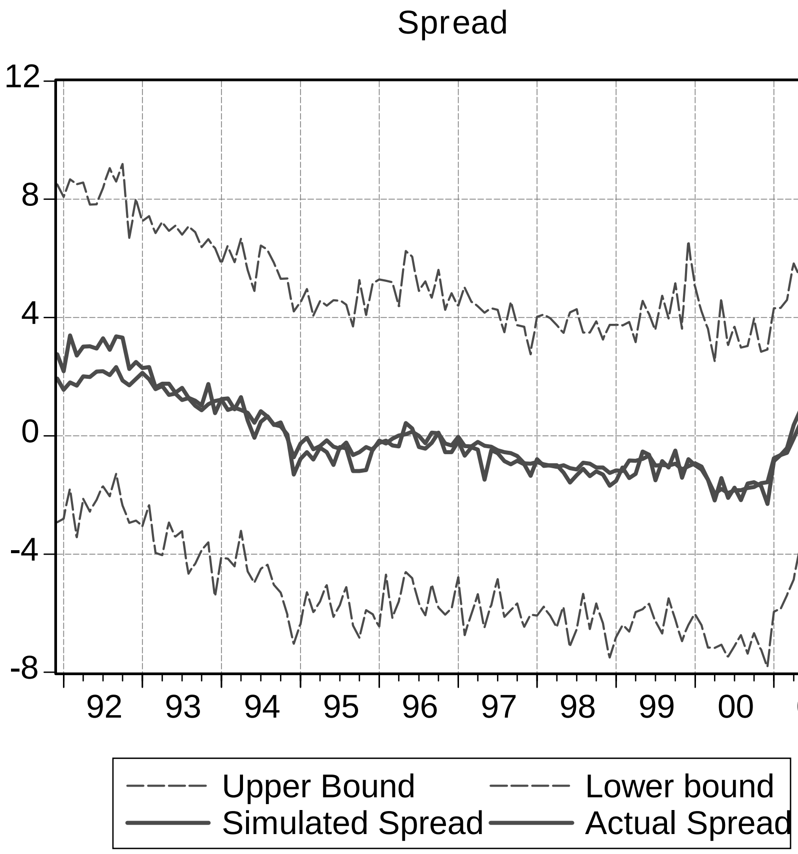


Figure 7: Out-of-sample simulated ET-consistent spread for the bi-variate CS Model, its 95% Confidence Interval and Actual Spread

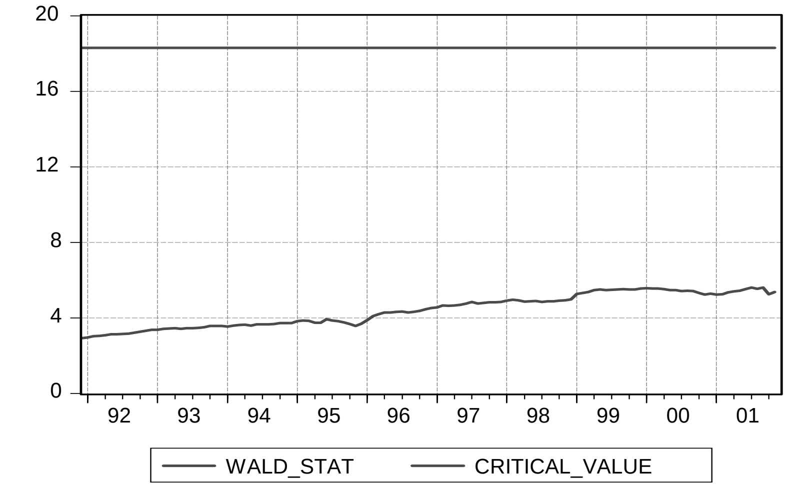


Figure 8: Recursive tests (and five per cent critical value) for the validity of the cross-equation restrictions implied by the Expectations Theory in a VAR with three financial factors (change in policy rates, spread and curvature) and two macroeconomic variables (unemployment rate and CPI inflation)

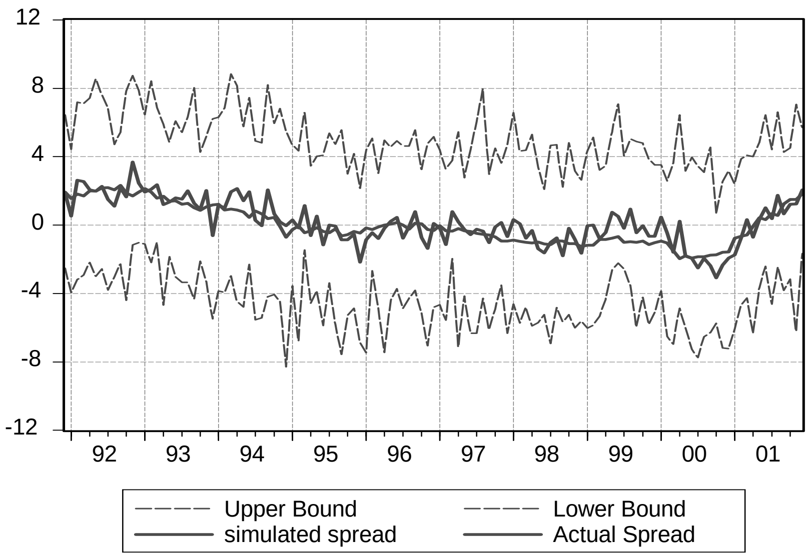
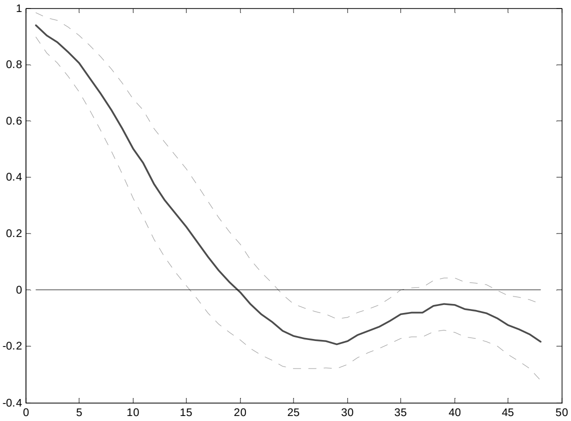
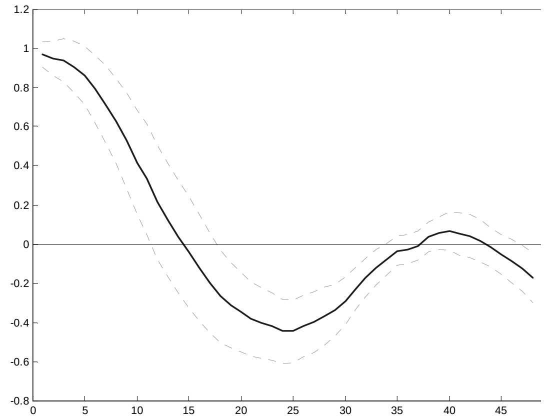


Figure 9: Out-of-sample simulated ET-consistent spread for the Macroeconomic Factors Model

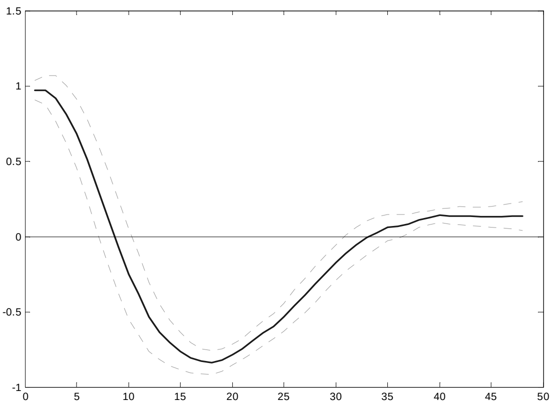
Figure 10: Forecasting performance of different model in predicting the n-step ahead policy rates



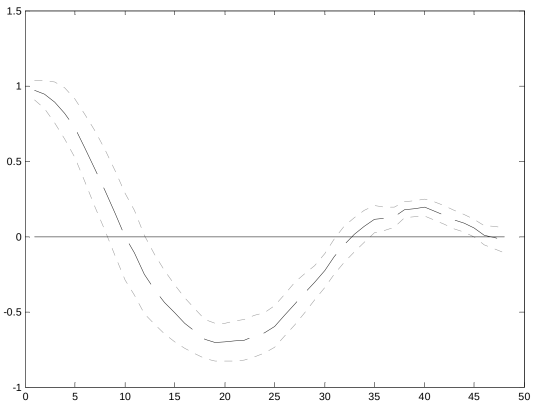
Random walk model



Model with two-financial factors



Model with three-financial factors



Model with three factors and two macro varia

The figures report the slope coefficients, $\hat{\beta}_{1j}$, for the following efficient prediction type of regressions

$$r_{t+j} = \hat{\beta}_{0j} + \hat{\beta}_{1j} E_t r_{t+j} + \hat{u}_{tj}$$

we consider forecasting horizon from one-month ahead to forty-eight months ahead. Forecasts are based on four different models: the simple random walk, the CS model with two financial factors, a model with three financial factors and a model with three financial factors and two macroeconomic variables, the rate of unemployment and CPI inflation.