The Term Structure of Euromarket Interest Rates: Some New Evidence*

Richard H. Clarida[a,b], Lucio Sarno[c,d], Mark P. Taylor[c,d] and Giorgio Valente[c]

a: Columbia University

b: National Bureau of Economic Research (NBER)

c: University of Warwick

d: Centre for Economic Policy Research (CEPR)

January 2004

Abstract

We examine the dynamic relationship between Euromarket interest rates at different maturities for the US and the UK over the period 1982-2000. We use a general, multivariate vector equilibrium correction modelling framework incorporating both asymmetric adjustments and regime shifts, which may be justified by a very general theoretical model of the term structure and which allows for both term premia and deviations from rational expectations. Our approach provides good in-sample fits to the data and has satisfactory out-of-sample forecasting properties. Further, the implied regime switches are found to be closely related to key state variables driving monetary policy decisions, namely inflation and business cycle indicators.

JEL classification: E43; E47.

Keywords: term structure of interest rates; nonlinearity; Markov switching; forecasting.

Word count: 5,462 words.

* *Acknowledgments:* This paper was partly written while Lucio Sarno and Mark Taylor were visiting the International Monetary Fund. Sarno and Valente acknowledge financial support from the Economic and Social Research Council (ESRC). The paper was written for the 2004 *Royal Economic Society* Annual Conference, University of Wales, Swansea, 5-7 April 2004. The authors alone are responsible for any errors that may remain and for the views expressed in the paper.

1 Introduction

In this paper we present some new evidence on the term structure of Euromarket interest rates and, in particular, show how information may be extracted from the term structure to provide out-of-sample interest rate forecasts. Our analysis is based on an examination of the dynamic relationship between interest rates of different maturities for the US and the UK using a very general, multivariate vector equilibrium correction modelling framework capable of simultaneously allowing for asymmetric adjustment and regime shifts, but which may be rationalized by reference to a very general model of the term structure capable of admitting both time-varying term premia and deviations of market expectations from rational expectations. The resulting nonlinear vector equilibrium correction models are shown not only to provide good in-sample fits to the data and economically interpretable regimes but also to have satisfactory out-of-sample forecasting properties.

In an early paper, Campbell and Clarida (1986) empirically investigate the predictability and comovement of risk premia in the term structure of Euromarket interest rates, demonstrating that risk premia in three Euromarket term structures and on uncovered foreign asset positions move together. Subsequently, following the seminal paper of Campbell and Shiller (1987) on cointegration and present value models, which, *inter alia*, demonstrates the cointegrating relationship between short- and long-term interest rates implied by the expectations model of the term structure, a large empirical literature developed which focused on building equilibrium correction models of the dynamic interaction between interest rates at different maturities (e.g. Campbell and Shiller, 1991; Hall, Anderson and Granger, 1992; Taylor, 1992).

Recently, an interesting strand of this literature has explored the role of asymmetric or nonlinear adjustment towards equilibrium in modelling interest rate movements. In this work, researchers have argued that the dynamics of the term structure of interest rates may be characterized by a nonlinear equilibrium correction model due to factors such as, for example, non-zero or asymmetric transactions costs or infrequent trading or the existence of regime shifts (e.g. Gray, 1996; Anderson, 1997; Enders and Granger, 1998; Bansal and Zhou, 2002; Coakley and Fuertes, 2002; Kim, 2003; Sarno and Thornton, 2003).[1] In addition to this growing amount of statistical evidence, there are sound

[1] It is interesting to note in this context that Hamilton's seminal paper on Markov switching (Hamilton, 1988) involved an application to the term structure of interest rates.

economic reasons to believe that regime shifts and asymmetries may improve our understanding of the behavior of the entire yield curve. For example, business cycle expansions and contractions may have statistically and economically important first-order effects on expectations of inflation, monetary policy and nominal interest rates. Further, on economic grounds, regime shifts and asymmetries may generate significant impacts not only on the short-term interest rate but also on the whole term structure of interest rates.

The research reported in this paper represents, to the best of the present authors' knowledge, the most general empirical model of the term structure to date. Using data for the sample period 1982-2000 for UK and US eurodeposit interest rates with five different maturities, we show first that a long-run equilibrium relationship between the five different interest rates consistent with the expectations theory of the term structure can be established. However, we find that conventional linear vector equilibrium correction models are easily rejected when tested against asymmetric regime-switching vector equilibrium correction models. Then, employing a Markov-switching, asymmetric vector equilibrium correction approach that allows for time-varying term premia, we are able to characterize satisfactorily the dynamic relationship between interest rates with different maturities for both countries. The regime-switching probabilities implied by the model appear to be intimately related to the key state variables driving monetary policy decisions—namely inflation and a business cycle indicator, the output gap—which has a natural economic interpretation (Clarida, Gali and Gertler, 1998, 1999, 2000; Ioannides and Peel, 2003). This model outperforms, both in-sample and out-of-sample, a range of alternative linear and nonlinear equilibrium correction models. Overall, these results show that, while allowing for both asymmetries and regime shifts is key to producing a satisfactory statistical representation of the term structure, asymmetries seem to play a particularly important role in enhancing the out-of-sample forecasting performance of the models.

The remainder of the paper is set out as follows. In Section 2 we provide a brief overview of the conventional theory of the term structure of interest rates and its basic statistical implications for the behavior of interest rates with different maturities, while in Section 3 we describe the recently developed econometric procedure which has allowed the extension of Markov-switching techniques to nonstationary systems and, in particular, to cointegrated vector autoregressions and their representation as time-varying vector equilibrium correction models. In Section 4 we describe the data set and report our empirical results from employing conventional unit root tests, cointegration and equilibrium correction analysis as well as from executing asymmetry and linearity tests. We also discuss the estimated Markov-switching vector equilibrium correction models in this section and provide an interpretation of their implied regime-switching probabilities in terms of monetary policy and business cycles. Section 5 discusses our forecasting results, and in a final section we briefly summarize our main results and conclude.

2 A general model of the term structure of interest rates

Let $i_{k,t}$ and $f_{k,t}$ be the yield to maturity of a k-period pure discount bond and the forward rate, defined as the contract rate of a one-period pure discount bond bought at time t that matures at time $t+k$. The relationship linking $i_{k,t}$ and $f_{k,t}$ may be described, according to the Fisher-Hicks recursive formulae, as $i_{k,t} = \frac{1}{k}\left(\sum_{j=1}^{k} f_{j,t}\right)$ for $k = 1, 2, 3, \ldots$. As is well known, the forward rate differs from the expected future yield to maturity because of term premia required by investors for risk considerations and preferences for liquidity. Assume that the relationship between forward rates and expected rates is characterized as $f_{j,t} = E_t\left(i_{k,t+j-1}\right) + \phi_{j,t}$, where E_t is the mathematical expectation operator conditioned on information available at time t, and $\phi_{j,t}$ is the term premium. We can then re-write the Fisher-Hicks formulae as follows:

$$i_{k,t} = \frac{1}{k}\left[\sum_{j=1}^{k} E_t\left(i_{1,t+j-1}\right)\right] + \gamma_{k,t}, \tag{1}$$

where $\gamma_{k,t} \equiv \frac{1}{k}\sum_{j=1}^{k}\phi_{j,t}$ denotes a variable capturing the effects of term premia components.

Equation (1) may be viewed as a general relationship linking yields at different maturities and shows clearly that yields having similar maturities move together. The Expectations Hypothesis (EH) of the term structure of interest rates focuses essentially on the properties of the premia components $\gamma_{k,t}$. According to the Pure Expectations Hypothesis the term premia are all identically equal to zero, $\gamma_{k,t} \equiv 0$, implying that the one-period holding yield of a k-period bond is equal to the yield to maturity of a one-period bond. A milder version of the EH asserts the less stringent proposition that the term premia are constant over time. In fact, in this paper we shall consider a very weak version of the EH which allows the term premia $\gamma_{k,t}$ to be time-varying and requires only that they be realizations of stationary stochastic processes. The EH has important and clear statistical implications (Hall, Anderson and Granger, 1992) which can easily be seen by rewriting equation (1) as follows:

$$i_{k,t} - i_{1,t} = \frac{1}{k}\left[\sum_{m=1}^{k-1}\sum_{j=1}^{m} E_t \Delta i_{1,t+j}\right] + \gamma_{k,t}, \tag{2}$$

where Δ is the first-difference operator. Assuming that the yields to maturity are realizations of stochastic processes integrated of order one, $I(1)$, then if the term premia components are stationary, all terms on the right-hand side of equation (2) must be stationary, which implies that the term on the left-hand side of (2) must be stationary also, i.e. $(i_{k,t} - i_{1,t}) \sim I(0)$. Hence, this model predicts that the yields to maturity are cointegrated with a cointegrating vector of the form $[1, -1]'$. This in turn implies that, given H different maturities, exactly $H - 1$ distinct cointegrating relationships must exist between the corresponding H yields, each given by the stationary spread $i_{k,t} - i_{1,t}$ for $k = 2, \ldots, H$. Moreover, given that cointegration between a set of variables implies, according to the Granger Representation Theorem (Engle and Granger, 1987), the existence of a statistical representation for the yields in the form of a vector equilibrium correction model (VECM), this provides a rationale for modelling the dynamic interrelationship between interest rates using a VECM approach.

It is, however, possible that the premia terms may induce important nonlinearities into this relationship—as suggested, for example, by Anderson (1997) and Kim (2003). Further, there is evidence that the dynamic adjustment of the term structure in response to deviations from equilibrium may in fact be asymmetric (Enders and Granger, 1998; Coakley and Fuertes, 2002; Sarno and Thornton, 2003) and characterized by regime shifts (Hamilton, 1988; Gray, 1996; Bansal and Zhou, 2002). In this paper, we therefore develop a VECM approach which is capable of allowing for all of these possibilities simultaneously.

While our analysis so far has assumed that agents form expectations which are equivalent to the true mathematical expectations given information at time t (i.e. rational expectations), it is worth noting that this assumption may be weakened to allow for non-rational market expectations where the deviations of the average market expectation from the mathematical conditional expectation are $I(0)$ (which seems a reasonable and minimal assumption to make). To see this, let $\exists_t x_{t+k}$ denote the market expectation at time t of a variable x k periods later, and replace the conditional expectations operator in equation (2):

$$i_{k,t} - i_{1,t} = \frac{1}{k}\left[\sum_{m=1}^{k-1}\sum_{j=1}^{m} \exists_t \Delta i_{1,t+j}\right] + \gamma_{k,t}. \tag{3}$$

This may be rewritten:

$$i_{k,t} - i_{1,t} = \frac{1}{k}\left[\sum_{m=1}^{k-1}\sum_{j=1}^{m} E_t \Delta i_{1,t+j}\right] + \delta_{k,t} + \gamma_{k,t}, \tag{4}$$

where

$$\delta_{k,t} \equiv \frac{1}{k}\left[\sum_{m=1}^{k-1}\sum_{j=1}^{m} \exists_t \Delta i_{1,t+j} - \sum_{m=1}^{k-1}\sum_{j=1}^{m} E_t \Delta i_{1,t+j}\right]. \tag{5}$$

Now, it seems intuitively reasonable to assume that deviations of the average market expectation from the true mathematical expectation are indeed $I(0)$ along with the term premia, since otherwise expectational errors would tend to grow without bound. But if this is the case, we have that $\delta_{k,t} + \gamma_{k,t} \sim I(0)$, and we can infer $(i_{k,t} - i_{1,t}) \sim I(0)$ and the cointegration properties of the term structure exactly as before.

3 Asymmetric Markov-switching equilibrium correction

In this section we outline the econometric procedure employed in order to model regime shifts in the dynamic relationship implied by the EH theory of the term structure of interest rates as discussed in the previous section. The procedure essentially extends Hamilton's (1988, 1989) Markov-switching regime framework to nonstationary systems, allowing us to apply it to cointegrated vector autoregressive (VAR) and VECM systems (Krolzig, 1997).

Consider y_t being a K-dimensional vector time series process, $y_t = [y_{1t}, y_{2t}, \ldots, y_{Kt}]'$ where y_t is nonstationary but first-difference stationary, i.e. $y_t \sim I(1)$. Then, given $y_t \sim I(1)$, there may be up to $K-1$ linearly independent cointegrating relationships, which represent the long-run equilibrium of the system (Granger, 1986; Engle and Granger, 1987). If indeed there is cointegration, the joint dynamics of the vector time series process can be written as a Markov-switching vector equilibrium correction model or MS-VECM of the form:

$$\Delta y_t = \nu(s_t) + \sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + \Pi y_{t-1} + u_t, \tag{6}$$

where $\nu(s_t) = [\nu_1(s_t), \nu_2(s_t), \ldots, \nu_K(s_t)]'$ is a K-dimensional column vector of regime-dependent intercept terms; the Γ_d's are $K \times K$ matrices of parameters, $\Pi = \alpha\beta'$ is the $K \times r$ cointegrating matrix with r being the number of cointegrating relationships; u_t is a K-dimensional vector of error terms, $u_t \sim NIID(\mathbf{0}, \Sigma_u(s_t))$. The regime-generating process is assumed to be an ergodic Markov chain with a finite number of states $s_t \in \{1, \ldots, M\}$ governed by the transition probabilities $p_{ij} = \Pr(s_{t+1} = j \mid s_t = i)$, and $\sum_{j=1}^{M} p_{ij} = 1 \; \forall i, j \in \{1, \ldots, M\}$.

The present application focuses on a multivariate model comprising, for both countries analyzed, the spot-next eurorate and the rates relative to one month (four weeks), three months (thirteen weeks), six months (twenty-six weeks) and twelve months (fifty-two weeks) to maturity so that $y_t = [i_{0,t}, i_{4,t}, i_{13,t}, i_{26,t}, i_{52,t}]'$, for which, following the reasoning of Section 2, four unique independent cointegrating relationships should exist.[2] In order to take into account the empirical evidence indicating that interest rates display asymmetric adjustment (e.g. Enders and Granger, 1998; Coakley and Fuertes, 2002; Sarno and Thornton, 2003), we allow the Markov-Switching-Intercept-Hetereskedastic VECM, or MSIH-VECM, (6) to display differing speeds of adjustment to equilibrium depending on whether there are positive or negative deviations from the equilibrium condition:

$$\Delta y_t = \nu\left(s_t\right) + \sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + I_t \Pi^{+} y_{t-1} + \left(1 - I_t\right) \Pi^{-} y_{t-1} + e_t, \tag{7}$$

where I_t is a $1 \times r$ vector whose $j - th$ element at time t, $\iota_{j,t}$ say, is the Heaviside indicator, taking the value zero or unity according to whether the lagged $j - th$ cointegrating combination—the $j - th$ element of $\beta'\, y_{t-1}$, $h_{j,t-1}$ say—is positive or negative:[3]

$$\iota_{j,t} = \begin{cases} 1 & \text{if} \;\; h_{j,t-1} > 0 \\ 0 & \text{if} \;\; h_{j,t-1} \leq 0 \end{cases} . \tag{8}$$

(In practice, as discussed above, we expect these cointegrating combinations to be the four term spreads.) Since the cointegrating parameter vectors are constant, the parameter matrices Π^{+} and Π^{-} must be partitioned as $\Pi^{+} = (\alpha^{+})\beta'$, $\Pi^{-} = (\alpha^{-})\beta'$, so that the equilibrium correction coefficients

[2] There is a slight shift in notation here from that employed in Section 2 in that $i_{0,t}$ represents the spot-next rate, which is an overnight rate rather than a "zero-period" interest rate as the index might suggest. Since we are working with weekly data, however, it makes sense to index the interest rate variables according to the number of weeks to maturity and we have therefore indexed the spot-next rate at 0. The discussion in Section 2 applies to these variables directly if we relabel the interest rates by the number of days to maturity. This is purely a notational issue.

[3] Strictly speaking, non-negative or negative.

shift according to whether the lagged equilibrium correction term to which it applies is positive or negative. As before, we also have a vector of error terms, $e_t \sim NIID(\mathbf{0}, \Sigma_e(s_t))$, and M states $s_t \in \{1, \ldots, M\}$. This procedure essentially extends Enders and Granger's (1998) M-TAR framework to nonlinear-nonstationary systems, allowing us to apply it to cointegrated VAR and VECM systems.[4]

We now turn to a brief discussion of our data set and then to our empirical analysis.

4 Empirical results

4.1 Data, unit root tests and cointegration analysis

Our data set comprises weekly observations of spot-next and 4-, 13-, 26- and 52-week eurorates for the UK and the US spanning the period from February 7 1982 to December 31 2000, a total of 987 observations for each series.[5] In our empirical work, we carried out our estimations over the period February 1982-December 1991, reserving the remaining data for out-of-sample forecasting tests.

As a preliminary exercise, we tested for evidence of unit root behavior in each of the interest rate time series examined for both countries under investigation by calculating standard augmented Dickey-Fuller (ADF) test statistics. In keeping with the large number of studies of unit root behavior for these time series, we were in each case unable to reject the unit root null hypothesis at conventional nominal levels of significance.[6] Hence, the results of the unit root tests suggest that testing for cointegration between the five interest rate series is the logical next step.[7]

[4] The (symmetric or asymmetric) MS-VECM can be estimated using a two-stage maximum likelihood procedure. The first stage of this procedure essentially consists of the implementation of the Johansen (1988, 1991) maximum likelihood cointegration technique in order to test for the number of cointegrating relationships in the system and to estimate the matrix of cointegrating parameters β. In fact, use of the conventional Johansen procedure is legitimate in the first stage without modelling the Markovian regime shifts explicitly (Saikkonen, 1992; Saikkonen and Luukkonen, 1997). The second stage then consists of the implementation of an expectation-maximization (EM) algorithm for maximum likelihood estimation which yields estimates of the remaining parameters of the model (Dempster, Laird and Rubin, 1977; Hamilton, 1993; Kim and Nelson, 1999).

[5] We are grateful to the Bank for International Settlements (BIS) for supplying the data.

[6] There is an apparent conflict between a large empirical literature on interest rates, which (at least since Engle and Granger, 1987) either assumes or finds that interest rates are nonstationary processes, and conventional economic and finance theory, which often assumes that interest rates are stationary processes. For example, see the vast finance literature assuming a Vasicek (1977) model of interest rates, which is simply a mean-reverting process representable as an Ornstein-Uhlenbeck process. We follow the empirical literature because very persistent series with a root at least very close (if not equal) to unity are better approximated by $I(1)$ processes than by stationary ones (see, for example, Stock, 1997).

[7] The results of the unit-root tests are not reported to save space, but are available from the authors on request.

We then employed the Johansen (1988, 1991) maximum likelihood procedure in a VAR for $y_t = [i_{0,t}, i_{4,t}, i_{13,t}, i_{26,t}, i_{52,t}]'$ and an unrestricted constant term.[8] On the basis of the Johansen likelihood ratio test statistics for the cointegrating rank, reported in Table 1, we could strongly reject the hypothesis of three independent cointegrating vectors against the alternative of four, but were not able to reject the hypothesis of exactly four cointegrating vectors for each of the countries examined at conventional nominal test sizes. Hence, we conclude that there are exactly four cointegrating relationships between the five rates examined, for both the UK and the US.

In order to identify the cointegrating vectors uniquely, we then tested the over-identifying restrictions on the β' matrix of cointegrating coefficients suggested by the framework discussed in Section 2:

$$\beta' y_t = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} i_t^0 \\ i_t^4 \\ i_t^{13} \\ i_t^{26} \\ i_t^{52} \end{bmatrix}. \tag{9}$$

For each country examined these restrictions were rejected by the data at standard significance levels. Nevertheless, we proceeded to examine whether the departures from the null hypothesis were large by imposing the following exactly-identifying restrictions:

$$\beta' y_t = \begin{bmatrix} -1 & \theta_4 & 0 & 0 & 0 \\ -1 & 0 & \theta_{13} & 0 & 0 \\ -1 & 0 & 0 & \theta_{26} & 0 \\ -1 & 0 & 0 & 0 & \theta_{52} \end{bmatrix} \begin{bmatrix} i_{0,t} \\ i_{4,t} \\ i_{13,t} \\ i_{26,t} \\ i_{52,t} \end{bmatrix}, \tag{10}$$

where θ_i, for $i = 4, 13, 26, 52$, is unrestricted. This estimation yielded the results reported in Table 2. These results suggest that the departure from the overidentifying restrictions, although statistically significant at conventional test sizes, is actually very small in magnitude. Indeed all of the estimated θ_i coefficients are in the range between 0.991 and 1.061 and are, therefore, very close indeed to the theoretical value of unity. Thus, rejection of the hypothesis $H_0 : \theta_i = 1 \; \forall i$ may be due to tiny departures from the null hypothesis (due, for example, to tiny data imperfections) which may not be

[8] We allowed for a maximum lag length of twelve and chose, for both countries, the appropriate lag length on the basis of conventional information criteria.

economically significant, but which appear as statistically significant given our large sample size.[9] In light of these results and given that the framework discussed in Section 2 provides strong economic priors in favor of the unity restrictions, we report below results obtained with the unity restrictions imposed.[10]

4.2 Asymmetry testing and MS-VECM estimation results

We next estimated a standard linear VECM using full-information maximum likelihood (FIML) methods:

$$\Delta y_t = \nu + \textstyle\sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + \Pi y_{t-1} + \varepsilon_t, \tag{11}$$

where $y_t = [i_{0,t}, i_{4,t}, i_{13,t}, i_{26,t}, i_{52,t}]'$, selecting the lag length on the basis of the Akaike Information Criterion, the Schwartz Information Criterion and the Hannan-Quinn Criterion. Employing the conventional general-to-specific procedure, we obtained fairly parsimonious models for each country, with no significant residual serial correlation.[11] We then investigated the presence of asymmetries in the adjustment towards equilibrium by executing standard likelihood ratio (LR) tests for the null hypothesis of symmetry. The results reported in Table 3 suggest rejection of the hypothesis of symmetry, providing clear empirical evidence that the linear VECM fails to capture significant asymmetries in the data generating process, as the restrictions imposed by the model without asymmetries are rejected with marginal significance levels (p-values) close to zero.[12]

We then proceeded to investigate the presence of nonlinearities further through the estimation of a fairly general Markov-switching model of the form:

$$\Delta y_t = \nu(s_t) + \sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + \Pi y_{t-1} + u_t, \tag{12}$$

[9] Leamer (1978, Chapter 4) points out that classical hypothesis testing will lead to rejection of any null hypothesis with a sufficiently large sample: 'Classical hypothesis testing at a fixed level of significance increasingly distorts the interpretation of the data against a null hypothesis as the sample size grows. The significance level should consequently be a decreasing function of sample size' (p. 114). See also Berkson (1938).

[10] We did, however, execute a fraction of the empirical analysis discussed below *without* imposing the unity restrictions and using instead the estimates of the cointegrating parameters reported in Table 3. The results were quantitatively extremely similar (virtually identical) and qualitatively identical to those reported below.

[11] Full details on these estimation results are available from the authors upon request, but are not reported to conserve space.

[12] We also compare below the forecasting performance of the linear VECM to that of an MS-VECM with and without asymmetries.

where $u_t \sim NIID(\mathbf{0}, \Sigma_u(s_t))$ and $s_t = 1, 2$. For parsimony considerations and consistent with previous research in this context (Gray, 1996; Ang and Bekaert, 2002; Bansal and Zhou, 2002), we limited ourselves to discriminating between linear models and Markov-switching models allowing for only two regimes in the VECM.

We applied the 'bottom-up' procedure designed to detect Markovian shifts in order to select the most adequate characterization of a 2-regime p-th order MS-VECM for Δy_t.[13] Specifically, we tested not only the hypothesis of no regime switching in the intercept but also the hypothesis of no regime switching in the variance-covariance matrix using likelihood ratio (LR) tests of the type suggested by Krolzig (1997, p. 135-6). The results indicated strong rejections of the null of no regime dependence in the intercept as well as in the variance-covariance matrix, clearly suggesting that an MS-VECM that allows for shifts in both the intercept *and* the variance-covariance matrix, namely an MSIH(2)-VECM(p), is the most appropriate model within its class in the present application. Further, in the same spirit of the test for regime-conditional intercept and homoskedasticity, we carried out a test in order to select the most parsimonious MSIH-VECM appropriately representing the dynamic relationship between the interest rates examined. In particular, considering a maximum lag length of 12 for the VAR in levels and hence a maximum lag length of 11 in the VECM formulation, we tested the null of MSIH(2)-VECM(1) against the alternative of MSIH(2)-VECM(11) and for both countries examined we were not able to reject this null hypothesis at standard significance levels.

Next, we tested each of the symmetric and asymmetric linear VECMs against their MSIH-VECM counterpart selected by means of the 'bottom-up' procedure. As shown by the LR tests in Table 4, which may be thought of as tests of the hypothesis of linearity of the VECM against the alternative of Markov-switching nonlinearity, the large test statistics indicate in each case the rejection of the (symmetric and asymmetric) VECM in favor of the alternative (symmetric and asymmetric) MSIH-VECM.

Hence, the final result of the selection procedure identifies for both countries an asymmetric

[13] Essentially, the bottom-up procedure consists of starting with a simple but statistically reliable Markov-switching model by restricting the effects of regime shifts on a limited number of parameters and checking the model against alternatives. In such a procedure, most of the structure contained in the data is not attributed to regime shifts, but explained by observable variables, consistent with the general-to-specific approach to econometric modelling. For a technical discussion of the bottom-up procedure, see Krolzig (1997).

MSIH-VECM governed by two different regimes and one autoregressive lag that may be written as follows:

$$\Delta y_t = \nu(s_t) + \Gamma_1 \Delta y_{t-1} + I_t \Pi^+ y_{t-1} + (1 - I_t)\Pi^- y_{t-1} + e_t, \tag{13}$$

where I_t is as defined above, $\Pi^+ = (\alpha^+)\beta'$, $\Pi^- = (\alpha^-)\beta'$, $e_t \sim NIID[\mathbf{0}, \Sigma_e(s_t)]$ and $s_t = 1, 2$. We estimated the MSIH-VECM (13), using an EM maximum-likelihood algorithm, for both the UK and the US.[14] The estimation yields fairly plausible estimates of the coefficients for the VECMs estimated, including the adjustment coefficients in α^+ and α^-, which were generally found to be statistically significantly different from zero.

This model is not as parsimonious as some other term structure models in the literature. However, evidence provided by Dai and Singleton (2000), Ahn, Dittmar and Gallant (2002), Bansal and Zhou (2002), Bansal, Tauchen and Zhou (2003) and Kim (2003) also indicates that a fairly rich characterization of the dynamics of the market price of risk is required to characterize satisfactorily the behavior of the term structure of interest rates. Therefore, our proposed model is consistent with this strand of the literature.

4.3 Implied regimes, the business cycle and monetary policy

For each country we employed an asymmetric MSIH-VECM with two regimes, which was found to provide an adequate characterization of the dynamics of the term structure. The regime shifts occur in the intercept and in the variance-covariance matrix. For both of the countries considered, the regime with higher variances corresponded to periods in which the average interest rate at each maturity was relatively high; this is also reflected in the fact that the high-variance regimes also had estimated intercept terms which were in every case greater than the intercept in the low-variance regimes. Thus, the two regimes may be seen as reflecting higher mean and variance in interest rates in one regime and as reflecting on average lower and less volatile interest rates in the other regime. Also, this characterization of the regimes appears to be in line with the extensive empirical literature investigating the time-varying nature of risk premia.

Having identified the two regimes as high interest rate and low interest rate regimes, the next issue we wished to investigate was whether or not the probability of switching between regimes was

[14] These estimations are not reported to save space but are available from the authors upon request.

related to macroeconomic fundamentals, in the spirit of recent research by Bansal and Zhou (2002) and Bansal, Tauchen and Zhou (2003): using a visual approach, these authors find that regime shifts in the term structure appear to be intimately associated with the business cycle. Building on these findings, and on the literature relating monetary policy to indicators of the business cycle and inflation (Taylor, 1993, 1999; Clarida, Gali and Gertler, 1998, 1999, 2000), we estimate logit models designed to explain the probability of being in either the high or the low interest rate regime using appropriate economic indicators.

In order to ensure consistency and comparability with previous research (e.g. Bansal and Zhou, 2002) and because the data on the explanatory variables we consider are not available at weekly frequency, we use monthly data. Hence, from the estimated MSIH-VECMs, we converted the weekly smoothed probabilities by monthly averaging. Further, in order to obtain a binary variable so as to be able to estimate a logit model, from the estimated average MSIH-VECM probabilities we defined a variable which is equal to zero when the average monthly probability of being in the high interest rate regime is smaller than 0.5 and equal to unity when this average probability is greater than or equal to 0.5. The explanatory variables we consider in the logit model are a business cycle indicator, namely the output gap measured as the deviation of industrial output from the Hodrick-Prescott trend, and the demeaned annualized inflation rate as a proxy for the deviation of inflation from its target or equilibrium level. Thus, the logit model may be written as follows:[15]

$$p_t\left(i^{HIGH}\right) = \frac{\exp^Z}{1+\exp^Z} \tag{14}$$

where $Z = \{\delta_0 + \delta_1 (x_t - x_t^*) + \delta_2 (\pi_t - \pi^*)\}$, $p_t\left(i^{HIGH}\right)$ is the implied MSIH-VECM probability of being in the high interest rate regime and $(x_t - x_t^*)$ and $(\pi_t - \pi^*)$ denote the measured output gap and the deviation of inflation from its mean level.[16]

As we have noted, the explanatory variables we have used are in line not only with recent empirical research on the regime-shifting behavior of interest rates (e.g. Bansal and Zhou, 2002; Bansal, Tauchen and Zhou, 2003) but also with the very large literature on monetary policy rules—so-called 'Taylor

[15] Clearly, the probability of being in the low interest rate regime is $1 - p_t\left(i^{HIGH}\right)$.

[16] Our monthly data for inflation is, for both countries examined, the (annualized) rate of change in the consumer price index (CPI). For the output gap measure, we use the Hodrick-Prescott filtered industrial production, which was used instead of gross domestic product since industrial production is available monthly. These time series were obtained from the International Monetary Fund's *International Financial Statistics* database.

rules' (e.g. see Taylor, 1993, 1999; Clarida, Gali and Gertler, 1998, 2000, and the references therein). The Taylor rule literature provides evidence that it is possible to characterize monetary policy as the minimization of inefficient economic fluctuations via the implementation of an interest rate rule. Such an interest rate rule relates the setting of short-term money market rates to the evolution of two key state variables: price inflation and a business cycle indicator, the output gap. If these state variables do in fact drive monetary policy decisions and hence, via movements in short-term interest rates, the whole term structure of interest rates, then it seems plausible that the same state variables may also impact on the probability of shifting from a low to a high interest rate regime. Given that a standard Taylor rule would relate movements in interest rates positively to deviations of both output and inflation above their equilibrium levels, we might expect both δ_1 and δ_2 to be positive.

The results of our logit estimations for both the UK and the US are reported in Table 5. Consistent with the findings reported in Bansal and Zhou (2002) and Bansal, Tauchen and Zhou (2003), we confirm that the business cycle (output gap) is indeed important in explaining the dynamics of the regime-switching probabilities. In fact, the estimated coefficient $\widetilde{\delta}_1$, associated with our proxy for the output gap, is found to be statistically significant for both countries examined. Furthermore, it is interesting to note that, consistent with our conjecture and with the literature on interest rate rules (Taylor, 1993; Clarida, Gali and Gertler, 1998, 2000), inflation is also important in explaining the behavior of the regime-switching probabilities, as evidenced by the fact that the estimated coefficient on inflation, $\widetilde{\delta}_2$, is found to be statistically significant at conventional significance levels. The sign of the estimated parameters, $\widetilde{\delta}_1$ and $\widetilde{\delta}_2$, are both significantly positive, confirming our economic priors that the probability of being in a regime with high interest rates is higher when an economy is in expansion or inflation is relatively high.

The estimated logit model presents a moderately satisfactory R^2 and, perhaps more importantly, displays a very interesting 'classification ratio': the ratio of correctly classified observations implied by the logit model to the total number of observations is higher than 77% for both countries, which seems extremely encouraging given the simplicity of the logit model considered.

Overall, our results suggest that the shifts in mean and variance of the term structure of interest rates may be intimately related to changes in the sort of economic fundamentals one would expect to play a role in driving interest rate regimes, in particular the state of the business cycle and fluctuations in inflation.

We now turn to our out-of-sample forecasting results.

5 Forecasting the term structure

The procedure we have applied so far allowed us to achieve a reliable in-sample representation of the dynamic relationship among eurorates with different maturities implied by the EH theory. In order to assess further the usefulness of our asymmetric-nonlinear VECM characterization of the term structure, dynamic out-of-sample forecasts of the term structure were constructed using the asymmetric MSIH(2)-VECM(1) estimated and described in the previous section. In particular, we performed forecasting exercises for the period January 1992-December 2000 with forecast horizons up to 52 weeks ahead. The out-of-sample forecasts for a given horizon $j = 1, \ldots, 52$ were constructed recursively, conditional only upon information up to the date of the forecast and with successive re-estimation as the date on which forecasts are conditioned moves through the data set.

Forecast accuracy is evaluated using absolute and square error criteria (e.g. see Bansal and Zhou, 2002); specifically, the average absolute cross-sectional pricing forecast error (APFE)

$$APFE_{t+j} = \frac{\sum_{k=1}^{H} \left| i_{k,t+j} - \widetilde{i}_{k,t+j} \right|}{H} \tag{15}$$

and the average square cross-sectional pricing forecast error (SPFE)

$$SPFE_{t+j} = \frac{\sum_{k=1}^{H} \left(i_{k,t+j} - \widetilde{i}_{k,t+j} \right)^2}{H}, \tag{16}$$

where H is the number of eurorates included in the system (i.e. $H = 5$) and $\widetilde{i}_{k,t+j}$ is the $j-$period-ahead forecast of $i_{k,t+j}$ based on information at time t.

We compared the forecasts produced by the asymmetric MSIH-VECM (13) to the forecasts generated by the (linear and nonlinear) VAR models comprising the same set of variables (i.e. VAR and MSVAR) as well as the forecasts generated by the symmetric (linear and nonlinear) term-structure VECMs (i.e. VECMS and MS-VECMS) and a linear VECM with asymmetry (VECMA).

In order to assess the relative accuracy of forecasts derived from two different models we employed the Diebold and Mariano (1995) test:

$$DM = \frac{\overline{d}}{\sqrt{\frac{2\pi \widehat{f}(0)}{T}}}, \tag{17}$$

where $\overline{d}$ is an average (over T observations) of a general loss differential function of the $APFE$ (or $SPFE$) and $\widehat{f}(0)$ is a consistent estimate of the spectral density of the loss differential function at frequency zero. Diebold and Mariano show that the DM statistic is distributed as standard normal under the null hypothesis of equal forecast accuracy. Consistent with a large literature (see, *inter alia*, Mark, 1995), the loss differential functions we consider are the difference between either the APFE for the two models, or the difference between the SPFEs. A consistent estimate of the spectral density at frequency zero $\widehat{f}(0)$ is obtained using the method of Newey and West (1987) where the optimal truncation lag has been selected using the Andrews (1991) AR(1) rule.[17]

Several problems may arise when using DM statistics in small samples and taking into account parameter uncertainty (West, 1996; West and McCracken, 1998; McCracken, 2000). In the present case, where we are dealing with nested competing forecasting (linear and nonlinear) models and with multi-step-ahead forecasts, the asymptotic distribution of the DM statistic is non-standard and unknown. Therefore, the marginal significance levels reported below should be interpreted with caution.

Table 6 gives detailed results of the accuracy of the forecasts for the UK (Panel A) and the US (Panel B) respectively, using APFE and SPFE criteria for forecast accuracy. The results generally provide evidence in favor of the predictive superiority of the asymmetric MSIH-VECMs against VAR models. In particular, comparing our results to those obtained using simple (linear and Markov-switching) VAR models we can see that the asymmetric MSIH-VECMs give more accurate forecasts. At the 4-week horizon, for example, we achieve improvements ranging between 19% and 77% using the APFE and between 32% and 93% using the SPFE. At the 52-week horizon we obtain improvements

[17] The rule is implemented as follows: we estimate an AR(1) model to the quantity $APFE_t$ (or $SPFE_t$) obtaining the autocorrelation coefficient $\widehat{\rho}$ and the innovation variance from the AR(1) process $\widehat{\sigma}^2$. Then the optimal truncation lag A for the Parzen window in the Newey-West estimator is given by the Andrews rule $A = 2.6614 \left[\widehat{\zeta}(2)\, T\right]^{1/5}$ where $\widehat{\zeta}(2)$ is a function of $\widehat{\rho}$ and $\widehat{\sigma}^2$. The Parzen window has been used because it minimizes the mean square error of the estimator (Gallant, 1987, p. 534).

ranging between 30% and 78% using the APFE and between 48% and 93% using the SPFE. The statistical significance of these results is confirmed by executing the DM tests.[18]

The gain in terms of accuracy of the predictive performance of the asymmetric MSIH-VECM is also impressive when compared with the symmetric (linear and Markov-switching) VECMs. In fact the asymmetric MSIH-VECM performs very well at short as well as longer horizons for both countries. A different pattern can be seen by looking at the relative performance of the asymmetric MSIH-VECM against its linear counterpart (VECMA). Both asymmetries and regime shifts are relevant in the case of the UK, while, in the case of the US, asymmetries seem to be more important than regime shifts in terms of forecasting performance.

Overall, these results suggest that, using a VECM framework for the term structure of interest rates, it is possible to generate satisfactory out-of-sample forecasts of the term structure. Moreover, by explicitly incorporating asymmetry and regime shifts into the modelling framework, we have in the present analysis been able to improve upon a standard linear symmetric VECM framework.

6 Conclusion

In this paper we have reported an analysis of the term structure of interest rates in a multivariate asymmetric Markov-switching framework, and in particular we have applied that framework to forecast out-of-sample the term structure of interest rates. Using weekly data on eurorates for the UK and the US over the period February 1982 through December 1991, we found strong evidence of the presence of nonlinearities and asymmetries in the term structure, which appeared to be modelled satisfactorily by a multivariate asymmetric two-regime Markov-switching VECM that allows for shifts in both the intercept and the covariance structure of the error terms. We then used this model to forecast dynamically out of sample over the period January 1992 through to December 2000. The forecasting results were very interesting. The asymmetric MSIH-VECM forecasts were found to be superior to the forecasts obtained from VAR models, comprising the same set of variables, at a range of forecasting horizons up to 52 weeks ahead, using standard forecasting accuracy criteria and on the

[18] Although, in the light of our earlier discussion concerning the asymptotic properties of the DM statistic, we have cautioned that the marginal significance levels reported should be interpreted with care, their extremely small magnitude is nevertheless quite striking.

basis of standard tests of significance. Moreover, the asymmetric nonlinear VECM outperformed, in general, a symmetric (linear or nonlinear) VECM, although the magnitude of the gain from using the asymmetric Markov-switching VECM relative to a linear and nonlinear VECM is generally smaller in magnitude.

Our research was motivated by encouraging results previously reported in the literature on the presence of nonlinearities and regime shifts (e.g. Hamilton, 1988; Gray, 1996; Anderson, 1997; Kim, 2003) and asymmetries (e.g. Enders and Granger, 1998; Coakley and Fuertes, 2002) as well as by the relative success of the linear VECM model of the term structure of interest rates (e.g. Hall, Anderson and Granger, 1992). The research was also inspired by the notion that, in addition to the statistical importance of asymmetries and regime shifts for fitting interest rate data, there are economic reasons for believing that the allowance for regime shifts and asymmetries can provide potentially important insights into the behavior of the entire yield curve. For example, business cycle expansions and contractions may have important first-order effects on expectations of inflation, monetary policy and nominal interest rates, so that regime shifts and asymmetries may generate significant impacts both on the short-term interest rate and on the entire term structure (see e.g. Bansal and Zhou, 2002). In fact, the estimated regime shifts appear to be related to the state of the business cycle and to inflation, as one would expect in economies where monetary policy decisions are implemented via changes in short-term interest rates in response to deviations of output and inflation from their respective equilibrium levels (Taylor, 1993, 1999; Clarida, Gali and Gertler, 1998, 2000). Overall, our results suggest that regime shifts and—to a greater extent—asymmetries have important statistical and economic effects in driving the behavior of the term structure of interest rates.

In this work, however, we were primarily concerned with providing sound forecasting models of the term structure of interest rates and we therefore explicitly adopted an 'agnostic' approach both with respect to the sources of the underlying departures from the expectations hypothesis and in the sources of the underlying nonlinearities. Future research might, therefore, usefully analyze the sources of these nonlinearities further and attempt to improve on the parametric nonlinear formulation proposed in this paper. Understanding more deeply the implications of regime shifts and asymmetries for the inflation expectations formation mechanism and monetary policy represents a logical next step to take forward this research agenda.

With regard to the evaluation of forecasting models, although the relevant literature has traditionally focused on accuracy evaluations based on point forecasts, several authors have recently emphasized the importance of evaluating the forecast accuracy of economic models on the basis of density—as opposed to point—forecasting performance (see, *inter alia*, Diebold, Gunther and Tay, 1998; Granger and Pesaran, 1999; Tay and Wallis, 2000; Timmerman, 2000). Especially when evaluating nonlinear models, which are capable of producing highly non-normal forecast densities, it would seem appropriate to consider a model's density forecasting performance. This is a further immediate avenue for future research.

Table 1. Johansen maximum likelihood cointegration procedure

LR tests based on maximum eigenvalue ($LR_{\max}$) and trace of the stochastic matrix (LR_{trace})

A) UK

H_0	$LR_{\max}$	5% CV	LR_{trace}	5% CV
$r = 0$	194.80	34.40	436.70	76.10
$r \leq 1$	113.00	28.10	241.90	53.10
$r \leq 2$	99.27	22.00	128.90	34.90
$r \leq 3$	21.78	15.70	29.64	20.00
$r \leq 4$	7.865	9.20	7.865	9.20

B) US

H_0	$LR_{\max}$	5% CV	LR_{trace}	5% CV
$r = 0$	112.20	34.40	287.20	76.10
$r \leq 1$	83.41	28.10	175.00	53.10
$r \leq 2$	60.09	22.00	91.58	34.90
$r \leq 3$	24.36	15.70	31.49	20.00
$r \leq 4$	7.12	9.20	7.128	9.20

Notes: The VAR being tested for cointegration is $y_t = [i_t^0, i_t^4, i_t^{13}, i_t^{26}, i_t^{52}]'$, allowing for an unconstrained intercept under the null hypothesis H_0. r denotes the number of cointegrating vectors. The 5 percent critical value (denoted 5% CV) reported is taken from Osterwald-Lenum (1992).

Table 2. Long-run cointegrating equilibrium parameters

k	UK	US
4 weeks	1.004 (0.01)	0.991 (0.03)
13 weeks	1.019 (0.02)	0.997 (0.03)
26 weeks	1.036 (0.03)	1.012 (0.04)
52 weeks	1.061 (0.04)	1.018 (0.05)

Notes: The table gives the estimated long-run slope parameter for the relevant interest rate at different maturities. Figures in parentheses denote asymptotic standard errors.

Table 3. Asymmetry tests

H_0	LR	$p-value$
UK	107.419	5.79×10^{-14}
US	63.585	5.11×10^{-13}

Notes: LR is a likelihood ratio test of the symmetry null hypothesis, where the restricted model being tested is the symmetric linear VECM (11) and the alternative VECM allows for asymmetric equilibrium correction. The test is constructed as $2(\ln L^* - \ln L)$, where L^* and L represent the unconstrained and the constrained maximum likelihood respectively. These test statistics are asymptotically distributed as $\chi^2(g)$ under the null hypothesis, where g is the number of restrictions.

Table 4. Markov-switching VECM estimation: linearity tests

H_0	LR_{S1}	LR_{S2}
UK	808.08	786.46
US	1072.782	1077.931

Notes: LR_{S1} and LR_{S2} are likelihood ratio tests where the restricted models being tested are the symmetric linear VECM in equation (11) and the asymmetric linear VECM respectively; the alternative models are the symmetric MSIH(2)-VECM(1) and the asymmetric MSIH(2)-VECM(1) respectively. The tests are constructed as $2(\ln L^* - \ln L)$, where L^* and L represent the unconstrained and the constrained maximum likelihood respectively. These test statistics are asymptotically distributed as $\chi^2(g)$ under the null hypothesis, where g is the number of restrictions. p-values are not reported as they are virtually zero in each case.

Table 5. MSIH-VECM regime interpretation: logit estimation

	$\widetilde{\delta}_1$	$\widetilde{\delta}_2$	Pseudo-R^2	CR
UK	0.4420	0.0439	0.077	0.784
	(0.091)	(0.019)		
US	0.3928	0.2724	0.198	0.775
	(0.168)	(0.117)		

Notes: $\delta_{1,}\delta_2$ are estimated parameters relative to output gap and inflation in the logit model (14), as discussed in the text. Pseudo-R^2 denotes Estrella's (1998) measure of goodness of fit for logit models. CR is the classification ratio, calculated as the ratio of correctly classified observations to the total number of observations used in the logit estimation. Estimates are obtained by Generalized Method of Moments (GMM) calculated by two-step nonlinear two-stage least square (Hansen, 1982). The optimal weighting matrix is obtained from the first step two-stage least squares parameter estimates; the instrument set includes 12 lags of each of inflation and the output gap. Standard errors are reported in parentheses.

Table 6. Out-of-sample forecasting exercises

A) UK

	Average Absolute Cross-sectional Pricing Errors (APFE)				
k	VAR	VECMS	VECMA	MSVAR	MSVECMS
4	0.8142	0.4319	0.4948	0.6048	0.6649
	$[3.33 \times 10^{-2}]$	$[4.74 \times 10^{-7}]$	$[1.48 \times 10^{-123}]$	$[1.08 \times 10^{-7}]$	$[2.50 \times 10^{-6}]$
13	0.7083	0.4509	0.4458	0.5200	0.6050
	$[6.89 \times 10^{-6}]$	$[4.83 \times 10^{-15}]$	$[1.35 \times 10^{-48}]$	$[2.04 \times 10^{-20}]$	$[7.74 \times 10^{-14}]$
26	0.6703	0.1289	0.4435	0.5007	0.5109
	$[1.93 \times 10^{-12}]$	$[1.17 \times 10^{-3}]$	$[1.88 \times 10^{-23}]$	$[2.71 \times 10^{-38}]$	$[6.70 \times 10^{-40}]$
52	0.7039	0.1678	0.4623	0.5360	0.4656
	$[4.94 \times 10^{-15}]$	$[4.41 \times 10^{-4}]$	$[7.74 \times 10^{-12}]$	$[1.95 \times 10^{-55}]$	$[1.05 \times 10^{-80}]$
	Average Square Cross-sectional Pricing Errors (SPFE)				
k	VAR	VECMS	VECMA	MSVAR	MSVECMS
4	0.6863	0.3961	0.2648	0.3902	0.4827
	$[1.43 \times 10^{-2}]$	$[1.55 \times 10^{-8}]$	$[6.98 \times 10^{-41}]$	$[8.57 \times 10^{-10}]$	$[1.89 \times 10^{-7}]$
13	0.5370	0.2843	0.2218	0.3209	0.4004
	$[2.66 \times 10^{-7}]$	$[1.92 \times 10^{-17}]$	$[3.72 \times 10^{-25}]$	$[8.77 \times 10^{-20}]$	$[1.44 \times 10^{-16}]$
26	0.4300	0.2735	0.1670	0.2622	0.2754
	$[3.98 \times 10^{-17}]$	$[8.71 \times 10^{-5}]$	$[2.20 \times 10^{-22}]$	$[1.09 \times 10^{-33}]$	$[6.46 \times 10^{-36}]$
52	0.5267	0.1254	0.1710	0.3276	0.2610
	$[1.12 \times 10^{-27}]$	$[2.90 \times 10^{-5}]$	$[9.90 \times 10^{-23}]$	$[2.43 \times 10^{-56}]$	$[2.17 \times 10^{-69}]$

(continued ...)

(... Table 6 continued)

B) US

Average Absolute Cross-sectional Pricing Errors (APFE)					
k	VAR	VECMS	VECMA	MSVAR	MSVECMS
4	0.2577	0.2658	0.5590	0.2306	0.2526
	$[2.02 \times 10^{-123}]$	$[2.73 \times 10^{-121}]$	$[0]$	$[2.62 \times 10^{-136}]$	$[8.36 \times 10^{-123}]$
13	0.4748	0.4289	1.4997	0.4357	0.4277
	$[1.62 \times 10^{-62}]$	$[2.22 \times 10^{-81}]$	$[3.05 \times 10^{-21}]$	$[6.66 \times 10^{-73}]$	$[1.70 \times 10^{-79}]$
26	0.3334	0.3104	1.2529	0.3083	0.3035
	$[7.16 \times 10^{-140}]$	$[6.58 \times 10^{-150}]$	$[5.92 \times 10^{-9}]$	$[1.64 \times 10^{-154}]$	$[1.04 \times 10^{-158}]$
52	0.2373	0.2308	0.7201	0.2222	0.2207
	$[3.67 \times 10^{-306}]$	$[7.32 \times 10^{-292}]$	$[1.66 \times 10^{-13}]$	$[9.22 \times 10^{-321}]$	$[8.55 \times 10^{-312}]$
Average Square Cross-sectional Pricing Errors (SPFE)					
k	VAR	VECMS	VECMA	MSVAR	MSVECMS
4	0.0741	0.0771	0.3354	0.0626	0.0693
	$[3.48 \times 10^{-156}]$	$[9.73 \times 10^{-117}]$	$[1.21 \times 10^{-265}]$	$[6.34 \times 10^{-165}]$	$[8.01 \times 10^{-139}]$
13	0.2207	0.1759	1.5643	0.1854	0.1765
	$[1.77 \times 10^{-110}]$	$[7.11 \times 10^{-108}]$	$[3.15 \times 10^{-11}]$	$[2.26 \times 10^{-124}]$	$[8.54 \times 10^{-122}]$
26	0.1211	0.1027	1.5570	0.1040	0.0997
	$[1.09 \times 10^{-180}]$	$[8.07 \times 10^{-161}]$	$[1.00 \times 10^{-8}]$	$[2.23 \times 10^{-191}]$	$[8.00 \times 10^{-179}]$
52	0.0778	0.0727	0.6119	0.0685	0.0670
	$[2.99 \times 10^{-235}]$	$[2.21 \times 10^{-204}]$	$[5.32 \times 10^{-8}]$	$[3.54 \times 10^{-247}]$	$[1.18 \times 10^{-223}]$

Notes: VAR, VECMS, VECMA, MSVAR, MSVECMS are the ratios of the average (absolute or square) cross-sectional pricing forecast errors (APFE and SPFE as defined in equations (15) and (16) respectively) obtained by the asymmetric MSIH-VECM to the ones obtained by the linear VAR, the linear symmetric VECM, the linear asymmetric VECM, the MSIH-VAR and the symmetric MSIH-VECM respectively. The average cross-sectional pricing forecast errors are obtained by recursive out-of-sample dynamic forecasting up to $k = 52$ steps ahead over the period 1992:1-2000:52. Figures in brackets are the Diebold-Mariano statistics comparing the average (absolute or square) cross-sectional pricing forecast errors of the asymmetric MSIH-VECM model to the ones obtained by each of the other competing models. The optimal truncation lag has been constructed according to Andrews (1991) AR(1) rule. For the Diebold-Mariano statistics only $p-$values are reported.

References

Ahn, D.-H., Dittmar, R.F. and Gallant, A.R. (2002), "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, 15, 243-288.

Anderson, H.M. (1997), "Transaction Costs and Nonlinear Adjustment Towards Equilibrium in the US Treasury Bill Market," *Oxford Bulletin of Economics and Statistics*, 59, 465-484.

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

Ang, A. and Bekaert, G. (2002), "Regime Switches in Interest Rates," *Journal of Business and Economic Statistics*, 20, 163-182.

Bansal, R. and Zhou, H. (2002) "Term Structure of Interest Rates with Regime Shifts," *Journal of Finance*, 57, 1997-2043.

Bansal, R., Tauchen, G. and Zhou, H. (2003), "Regime-Shifts, Risk Premiums in the Term Structure, and the Business Cycle," *Journal of Business and Economics Statistics*, forthcoming.

Berkson, J. (1938), "Some Difficulties of Interpretation Encountered in the Application of the Chi-Square Test," *Journal of the American Statistical Association*, 53, 28-38.

Campbell, J.Y. and Clarida, R.H. (1986), "The Term Structure of Euromarket Interest Rates: An Empirical Investigation," *Journal of Monetary Economics*, 19, 25-44.

Campbell, J.Y. and Shiller, R.J. (1987), "Cointegration and Tests of Present Value Models," *Journal of Political Economy*, 95, 1062-1088.

Campbell, J.Y. and Shiller, R.J. (1991), "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495-514.

Clarida, R.H., Gali, J. and Gertler, M. (1998), "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review*, 42, 1033-1067.

Clarida, R.H., Gali, J. and Gertler, M. (1999), "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37, 1661-1707.

Clarida, R.H., Gali, J. and Gertler, M. (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147-180.

Coakley, J. and Fuertes, A.M. (2002), "Asymmetric Dynamics in UK Real Interest Rates", *Applied Financial Economics*, 12, 379-87.

Dai, Q. and Singleton, K.J. (2000), "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, 55, 1943-78.

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977), "Maximum Likelihood Estimation from Incomplete Data Via the EM Algorithm," *Journal of the Royal Statistical Society*, 39, Series B, 1-38.

Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, pp. 863-883.

Diebold, F.X. and Mariano, R.S. (1995), “Comparing Predictive Accuracy,” *Journal of Business and Economics Statistics*, 13, 253-263.

Enders, W. and Granger, C.W.J. (1998), “Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates,” *Journal of Business and Economic Statistics*, 16, 304-311.

Engle, R.E. and Granger, C.W.J. (1987), “Co-integration and Equilibrium Correction Representation, Estimation and Testing,” *Econometrica*, 55, 251-276.

Estrella, A. (1998), “A New Measure of Fit for Equations with Dichotomous Dependent Variables,” *Journal of Business and Economic Statistics*, 16, 198-205.

Gallant A.R. (1987), *Nonlinear Statistical Models*, New York, NY: John Wiley.

Granger, C.W.J. (1986), “Developments in the Study of Cointegrated Variables,” *Oxford Bulletin of Economics and Statistics*, 48, 213-228.

Granger, C.W.J. and Pesaran, M.H. (1999), “A Decision Theoretic Approach to Forecast Evaluation,” in Chan, W.S., Lin, W.K. and Tong, H. (eds.), *Statistics and Finance: An Interface*, London: Imperial College Press.

Gray, S.F. (1996), “Modelling the Conditional Distribution of Interest Rates as a Regime-Switching Process,” *Journal of Financial Economics*, 42, 27-62.

Hall, A.D., Anderson, H.M. and Granger, C.W.J. (1992), “A Cointegration Analysis of Treasury Bill Yields”, *Review of Economics and Statistics*, 74, 116-126.

Hamilton, J.D. (1988), “Rational Expectations Econometric Analysis of Changes in Regime. An Investigation of the Term Structure of Interest Rates,” *Journal of Economics Dynamics and Control*, 12, 385-423.

Hamilton, J.D. (1993), “Estimation, Inference and Forecasting of Time Series Subject to Changes in Regime,” in Maddala, G.S., Rao, C.R. and Vinod, H.D. (eds.), *Handbook of Econometrics*, Vol. 4, Elsevier, Amsterdam.

Hansen, L. P. (1982), “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, 50, 1029-1054.

Ioannides, C. and Peel, D.A. (2003), “Empirical Evidence on the Relationship between the Term Structure of Interest Rates and Future Real Output Changes when There Are Changes in Policy Regimes,” *Economics Letters*, 147-52.

Johansen, S. (1988), “Statistical Analysis of Cointegrating Vectors,” *Journal of Economic Dynamics and Control*, 12, 231-254.

Johansen, S. (1991), “Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models,” *Econometrica*, 59, 1551-1580.

Kim, D.H. (2003), “Nonlinearity in the Term Structure,” University of Manchester, Department of Economics, mimeo.

Kim, C.-J. and Nelson, C.R. (1999), "State-Space Models with Regime Switching," Cambridge, Mass. and London, UK: MIT Press.

Krolzig, H.-M. (1997), *Markov-Switching Vector Autoregressions*, New York, NY: Springer.

Leamer, E.E. (1978), *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: John Wiley and Sons.

Mark, N. (1995), "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, 201-218.

McCracken, M.W. (2000), "Robust Out-of-Sample Inference," *Journal of Econometrics*, 99, 195-223.

Newey, W.K. and West, K.D. (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 3, 703-708.

Osterwald-Lenum, M. (1992), "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics*, 54, 461-472.

Saikkonen, P. (1992), "Estimation and Testing of Cointegrated Systems by an Autoregressive Approximation," *Econometric Theory*, 8, 1-27.

Saikkonen, P. and Luukkonen, R. (1997), "Testing Cointegration in Infinite Order Vector Autoregressive Processes," *Journal of Econometrics*, 81, 93-126.

Sarno, L. and Thornton, D.L. (2003), "The Dynamic Relationship Between the Federal Funds Rate and the Treasury Bill Rate: An Empirical Investigation," *Journal of Banking and Finance*, 27, 1079-1110.

Stock, J.H. (1997), "Cointegration, Long-run Movements, and Long-Horizon Forecasting," in Kreps, D.M., and Wallis, K.F. (eds.), *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, Vol. III, Cambridge, UK: Cambridge University Press.

Tay, A.S. and Wallis, K.F. (2000), "Density Forecasting: A Survey," *Journal of Forecasting*, 19, pp. 235-254.

Taylor, J.B. (1993), "Dicretion Versus Policy Rules in Practice," *Carnegie Rochester Conference Series on Public Policy*, 39, 195-214.

Taylor, J.B. (1999), *Monetary Policy Rules*, National Bureau of Economic Research Conference Report series. Chicago and London: University of Chicago Press.

Taylor, M.P. (1992), "Modeling the Yield Curve," *Economic Journal*, 102, 524-537.

Timmermann, A. (2000), "Density Forecasting in Economics and Finance: Editorial," *Journal of Forecasting*, 19, 231-234.

Vasicek, O. (1977), "An Equilibrium Characterisation of the Term Structure," *Journal of Financial Economics*, 5, 177-188.

West, K.D. (1996), "Asymptotic Inference about Predictive Ability," *Econometrica*, 64, 1067-1084.

West, K.D. and McCracken, M. W. (1998), "Regression-based Tests of Predictive Ability," *International Economic Review*, 39, 817-840.