Richard H. Clarida
Columbia University and National Bureau of Economic Research

Lucio Sarno
University of Warwick and Centre for Economic Policy Research

Mark P. Taylor
University of Warwick and Centre for Economic Policy Research

Giorgio Valente
Chinese University of Hong Kong

The Role of Asymmetries and Regime Shifts in the Term Structure of Interest Rates*

We examine the term structure of interest rates for the United States, Germany, and Japan over the period 1982–2000, using a nonlinear multivariate vector equilibrium correction-modeling framework that allows for asymmetric adjustment and regime shifts. The model has a very general underlying theoretical rationale that allows for time-varying term premia and other short-run deviations from the expectations model of the term structure. The empirical models fit well, display regime switches closely correlated with key monetary policy variables, and have good forecasting properties.

I. Introduction

In this paper we reexamine the dynamic relationship between interest rates of different maturities for three countries using a very general, multivariate vector equilibrium correction-modeling framework capable of simultaneously allowing for asymmetric adjustment and regime shifts. Our model has an underlying theoretical rationale based on the expectations model of the term structure, allowing for time-varying term premia. The resulting nonlinear vector equilibrium correction models are shown not only to provide good

* This paper was partly written while Sarno and Taylor were visiting the International Monetary Fund. Sarno and Valente acknowledge financial support from the Economic and Social Research Council (grant L138251044). The authors are indebted for their constructive comments to Albert Madansky (editor), an anonymous referee, and to participants in seminars at the University of Oxford and the University of Warwick. The authors alone are responsible for any errors and for the views expressed in the paper. Contact the corresponding author, Mark Taylor, at Mark.Taylor@warwick.ac.uk.

[*Journal of Business*, 2006, vol. 79, no. 3]

0021-9398/2006/7903-0007$10.00

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 during the 1990s that focused on the cointegrating properties of the term structure and 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 developed that allows for asymmetric or nonlinear adjustment toward equilibrium in modeling 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, nonzero 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; Sarno and Thornton 2003).[1] In addition to this growing amount of statistical evidence, there are sound economic reasons to believe that regime shifts and asymmetries 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 German,[2] Japanese, and U.S. eurodeposit interest rates with five different maturities, we show, first, that while a long-run equilibrium relationship between the five different interest rates consistent with the expectations theory of the term structure can be established, 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 pre-

1. It is interesting to note in this context that Hamilton's (1988) seminal paper on Markov switching involved an application to the term structure of interest rates. See also Clarida et al. (2003) and Sarno and Valente (2005).

2. The German mark interest rate data are spliced with euro interest rate data at the inception of European Monetary Union in 1999.

mia, we are able to characterize satisfactorily the dynamic relationship between interest rates with different maturities for each country. 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). This model outperforms, both in sample and out of sample, a range of alternative linear and nonlinear equilibrium correction models for Japan and the United States, although for Germany the linear asymmetric vector equilibrium correction model (without regime shifts) emerges as the best forecasting model among the competing models considered. 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 II 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, and in Section III we describe the recently developed econometric procedure that 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 IV 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 report 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. In Section V we report and discuss our forecasting results, and Section VI contains the results of several robustness checks. In Section VII we briefly summarize our main results and conclude. In Appendices A and B we report further estimation results and give details of our robustness checks.

II. 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 $t + k$. With the conventional Fisher-Hicks recursive formulas, the relationship linking $i_{k,t}$ and $f_{k,t}$ may be described as follows:

$$i_{k,t} = \frac{1}{k}\left(\sum_{j=1}^{k} f_{j,t}\right) \quad \text{for } k = 1, 2, 3, \ldots. \tag{1}$$

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 $f_{j,t} = E_t(i_{k,t+j-1}) + \phi_{j,t}$, 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 rewrite (1) as follows:

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

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

Equation (2) 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 allow a very weak version of the EH that allows the term premia $\gamma_{k,t}$ to be time-varying and requires only that they be realizations of stationary stochastic processes.

Even this very weak form of the EH, however, has important and clear statistical implications (Hall et al. 1992). To see this, note that we can rewrite equation (2) as follows:

$$i_{k,t} - i_{1,t} = \frac{1}{k}\left(\sum_{m=1}^{k-1}\sum_{j=1}^{m} E_t\Delta i_{1,t+j}\right) + \gamma_{k,t}, \tag{3}$$

where Δ is the first-difference operator. Under the assumption that the yields to maturity are realizations of stochastic processes integrated of order one, I(1), if the term premia components are stationary, all terms on the right-hand side of equation (3) must be stationary, which implies that the term on the left-hand side of (3) must be stationary also, that is, $i_{k,t} - i_{1,t} \sim \mathrm{I}(0)$. 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 modeling 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). 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; 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 that is capable of allowing for all these possibilities simultaneously.

III. 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, 1999).

Consider the following M-regime pth-order Markov-switching vector autoregression (MS(M)-VAR(p)), which allows for regime shifts in the intercept term:[3]

$$y_t = \nu(z_t) + \sum_{i=1}^{p} \Pi_i y_{t-i} + \varepsilon_t, \tag{4}$$

where y_t is a K-dimensional vector time-series process, $y_t = [y_{1t}, y_{2t}, \ldots, y_{Kt}]'$; $\nu(z_t)$ is a K-dimensional column vector of regime-dependent intercept terms, $\nu(z_t) = [\nu_1(z_t), \nu_2(z_t), \ldots, \nu_K(z_t)]'$; the Π_i's are $K \times K$ matrices of parameters; and $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{Kt}]'$ is a K-dimensional vector of Gaussian white-noise processes with covariance matrix Σ, $\varepsilon_t \sim \text{NID}(\mathbf{0}, \Sigma_\varepsilon)$. The regime-generating process is assumed to be an ergodic Markov chain with a finite number of states $z_t \in \{1, \ldots, M\}$ governed by the transition probabilities $p_{ij} = \Pr(z_{t+1} = j \mid z_t = i)$, and $\sum_{j=1}^{M} p_{ij} = 1$ for all $i, j \in \{1, \ldots, M\}$.[4]

A standard case in economics and finance is one in which y_t is nonstationary but first-difference stationary, that is, $y_t \sim \text{I}(1)$. Then, given $y_t \sim \text{I}(1)$, there may be up to $K - 1$ linearly independent cointegrating relationships, which represent the long-run equilibrium of the system, and the equilibrium error

3. Although, for expositional simplicity, this section focuses on eq. (4), clearly a more general formulation of (4) may be considered that allows for other parameters of the model to be conditioned on the state z_t, as illustrated below.

4. To be precise, z_t is assumed to follow an ergodic M-state Markov process with transition matrix

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix},$$

where $p_{iM} = 1 - p_{i1} - \cdots - p_{i,M-1}$ for $i \in \{1, \ldots, M\}$.

(the deviation from the long-run equilibrium) is measured by the stationary stochastic process $h_t = \beta' y_t$ (Granger 1986; Engle and Granger 1987). If indeed there is cointegration, the cointegrated MS-VAR (4) implies a Markov-switching vector equilibrium correction model or MS-VECM of the form

$$\Delta y_t = \nu(z_t) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + \varepsilon_t, \tag{5}$$

where $\varepsilon_t \sim \text{NIID}(\mathbf{0}, \Sigma_\varepsilon)$, $\Gamma_i = -\sum_{j=i+1}^{p} \Pi_j$ are $K \times K$ matrices of parameters, and $\Pi = \sum_{i=1}^{p} \Pi_i - \mathbf{I}$ (where $\mathbf{I}$ is the identity matrix) is the long-run impact matrix whose rank r determines the number of cointegrating vectors (e.g., Johansen 1995; Krolzig 1999). The matrix Π may be partitioned into the $K \times r$ matrix β forming a basis for the space spanned by the $r \leq K - 1$ linearly independent cointegrating vectors and a $K \times r$ matrix α containing the adjustment or equilibrium correction coefficients: $\Pi = \alpha\beta'$.

Although, for expositional purposes, we have outlined the MS-VECM framework for the case of regime shifts in the intercept alone, shifts may be allowed for elsewhere. The present application focuses on a multivariate model comprising, for each of the three countries analyzed, the spot-next eurorate and the rates relative to one month (4 weeks), three months (13 weeks), six months (26 weeks), and 12 months (52 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 II, four unique independent cointegrating relationships should exist.[5] As discussed in Section IV below, in our empirical work, after considerable experimentation, we selected a specification of the MS-VECM that allows for regime shifts in the intercept as well as in the variance-covariance matrix. This model, the Markov-switching-intercept-heteroskedastic-VECM (MSIH-VECM) may be written as follows:

$$\Delta y_t = \nu(z_t) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + u_t, \tag{6}$$

where $\Pi = \alpha\beta'$, $u_t \sim \text{NIID}(\mathbf{0}, \Sigma_u(z_t))$, and $z_t \in \{1, \dots, M\}$.

In order to take into account the empirical evidence indicating that interest rates display asymmetric adjustment (e.g., Enders and Granger 1998; Clarida and Taylor 2003; Sarno and Thornton 2003), we allow the MSIH-VECM (6)

5. There is a slight shift in notation here from that employed in Sec. II 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; therefore, we have indexed the spot-next rate at zero. The discussion in Sec. II applies to these variables directly if we relabel the interest rates by the number of days to maturity. This is purely a notational issue.

to display differing speeds of adjustment to equilibrium depending on whether there are positive or negative deviations from the equilibrium condition:

$$\Delta y_t = \nu(z_t) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + I_t \Pi^+ y_{t-1} + (1 - I_t)\Pi^- y_{t-1} + e_t, \qquad (7)$$

where I_t is a $1 \times r$ vector whose jth element at time t, $\iota_{j,t}$, say, is the Heaviside indicator, taking the value zero or unity according to whether the lagged jth cointegrating combination—the jth element of $\beta' y_{t-1}$, $h_{j,t-1}$, say—is positive or negative:[6]

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

(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 $\Pi^+ = (\alpha^+)\beta'$ and $\Pi^- = (\alpha^-)\beta'$, so that the equilibrium correction coefficients shift according to whether the lagged equilibrium correction term to which it applies is positive or negative. As before, we also have Gaussian error terms, $e_t \sim \text{NIID}(\mathbf{0}, \Sigma_e(z_t))$, and M states $z_t \in \{1, \ldots, M\}$. 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.

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 modeling 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 that yields estimates of the remaining parameters of the model (Dempster, Laird, and Rubin 1977; Hamilton 1993; Kim and Nelson 1999; Krolzig 1999).

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

IV. Empirical Results

A. 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 Germany, Japan, and the United States spanning the

6. Strictly speaking, nonnegative or negative.

TABLE 1 Unit Root Tests

	Germany	Japan	United States
		A. Levels	
$i_{0,t}$	−.781	−1.498	−2.506
$i_{4,t}$	−.724	−1.571	−2.802
$i_{13,t}$	−.463	−1.828	−2.396
$i_{26,t}$	−.627	−1.809	−2.240
$i_{52,t}$	−.673	−1.705	−1.923
		B. First Differences	
$\Delta i_{0,t}$	−6.800	−5.934	−4.551
$\Delta i_{4,t}$	−5.146	−4.965	−4.592
$\Delta i_{13,t}$	−14.917	−5.221	−4.779
$\Delta i_{26,t}$	−13.637	−5.982	−5.846
$\Delta i_{52,t}$	−11.165	−5.961	−9.931

NOTE.—Statistics are augmented Dickey-Fuller test statistics for the null hypothesis of a unit-root process; $i_{0,t}$, $i_{4,t}$, $i_{13,t}$, $i_{26,t}$, and $i_{52,t}$ are spot-next, one-month, three-month, six-month, and one-year eurorates, respectively; and Δ is the first-difference operator. The critical value at the 1% (5%) significance level is −3.446 (−2.868) to three decimal places (MacKinnon 1991).

period from February 7, 1982, to December 31, 2000, a total of 987 observations for each series.[7] 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 each of the three countries under investigation by calculating standard augmented Dickey-Fuller test statistics. In each case the number of lags was chosen such that no residual autocorrelation was evident in the auxiliary regressions. As shown in table 1, in keeping with the large number of studies of unit root behavior for these time series, in each case we were unable to reject the unit root null hypothesis at conventional nominal levels of significance. On the other hand, differencing the series did appear to induce stationarity in each case.[8] Hence, the unit root tests clearly indicate that each of the time series examined is a realization from a stochastic process integrated of order one, which suggests that testing for cointegration between the five interest rate series is the logical next step.

7. We are grateful to the Bank for International Settlements (BIS) for supplying the data. The German mark interest rate data were spliced with euro interest rate data at the inception of European Monetary Union in 1999. The start date was chosen since it was the earliest date for which BIS data for *all* three countries examined are available; specifically, for Japan, the BIS does not hold weekly data on spot-next eurorates prior to 1982. While these data are available for longer samples for the other two countries examined, we preferred to investigate the same sample period for each country for consistency and comparability purposes.

8. 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. See, e.g., 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, e.g., Stock 1997).

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.[9] On the basis of the Johansen likelihood ratio test statistics for the cointegrating rank reported in table 2 (based on the maximal eigenvalue and on the trace of the stochastic matrix), 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.[10] Hence, we conclude that there are exactly four cointegrating relationships between the five rates examined, for each of Germany, Japan, and the United States.

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

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

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:

$$\beta' y_t = \begin{bmatrix} -1 & \phi_4 & 0 & 0 & 0 \\ -1 & 0 & \phi_{13} & 0 & 0 \\ -1 & 0 & 0 & \phi_{26} & 0 \\ -1 & 0 & 0 & 0 & \phi_{52} \end{bmatrix} \begin{bmatrix} i_{0,t} \\ i_{4,t} \\ i_{13,t} \\ i_{26,t} \\ i_{52,t} \end{bmatrix}, \tag{10}$$

where the ϕ_i parameters are unrestricted. This yielded the results reported in table 3. 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 ϕ_i coefficients are in the range between 0.991 and 1.028 and are, therefore, very close indeed to the theoretical value of unity. Thus rejection of the hypothesis $H_0 : \phi_i = 1$ for all i may be due to tiny departures from the null hypothesis (due, e.g., to tiny data imperfections), which may not be economically significant but

9. We allowed for a maximum lag length of 12 and chose, for each country, the appropriate lag length on the basis of conventional information criteria.

10. The choice of exactly four independent cointegrating vectors was also confirmed by the Hansen-Johansen (1999) recursive procedure for the cointegrating rank. Hence, our cointegration results are robust to the presence of possible structural breaks in the cointegrating rank, as allowed for in the Hansen-Johansen procedure. These cointegration test statistics are not reported here in order to conserve space but are available from the authors on request.

TABLE 2 Johansen Maximum Likelihood Cointegration Procedure

H_0	LR_{max}	5% Critical Value	LR_{trace}	5% Critical Value
		A. Germany		
$r = 0$	141.70	34.40	270.70	76.10
$r \leq 1$	60.68	28.10	129.10	53.10
$r \leq 2$	52.57	22.00	68.39	34.90
$r \leq 3$	18.15	15.70	20.82	20.00
$r \leq 4$	1.67	9.20	1.67	9.20
		B. Japan		
$r = 0$	149.80	34.40	345.70	76.10
$r \leq 1$	106.10	28.10	195.90	53.10
$r \leq 2$	62.56	22.00	89.88	34.90
$r \leq 3$	23.25	15.70	27.32	20.00
$r \leq 4$	4.06	9.20	4.06	9.20
		C. United States		
$r = 0$	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

NOTE.—LR tests based on maximum eigenvalue (LR_{max}) and trace of the stochastic matrix (LR_{trace}). The VAR being tested for cointegration is $y_t = [i_{0,t}, i_{4,t}, i_{13,t}, i_{26,t}, i_{52,t}]'$, allowing for an unconstrained intercept under the null hypothesis H_0. r denotes the number of cointegrating vectors. The 5% critical value reported is taken from Osterwald-Lenum (1992).

appear as statistically significant given our large sample size.[11] In light of these results and given that the framework discussed in Section II provides strong economic priors in favor of the unity restrictions, we report below results obtained with the unity restrictions imposed.[12]

B. Asymmetry Testing and MS-VECM Estimation Results

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

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

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

11. Leamer (1978, chap. 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" (114). See also Berkson (1938).

12. We did, however, execute all 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.

TABLE 3 **Long-Run Cointegrating Equilibrium Parameters**

k	Germany	Japan	United States
4 weeks	.998 (.01)	.999 (.01)	.991 (.03)
13 weeks	.993 (.02)	1.007 (.03)	.997 (.03)
26 weeks	1.001 (.03)	1.017 (.04)	1.012 (.04)
52 weeks	1.022 (.05)	1.028 (.05)	1.018 (.05)

NOTE.—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 4 **Asymmetry Tests**

H_0	LR	p-Value
Germany	36.846	1.93×10^{-7}
Japan	52.913	8.88×10^{-11}
United States	63.585	5.11×10^{-13}

NOTE.—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.

Hannan-Quinn criterion. Employing the conventional general-to-specific procedure, we obtained fairly parsimonious models for each country, with no significant residual serial correlation.[13] We then investigated the presence of asymmetries in the adjustment toward the equilibrium condition by executing standard likelihood ratio (LR) tests for the null hypothesis of symmetry. The results reported in table 4 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, since the restrictions imposed by the model without asymmetries are rejected with marginal significance levels (p-values) close to zero.[14]

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

$$\Delta y_t = \nu(z_t) + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \Pi y_{t-1} + u_t, \tag{12}$$

where $u_t \sim \text{NIID}(\mathbf{0}, \Sigma_u(z_t))$ and $z_t = 1, 2$. 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 conventional “bottom-up” procedure designed to detect Markovian shifts in order to select the most adequate characterization of a

13. Full details on these estimation results are available from the authors on request but are not reported to conserve space.

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

TABLE 5 Markov-Switching VECM Estimation: Linearity Tests

H_0	LR_{S1}	LR_{S2}
Germany	650.367	648.264
Japan	611.039	621.346
United States	1,072.782	1,077.931

NOTE.—LR_{S1} and LR_{S2} are likelihood ratio tests in which the restricted models being tested are the symmetric linear VECM in eq. (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 since they are virtually zero in each case.

two-regime pth-order MS-VECM for Δy_t.[15] 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 LR tests of the type suggested by Krolzig (1997, 135–36). The results (see table A1 in App. A) indicated strong rejection of the null of no regime dependence in the intercept (LR_1) as well as in the variance-covariance matrix (LR_2), 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, as may be seen from inspection of the LR_3 tests in Appendix table A1, for all 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 5, 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 all countries

15. 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 is explained by observable variables, consistent with the general-to-specific approach to econometric modeling. For a technical discussion of the bottom-up procedure, see Krolzig (1997).

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

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

where I_t is as defined above, $\Pi^+ = (\alpha^+)\beta'$, $\Pi^- = (\alpha^-)\beta'$, $e_t \sim \text{NIID}[\mathbf{0}, \Sigma_e(z_t)]$, and $z_t = 1, 2$. We estimated the MSIH-VECM (13), using an EM maximum-likelihood algorithm, for each of Germany, Japan, and the United States.[16] 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.[17]

This model is not as parsimonious as some other term structure models in the literature. However, evidence provided by Dai and Singleton (2000), Jagannathan, Kaplin, and Sun (2000), Ahn, Dittmar, and Gallant (2002), and Bansal and Zhou (2002) 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.

C. 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 each 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 that in virtually every case were 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

16. These estimation results for a representative country, namely the United States, are reported in App. A. The full set of results are available from the authors on request.

17. We also examined graphs of the standardized residuals, the smoothed residuals, and the one-step prediction errors from each estimated MSIH-VECM in order to check for evidence of misspecification; the differences between these three residual measures depend on the way in which the residuals in each regime are weighted in order to form an overall measure. Loosely speaking, the smoothed residuals are the closest to the sample residuals from a linear regression model; however, they overestimate the explanatory power of the Markov-switching model because of the use of the full-sample information covered in the smoothed regime vector. The standardized residuals are conditional residuals. The one-step prediction errors are based on the predicted regime probabilities. Unfortunately, many conventional diagnostic tests, such as standard residual serial correlation tests, may not have their conventional asymptotic distribution when the residuals come from Markov-switching models and are therefore not reported here. However, the graphs of standardized residuals, the smoothed residuals, and the one-step prediction errors provided no visual evidence of residual serial correlation in any of the residuals series plotted.

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 related to macroeconomic fundamentals, in the spirit of recent research by Bansal and Zhou (2002) and Bansal, Tauchen, and Zhou (2004): using a visual approach, these authors find that regime shifts in the term structure appear to be intimately related to 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 relating the probability of being in either the high– or the low–interest rate regime to 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 a 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 that is equal to zero when the average monthly probability of being in the high–interest rate regime is smaller than one-half and equal to unity when this average probability is greater than or equal to one-half. 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:[18]

$$p_t(i^{\mathrm{HIGH}}) = \frac{\exp[\delta_0 + \delta_1(x_t - x_t^*) + \delta_2(\pi_t - \pi^*)]}{1 + \exp[\delta_0 + \delta_1(x_t - x_t^*) + \delta_2(\pi_t - \pi^*)]}, \tag{14}$$

where $p_t(i^{\mathrm{HIGH}})$ 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.[19]

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 et al. 2004) but also with the very large literature on monetary policy rules—so-called Taylor rules (see, e.g., Taylor [1993, 1999] and Clarida, Gali, and Gertler [1998, 2000] and the references therein). The Taylor rule literature provides evidence that it is

18. Clearly, the probability of being in the low–interest rate regime is $1 - p_t(i^{\mathrm{HIGH}})$.

19. Our monthly data for inflation are, for all three countries examined, the (annualized) rate of change in the consumer price index. 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.

TABLE 6 **MSIH-VECM Regime Interpretation: Logit Estimation**

	$\tilde{\delta}_1$	$\tilde{\delta}_2$	Pseudo-R^2	Classification Ratio
Germany	.1905 (.080)	.1803 (.056)	.255	.715
Japan	.0878 (.033)	.0398 (.012)	.218	.629
United States	.3928 (.168)	.2724 (.117)	.198	.775

NOTE.—δ_1 and δ_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. The classification ratio is 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 calculated by two-step nonlinear two-stage least squares (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.

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 affect 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 should probably expect both δ_1 and δ_2 to be positive.

The results of our logit estimations for each of the countries examined are reported in table 6. Consistent with the findings reported in Bansal and Zhou (2002) and Bansal et al. (2004), 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 $\tilde{\delta}_1$, associated with our proxy for the output gap, is found to be statistically significant for all 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, $\tilde{\delta}_2$, is found to be statistically significant at conventional significance levels. The estimated parameters, $\tilde{\delta}_1$ and $\tilde{\delta}_2$, are both significantly positive, thus 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 (in the range between 0.20 and 0.25) and, perhaps more important, displays a very interesting "classification ratio": the ratio of correctly classified observations implied by the logit model to the total number of observations is in the range between 63% for Japan and 77% for the United States, 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.

V. Forecasting the Term Structure of Interest Rates out of Sample with the MSIH-VECM

The procedure we have applied so far allowed us to achieve a reliable in-sample representation of the dynamic relationship implied by the EH theory of the term structure of interest rates. 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 $j = 1, \dots, 52$ were constructed recursively, conditional only on information up to the date of the forecast and with successive reestimation as the date on which forecasts are conditioned moves through the data set.

It is well known in the literature that forecasting with nonlinear models raises special technical problems (see Brown and Mariano [1984, 1989], Hamilton [1993, 1994], and, for general discussions, Granger and Teräsvirta [1993, chap. 8] and Franses and van Dijk [2000, chaps. 3, 4]). We therefore adopted a very general forecasting procedure based on Monte Carlo integration that is capable of producing forecasts virtually identical to the analytical forecasts for a wide range of models. In particular, we forecast the path for y_{t+j} for $j = 1, \dots, 52$ using Monte Carlo simulations calibrated on the estimated MSIH-VECMs. The simulation procedure is repeated with identical random numbers 10,000 times, and the average of the 10,000 realizations of the sequence of forecasts is taken as the point forecast. Since we use a large number of simulations, by the law of large numbers this procedure should produce results virtually identical to those that would result from calculating the exact forecasts analytically (Brown and Mariano 1984, 1989; Gallant, Rossi, and Tauchen 1993).

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

$$\mathrm{APFE}_{t+j} = \frac{\sum_{k=1}^{N} |i_{k,t+j} - \tilde{i}_{k,t+j}|}{N}, \tag{15}$$

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

$$\mathrm{SPFE}_{t+j} = \frac{\sum_{k=1}^{N} (i_{k,t+j} - \tilde{i}_{k,t+j})^2}{N}, \tag{16}$$

where N is the number of eurorates included in the system (i.e., $N = 5$) and $\tilde{i}_{k,t+j}$ 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 (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:

$$\mathrm{DM} = \frac{\bar{d}}{\sqrt{2\pi \hat{f}(0)/T}}, \tag{17}$$

where $\bar{d}$ is an average (over T observations) of a general loss differential function of the APFE (or SPFE), and $\hat{f}(0)$ 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, e.g., Mark 1995), the loss differential functions we consider are either the difference between the APFE for the two models or the difference between the SPFEs. A consistent estimate of the spectral density at frequency zero, $\hat{f}(0)$, 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.[20]

Several problems may arise when one uses DM statistics in small samples and takes 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 nonstandard and

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

unknown. Therefore, the marginal significance levels reported below should be interpreted with caution.[21]

Table 7 gives detailed results of the accuracy of the forecasts for Germany, Japan, and the United States 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, across countries, the asymmetric MSIH-VECMs give more accurate forecasts. At the four-week horizon, for example, we achieve improvements ranging between 8% and 77% across countries using the APFE and between 6% and 93% using the SPFE. At the 52-week horizon, we obtain improvements ranging between 17% and 78% using the APFE and between 28% and 93% using the SPFE. The statistical significance of these results is confirmed by executing the DM tests.[22]

The gain in terms of accuracy of the predictive performance of the asymmetric MSIH-VECM is less impressive when compared with the symmetric (linear and Markov-switching) VECMs. In fact, while the asymmetric MSIH-VECM performs very well at longer horizons, at the one- and four-week horizons there are improvements only for Japan and the United States. A similar 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 Japan and the United States, whereas in the case of Germany, only asymmetries seem to be key to improving forecasting performance.

Overall, these results suggest that using a VECM framework for the term structure of interest rates, one can generate satisfactory out-of-sample forecasts of the term structure. Moreover, by explicitly incorporating asymmetry and, for two out of three countries examined, regime shifts into the modeling framework, in the present analysis we have been able to improve on a standard linear VECM framework. The gain from using an MSIH-VECM rather than a linear VECM may be relatively small at short horizons; however, this gain

21. Clark and McCracken (2001) derive the asymptotic distributions of two standard tests in this context for one-step-ahead forecasts from nested linear models. Their results are, unfortunately, not directly applicable to our case because we are dealing with multi-step-ahead forecasts from nested models and because one of the competing models is a Markov-switching model. Therefore, our case is one for which the asymptotic theory of the DM statistic is unknown at the present time. A possible solution would involve calculating the marginal significance levels by bootstrap methods using a variant of the bootstrap procedure proposed by Kilian (1999), although this procedure is computationally very demanding and it is unknown whether it is valid in the context of symmetric and asymmetric MSIH-VECMs.

22. 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.

generally increases with the forecast horizon and becomes very substantial indeed at the one-year horizon.[23]

VI. Robustness Checks

In this section we report the results of robustness checks that were carried out in order to evaluate the sensitivity of some of the estimation results reported in Section IV and the forecasting results reported in Section V. Specifically, we first report robustness checks designed to assess whether our choice of an asymmetric MSIH(2)-VECM(1), as suggested by the asymmetry tests and the bottom-up procedure over the sample period 1982–91, would change if we used either data for the full sample from 1982 to 2000 or the subsample from 1992 to 2000. We then also investigate in greater detail the robustness of the forecasting results by examining the forecasting performance of the asymmetric MSIH(2)-VECM(1) recursively year after year over the forecast period in order to shed light on whether the forecasts from this model work particularly well (or poorly) over certain periods.

The robustness results are reported in Appendix B and figure 1. Tables B1–B4 in Appendix B suggest that (i) the long-run cointegrating parameters are relatively stable and qualitatively identical to the ones reported in table 3 when we use either the full sample from 1982 to 2000 or the subsample from 1992 to 2000 (see table B1); (ii) for these two samples, the hypothesis of symmetry of the term structure VECM is strongly rejected using the appropriate LR test, providing clear empirical evidence that the linear VECM fails to capture significant asymmetries in the data-generating process (see table B2, compared to table 4); (iii) for these two samples the bottom-up identification procedure suggests that an MSIH(2)-VECM(1) is the most adequate model, within its class for our data (see table B3, compared to App. table A1); and (iv) for these two samples the hypothesis of linearity of either the symmetric VECM or the asymmetric VECM is rejected when each of these two models is tested against their MSIH-VECM counterparts (see table B4, compared to table 5). Overall, therefore, the results in tables B1–B4 suggest that our choice of an MSIH(2)-VECM(1) is robust to our choice of the sample period (1982–91) that precedes the forecasting exercise since, whether we had used the full sample beginning in 1982 or the subsample beginning in 1992,

23. The superior long-horizon forecasting performance of the regime-switching model relative to its linear counterpart may at first sight seem puzzling, since the steady-state or long-run equilibria will be the same for each of the models—namely, the cointegrating relationships linking the five interest rates examined. However, it seems plausible that this arises because the Markov-switching model generates a better estimate of the intercept term in the VECM. If the world truly does approximate to the kind of Markov-switching that we have modeled, then this should provide a better estimate of the intercept term than can be obtained by estimating it without accounting for regime switching. The importance of allowing for structural shifts in intercept terms when conducting forecasting exercises has been emphasized by Clements and Hendry (1996).

TABLE 7 **Out-of-Sample Forecasting Exercises**

k	VAR	VECMS	VECMA	MSVAR	MSVECMS
			A. Germany: Average APFE		
1	.9294	1.8772	3.1519	.8241	1.0475
	$[2.32 \times 10^{-22}]$	[0]	[0]	$[2.66 \times 10^{-98}]$	$[8.68 \times 10^{-45}]$
4	.9192	1.6191	3.0603	.8727	1.1181
	$[4.87 \times 10^{-5}]$	[0]	$[4.77 \times 10^{-319}]$	$[4.23 \times 10^{-112}]$	$[3.24 \times 10^{-99}]$
13	.9751	1.4572	1.8367	.9293	1.1013
	$[7.90 \times 10^{-6}]$	[0]	[0]	$[2.60 \times 10^{-41}]$	$[3.99 \times 10^{-182}]$
26	.8912	1.1060	1.7791	.85103	.97438
	$[4.69 \times 10^{-98}]$	[0]	[0]	$[1.27 \times 10^{-178}]$	$[9.78 \times 10^{-18}]$
52	.8355	.9737	2.5175	.7922	.8983
	$[5.86 \times 10^{-172}]$	$[7.04 \times 10^{-36}]$	$[4.61 \times 10^{-255}]$	$[8.47 \times 10^{-256}]$	$[4.64 \times 10^{-138}]$
			Average SPFE		
1	.9273	3.3897	7.4424	.7502	1.0861
	$[5.75 \times 10^{-12}]$	$[1.74 \times 10^{-218}]$	$[7.53 \times 10^{-239}]$	$[9.82 \times 10^{-133}]$	$[8.83 \times 10^{-148}]$
4	.9371	2.2721	7.9361	.83698	1.1160
	$[7.75 \times 10^{-10}]$	$[2.26 \times 10^{-195}]$	$[1.13 \times 10^{-191}]$	$[1.63 \times 10^{-74}]$	$[3.02 \times 10^{-37}]$
13	.9695	1.9388	2.7786	.8745	1.1850
	$[2.02 \times 10^{-4}]$	$[3.69 \times 10^{-280}]$	[0]	$[2.80 \times 10^{-50}]$	$[4.91 \times 10^{-100}]$
26	.8119	1.1987	2.5704	.7345	.9545
	$[2.80 \times 10^{-107}]$	[0]	[0]	$[2.24 \times 10^{-205}]$	$[1.29 \times 10^{-15}]$
52	.7191	.9657	5.1224	.6383	.8336
	$[5.60 \times 10^{-182}]$	$[2.90 \times 10^{-20}]$	$[2.12 \times 10^{-197}]$	$[7.24 \times 10^{-266}]$	$[2.99 \times 10^{-121}]$
			B. Japan: Average APFE		
1	.8840	.6758	.7500	.8568	.8093
	$[7.97 \times 10^{-72}]$	$[2.15 \times 10^{-187}]$	$[5.92 \times 10^{-248}]$	$[7.52 \times 10^{-108}]$	$[3.88 \times 10^{-182}]$
4	.6948	.5790	.5374	.6853	.8044
	[0]	$[2.18 \times 10^{-239}]$	$[2.48 \times 10^{-251}]$	[0]	$[3.37 \times 10^{-308}]$
13	.9201	.5567	.5298	.9103	.9920
	$[2.77 \times 10^{-12}]$	$[1.34 \times 10^{-103}]$	$[1.56 \times 10^{-262}]$	$[1.20 \times 10^{-2}]$	$[3.62 \times 10^{-1}]$
26	.8832	.6407	.5884	.8703	.7550
	$[1.92 \times 10^{-190}]$	$[2.97 \times 10^{-152}]$	[0]	$[1.10 \times 10^{-280}]$	$[4.00 \times 10^{-229}]$
52	.4947	.4151	.8069	.4847	.5167
	$[7.88 \times 10^{-250}]$	$[6.41 \times 10^{-240}]$	$[3.14 \times 10^{-67}]$	$[1.52 \times 10^{-243}]$	$[5.71 \times 10^{-171}]$
			Average SPFE		
1	1.2591	.3574	.5840	1.1814	.6441
	$[4.12 \times 10^{-40}]$	$[1.74 \times 10^{-127}]$	$[6.81 \times 10^{-169}]$	$[2.44 \times 10^{-22}]$	$[1.88 \times 10^{-135}]$
4	.6303	.3509	.2819	.6052	.6847
	$[3.54 \times 10^{-198}]$	$[1.66 \times 10^{-175}]$	$[4.96 \times 10^{-175}]$	$[1.36 \times 10^{-225}]$	$[2.26 \times 10^{-210}]$
13	.9117	.4069	.3335	.8865	1.0138
	$[7.24 \times 10^{-12}]$	$[2.83 \times 10^{-72}]$	$[9.91 \times 10^{-184}]$	$[2.50 \times 10^{-2}]$	$[3.54 \times 10^{-1}]$
26	.9050	.4400	.3706	.8657	.6222
	$[4.13 \times 10^{-37}]$	$[2.16 \times 10^{-149}]$	$[1.08 \times 10^{-229}]$	$[2.44 \times 10^{-103}]$	$[2.87 \times 10^{-187}]$
52	.3022	.2176	.6892	.2856	.3127
	$[1.01 \times 10^{-133}]$	$[4.06 \times 10^{-184}]$	$[1.37 \times 10^{-59}]$	$[4.94 \times 10^{-144}]$	$[6.43 \times 10^{-113}]$
			C. United States: Average APFE		
1	.3885	.4784	.6838	.3584	.4368
	$[9.12 \times 10^{-114}]$	$[3.65 \times 10^{-134}]$	[0]	$[2.25 \times 10^{-123}]$	$[1.57 \times 10^{-106}]$
4	.2577	.2658	.5590	.2306	.2526
	$[2.02 \times 10^{-123}]$	$[2.73 \times 10^{-121}]$	[0]	$[2.62 \times 10^{-136}]$	$[8.36 \times 10^{-123}]$
13	.4748	.4289	1.4997	.4357	.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	.3334	.3104	1.2529	.3083	.3035
	$[7.16 \times 10^{-140}]$	$[6.58 \times 10^{-150}]$	$[5.92 \times 10^{-9}]$	$[1.64 \times 10^{-154}]$	$[1.04 \times 10^{-158}]$

TABLE 7 (*Continued*)

k	VAR	VECMS	VECMA	MSVAR	MSVECMS
52	.2373	.2308	.7201	.2222	.2207
	[3.67×10^{-306}]	[7.32×10^{-292}]	[1.66×10^{-13}]	[9.22×10^{-321}]	[8.55×10^{-312}]
			Average SPFE		
1	.1494	.1982	.4579	.1278	.1447
	[3.14×10^{-146}]	[6.35×10^{-109}]	[2.76×10^{-267}]	[7.57×10^{-155}]	[1.28×10^{-126}]
4	.0741	.0771	.3354	.0626	.0693
	[3.48×10^{-156}]	[9.73×10^{-117}]	[1.21×10^{-265}]	[6.34×10^{-165}]	[8.01×10^{-139}]
13	.2207	.1759	1.5643	.1854	.1765
	[1.77×10^{-110}]	[7.11×10^{-108}]	[3.15×10^{-11}]	[2.26×10^{-124}]	[8.54×10^{-122}]
26	.1211	.1027	1.5570	.1040	.0997
	[1.09×10^{-180}]	[8.07×10^{-161}]	[1.00×10^{-8}]	[2.23×10^{-191}]	[8.00×10^{-179}]
52	.0778	.0727	.6119	.0685	.0670
	[2.99×10^{-235}]	[2.21×10^{-204}]	[5.32×10^{-8}]	[3.54×10^{-247}]	[1.18×10^{-223}]

NOTE.—VAR, VECMS, VECMA, MSVAR, and MSVECMS are the ratios of the average (absolute or square) cross-sectional pricing forecast errors (APFE and SPFE as defined in eqq. [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 the Andrews (1991) AR(1) rule. For the Diebold-Mariano statistics, only p-values are reported (0 indicates a p-value below 10^{-350}).

our testing procedure would have yielded the same outcome; that is, it would have indicated that an MSIH(2)-VECM(1) is an appropriate characterization of our interest rate data.

With respect to the forecasting results, the main concern is whether they display large variability over different subperiods of the full forecast period we have studied, which goes from 1992 through 2000. In figure 1 we graph the ratios of the average APFE (again as defined in eq. [15]) obtained by the asymmetric MSIH-VECM relative to the APFE obtained by the asymmetric VECM without regime switching; these two models are indeed the two best-performing models on the basis of the evidence reported in Section V. The ratios were obtained recursively, and we plot their evolution over the forecast period year after year, until they converge to the results for the full forecast period ending in 2000, which we report in table 7. The results are reported for each of the forecast horizons studied, namely 1, 4, 13, 26, and 52 weeks ahead, and for each of the countries examined.[24] The bar charts in figure 1 suggest that there is some degree of variability of the relative performance of the MSIH-VECM versus its more parsimonious linear asymmetric counterpart. However, with the possible exception of the 13- and 26-week-ahead forecasts for the United States, the bar charts become rather flat after two to three years, the ratios do not vary drastically over time, and, qualitatively,

24. We restrict ourselves to reporting only the results for the APFE (not the SPFE) and for the two best-performing models because of space constraints. Investigation of the SPFE and of other models did not change qualitatively any of the results discussed below.

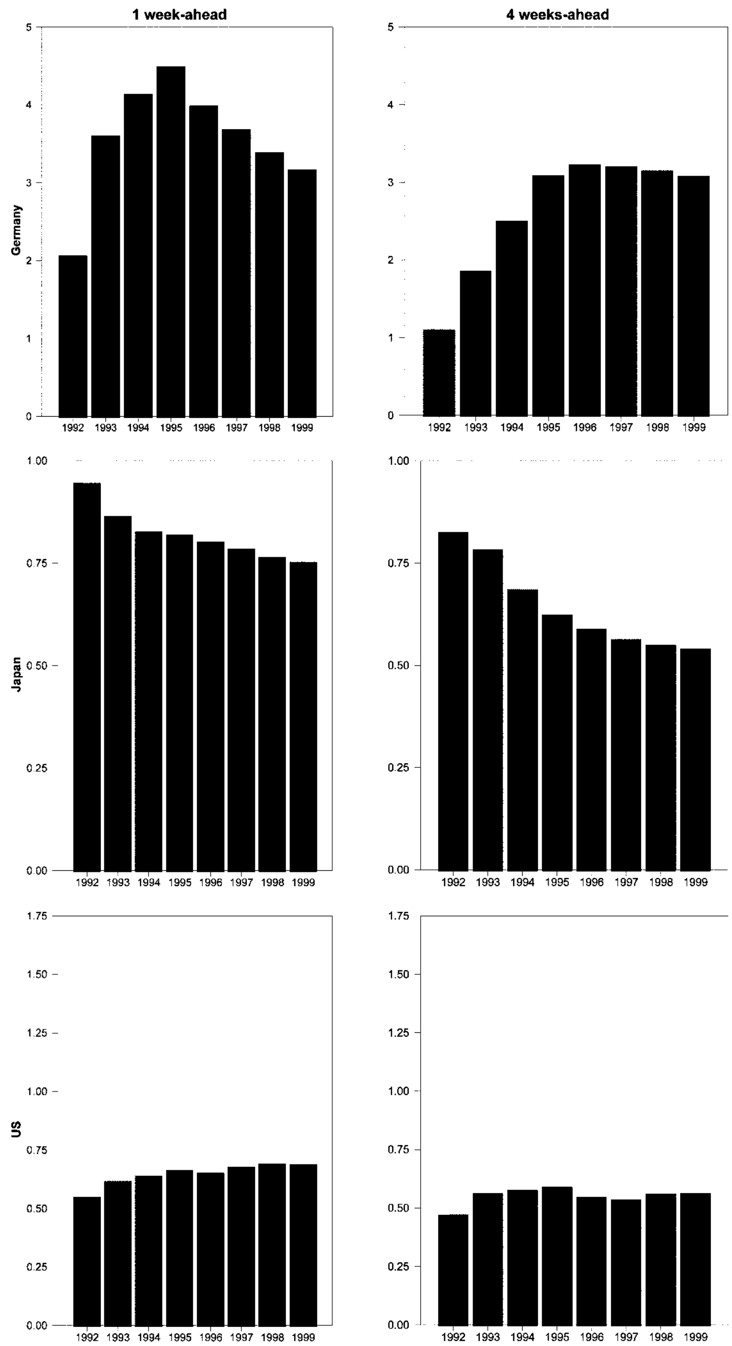
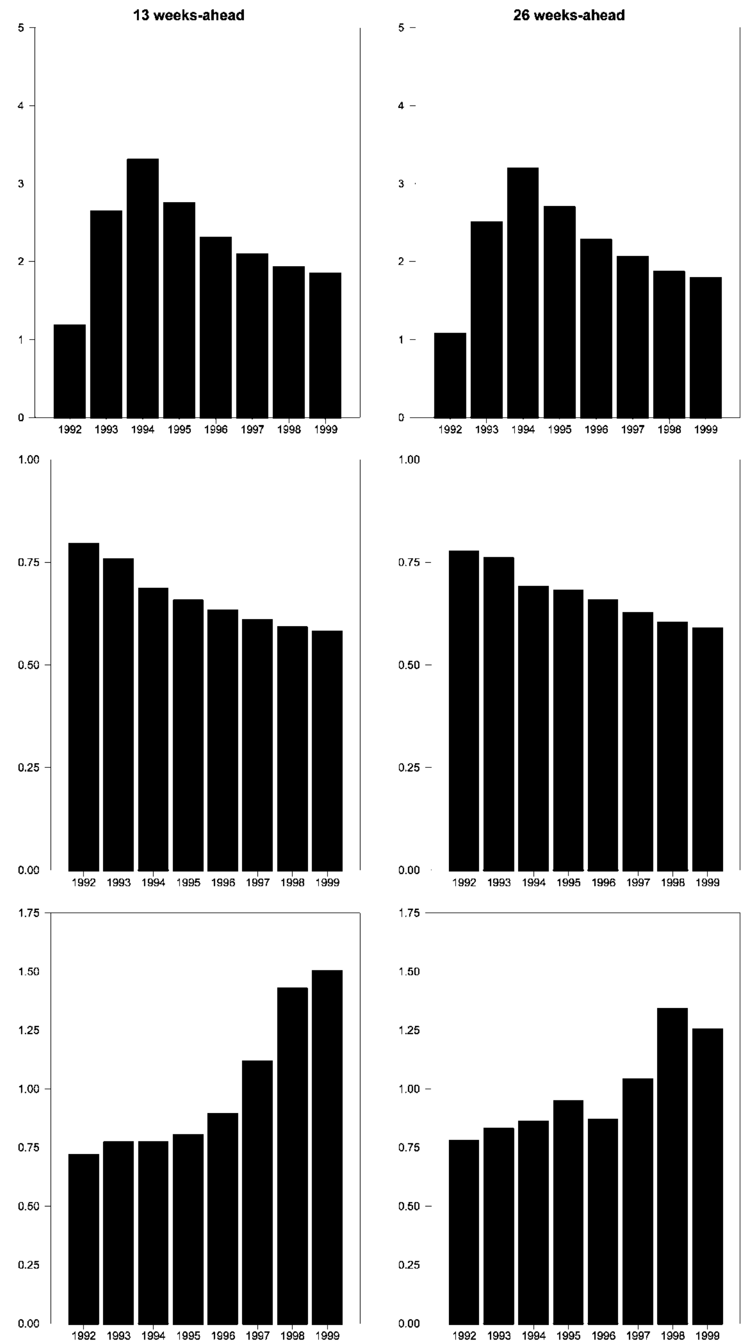


FIG. 1.—Average absolute pricing errors ratio


13 weeks-ahead
26 weeks-ahead
0
1
2
3
4
5
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
1992
1993
1994
1995
1996
1997
1998
1999

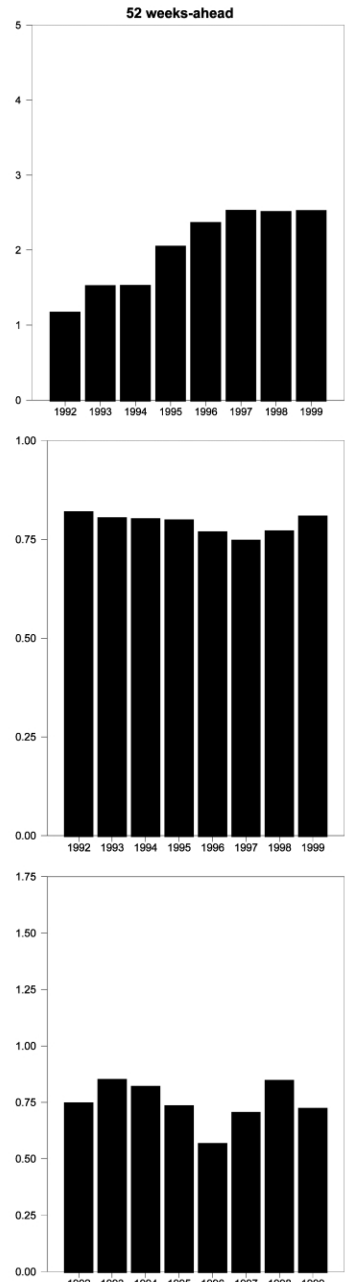


FIG. 1.—*Continued*

the issue as to which model does better is not strongly dependent on the period chosen for the forecasting exercise. We view this evidence as suggesting that our conclusion that the asymmetric MSIH-VECM is the best-performing model for the United States and Japan, whereas the asymmetric VECM without regime shifts is the best model for German interest rates, is fairly robust to the length of the forecast period considered.

VII. 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 Germany, Japan, and the United States 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 modeled satisfactorily by a multivariate asymmetric two-regime Markov-switching VECM that allows for shifts in both the intercept and the covariance structure. We then used this model to forecast dynamically out of sample over the period January 1992 through December 2000. The forecasting results were extremely 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 basis of standard tests of significance. Moreover, the asymmetric nonlinear VECM, in general, outperformed 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.[25]

Our research was motivated by encouraging results previously reported in the literature on the presence of regime shifts (e.g., Hamilton 1988; Gray 1996; Anderson 1997) and asymmetries (e.g., Enders and Granger 1998) as well as by the relative forecasting success of the linear VECM model of the term structure of interest rates (e.g., Hall et al. 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

25. The only exception is Germany, for which a VECM that allows for asymmetries but not regime shifts emerges as the best forecasting model.

appear to be related to the state of the business cycle and to inflation, as one would expect in economies in which 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, e.g., 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 nonnormal forecast densities, it would seem appropriate to consider a model's density forecasting performance. This is a further immediate avenue for future research.

Appendix A

Model Estimation Results

TABLE A1 Bottom-Up Identification Procedure

	LR_1	LR_2	LR_3
Germany	2.65×10^{-67}	1.18×10^{-112}	.073
Japan	8.04×10^{-38}	3.43×10^{-104}	.409
United States	5.72×10^{-74}	3.52×10^{-158}	.239

NOTE.—LR_1, LR_2, and LR_3 are test statistics of the null hypothesis of no regime-dependent intercept, no regime-dependent variance-covariance matrix, and of MSIH(2)-VECM(1) vs. MSIH(2)-VECM(11), respectively. Each of LR_1, LR_2, and LR_3 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 $\chi^2(g)$, where g is the number of restrictions. Figures reported denote p-values.

Asymmetric MSIH(2)-VECM(1) Estimation Results: United States

A.

$$\tilde{\Gamma}_1 = \begin{bmatrix} -.2444 & -.0865 & .3088 & -.2249 & -.0008 \\ (.039) & (.081) & (.161) & (.172) & (.116) \\ -.0638 & .0805 & -.0575 & -.0065 & .0406 \\ (.034) & (.078) & (.146) & (.155) & (.100) \\ -.0576 & .0873 & -.0848 & -.0391 & .1141 \\ (.037) & (.082) & (.163) & (.169) & (.110) \\ -.0556 & .0621 & .0220 & -.1743 & .1945 \\ (.040) & (.088) & (.178) & (.188) & (.123) \\ -.0565 & -.0006 & .0020 & .0630 & .0413 \\ (.044) & (.094) & (.192) & (.208) & (.136) \end{bmatrix};$$

$$\tilde{\nu}(z_1) = \begin{bmatrix} -.0479 \\ (.054) \\ -.0480 \\ (.044) \\ .0011 \\ (.041) \\ .0102 \\ (.042) \\ .0002 \\ (.041) \end{bmatrix}; \quad \tilde{\nu}(z_2) = \begin{bmatrix} .0201 \\ (.012) \\ .0001 \\ (.010) \\ .0113 \\ (.012) \\ .0027 \\ (.014) \\ .0053 \\ (.016) \end{bmatrix};$$

$$\tilde{\alpha}^+ = \begin{bmatrix} -.7300 & .0675 & .0099 & .0147 \\ (.100) & (.214) & (.085) & (.112) \\ .1470 & -.1543 & .1438 & -.0962 \\ (.103) & (.192) & (.090) & (.098) \\ -.0204 & .2146 & -.0732 & -.0856 \\ (.105) & (.181) & (.097) & (.111) \\ -.0005 & -.1591 & -.4854 & -.2963 \\ (.109) & (.248) & (.094) & (.125) \\ .0370 & -.0677 & .0656 & -.1191 \\ (.108) & (.276) & (.092) & (.139) \end{bmatrix};$$

$$\tilde{\alpha}^- = \begin{bmatrix} .0250 & -.0440 & .0207 & -.0749 \\ (.107) & (.06) & (.181) & (.082) \\ -.3183 & -.2694 & .0757 & -.0140 \\ (.093) & (.073) & (.158) & (.072) \\ .0606 & .0484 & -.1098 & .0591 \\ (.102) & (.081) & (.189) & (.081) \\ .0775 & -.0348 & .1865 & -.0223 \\ (.104) & (.095) & (.196) & (.087) \\ .0555 & -.0707 & .0172 & 0.0773 \\ (.106) & (.056) & (.204) & (.093) \end{bmatrix}$$

B.

$$\tilde{\Sigma}_e(z_1) = \begin{bmatrix} .0177 & & & & \\ .0102 & .0149 & & & \\ .0096 & .0154 & .0202 & & \\ .0099 & .0168 & .0223 & .0275 & \\ .0097 & .0183 & .0249 & .0309 & .0378 \end{bmatrix};$$

$$\tilde{\Sigma}_e(z_2) = \begin{bmatrix} .3423 & & & & \\ .1737 & .2167 & & & \\ .1474 & .1700 & .1774 & & \\ .1356 & .1526 & .1686 & .1706 & \\ .1215 & .1373 & .1520 & .1545 & .1488 \end{bmatrix};$$

$$\tilde{\mathbf{P}} = \begin{bmatrix} .688 & .118 \\ .312 & .882 \end{bmatrix}; \ \tilde{\xi} = \begin{bmatrix} .275 \\ .725 \end{bmatrix}$$

$$\rho(A) = .22510$$

Tildes denote estimated values obtained using the EM algorithm for maximum likelihood (Dempster et al. 1977). Figures in parentheses are asymptotic standard errors. Symbols are defined as in eq. (13). P and ξ denote the $M \times M$ transition matrix and the M-dimensional ergodic probabilities vector, respectively. $\rho(A)$ is the "spectral radius" calculated as in Karlsen (1990), which can be thought as a measure of stationarity of the MSIH-VECM; stationarity requires $0 \leq \rho(A) < 1$.

Appendix B

Robustness Results

TABLE B1 Long-Run Cointegrating Equilibrium Parameters

k	Germany	Japan	United States
		A. January 1992–December 2000	
4 weeks	.985 (.006)	1.000 (.006)	.925 (.010)
13 weeks	.986 (.013)	.994 (.014)	.959 (.020)
26 weeks	.992 (.025)	.978 (.024)	1.001 (.030)
52 weeks	1.011 (.047)	.924 (.040)	1.029 (.056)
		B. February 1982–December 2000	
4 weeks	.972 (.006)	.978 (.005)	.935 (.007)
13 weeks	.955 (.013)	.975 (.009)	.951 (.015)
26 weeks	.948 (.023)	.984 (.013)	.966 (.023)
52 weeks	.962 (.038)	.998 (.018)	.979 (.034)

NOTE.—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 B2 Asymmetry Tests

H_0	January 1992–December 2000 LR	January 1992–December 2000 *p*-Value	February 1982–December 2000 LR	February 1982–December 2000 *p*-Value
Germany	29.977	7.02×10^{-6}	35.621	1.70×10^{-6}
Japan	117.352	8.80×10^{-16}	105.235	1.44×10^{-13}
United States	50.885	1.65×10^{-4}	100.837	8.92×10^{-13}

NOTE.—LR is a likelihood ratio test of the symmetry null hypothesis, where the restricted model being tested is the symmetric linear VECM in eq. (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 B3 Bottom-Up Identification Procedure

	LR_1	LR_2	LR_3
	A. January 1992–December 2000		
Germany	2.37×10^{-69}	1.60×10^{-101}	.69
Japan	5.58×10^{-49}	9.62×10^{-206}	.83
United States	1.29×10^{-78}	1.41×10^{-191}	.25
	B. February 1982–December 2000		
Germany	8.93×10^{-76}	0	.74
Japan	1.61×10^{-51}	0	.61
United States	3.30×10^{-86}	0	.27

NOTE.—LR_1 and LR_2 are test statistics of the null hypothesis of no regime dependent variance-covariance matrix and of MSIH(2)-VECM(1) vs. MSIH(2)-VECM(11), respectively. Each of LR_1 and LR_2 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 $\chi^2(g)$, where g is the number of restrictions. Figures in braces denote p−values, and {0} indicates *p*-values below 10^{-350}, which are considered virtually zero.

TABLE B4 Markov-Switching VECM Estimation: Linearity Tests

H_0	January 1992–December 2000 LR_{S1}	January 1992–December 2000 LR_{S2}	February 1982–December 2000 LR_{S1}	February 1982–December 2000 LR_{S2}
Germany	789.61	792.15	1,653.78	1,662.08
Japan	1,245.52	1,222.09	2,038.10	2,030.85
United States	995.28	1,028.15	2,796.47	2,807.86

NOTE.—LR_{S1} and LR_{S2} are likelihood ratio tests in which the restricted models being tested are the symmetric linear VECM in eq. (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 since they are virtually zero in each case.

References

Ahn, D.-H., R. F. Dittmar, and A. R. Gallant. 2002. Quadratic term structure models: Theory and evidence. *Review of Financial Studies* 15:243–88.

Anderson, H. M. 1997. Transaction costs and nonlinear adjustment towards equilibrium in the US Treasury bill market. *Oxford Bulletin of Economics and Statistics* 59:465–84.

Andrews, D. W. K. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59:817–58.

Ang, A., and G. Bekaert. 2002. Regime switches in interest rates. *Journal of Business and Economic Statistics* 20:163–82.

Bansal, R., G. Tauchen, and H. Zhou. 2004. Regime shifts, risk premiums in the term structure, and the business cycle. *Journal of Business and Economics Statistics* 22:396–409.

Bansal, R., and H. Zhou. 2002. Term structure of interest rates with regime shifts. *Journal of Finance* 57:1997–2043.

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.

Brown, B. Y., and R. S. Mariano. 1984. Residual based stochastic predictors and estimation in nonlinear models. *Econometrica* 52:321–43.

———. 1989. Predictors in dynamic nonlinear models: Large sample behaviour. *Econometric Theory* 5:430–52.

Campbell, J. Y., and R. H. Clarida. 1986. The term structure of euromarket interest rates: An empirical investigation. *Journal of Monetary Economics* 19:25–44.

Campbell, J. Y., and R. J. Shiller. 1987. Cointegration and tests of present value models. *Journal of Political Economy* 95:1062–88.

———. 1991. Yield spreads and interest rate movements: A bird's eye view. *Review of Economic Studies* 58:495–514.

Clarida, R. H., J. Gali, and M. Gertler. 1998. Monetary policy rules in practice: Some international evidence. *European Economic Review* 42:1033–67.

———. 1999. The science of monetary policy: A new Keynesian perspective. *Journal of Economic Literature* 37:1661–1707.

———. 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics* 115:147–80.

Clarida, R. H., L. Sarno, M. P. Taylor, and G. Valente. 2003. The out-of-sample success of term structure models as exchange rate predictors: A step beyond. *Journal of International Economics* 60:61–83.

Clarida, R. H., and M. P. Taylor. 2003. Nonlinear permanent-temporary decompositions in macroeconomics and finance. *Economic Journal* 113:C125–C139.

Clark, T. E., and M. W. McCracken. 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105:85–110.

Clements, M. P., and D. F. Hendry. 1996. Intercept corrections and structural change. *Journal of Applied Econometrics* 11:475–94.

Dai, Q., and K. J. Singleton. 2000. Specification analysis of affine term structure models. *Journal of Finance* 55:1943–78.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, ser. B, 39: 1–38.

Diebold, F. X., T. A. Gunther, and A. S. Tay. 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39:863–83.

Diebold, F. X., and R. S. Mariano. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13:253–63.

Enders, W., and C. W. J. Granger. 1998. Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates. *Journal of Business and Economic Statistics* 16:304–11.

Engle, R. E., and C. W. J. Granger. 1987. Co-integration and equilibrium correction representation, estimation and testing. *Econometrica* 55:251–76.

Estrella, A. 1998. A new measure of fit for equations with dichotomous dependent variables. *Journal of Business and Economic Statistics* 16:198–205.

Franses, P. H., and D. van Dijk. 2000. *Non-linear time series models in empirical finance*. Cambridge: Cambridge University Press.

Gallant A. R. 1987. *Nonlinear statistical models*. New York: Wiley.

Gallant, A. R., P. E. Rossi, and G. Tauchen. 1993. Nonlinear dynamic structures. *Econometrica* 61:871–907.

Granger, C. W. J. 1986. Developments in the study of cointegrated variables. *Oxford Bulletin of Economics and Statistics* 48:213–28.

Granger, C. W. J., and M. H. Pesaran. 1999. A decision theoretic approach to forecast evaluation.

In *Statistics and finance: An interface*, ed. W. S. Chan, W. K. Lin, and H. Tong. London: Imperial College Press.

Granger, C. W. J., and T. Teräsvirta. 1993. *Modelling nonlinear economic relationships*. Oxford: Oxford University 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., H. M. Anderson, and C. W. J. Granger. 1992. A cointegration analysis of Treasury bill yields. *Review of Economics and Statistics* 74:116–26.

Hamilton, J. D. 1988. Rational expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control* 12:385–423.

———. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57:357–84.

———. 1993. Estimation, inference and forecasting of time series subject to changes in regime. In *Handbook of econometrics*, vol. 4, ed. G. S. Maddala, C. R. Rao, and H. D. Vinod. Amsterdam: Elsevier.

———. 1994. *Time series analysis*. Princeton, NJ: Princeton University Press.

Hansen, H., and S. Johansen. 1999. Some tests for parameter constancy in cointegrated VAR-models. *Econometrics Journal* 2:306–33.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50:1029–54.

Jagannathan, R., A. Kaplin, and G. Sun. 2000. An evaluation of multi-factor CIR models using LIBOR, swap rates and cap and swaption prices. Unpublished manuscript, Northwestern University, Kellogg Graduate School of Management.

Johansen, S. 1988. Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control* 12:231–54.

———. 1991. Estimation and hypothesis testing of cointegrating vectors in Gaussian vector autoregressive models. *Econometrica* 59:1551–80.

———. 1995. Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of Econometrics* 69:111–32.

Karlsen, H. A. 1990. A class of non-linear time series models. PhD diss., University of Bergen, Department of Mathematics.

Kilian, L. 1999. Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? *Journal of Applied Econometrics* 14:491–510.

Kim, C.-J., and C. R. Nelson. 1999. *State-space models with regime switching*. Cambridge, MA: MIT Press.

Krolzig, H.-M. 1997. *Markov-switching vector autoregressions*. New York: Springer.

———. 1999. Statistical analysis of cointegrated VAR processes with Markovian regime shifts. Mimeo, Nuffield College and University of Oxford, Department of Economics.

Leamer, E. E. 1978. *Specification searches: Ad hoc inference with nonexperimental data*. New York: Wiley.

MacKinnon, J. G. 1991. Critical values for cointegration tests. In *Long-run economic relationships: Readings in cointegration*, ed. R. F. Engle and C. W. J. Granger. Oxford: Oxford University Press.

Mark, N. 1995. Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review* 85:201–18.

McCracken, M. W. 2000. Robust out-of-sample inference. *Journal of Econometrics* 99:195–223.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

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: 461–72.

Saikkonen, P. 1992. Estimation and testing of cointegrated systems by an autoregressive approximation. *Econometric Theory* 8:1–27.

Saikkonen, P., and R. Luukkonen. 1997. Testing cointegration in infinite order vector autoregressive processes. *Journal of Econometrics* 81:93–126.

Sarno, L., and D. L. Thornton. 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.

Sarno, L., and G. Valente. 2005. Modelling and forecasting stock returns: Exploiting the futures market, regime shifts and international spillovers. *Journal of Applied Econometrics* 20:345–76.
Stock, J. H. 1997. Cointegration, long-run movements, and long-horizon forecasting. In *Advances in economics and econometrics: Theory and applications*, Seventh World Congress, vol. 3, ed. D. M. Kreps and K. F. Wallis. Cambridge: Cambridge University Press.
Tay, A. S., and K. W. Wallis. 2000. Density forecasting: A survey. *Journal of Forecasting* 19: 235–54.
Taylor, J. B. 1993. Discretion versus policy rules in practice. *Carnegie Rochester Conference Series on Public Policy* 39:195–214.
———. 1999. *Monetary policy rules*. National Bureau of Economic Research Conference Report series. Chicago: University of Chicago Press.
Taylor, M. P. 1992. Modeling the yield curve. *Economic Journal* 102:524–37.
Timmermann, A. 2000. Density forecasting in economics and finance: Editorial. *Journal of Forecasting* 19:231–34.
Vasicek, O. 1977. An equilibrium characterisation of the term structure. *Journal of Financial Economics* 5:177–88.
West, K. D. 1996. Asymptotic inference about predictive ability. *Econometrica* 64:1067–84.
West, K. D., and M. W. McCracken. 1998. Regression-based tests of predictive ability. *International Economic Review* 39:817–40.