

Empirical Exchange Rate Models and Currency Risk: Some Evidence from Density Forecasts*

Lucio Sarno

University of Warwick

Giorgio Valente

University of Warwick

and

Centre for Economic Policy Research (CEPR)

First version: January 2004 - This version: May 2004

Abstract

A large literature in exchange rate economics has investigated the ability of empirical exchange rate models to forecast nominal exchange rates using conventional point forecast accuracy criteria. However, in the context of managing exchange rate risk, interest centers on more than just point forecasts. This paper provides a formal evaluation of recent empirical exchange rate models based on the term structure of forward exchange rates, which previous research has shown to be satisfactory in point forecasting, in terms of density forecasting performance. The economic value of the exchange rate density forecasts is investigated in the context of an application to a simple risk management exercise.

JEL classification: F31; F37.

Keywords: foreign exchange; forecasting; term structure; density forecast; nonlinearity.

**Acknowledgments:* This paper was partly written while Lucio Sarno was 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 *JIMF* Special Issue on "Exchange Rate Economics" and the 8th International Conference on Macroeconomic Analysis and International Finance, University of Crete, May 27-29 2004. The authors are grateful for useful conversations and comments to Mark Salmon and Mark Taylor. The authors alone are responsible for any errors that may remain and for the views expressed in the paper.

1 Introduction

Exchange rate risk plays a major role in international portfolio diversification and in several aspects of economic policy, including the assessment of the uncertainty surrounding prices of exports and imports, the value of international reserves and open positions in foreign currency, and the domestic currency value of debt payments and workers' remittances which, in turn, may affect domestic wages, prices, output and employment. In international financial markets, expectations of future exchange rates affect agents' decisions in a number of respects, including their investment, hedging, and borrowing and lending decisions. It is not surprising, therefore, that an enormous empirical literature has developed which focuses on modelling and forecasting nominal exchange rates.

The vast majority of the empirical literature on forecasting exchange rates has centered on forecasting the *level* of nominal exchange rates. This literature is highly influenced by the seminal work of Meese and Rogoff (1983a,b), who first documented that empirical exchange rate models, based on conventional macroeconomic fundamentals suggested by international macroeconomics theory, cannot outperform a simple no-change or random walk forecast of exchange rates in terms of standard measures of point forecast accuracy. After over twenty years of research since the publication of the Meese-Rogoff studies, their findings remain, with a few exceptions (e.g. Mark, 1995), very robust.¹

While macroeconomic fundamentals do not appear to be useful in forecasting exchange rates, however, models which exploit the information in the term structure of forward exchange rates and forward premia have produced satisfactory results. Clarida and Taylor (1997) first argued that, although the forward exchange rate is not an optimal predictor of the future spot exchange rate (e.g. see Lewis, 1995; Engel, 1996; and the references therein), forward rates may still contain valuable information for forecasting future spot exchange rates. Under fairly general assumptions, Clarida and Taylor (1997) derive a linear vector equilibrium correction model (VECM) of spot and forward exchange rates. Using this linear VECM, Clarida and Taylor show that it is possible to extract sufficient information from the term structure of forward premia to outperform the random walk model for several exchange rates in out-of-sample forecasting. Then, following the large literature showing the existence of regime-switching behavior in exchange rate movements (e.g. Engel and Hamilton, 1990; LeBaron, 1992; Engel, 1994; Engel and Hakkio, 1996), Clarida, Sarno, Taylor and Valente (2003) generalize the linear VECM of spot and forward exchange rates to a nonlinear, three-regime Markov-switching VECM (MS-VECM) which is found to outperform a random walk as well as to improve on

¹See, for example, the papers published in the special issue of the *Journal of International Economics* (May 2003) on "Empirical Exchange Rate Models."

the linear VECM in terms of out-of-sample forecasting.

While an extensive body of literature has investigated the performance of exchange rate models in forecasting the level of the exchange rate, surprisingly little attention has been devoted to forecasting the *density* of exchange rates. In a decision-theoretical context, the need to consider the density forecast of a variable² - as opposed to considering only its conditional mean and variance - seems fairly accepted on the basis of the argument that economic agents may have loss functions that do not depend symmetrically on the realizations of future values of potentially non-Gaussian variables (Satchell and Timmermann, 1995; Granger, 2003). In this case, agents are interested in forecasting not only the mean and variance of the variables in question, but their full predictive densities. In various contexts in economics and finance - among which the recent boom in financial risk management represents an obvious case - there is an increasingly strong need to provide and evaluate density forecasts.

Several researchers have recently proposed methods for evaluating density forecasts. These methods allow us to evaluate a model-based predictive density by measuring the closeness of two density functions or by testing the hypothesis that the predictive density generated by a particular model corresponds to the true predictive density (e.g. Diebold, Gunther and Tay, 1998; Berkowitz, 2001).³ More recently, this line of research has also provided a test statistic to formally evaluate the relative ability of competing models in matching the true predictive density (Corradi and Swanson, 2003, 2004). However, a literature focusing on the performance of empirical exchange rate models to forecast the density of exchange rates has not emerged to date, and the main focus of the relevant literature remains on point forecasting of the nominal exchange rate.⁴

The present paper contributes to the relevant literature in that we re-examine the forecasting

²A density forecast (or predictive density) of the realization of a random variable at some future time is an estimate of the probability distribution of the possible future values of that variable. It therefore provides a full description of the uncertainty associated with a forecast, in contrast with a point forecast, which contains no description of the associated uncertainty. For a survey of the literature on density forecasting and a discussion of its applications in macroeconomics and finance, see Tay and Wallis (2000). See also Granger and Pesaran (1999), Li and Tkacz (2001), Granger (2003), Wallis (2003) and Sarno and Valente (2004a,b).

³By 'true' predictive density of the data we mean the density of the data over the chosen forecast period. Therefore, no forecast is in fact carried out in this case, and the term 'predictive' simply refers to the fact that the density in question does not refer to the full sample but only to the forecast period. Also note that we use the terms 'predictive density' and 'forecast density' interchangeably below.

⁴An exception is the paper by Bos, Mahieu and van Dijk (2000), who analyze, using Bayesian methods, the implications of basic time series patterns for daily exchange rate risk management, with specific emphasis on hedging. See also Clements and Smith (2001), who consider the forecasting performance of two nonlinear exchange rate models and show that whether the nonlinearities inherent in the data can be exploited to forecast better than a random walk depends on how forecast accuracy is assessed. Evaluation based on traditional measures, such as (root) mean squared forecast errors, may mask the superiority of the nonlinear models (see also Satchell and Timmermann, 1995). Other related studies of interest include Peel and Speight (1994), Chappell, Padmore, Mistry and Ellis (1996), Dacco and Satchell (1999), Ma and Kanas (2000a,b) and Boero and Marrocu (2002).

performance of term structure models of exchange rates, which were shown to outperform a random walk in out-of-sample point forecasting by Clarida and Taylor (1997) and Clarida *et al.* (2003). However, we assess the ability of these models to forecast out-of-sample the one-step-ahead density of nominal exchange rates, hence filling, to some extent, the important gap in the literature described above.⁵ Our analysis is carried out using the recent techniques on evaluating density forecasts mentioned above as well as on Value-at-Risk (VaR) calculations.

In particular, using weekly data for eight bilateral dollar exchange rates from January 1985 to December 2003, we focus on the ability of both the linear VECM and the MS-VECM to forecast the one-week-ahead exchange rate density. To anticipate our main results, we find that Markov-switching term structure models of exchange rates produce satisfactory density forecasts of exchange rates. In particular, the MS-VECM of the term structure convincingly outperforms a random walk forecast and a linear term structure VECM in our density forecasting exercise, suggesting that the allowance for nonlinearity in these models may be particularly important to produce satisfactory out-of-sample density forecasting performance.

Finally, we illustrate the practical importance of our results on density forecasting with a simple application to a risk management exercise. In recent years, trading accounts at large financial institutions have shown a dramatic growth and become increasingly more complex. Partly in response to this trend, major trading institutions have developed risk measurement models designed to manage risk. The most common approach employed in this context is based on the VaR methodology, where VaR is defined as the expected maximum loss over a target horizon within a given confidence interval (Jorion, 2001).⁶ In our simple application we analyze the out-of-sample forecasting performance of term structure models of exchange rates, investigating the implications of these forecasts for a risk manager who has to quantify the risk associated with holding a currency portfolio over a one-week horizon. This application further illustrates how the MS-VECM captures satisfactorily the higher moments of the predictive density of exchange rates, generating VaRs that estimate the probability of large losses better than the other two competing models. Put another way, our findings indicate that better density forecasts of exchange rates, of the type recorded by the regime-switching model considered in this paper, can potentially lead to substantial improvements in risk management and,

⁵We focus on one-step-ahead forecast densities since the econometric techniques developed to date are designed for one-step-ahead density forecasting tests alone. Extensions of these tests to multi-step-ahead forecasts are awaited (e.g. Tay and Wallis, 2000).

⁶More formally, VaR is an interval forecast, typically a one-sided 95% or 99% interval of the distribution of expected wealth or returns.

more precisely, to better estimates of downside risk.

The remainder of the paper is set out as follows. In Section 2, we briefly review the literature employing term structure models of forward premia to forecast exchange rates, both in a linear framework and in a Markov-switching framework. In Section 3, we describe the data and carry out a preliminary unit root and cointegration analysis of the spot and forward exchange rates data. We also report linearity tests and other tests designed to select the most adequate MS-VECM in our context. We then forecast from these models in Section 4, where we report tests designed to assess the performance of the random walk model, the linear VECM and the MS-VECM in terms of density forecasting. We also explore the implications of the density forecasting results in a risk management exercise. A final section briefly summarizes and concludes.

2 Term Structure Forecasting Models of Exchange Rates: A Brief Overview

Let s_t and f_t^k be, respectively, the spot exchange rate and the k -period forward exchange rate at time t . Under the mild assumptions that (i) each of s_t and f_t^k are well described by unit root processes and that (ii) departures from the risk-neutral efficient markets hypothesis - namely expected foreign exchange excess returns, $f_t^k - E_t(s_{t+k}|\Omega_t)$, defined with respect to a given information set Ω_t - are stationary, it is straightforward to derive an expression which implies that the forward premium, $f_t^k - s_t$ is stationary (Clarida and Taylor, 1997). In turn, this result implies that forward and spot exchange rates have a common stochastic trend and are cointegrated with cointegrating vector $[1, -1]$. This also implies that, since this is true for any k , if we consider the vector of forward rates of tenor 1 to m periods, together with the current spot rate, $[s_t, f_t^1, f_t^2, f_t^3, \dots, f_t^m]'$, then this must be cointegrated with m unique cointegrating vectors, each given by a row of the matrix $[-\iota, I_m]$, where I_m is an m -dimensional identity matrix and ι is an m -dimensional column vector of ones. Finally, by the Granger Representation Theorem (Granger, 1986; Engle and Granger, 1987) this vector of forward and spot rates must possess a VECM representation in which the term structure of forward premia plays the part of the equilibrium errors.

The linear VECM used by Clarida and Taylor (1997) may be written as follows:

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

where $y_t = [s_t, f_t^4, f_t^{13}, f_t^{26}, f_t^{52}]'$, with the superscript denoting the number of weeks corresponding to the maturity of the forward contract; $\Pi = \alpha\beta'$ is the long-run impact matrix whose rank determines the

number of cointegrating vectors linking spot and forward rates (equal to four in this specific VECM); and u_t is a vector of Gaussian error terms (Johansen, 1988, 1991). Clarida and Taylor (1997) exploit this linear VECM representation to show that sufficient information may be extracted from the term structure in order to forecast the spot dollar exchange rate during the recent floating exchange rate regime. Their dynamic out-of-sample forecasts suggest that the linear VECM is superior to a range of alternative forecasts, including a random walk and standard spot-forward regressions.

Clarida *et al.* (2003) then generalize the linear VECM in equation (1) to a multivariate Markov-switching framework and examine the performance of such a model in out-of-sample exchange rate forecasting. This generalized term structure model was inspired by encouraging results previously reported in the literature - cited in the introduction - on the presence of nonlinearities (and particularly by the success of Markov-switching models) in the context of exchange rate modelling. Using weekly data on spot and forward dollar exchange rates for the G5 countries over the period 1979 through 1995, Clarida *et al.* report evidence of the presence of nonlinearities in the term structure, which appeared to be well characterized by a three-regime MS-VECM that allows for shifts in both the intercept and the covariance structure of the error terms. This MS-VECM may be written as follows:

$$\Delta y_t = \nu(z_t) + \Pi y_{t-1} + \sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + \varepsilon_t, \quad (2)$$

where y_t is as defined in equation (1); $\nu(z_t)$ is a 3-dimensional column vector of regime-dependent intercept terms, $\nu(z_t) = [\nu_1(z_t), \nu_2(z_t), \nu_3(z_t)]'$; the Π_i 's are 3×3 matrices of parameters; $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}]'$ is a 3-dimensional vector of error terms, $\varepsilon_t \sim NID(\mathbf{0}, \Sigma_\varepsilon(z_t))$. The regime-generating process is assumed to be an ergodic Markov chain with three states $z_t \in \{1, 2, 3\}$ governed by the transition probabilities $p_{ij} = \Pr(z_{t+1} = j \mid z_t = i)$, and $\sum_{j=1}^3 p_{ij} = 1 \forall i, j \in \{1, 2, 3\}$. This MS-VECM is termed Markov-Switching-Intercept-Heteroskedastic-VECM or MSIH-VECM. In order to reflect the fact that the model has three regimes and one lag of the dependent variable in each equation, the model is termed MSIH(3)-VECM(1).⁷

Clarida *et al.* (2003) use the MSIH(3)-VECM(1) to forecast dynamically out of sample over the period 1996 through to 1998. The results suggest that the MS-VECM forecasts are strongly superior to the random walk forecasts at a range of forecasting horizons up to 52 weeks ahead, using standard forecasting accuracy criteria. Moreover, the MS-VECM also outperforms, in general, a linear VECM for spot and forward rates in out-of-sample forecasting of the spot rate, although the magnitude of

⁷Estimation of an MS-VECM can be carried out using an expectations maximization (EM) algorithm for maximum likelihood (Dempster, Laird and Rubin, 1977; Krolzig, 1997).

the gain, in point forecasting, from using an MS-VECM relative to a linear VECM is rather small in magnitude at short horizons (about 10% on average at the 4-week forecast horizon). Nevertheless, it is possible that traditional measures of forecast accuracy mask somehow the potential superiority of nonlinear models (Satchell and Timmermann, 1995; Clements and Smith, 2001; Granger, 2003).

Overall, this literature suggests that not only the term structure of forward premia contains valuable information about the future spot exchange rate but also that the allowance for nonlinear dynamics in the form of regime-switching behavior enhances somewhat this information to produce a satisfactory forecasting model of the exchange rate. While previous research on term structure forecasting models of the exchange rate has analyzed forecasting performance focusing primarily on accuracy evaluations based on point forecasts, several authors, cited in the introduction, have recently emphasized the importance of evaluating the forecast accuracy of economic models on the basis of density - as opposed to point - forecasting performance. Especially when evaluating nonlinear models, which are capable of producing non-normal forecast densities, it would seem appropriate to consider a model's density forecasting performance. This is indeed the primary objective of the empirical work undertaken in this paper, where we carry out density forecasting tests on the linear VECM and the MS-VECM of the term structure of forward premia as well as on a random walk exchange rate model. We then investigate some of the implications of our density forecasting results for exchange rate risk management.

3 Empirical Analysis I: Modelling

In this section, we describe the data and carry out a preliminary unit root and cointegration analysis of the spot and forward exchange rates data. We also report linearity tests and other tests designed to select the most adequate MS-VECM in our context, which we then estimate.⁸

3.1 Data and Preliminaries⁹

Our data set comprises weekly observations of eight bilateral spot and 4-, 13-, 26- and 52-week forward US dollar exchange rates (*vis-à-vis* the UK sterling, Swiss franc, Japanese yen, Canadian dollar, New Zealand dollar, Swedish krona, Norwegian krona, Danish krona). The sample period spans from

⁸Throughout our discussion of the empirical results, we employ a nominal significance level of 5% percent unless explicitly stated otherwise.

⁹Full details on the preliminary empirical analysis discussed in this sub-section are available from the authors upon request, but are not reported to conserve space.

January 4 1985 to December 31 2003 for all exchange rates.¹⁰ Following previous literature (e.g. Hansen and Hodrick, 1980, p. 852), data are Tuesdays of every week, taken from *Datastream*. From this data set, we constructed the time series of interest, namely the logarithm of the spot exchange rate, s_t and the logarithm of the k -week forward exchange rate, f_t^k for $k = 4, 13, 26, 52$. In our empirical work, we carried out our estimations over the period January 1985-December 1995, reserving the last eight years of data for out-of-sample forecasting tests.

As a preliminary exercise, we tested for unit root behavior of the spot rate and the four forward rate time series examined for each of the exchange rates under investigation by calculating standard augmented Dickey-Fuller (ADF) test statistics.¹¹ Consistent with the large literature on unit root behavior for these time series, our 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 $s_t, f_t^4, f_t^{13}, f_t^{26}$, and f_t^{52} , is the logical next step.

We then employed the Johansen (1988, 1991) maximum likelihood procedure in a vector autoregression for $y_t = [s_t, f_t^4, f_t^{13}, f_t^{26}, f_t^{52}]'$ and an unrestricted constant term. On the basis of the Johansen likelihood ratio test statistics for cointegrating rank, 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 exchange rate examined at conventional nominal test sizes.

When testing the hypothesis that the cointegrating vectors linking spot and forward rates are of the form $[1, -1]$, we rejected this hypothesis but - consistent with Naka and Whitney (1995), Luintel and Paudyal (1998) and Clarida *et al.* (2003) - the departure from the overidentifying restrictions, albeit statistically significant at conventional test sizes, was found to be very small in magnitude. Following Clarida *et al.* (2003), we therefore interpreted the rejection of the unity restrictions on the cointegration space as due to tiny departures from the null hypothesis (due, for example, to tiny data imperfections) which are not economically significant, but which appear as statistically significant given our large sample size. Given the theoretical economic priors in favor of the unity restrictions and the fact that, under covered interest parity, coefficients different from unity would have the implausible implication of a unit root in international interest rate differentials, we carried out our empirical work employing the unity restrictions.¹²

¹⁰The start date was chosen since it was the earliest date for which data for *all* exchange rates examined are available.

¹¹In each case, the number of lags was chosen such that no residual autocorrelation was evident in the auxiliary regressions.

¹²We did, however, carry out a fraction of the empirical analysis discussed below *without* imposing the unity restric-

We next estimated a standard linear VECM, as given in equation (1), using full-information maximum likelihood (FIML) methods, assuming a maximum lag length of three, as suggested by both the Akaike Information Criterion and the Schwartz Information Criterion. Employing the conventional general-to-specific procedure, we obtained fairly parsimonious models for each exchange rate.

3.2 Linearity Tests and Identification of the MS-VECM

We proceeded to investigate the presence of nonlinearities further through the estimation of a very general MS-VECM of the form:

$$\Delta y_t = \nu(z_t) + \Pi(z_t) y_{t-1} + \sum_{d=1}^{p-1} \Gamma_d(z_t) \Delta y_{t-d} + \eta_t, \quad (3)$$

where all parameters are as defined in equation (2), except for the autoregressive parameters Γ_d 's and the long-run matrix Π which are also allowed to be regime-shifting, i.e. $\Gamma_d(z_t)$ and $\Pi(z_t) = \alpha(z_t) \beta'$; η_t is a vector of error terms, $\eta_t \sim NID(\mathbf{0}, \Sigma_\eta(z_t))$. The number of regimes, z_t - for which we consider a maximum of three regimes, i.e. $z_t \in \{1, 2, 3\}$ - is identified using a likelihood ratio test specifically designed for this purpose and described below. The MS-VECM in equation (3) is indeed slightly more general than the MS-VECM used by Clarida *et al.* (2003) in that, except for the long-run cointegrating matrix β' , which is restricted to be consistent with stationary forward premia, every other parameter of the model is now allowed to be regime shifting. In essence, this MS-VECM allows for each of the intercepts, the variance-covariance matrix and the autoregressive structure to be regime dependent. To reflect the fact that model (3) allows for each of the parameters on the autoregressive lags of Δy to be regime switching, in addition to regime-switching intercept vector and covariance matrix, this VECM is termed Markov-Switching-Intercept-Autoregressive-Heteroskedastic-VECM or MSIAH-VECM.

Next we applied the 'bottom-up' procedure designed to detect Markovian shifts in order to select the most adequate characterization of an MS-VECM for Δy_t .¹³ The linearity tests led us, for each exchange rate, to reject the linear VECM against an MS-VECM. These tests provide strong empirical evidence that the linear VECM fails to capture important nonlinearities in the data generating process, tions in the cointegration space and using instead the unrestricted estimates of the cointegrating parameters. The results were virtually identical to those reported below.

¹³Essentially, 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).

since linearity is rejected with marginal significance levels (p -values) of virtually zero - see the last column in Table 1.

For each MS-VECM estimated, we tested the hypothesis of no regime shifts against the alternative hypothesis of regime shifts in each of the vector of intercept terms, the variance-covariance matrix, and the autoregressive structure respectively. The likelihood ratio (LR) tests $LR1$, $LR2$ and $LR3$, constructed as suggested by Krolzig (1997, p. 135-6) and reported in Table 1, suggest in each case massive rejections of the null hypotheses tested, clearly indicating that an MS-VECM that allows for shifts in each of the intercept, the variance-covariance matrix and the autoregressive structure, namely an MSIAH-VECM, is the most appropriate model within its class in the present application.¹⁴

Finally, in order to discriminate between models allowing for two regimes against models governed by three regimes we also calculated an LR test for this null hypothesis. The results produced (see $LR4$, Table 1) very low p -values, suggesting that three regimes are appropriate in all cases. Therefore, in spite of parsimony considerations, we allowed for three regimes in the MSIAH-VECM.

4 Empirical Analysis II: Forecasting and Risk Assessment

In this section we compute tests designed to assess the performance of the random walk model, the linear VECM and the MS-VECM in terms of exchange rate density forecasting. We then explore the implications of the density forecasting results in a risk management exercise.

4.1 Out-of-Sample Density Forecasts

Following the results from the identification procedure reported in Table 1, we estimated the MSIAH-VECM (3) for each of the eight exchange rates examined. The estimation yielded fairly plausible estimates of the coefficients, including the adjustment coefficients in α , which were generally found to be strongly statistically significantly different from zero.¹⁵ ¹⁶

¹⁴Further, in the same spirit of the above LR tests, we executed an LR test in order to select the most parsimonious MS-VECM appropriately representing the dynamic relationship between spot and forward exchange rates. In particular, we tested the null of MSIAH-VECM with one lag against the alternative of MSIAH-VECM with three lags and, for all exchange rates examined, we were not able to reject this null hypothesis at standard significance levels. The tests are not reported to conserve space but are available upon request.

¹⁵We also checked the MSIAH-VECM residuals. 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 the residuals provided no visual evidence of serial correlation in any of the residuals series.

¹⁶For each exchange rate we find that three regimes are appropriate in describing the data. Shifts from one regime to another appeared to be due largely to shifts in the variance of the term structure equilibrium. On the other hand, shifts in the intercept terms and, to a lesser extent, in the autoregressive structure were found to be relatively smaller in magnitude, albeit strongly statistically significant. This seems consistent with the large empirical literature

To investigate the forecasting ability of our linear and Markov-switching VECMs, we attempt to exploit the whole information provided by the out-of-sample predictions of the models in the context of a density forecasting approach. In particular, it is interesting to examine whether the MSIAH-VECM (3), which was chosen on the basis of the identification procedure which rejected the linear VECM, performs satisfactorily in terms of ‘closeness’ of the predicted moments of the forecast density of the model relative to the true moments of exchange rate movements obtained from the data over the forecast period. This question cannot be addressed fully by using conventional methods of point forecast accuracy evaluation, since these methods consider only the first two moments of the distribution of exchange rates.

A large and growing literature has recently focused on evaluating the forecast accuracy of empirical models on the basis of density, as opposed to point, forecasting performance (see, *inter alia*, Diebold *et al.*, 1998; Granger and Pesaran, 1999; Tay and Wallis, 2000; Timmermann, 2000; Sarno and Valente, 2004a,b). Several researchers have proposed methods for evaluating density forecasts. For example, Diebold *et al.* (1998) extend previous work on the probability integral transform and show how it is possible to evaluate a model-based predictive density. Diebold *et al.* propose the calculation of the probability integral transforms of the actual realizations of the variables (i.e. exchange rate changes for each country under investigation) over the forecast period, $\{\Delta s_{t+1}\}_{t=1}^n$ with respect to the models’ forecast densities, denoted by $\{p_t(\Delta s_{t+1})\}_{t=1}^n$:

$$w_t = \int_{-\infty}^{\Delta s_{t+1}} p_t(\varphi) d\varphi \quad t = 1, \dots, n. \quad (4)$$

When the model forecast density corresponds to the true predictive density, then the sequence of $\{w_t\}_{t=1}^n$ is iid $U[0, 1]$. The idea is therefore to evaluate whether the realizations of the data over the forecast period come from the selected forecast density by testing whether the $\{w_t\}$ series departs from the iid uniformity assumption.

Berkowitz (2001) suggests that rather than working with the $\{w_t\}$ series it may be fruitful to take the inverse normal cumulative distribution function (CDF) transform of the series $\{w_t\}$, denoted by $\{x_t\}$. Under the null hypothesis of equality of the model density and the true predictive density, $\{x_t\}$ is distributed as standard normal, and Berkowitz proposes an LR test for zero mean and unit variance, under the maintained hypothesis of iid normality. We rely on the test of Berkowitz (2001)

investigating the time-varying nature of exchange rates risk premia. One tentative interpretation of this MSIAH-VECM is, in fact, in terms of shifts in the mean and variance of foreign exchange returns consistent with deviations from the equilibrium levels implied by conventional macroeconomic fundamentals that may be caused, for example, by ‘peso problems’ or by other kinds of departures from the efficient markets hypothesis (see Engel and Hamilton, 1990; Clarida *et al.*, 2003).

in our empirical work, since Berkowitz shows that working with the inverse normal transform of the series $\{w_t\}$ drastically increases the power of the test relative the version based on uniformity.

While, under general conditions, the linear VECMs forecast densities are easy to calculate analytically (they are in fact multivariate normal distributions with means and variances given by functions of the estimated parameters), the MSIAH-VECM forecast densities can, in general, be obtained analytically only for one-step-ahead forecasts. The MSIAH-VECM forecast densities are mixtures of multivariate normal distributions with weights given by the predicted regime probabilities. In general the MSIAH-VECM forecast densities are non-normal, asymmetric and heteroskedastic. In this paper we focus on the one-step-ahead forecast density of the MSIAH(3)-VECM(1), which is given by:

$$p_{t+1}(\Delta y_{t+1}) = \sum_{j=1}^3 \left\{ \sum_{i=1}^3 p_{ij} \mathbf{P} \right\} p_{t+1}(\Delta y_{t+1} | z_{t+1} = j, \Omega_t), \quad (5)$$

where $p_{ij} = \Pr(z_{t+1} = j | z_t = i)$ are the transition probabilities; \mathbf{P} is the transition matrix conditional on the information set at time t , Ω_t ; and $p_{t+1}(\Delta y_{t+1} | z_{t+1} = j, \Omega_t)$ is the regime-conditional forecast density.

We now turn to the evaluation of the probability integral transforms. The null of iid normality is a joint hypothesis and, in the spirit of Diebold *et al.* (1998), we consider each part of the hypothesis in turn. The iid assumption is tested by calculating the Ljung-Box (1978) test for no serial correlation. In order to take into account the dependence occurring in the higher moments, we consider $(x - \bar{x})^q$ for q up to four. In our forecasting exercise, we compare the density forecasting performance of the linear VECM in equation (1) and the MSIAH-VECM in equation (3) with the standard benchmark in the literature on exchange rate forecasting, namely the random walk model.¹⁷ The p -values from carrying out these tests are reported in Panel a) of Table 2. The results suggest that, for each exchange rate, we are unable to reject the null hypothesis of no serial correlation at conventional significance levels. This finding holds for each of three models examined, namely the random walk model, the linear VECM, and the MSIAH-VECM. We then carry out tests for normality to verify whether the $\{x_t\}$ series displays any statistically significant skewness or excess kurtosis. The p -values from these tests, reported in Panel b) of Table 2, suggest that the null hypotheses of no skewness and no excess kurtosis cannot be rejected for any of the exchange rates and any of the three competing models considered. Taken together, the results in Panel a) and Panel b) of Table 2 indicate that, for each of the three models considered and for each exchange rate, the null hypothesis of iid normality

¹⁷Obviously, the random walk model refers to the level of the exchange rate, implying that in the case of exchange rate changes the model becomes a model of normally distributed exchange rate changes with mean equal to the drift term and variance equal to the residual variance estimated over the sample period available.

of the $\{x_t\}$ series cannot be rejected.

Given these findings, the maintained hypothesis which is required to carry out the LR test proposed by Berkowitz (2001) is validated by the data. Hence, we calculate the LR tests of zero mean and unit variance proposed by Berkowitz (2001), which we report in Panel a) of Table 3. The results are interesting. On the basis of these LR tests, the only model for which we cannot reject the null hypothesis is the MSIAH-VECM. Indeed, the forecast densities of the random walk model and the linear VECM lead to rejection of the null hypothesis for each exchange rate except for the New Zealand dollar. In other words, except for the New Zealand dollar (where each of three competing models performs satisfactorily), the random walk model and the linear VECM produce density forecasts that are statistically significantly different from the actual density of exchange rates data over the forecast period. On the other hand, the MSIAH-VECM generates, for each exchange rate considered, density forecasts that are statistically identical to the true predictive densities.

One of the limitations of the testing procedure employed above is the fact that, while it allows us to measure how well a model's predictive density approximates the true predictive density of the data, it does not allow a formal test of which of the competing models considered performs best in terms of density forecasting performance. Heuristically this can be investigated by inspecting the p -values of the Berkowitz tests for each individual model (i.e. higher p -values presumably reflect a better density forecast) but it is not possible to use the Berkowitz testing procedure to obtain a p -value for the null hypothesis that two models perform equally well in forecasting the true predictive distribution. A solution to this problem has recently been proposed by Corradi and Swanson (2003, 2004), who derive a test statistic for the null hypothesis that two models have equal density forecast accuracy. In some sense, this test evaluates competing forecasting models in terms of density forecasting in the same spirit of the Diebold and Mariano (1995) test for equal point forecast accuracy of competing models. Using a measure that may be seen as the analogue of the mean square error in the contest of density forecasts, Corradi and Swanson (2003, 2004) test a null hypothesis that can be expressed as

$$H_0 : \int_U E \left\{ [F_{1,t+1}(\Delta s_{t+1}|u) - F_0(\Delta s_{t+1}|u)]^2 - [F_{2,t+1}(\Delta s_{t+1}|u) - F_0(\Delta s_{t+1}|u)]^2 \right\} \phi(u) du, \quad (6)$$

where $u \in U \subset \Re$, $\phi(u) \geq 0$ is a possible unbounded set on the real line; $\int_U \phi(u) du = 1$; $F_{q,t+1}(\Delta s_{t+1}|u)$ is the predictive cumulative density function (CDF) implied by model $q = 1, 2$ at time $t + 1$ for a given u ; and $F_0(\Delta s_{t+1}|u)$ is the value of the true CDF for a given u . The test statistic for the null hypothesis (6) takes the form of

$$Z_T = \int_U Z_{T,u} \phi(u) du, \quad (7)$$

where $Z_{T,u} = \frac{1}{\sqrt{n}} \sum_{t=0}^{n-1} \left\{ \left[\widetilde{F}_{1,t+1}(\Delta s_{t+1}|u) - F_0(\Delta s_{t+1}|u) \right]^2 - \left[\widetilde{F}_{2,t+1}(\Delta s_{t+1}|u) - F_0(\Delta s_{t+1}|u) \right]^2 \right\}$; and $\widetilde{F}_{q,t+1}$ is the estimated counterpart of $F_{q,t+1}$ for $q = 1, 2$. The test statistic Z_T is then calculated by averaging $Z_{T,u}$ over $u \in U$. In practice, following Corradi and Swanson (2003, 2004), this can be done by generating a fine grid of u whose values are equally spaced across the range determined by the minimum and maximum value of Δs_{t+1} over the sample period. Hence, if we assume a grid with S points, the tests statistic (7) becomes:

$$Z_T = \frac{1}{S} \sum_{\varrho=1}^S \left\{ \frac{1}{\sqrt{R}} \sum_{t=0}^{n-1} \left\{ \left[\widetilde{F}_{1,t+1}(\Delta s_{t+1}|u_{\varrho}) - F_0(\Delta s_{t+1}|u_{\varrho}) \right]^2 - \left[\widetilde{F}_{2,t+1}(\Delta s_{t+1}|u_{\varrho}) - F_0(\Delta s_{t+1}|u_{\varrho}) \right]^2 \right\} \right\}. \quad (8)$$

Differently from the previous testing procedures based on the probability integral transform, the limiting distribution of the test statistic (8) is a functional of a Gaussian process whose covariance kernel is not a martingale difference sequence in the presence of model misspecification. This implies that its limiting distribution is not nuisance-parameter free and therefore cannot be tabulated. Corradi and Swanson (2003, 2004) show how to obtain critical values for the distribution of the test statistic (8) by bootstrap.^{18 19}

The results from calculating the Corradi-Swanson (2004) test are reported in Panel b) of Table 3. In general, the null hypothesis of equal density forecast accuracy is rejected for most exchange rates at the 5% significance level. In particular, there is clear evidence that the MSIAH-VECM is able to generate predictive densities which are better approximations of the true predictive density than the ones implied by the random walk model and the linear VECM. Indeed, by inspecting the second and third columns in Panel b) of Table 3, the Corradi-Swanson test statistics are all positive and statistically significant. This means that the distance between the true predictive density implied by the data and the predictive densities generated by the random walk model (column 2) or the linear VECM (column 3) is systematically larger than the distance between the true predictive density and the predictive density generated by the MSIAH-VECM. Further, while it is straightforward to establish, on the basis of these results, that the MSIAH-VECM is the best model in terms of density forecasting performance, it is difficult to discriminate between the linear VECM and the random walk

¹⁸Full details of the bootstrap procedure are provided in Corradi and Swanson (2004, Section 3), which we follow in computing these tests below. In our density forecasting exercise we implement the bootstrap designed to calculate critical values using subsampling, as shown in Politis, Romano and Wolf (1999, Ch. 3) and recently employed by Linton, Maasoumi and Whang (2003).

¹⁹Note that another test for equal density forecast accuracy available in the literature is due to Sarno and Valente (2004a). While this test has the advantage of being easily applicable since it has a known limiting distribution, we prefer to use the Corradi-Swanson test since this is more general and works under less stringent assumptions and regularity conditions.

model since there is no clear pattern in the sign of the test statistics.

Summing up, the forecasting results in this section suggest that, in terms of density forecasting performance, the general MSIAH-VECM that allows for multiple regimes performs better than the linear VECM and the random walk model in terms of explaining the out-of-sample behavior of exchange rate movements. Clearly, this finding, obtained by both testing procedures employed - namely the Berkowitz test and the Corradi-Swanson test given in Table 3 - is due to the allowance for multiple regimes in the MSIAH-VECM, which enhances the information embedded in the forward premia and generates density forecasts that are closer to the true predictive densities, providing a better characterization of the uncertainty surrounding the exchange rate forecasts than the other two competing models.

4.2 The Economic Value of Density Forecasts: A Simple Example of VaR Analysis

Under the 1997 Amendment to the Basle Accord, banks may seek approval for the adoption of their own in-house risk models in order to calculate the minimum required capital to cover their market risk. Given that banks are permitted to develop different risk models, it is very important to assess the relative performance of alternative models. In particular, it is interesting to further investigate the practical implications of the density forecasting results reported in the previous sub-section in the context of a simple risk management exercise. Given the predictions of the three competing models examined here, assume that a US risk manager wishes to quantify the one-week-ahead risk associated with holding a position in foreign currency. Assume that the position in question is a naive diversified portfolio comprising a domestic asset and N foreign assets which are identical in all respects except for the currency of denomination, e.g. euro-deposit rates. Each asset delivers yields in local currency, and given that diversification of this portfolio is assumed to be naive, the weight on each asset is $\frac{1}{N+1}$. Define i_t the domestic (US) one-period interest rate and i_t^c the foreign one-period interest rate associated with country c . Then, given a certain level of wealth at time t invested in this portfolio, say W_t , the law of motion of wealth is as follows:

$$W_{t+1} = \left[\frac{1}{N+1} \exp(i_t) + \sum_{c=1}^N \frac{1}{N+1} \exp(i_t^c + \Delta s_{t+1}^c) \right] W_t \quad (9)$$

where Δs_{t+1}^c is the first difference in the dollar log-exchange rate *vis-à-vis* the currency of country c (e.g. Mark, 2001; Elton, Gruber, Brown and Goetzmann, 2003). Given that i_t and i_t^c are known at time t , the only source of risk to be taken in consideration by the risk manager at time t is the future

nominal exchange rate Δs_{t+1}^c for $c = 1, \dots, N$. Normalizing $W_t = 1$ for simplicity, the risk manager can use a specific model of exchange rates at time t to produce one-week-ahead density forecasts of Δs_{t+1}^c , which, in turn, imply one-week-ahead density forecasts for the one-week-ahead wealth W_{t+1} . On the basis of these densities the risk manager calculates, for each of the three competing models - the random walk, the linear VECM and the MSIAH-VECM - the VaR as a confidence interval for losses such that

$$\Pr(W_{t+1} < VaR_{t+1}) = 1 - \phi. \quad (10)$$

In our example the VaR is a 99% confidence level for losses (i.e. $\phi = 0.99$), for all models; and $N = 8$, corresponding to the eight dollar exchange rates studied in this paper. Equation (10) simply states that the probability that one-week-ahead wealth, W_{t+1} is less than the VaR is equal to the significance level $(1 - \phi)$.²⁰

Summary statistics are reported in Panel a) of Table 4. In order to assess the relative size and relative variability of the VaR estimates across the competing models we use the mean relative bias statistic (*MRB*) and root mean squared relative bias statistic (*RMSRB*), suggested by Hendricks (1996). The *MRB* statistic is calculated as:

$$MRB_g = \frac{1}{n} \sum_{h=1}^n \frac{VaR_{g,t+h} - \overline{VaR}_{t+h}}{VaR_{g,t+h}} \quad (11)$$

where $VaR_{g,t}$ is the estimated VaR from the specific model g used at time t , \overline{VaR}_t is the cross-sectional average VaR at time t over the three competing models, and n is the number of out-of-sample observations. This statistic gives a measure of size for each estimated VaR relative to the average of all competing models.

The *RMSRB* statistic is calculated as:

$$RMSRB_g = \sqrt{\frac{1}{n} \sum_{h=1}^n \left(\frac{VaR_{g,t+h} - \overline{VaR}_{t+h}}{VaR_{g,t+h}} \right)^2}. \quad (12)$$

This measure provides us with information about the extent to which the estimated VaR tends to vary around the average VaR at time t . The results from calculating the *MRB* and *RMSRB*, reported in Panel a) of Table 4, suggest that the MSIAH-VECM produces higher VaRs (compared to the average VaR produced by the other two competing models) and it also produces more volatile VaRs (around the average VaR). The higher VaR produced by the MSIAH-VECM is indicative that this model is less conservative, on average, than the other two competing models.

²⁰For this sub-section, we obtained weekly observations of 1-week eurorates for each country examined during the sample period January 1996-December 2003 from *Datastream*.

We also report in Panel a) of Table 4 the average distance (AD) between the realized data and the VaR implied by the random walk model, the linear VECM and the MSIAH-VECM, standardized by the average distance of the random walk model; and, finally, we report the correlation coefficient between the estimated VaR from each model and the realized data, calculated as in Hendricks (1996) and termed $corr_{VaR,W}$. These calculations indicate, on the basis of the AD , that the MSIAH-VECM produces the “closest” VaR to the realized data by some 13% relative to the random walk model and 12% relative to the linear VECM. Also, on the basis of $corr_{VaR,W}$, the MSIAH-VECM generates the VaR most highly correlated with the realized data across the three models considered.²¹

In Panel b) of Table 4, we report some VaR backtests. In particular, we report (as V) the number of times that $W_{t+1} < VaR_{t+1}$, and the implied estimate of the violation rate (i.e. V divided by the number of out-of-sample observations), say VR , for each of the random walk model, the linear VECM and the MSIAH-VECM respectively. We also test of the null hypothesis that the violation rate VR does not exceed the theoretical 1% violation rate considered in this VaR application, calculated as in Kupiec (1995). Finally, we report the Christoffersen and Diebold (2000) test for the sample first-order autocorrelation of a binary variable which is equal to unity if a violation occurs and zero otherwise.²² The results in Panel b) indicate that the random walk model and the linear VECM are too conservative in that they both exhibit one violation, whereas the number of violations under the VaR estimated for the MSIAH-VECM is three, which implies an estimated violation rate of 0.72%. Such estimated violation rate is indeed insignificantly different from the theoretical violation rate of 1%, as confirmed by the Kupiec test statistic. On the other hand, the estimated violation rate for the other two competing models is, in each case, equal to about 0.23%, and is statistically significantly different from the theoretical violation rate of 1%. The CD test also suggests that there is no first-order serial correlation in the VaR violations for each of the three models under examination, indicating that these violations are, for each model, non-systematic.

Summing up, this simple application further illustrates the satisfactory out-of-sample forecasting performance of the MSIAH-VECM relative to the random walk benchmark and to the linear VECM. We find that the MSIAH-VECM provides a violation rate which is statistically insignificantly different from the theoretical violation rate of 1%, whereas the other two competing models provide estimated VaRs which are too conservative and, generally, poor estimates of the risk of the portfolio under

²¹The results show the poor performance of the other two models, which both display statistically insignificant correlation coefficients and, for the random walk model, a negatively signed one.

²²A significant autocorrelation coefficient denotes a persistent series of violations, which in turn implies unsatisfactory performance of a model in estimating the VaR.

examination. This may be seen as evidence that the most general model, the MSIAH-VECM does better than the other two competing models at matching the moments of the predictive distribution of exchange rate changes, generating VaRs that are more in line with the theoretical violation rate of 1%, confirming the findings of the previous sub-section on density forecasting performance and illustrating the practical importance of such results.

5 Conclusion

This article has re-examined the performance of some empirical exchange rate models in terms of out-of-sample forecasting of nominal exchange rates. In particular, inspired by the success of recently developed models of the term structure of forward exchange rates in terms of point forecasting, we have carried out a density forecasting analysis, applied to both linear and regime-switching versions of these term structure models. This exercise was aided by the recent developments of sophisticated econometric techniques which allow us to formally evaluate the performance of time series models in terms of density forecasting.

Our main result, using weekly data for eight US dollar exchange rates during the recent floating exchange rate regime, is that a Markov-switching VECM for spot and forward exchange rates that explicitly takes into account the mounting evidence that the conditional distribution of exchange rates is well characterized by a mixture of normal distributions produces very satisfactory one-week-ahead density forecasts. This model was found to outperform its more parsimonious linear counterpart as well as the standard benchmark in the exchange rate forecasting literature, namely the random walk model.

The implication of our findings were further investigated in the context of a simple application of Value-at-Risk methods. In our application we specifically examined the implications of our exchange rate density forecasts for a risk manager who has to quantify the risk associated with holding a currency portfolio over a one-week horizon. This application further illustrated how the Markov-switching VECM captures satisfactorily the moments of the predictive density of exchange rates, generating VaRs that measure the probability of large losses more accurately than the other two competing models.

Overall, the pecking order implied by our density forecasting results - Markov-switching term structure model, linear term structure model, random walk model²³ - is the same as recorded in

²³Specifically, this is the result suggested by the Berkowitz test statistic, while using the Corradi-Swanson test we were not able to discriminate between the linear VECM and the random walk model. All tests indicated, however,

previous work based on conventional point forecasting criteria. However, when evaluating these models in terms of density forecasting results, the superiority of the Markov-switching VECM relative to the linear VECM becomes much clearer than previously recorded in the literature using point forecast evaluation. Overall, our findings highlight how better density forecasts of exchange rates, of the type recorded in this paper using Markov-switching models of the term structure, can potentially lead to substantial improvements in risk management and, more precisely, to better estimates of downside risk.

that the best performing model is the Markov-switching VECM.

Table 1. ‘Bottom up’ identification procedure

	<i>LR1</i>	<i>LR2</i>	<i>LR3</i>	<i>LR4</i>	Linearity
UK	0	1.67×10^{-16}	1.59×10^{-14}	2.61×10^{-22}	0
Switzerland	0	4.13×10^{-14}	7.36×10^{-18}	2.23×10^{-24}	0
Japan	0	3.02×10^{-12}	2.61×10^{-18}	2.83×10^{-37}	0
Canada	3.98×10^{-243}	5.93×10^{-3}	9.77×10^{-15}	5.85×10^{-46}	2.24×10^{-275}
New Zealand	0	2.08×10^{-8}	2.86×10^{-8}	4.01×10^{-82}	0
Sweden	0	0	4.77×10^{-9}	1.93×10^{-17}	0
Norway	0	4.23×10^{-27}	6.64×10^{-10}	2.64×10^{-32}	0
Denmark	0	1.53×10^{-6}	2.79×10^{-13}	1.09×10^{-50}	0

Notes: *LR1* is a test statistic of the null hypothesis of MSI(3)-VECM(1) versus MSIAH(3)-VECM(1). *LR2* is a test statistic of the null hypothesis of MSH(3)-VECM(1) versus MSIAH(3)-VECM(1). *LR3* is a test statistic of the null hypothesis of MSIH(3)-VECM(1) versus MSIAH(3)-VECM(1). *LR4* is the likelihood ratio test for the null hypothesis that the MSIAH-VECM(1) with 2 regimes is equivalent to the MSIAH-VECM(1) with 3 regimes. These LR tests are constructed as $2(\ln L^* - \ln L)$, where L^* and L represent the unconstrained and the constrained maximum likelihood respectively. These tests are distributed as $\chi^2(r)$ where r is the number of restrictions. ‘Linearity’ is a linearity test for the null hypothesis that the selected MSIAH-VECM(1) is equivalent to a linear Gaussian VECM(1). p -values relative to the *LR4* and linearity tests are calculated as in Ang and Bekaert (1998). For all test statistics only p -values are reported; p -values below 10^{-350} are reported as 0.

Table 2. Out-of-sample performance: tests for iid and normality

Panel a) Tests for iid based upon inverse normal probability transform

	$(x - \bar{x})$			$(x - \bar{x})^2$		
	RW	VECM	MSIAH-VECM	RW	VECM	MSIAH-VECM
UK	0.6418	0.7148	0.5057	0.8343	0.8589	0.4124
Switzerland	0.2083	0.3228	0.2938	0.4420	0.4007	0.2085
Japan	0.1297	0.3857	0.6756	0.8186	0.6971	0.3731
Canada	0.9946	0.9999	0.9996	0.3179	0.2214	0.2993
New Zealand	0.1137	0.2823	0.2758	0.2287	0.0826	0.2010
Sweden	0.5129	0.7465	0.8697	0.7254	0.7170	0.6521
Norway	0.2161	0.4233	0.4113	0.4269	0.4719	0.2673
Denmark	0.2266	0.3691	0.4011	0.4258	0.4131	0.3374

	$(x - \bar{x})^3$			$(x - \bar{x})^4$		
	RW	VECM	MSIAH-VECM	RW	VECM	MSIAH-VECM
UK	0.5712	0.6276	0.5653	0.6671	0.6519	0.6486
Switzerland	0.4365	0.3774	0.2621	0.5964	0.5721	0.4772
Japan	0.6344	0.7032	0.9260	0.7047	0.6953	0.3668
Canada	0.9997	0.9898	0.9927	0.1012	0.4297	0.2364
New Zealand	0.0927	0.0564	0.1713	0.3145	0.1416	0.1795
Sweden	0.6938	0.7875	0.7801	0.7622	0.7148	0.6689
Norway	0.6198	0.6643	0.5724	0.6167	0.6197	0.4444
Denmark	0.4958	0.5058	0.5879	0.5060	0.4802	0.3604

(continued ...)

(... Table 2 continued)

Panel b) Tests for normality based upon inverse normal probability transform

	RW		VECM		MSIAH-VECM	
	Sk	Ku	Sk	Ku	Sk	Ku
UK	0.4021	0.2618	0.3431	0.2669	0.5604	0.3330
Switzerland	0.1823	0.4309	0.1921	0.4089	0.2324	0.3615
Japan	0.0951	0.4780	0.0817	0.4511	0.2718	0.4734
Canada	0.2815	0.5165	0.4250	0.4801	0.4418	0.5155
New Zealand	0.7336	0.4482	0.6906	0.4043	0.6684	0.4273
Sweden	0.4482	0.4069	0.4624	0.4442	0.5021	0.3872
Norway	0.3649	0.2717	0.4698	0.2537	0.5232	0.3212
Denmark	0.4204	0.4567	0.4156	0.4464	0.5525	0.2804

Notes: *Panel a)*: Figures denote p -values for the Ljung and Box (1978) χ^2 test of no first-order serial correlation of the normal inverse probability transform series, x_t , as defined in Section 4.1. *Panel b)*: Figures denote p -values for the tests of the null hypothesis that the skewness (Sk) and excess kurtosis (Ku) of the distribution of the normal inverse probability transform series x_t are equal to 0.

Table 3. Out-of-sample performance: density forecasting tests

Panel a) Berkowitz (2001) LR test

	RW	VECM	MSIAH-VECM
UK	2.60×10^{-2}	3.08×10^{-2}	6.92×10^{-1}
Switzerland	5.47×10^{-6}	7.03×10^{-5}	7.03×10^{-1}
Japan	1.27×10^{-10}	4.25×10^{-8}	3.19×10^{-1}
Canada	2.91×10^{-7}	6.86×10^{-5}	3.11×10^{-1}
New Zealand	6.41×10^{-1}	6.35×10^{-1}	7.04×10^{-1}
Sweden	3.46×10^{-9}	3.19×10^{-8}	3.44×10^{-1}
Norway	4.88×10^{-6}	1.58×10^{-5}	8.12×10^{-1}
Denmark	5.70×10^{-4}	3.42×10^{-4}	3.89×10^{-1}

Panel b) Corradi and Swanson (2004) test

	RW vs VECM	RW vs MSIAH-VECM	VECM vs MSIAH-VECM
UK	-0.0013 [0.0274]	0.0201 [0.0032]	0.0214 [0.0030]
Switzerland	-0.0071 [0.0221]	0.0077 [0.0237]	0.0148 [0.0212]
Japan	0.0183 [0.0124]	0.0102 [0.0194]	0.0284 [0.0122]
Canada	-0.0211 [0.0101]	0.0264 [0.0073]	0.0475 [0.0037]
New Zealand	0.0074 [0.0220]	0.0285 [0.0092]	0.0211 [0.0129]
Sweden	0.0010 [0.0299]	0.0117 [0.0280]	0.0107 [0.0307]
Norway	0.0085 [0.0211]	0.0122 [0.0188]	0.0037 [0.0249]
Denmark	-0.0107 [0.0203]	0.0047 [0.0013]	0.0154 [0.0009]

Notes: *Panel a)*: Figures denote p -values from computing the LR test statistic of Berkowitz (2001), which is distributed as $\chi^2(3)$ under the null hypothesis. *Panel b)*: Figures reported are Corradi and Swanson (2004) test statistics for the null hypothesis that the conditional densities from two competing models are equally accurate relative to the true density. The test statistics are constructed by selecting $u \in [\Delta s_{\min}, \Delta s_{\max}]$ where Δs is the time series of the log-difference of the nominal exchange rate over the entire sample period and 100 equally spaced values of u across this range were used. Values in brackets are p -values calculated by bootstrap based on subsampling as in Politis, Romano and Wolf (1999), where the number of bootstrap replications is set to 100.

Table 4. Value-at-Risk calculations

Panel a) Summary statistics

	RW	VECM	MSIAH-VECM
<i>MRB</i>	-0.0014	-0.0010	0.0024
<i>RMSRB</i>	2.43×10^{-6}	1.45×10^{-6}	6.65×10^{-6}
<i>AD</i>	-	0.0132	0.1347
<i>corr_{VaR,W}</i>	-0.0121	0.0314	0.1553**

Panel b) VaR backtests

	RW	VECM	MSIAH-VECM
<i>V</i>	1	1	3
<i>VR</i>	0.23%*	0.23%*	0.72%
<i>CD</i>	0.998	0.996	0.994

Notes: *Panel a):* *MRB* and *RMSRB* are the mean relative bias and the root mean square relative bias, calculated as in Hendricks (1996) and given in equations (11) and (12) respectively. *AD* is the average distance between the realized data and the VaR implied by each of the random walk (RW) model, linear VECM and MSIAH-VECM, standardized by the average distance of the RW model. *corr_{VaR,W}* is the correlation coefficient between the estimated VaR and the realized data, calculated as in Hendricks (1996). ** indicates statistical significance at the 1% significance level. *Panel b):* *V* denotes the number of times that $W_{t+1} < VaR_{t+1}$, while *VR* denotes the estimated violation rate (i.e. *V* divided by the number of out-of-sample observations), for each of the three competing models. We test the null hypothesis that the violation rate *VR* does not exceed the theoretical violation rate of 1% using the Kupiec (1995) test; the asterisk * indicates statistical significance at the 5% significance level, while the absence of an asterisk next to *VR* implies that the Kupiec test statistic does not reject the null hypothesis of equality of the model's estimated violation rate and 1%. *CD* is the Christoffersen and Diebold (2000) test for the sample first-order autocorrelation of a binary variable which is equal to unity if a violation occurs and zero otherwise. For the *CD* tests only *p*-values are reported.

References

- Ang, A. and Bekaert, G. (1998), "Regime Switches in Interest Rates," *National Bureau of Economic Research Working Paper No. 6508*.
- Berkowitz, J. (2001), "Testing Density Forecasts, with Applications to Risk Management," *Journal of Business and Economic Statistics*, 19, pp. 465-474.
- Boero, G. and Marrocu, E. (2002), "The Performance of Non-Linear Exchange Rate Models: A Forecasting Comparison," *Journal of Forecasting*, 21, pp. 513-542.
- Bos, C.S., Mahieu, R.J., and van Dijk, H.K. (2000), "Daily Exchange Rate Behaviour and Hedging of Currency Risk," *Journal of Applied Econometrics*, 15, pp. 671-696.
- Chappel, D., Padmore, J., Minstry, P. and Ellis, C. (1996), "A Threshold Model for the French Franc/Deutschemark Exchange Rate," *Journal of Forecasting*, 15, pp. 155-164.
- Christoffersen, P.F. and Diebold, F.X. (2000), "How Relevant is Volatility Forecasting for Financial Risk Management?," *Review of Economics and Statistics*, 82, pp. 12-22.
- Clarida, R.H., Sarno, L., Taylor, M.P., and Valente, G. (2003), "The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond," *Journal of International Economics*, 60, pp. 61-83.
- Clarida, R.H. and Taylor, M.P. (1997), "The Term Structure of Forward Exchange Premiums and The Forecastability of Spot Exchange Rates: Correcting the Errors," *Review of Economics and Statistics*, 89, pp. 353-361.
- Clements, M.P. and Smith, J. (2000), "Evaluating the Forecast Densities of Linear and Non-linear Models: Applications to Output Growth and Unemployment," *Journal of Forecasting*, 19, pp. 255-276.
- Corradi, V. and Swanson, N.R. (2003), "A Test for Comparing Multiple Misspecified Conditional Distributions," *Journal of Econometrics*, forthcoming.
- Corradi, V. and Swanson, N.R. (2004), "Predictive Density Accuracy Tests," Queen Mary University and Rutgers University, mimeo.
- Dacco, R. and Satchell, S. (1999), "Why Do Regime-Switching Models Forecast So Badly?," *Journal of Forecasting*, 18, pp. 1-16.
- 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, Series B*, 39, pp. 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, pp. 253-263.
- Elton, E.J., Gruber, M.J., Brown, S.J., and Goetzmann, W.N. (2003), *Modern Portfolio Theory and Investment Analysis*, New York, NY: Wiley and Sons.
- Engel, C. (1994), "Can the Markov Switching Model Forecast Exchange Rates?," *Journal of International Economics*, 36, pp. 151-165.
- Engel, C. (1996), "The Forward Discount Anomaly and the Risk Premium: A Survey of Recent Evidence," *Journal of Empirical Finance*, 3, pp. 123-192.
- Engel, C. and Hakkio, C.S. (1996), "The Distribution of Exchange Rates in the EMS," *International Journal of Finance and Economics*, 1, pp. 55-67.
- Engel, C. and Hamilton, J.D. (1990), "Long Swings in the Dollar: Are They in the Data and Do Markets Know It?," *American Economic Review*, 80, pp. 689-713.
- Engle, R.E. and Granger, C.W.J. (1987), "Co-integration and equilibrium correction Representation, Estimation and Testing," *Econometrica*, 55, pp. 251-276.
- Granger, C.W.J. (1986), "Developments in the Study of Cointegrated Variables," *Oxford Bulletin of Economics and Statistics*, 48, pp. 213-228.
- Granger, C.W.J. (2003), "Time Series Concepts for Conditional Distributions," *Oxford Bulletin of Economics and Statistics*, 65, pp. 689-701.
- 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.
- Hansen, L.P. and Hodrick, R.J. (1980), "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy*, 88, pp. 829-853.
- Hendricks, D. (1996), "Evaluation of Value-at-Risk Model Using Historical Data," *Federal Reserve Bank of New York Economic Policy Review*, April, pp. 29-69.
- Johansen, S. (1988), "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control*, 12, pp. 231-254.
- Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, pp. 1551-1580.
- Jorion, P. (2001), *Value at Risk: The New Benchmark for Managing Financial Risk*, London and New York: McGraw-Hill.
- Krolzig, H.-M. (1997), *Markov-Switching Vector Autoregressions*, New York, NY: Springer.
- Kupiec, P.H. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models," *Journal of Derivatives*, Winter, pp. 73-84.

- LeBaron, B. (1992), "Forecast Improvements Using a Volatility Index," *Journal of Applied Econometrics*, Supplement, 7, pp. 137-149.
- Lewis, K.K. (1995), "Puzzles in International Financial Markets," in Grossman, G.M. and Rogoff, K., (eds.), *Handbook of International Economics*, Vol. 3, pp. 1913-1971.
- Li, F. and Tkacz, G. (2001), "A Consistent Bootstrap Test for Conditional Density Functions with Time-Dependent Data," *Bank of Canada, Working Paper No. 2001-21*.
- Linton, O., Maasoumi, E. and Whang, Y-J. (2003), "Consistent Testing for Stochastic Dominance under General Sampling Schemes," London School of Economics, Southern Methodist University and Korea University, mimeo.
- Ljung, G. and Box, G. (1979), "On a Measure of Lack of Fit in Time Series Models," *Biometrika*, 66, pp. 265-270.
- Luintel, K.B. and Paudyal, K. (1998), "Common Stochastic Trends between Forward and Spot Exchange Rates," *Journal of International Money and Finance*, 17, pp. 279-297.
- Ma, Y. and Kanas, A. (2000a), "Testing for a Nonlinear Relationship among Fundamentals and Exchange Rates in the ERM," *Journal of International Money and Finance*, 19, pp. 135-152.
- Ma, Y. and Kanas, A. (2000b), "Testing for Nonlinear Granger Causality from Fundamentals to Exchange Rates in the ERM," *Journal of International Financial Markets, Institutions and Money*, 10, pp. 69-82.
- Mark, N.C. (1995), "Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability," *American Economic Review*, 85, pp. 201-218.
- Mark, N.C. (2001), *International Macroeconomics and Finance*, Oxford: Blackwell.
- Meese, R.A. and Rogoff, K. (1983a), "Empirical Exchange Rate Model of the Seventies," *Journal of International Economics*, 14, pp. 3-24.
- Meese, R.A. and Rogoff, K. (1983b), "The Out-of-sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification?", in Frenkel, J.A. (ed.), *Exchange Rate and International Economics*, Chicago: University of Chicago Press.
- Naka, A. and Whitney, G. (1995), "The Unbiased Forward Rate Hypothesis Re-examined," *Journal of International Money and Finance*, 14, pp. 857-867.
- Peel, D.A., and Speight, A.E.H. (1994), "Testing for Nonlinear Dependence in Inter-war Exchange Rates," *Weltwirtschaftliches Archiv*, 130, pp. 391-417.
- Politis, D.N., Romano, J.P. and Wolf, M. (1999), *Subsampling*, New York, NY: Springer Verlag.
- Sarno, L. and Valente, G. (2004a), "Comparing the Accuracy of Density Forecasts from Competing Models," *Journal of Forecasting*, forthcoming.

- Sarno, L. and Valente, G. (2004b), "Modeling and Forecasting Stock Returns: Exploiting the Futures Market, Regime Shifts and International Spillovers," *Journal of Applied Econometrics*, forthcoming.
- Satchell, S. and Timmermann, A. (1995), "An Assessment of the Economic Value of Non-Linear Foreign Exchange Rates Forecasts," *Journal of Forecasting*, 14, pp. 477-497.
- Tay, A.S. and Wallis, K.F. (2000), "Density Forecasting: A Survey," *Journal of Forecasting*, 19, pp. 235-254.
- Timmermann, A. (2000), "Density Forecasting in Economics and Finance: Editorial," *Journal of Forecasting*, 19, pp. 231-234.
- Wallis, K. (2003), "Chi-squared Tests of Interval and Density Forecasts, and the Bank of England Fan Charts," *International Journal of Forecasting*, 19, pp. 165-175.