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

Lucio Sarno
University of Warwick
and
Centre for Economic Policy Research (CEPR)

Giorgio Valente
University of Warwick

January 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, 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 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 was 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 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.\footnote{See, for example, the papers published in the special issue of the Journal of International Economics (May 2003) on “Empirical Exchange Rate Models.”}

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 then 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 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 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 distribution of exchange rates. In a decision-theoretical context, the need to consider the density forecast of a variable - as opposed to considering only its conditional mean and variance - seems fairly accepted on the basis of the argument that economic agents may not have loss functions that depend symmetrically on the realizations of future values of potentially non-Gaussian variables (Satchell and Timmerman, 1995). 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. For example, Diebold, Gunther and Tay (1998) build on previous work on the probability integral transform and show how it is possible to evaluate a model-based predictive density. While the test proposed by Diebold et al. (1998) appears to have low power, Berkowitz (2001) has recently extended this work and has suggested a more powerful test of density forecasting. Indeed, this recent line of research has provided researchers with several methods either to measure the closeness of two density functions or to test the hypothesis that the predictive density generated by a particular model corresponds to the true predictive density. 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.

---

2 A 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), Wallis (2003) and Sarno and Valente (2004).

3 By ‘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.

4 An 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 a special emphasis on hedging.
The present paper contributes to the relevant literature in that we re-examine the forecasting 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 the important gap in the literature described above. 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.

We then 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. These models often employ the VaR methodology, where VaR is defined as the expected maximum loss over a target horizon within a given confidence interval (Jorion, 2001). In our simple application we analyze the out-of-sample forecasting performance of our proposed empirical 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

---

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 superior performance of the nonlinear models (see also Satchell and Timmerman, 1995). Other related studies on forecasting exchange rate volatility 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).

5 We 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.

6 More formally, VaR is an interval forecast, typically a one-sided 95 or 99 percent interval of the distribution of expected wealth or returns.
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, 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 as forecasting exchange rate models, 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, which we then estimate and forecast from. Finally 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, and explore the implications of the density forecasting results in a simple 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 $f^k_t$ be, respectively, the spot exchange rate and the $k$-period forward exchange rate, each at time $t$. Under the mild assumptions that (i) each of $s_t$ and $f^k_t$ are well described by unit root processes and that (ii) departures from the risk-neutral efficient markets hypothesis - namely expected foreign exchange excess returns, $f^k_t - E_t(s_{t+k}|\Omega_t)$, defined with respect to a given information set $\Omega_t$ - are stationary, it is straightforward to derive an expression which implies that the forward premium, $f^k_t - s_t$ 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, $[s_t, f^1_t, f^2_t, f^3_t, \ldots, f^m_t]$, then this must be cointegrated with $m$ unique cointegrating vectors, each given by a row of the matrix $[-i, I_m]$, where $I_m$ is an $m$-dimensional identity matrix and $i$ 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:

\[
\Delta y_t = \nu + \sum_{d=1}^{p-1} \Gamma_d \Delta y_{t-d} + \Pi y_{t-1} + u_t
\]

where \( y_t = [s_t, f^{13}_t, f^{26}_t, f^{52}_t]' \), 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 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 (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 modelled well by a multivariate 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:

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

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 \( \Pi_i \)'s are \( 3 \times 3 \) matrices of parameters; \( \varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}]' \) is a 3-dimensional vector of error terms, \( \varepsilon_t \sim NID(0, \Sigma(z_t)) \). 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 the gain from using an MS-VECM relative to a linear VECM is rather small in magnitude at short horizon (about 10% on average at the 4-week forecast horizon).

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.

\(^7\) Estimation of an MSIH-VECM can be carried out using an expectations maximization (EM) algorithm for maximum likelihood (Dempster, Laird and Rubin, 1977).
3 Empirical Analysis

3.1 Data and Preliminaries

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). Due to data availability considerations, the sample period spans from January 4 1985 to December 31 2003 for all exchange rates. 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, $f^k_t$ 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. 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 $s_t, f^4_t, f^{13}_t, f^{26}_t, a n d \ f^{52}_t$, is the logical next step.

We then employed the Johansen (1988, 1991) maximum likelihood procedure in a vector autoregression for $y_t = [s_t, f^4_t, f^{13}_t, f^{26}_t, f^{52}_t]'$ 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,
although statistically significant at conventional test sizes, was actually 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.\footnote{We did, however, carry out a fraction of the empirical analysis discussed below without imposing the unity restrictions in the cointegration space and using instead the unrestricted estimates of the cointegrating parameters. The results were virtually identical to those reported below.}

We next estimated a standard linear VECM of the form (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 with no significant residual serial correlation.

### 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:

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

where all parameters are as defined in equation (2), except for the autoregressive parameters \(\Gamma_d\)'s and the long-run matrix \(\Pi\) which are also allowed to be regime-shifting, i.e. \(\Gamma_d(z_t)\) and \(\Pi(z_t) = \alpha(z_t) \beta';\) and \(\eta_t\) is a vector of error terms, \(\eta_t \sim NID(0, \Sigma_\eta(z_t))\). 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 \(\beta'\), 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.

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 $\Delta y_t$.\(^{12}\) The linearity tests lead us, for each exchange rate, to reject the linear VECM against an MS-VECM. These tests provided strong empirical evidence that the linear VECM fails to capture important nonlinearities in the data generating process, as linearity is rejected with marginal significance levels ($p$-values) of virtually zero - see 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 $LR_1$, $LR_2$ and $LR_3$, 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 a Markov-Switching-Intercept-Autoregressive-Heteroskedastic-VECM or MSIAH-VECM, is the most appropriate model within its class in the present application.\(^{13}\)

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 $LR_4$, 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 our MSIAH-VECM.

### 3.3 Out-of-Sample Density Forecasts

Following the results from the identification procedure reported in Table 1, we estimated the MSIAH-VECM (3) using an EM algorithm for maximum likelihood (Dempster, Laird and Rubin, 1977), for each of the eight exchange rates examined. The estimation yielded fairly plausible estimates of the

\(^{12}\)Essentially, the bottom-up procedure consists of starting with a simple but statistically reliable Markov-switching model by restricting the effects of regime shifts on a limited number of parameters and checking the model against alternatives. In such a procedure, most of the structure contained in the data is not attributed to regime shifts, but explained by observable variables, consistent with the general-to-specific approach to econometric modelling. For a technical discussion of the bottom-up procedure, see Krolzig (1997).

\(^{13}\)Further, 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.
coefficients of the VECMs estimated, including the adjustment coefficients in $\alpha$, which were generally found to be strongly statistically significantly different from zero.\textsuperscript{14} \textsuperscript{15}

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. The answer to this question cannot be given 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 body of 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 \textit{et al.}, 1998; Granger and Pesaran, 1999; Tay and Wallis, 2000; Timmermann, 2000; Sarno and Valente, 2004). Several researchers have proposed methods for evaluating density forecasts. For example, Diebold \textit{et al.} (1998) extend previous work on the probability integral transform and show how it is possible to evaluate a model-based predictive density. In general, they 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 ${p_t (\Delta s_{t+1})}_{t=1}^{n}$:

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

When the model forecast density corresponds to the true predictive density, then the sequence of

\textsuperscript{14}We also checked the MS-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.

\textsuperscript{15}For 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 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 standard efficient markets hypothesis (see Engel and Hamilton, 1990; Clarida \textit{et al.}, 2003).
\( \{w_t\}_{t=1}^{n} \) 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) 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)\(^{16} \), which is given by:

\[
\begin{align*}
  p_{t+1}(\Delta y_{t+1}) &= \sum_{j=1}^{3} \left\{ \sum_{i=1}^{3} p_{ij} P \right\} p_{t+1}(\Delta y_{t+1} \mid z_{t+1} = j, \Omega_t), \\
  \end{align*}
\]

where \( p_{ij} = \Pr(z_{t+1} = j \mid z_t = i) \) are the transition probabilities; \( P \) is the transition matrix conditional on the information set at time \( t, \Omega_t \); and \( p_{t+1}(\Delta y_{t+1} \mid z_{t+1} = j, \Omega_t) \) 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 \((x - \overline{x})^q\) for \( q \) up to four. In our forecasting exercise, we compare the density forecasting performance of the linear VECM (1) and the MSIAH-VECM (3) with the standard benchmark in the literature on

\(^{16}\)This is primarily because the density-forecast testing procedures discussed above are only applicable to one-step-ahead density forecasts.
exchange rate forecasting, namely the random walk model. The resulting p-values are reported in Panel a) of Table 2. The results tell us that, in each case, 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 (1), and the MSIAH-VECM (3). 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 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 c) of Table 2. 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 (3). 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 (3) generates, for each exchange rate considered, density forecasts that are statistically identical to the true predictive densities.

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 is due to the allowance for multiple regimes in our model, 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.18

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.

18These results should be taken with some degree of caution, however, since we are not directly testing one model...
3.4 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 necessary to assess the relative performance of alternative models. Therefore, 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 \(i^c_t\) 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:

\[
W_{t+1} = \left[ \frac{1}{N+1} \exp(i_t) + \sum_{c=1}^{N} \frac{1}{N+1} \exp\left(i^c_t + \Delta s^c_{t+1}\right) \right] W_t
\]

(6)

where \(\Delta s^c_{t+1}\) is the first difference in the dollar log-exchange rate \(\text{vis-à-vis}\) the currency of country \(c\) (e.g. Mark, 2001; Elton, Gruber, Brown and Goetzmann, 2003). Given that \(i_t\) and \(i^c_t\) 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 \(\Delta s^c_{t+1}\) for \(c = 1, \ldots, N\). Normalizing, without loss of generality, \(W_t = 1\), the risk manager can use a specific model of exchange rates at time \(t\) to produce one-week-ahead density forecasts of \(\Delta s^c_{t+1}\), 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

$$\Pr(W_{t+1} < \text{VaR}_{t+1}) = 1 - \phi.$$  \hspace{1cm} (7)

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 (7) simply states that the probability that one-week-ahead wealth, $W_{t+1}$, is less than the VaR is equal to the significance level $(1 - \phi)$.

Summary statistics are reported in Panel a) of Table 3. 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:

$$MRB_g = \frac{1}{n} \sum_{h=1}^{n} \frac{VaR_{g,t+h} - VaR_{t+h}}{VaR_{g,t+h}}$$  \hspace{1cm} (8)

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:

$$RMSRB_g = \sqrt{\frac{1}{n} \sum_{h=1}^{n} \left( \frac{VaR_{g,t+h} - VaR_{t+h}}{VaR_{g,t+h}} \right)^2}.$$  \hspace{1cm} (9)

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 3, 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 produced by the other two competing models). The higher VaR produced by the MSIAH-VECM is indicative that this model is less conservative, on average, than the other two competing models.

We also report in Panel a) of Table 3 the average distance ($AD$) between the realized data and the VaR implied by the random walk model, the linear VECM (1) and the MSIAH-VECM (3), standardized by the average distance of the random walk model; and, finally, we report the correlation

\footnote{For 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.}
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 (1). Also, the MSIAH-VECM generates the VaR most highly correlated with the realized data across the three models considered.\(^{20}\)

In Panel b) of Table 3, 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 (1) and the MSIAH-VECM (3) 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.\(^{21}\) The results in Panel b) indicate that the random walk model and the linear VECM (1) are far too conservative in that they both exhibit one violation, whereas the number of violations under the VaR estimated for the MSIAH-VECM (3) 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 (3) 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 models provide estimated VaRs which are far too conservative and, generally, poor estimates of the risk of the portfolio under examination. This may be seen as evidence that the most general model, the MSIAH-VECM does

\(^{20}\)The results show the poor performance of the other two models, which both display statistically insignificant correlation coefficients and, for the random walk model, an incorrectly signed one.

\(^{21}\)A significant autocorrelation coefficient denotes a persistent series of violations, which in turn implies unsatisfactory performance of a model in estimating the VaR.
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.

4 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 premia 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 vector equilibrium correction model 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 to Value-at-Risk calculations. 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 MS-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. Therefore, the pecking order implied by our density forecasting results - Markov-switching term structure model, linear term structure model, random model - is the same as recorded in 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.
Table 1. ‘Bottom up’ identification procedure

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

Notes: $LR_1$ is a test statistic of the null hypothesis of MSI(3)-VECM(1) versus MSIAH(3)-VECM(1). $LR_2$ is a test statistic of the null hypothesis of MSH(3)-VECM(1) versus MSIAH(3)-VECM(1). $LR_3$ is a test statistic of the null hypothesis of MSIH(3)-VECM(1) versus MSIAH(3)-VECM(1). $LR_4$ is the likelihood ratio test for the null hypothesis that the MSIAH-VECM(1) with 3 regimes is equivalent to the MSIAH-VECM(1) with 2 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 $LR_4$ 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: density forecasting

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

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

<table>
<thead>
<tr>
<th></th>
<th>$(x - \overline{x})^3$</th>
<th>$(x - \overline{x})^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RW</td>
<td>VECM</td>
</tr>
<tr>
<td>UK</td>
<td>0.5712</td>
<td>0.6276</td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.4365</td>
<td>0.3774</td>
</tr>
<tr>
<td>Japan</td>
<td>0.6344</td>
<td>0.7032</td>
</tr>
<tr>
<td>Canada</td>
<td>0.9997</td>
<td>0.9898</td>
</tr>
<tr>
<td>New Zealand</td>
<td>0.0927</td>
<td>0.0564</td>
</tr>
<tr>
<td>Sweden</td>
<td>0.6938</td>
<td>0.7875</td>
</tr>
<tr>
<td>Norway</td>
<td>0.6198</td>
<td>0.6643</td>
</tr>
<tr>
<td>Denmark</td>
<td>0.4958</td>
<td>0.5058</td>
</tr>
</tbody>
</table>

(continued ...)

21
Panel b) Tests for normality based upon inverse normal probability transform

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

Panel c) Berkowitz (2001) LR test

<table>
<thead>
<tr>
<th>Country</th>
<th>RW</th>
<th>VECM</th>
<th>MSIAH-VECM</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>$2.60 \times 10^{-2}$</td>
<td>$3.08 \times 10^{-2}$</td>
<td>$6.92 \times 10^{-1}$</td>
</tr>
<tr>
<td>Switzerland</td>
<td>$5.47 \times 10^{-6}$</td>
<td>$7.03 \times 10^{-5}$</td>
<td>$7.03 \times 10^{-1}$</td>
</tr>
<tr>
<td>Japan</td>
<td>$1.27 \times 10^{-10}$</td>
<td>$4.25 \times 10^{-8}$</td>
<td>$3.19 \times 10^{-1}$</td>
</tr>
<tr>
<td>Canada</td>
<td>$2.91 \times 10^{-7}$</td>
<td>$6.86 \times 10^{-5}$</td>
<td>$3.11 \times 10^{-1}$</td>
</tr>
<tr>
<td>New Zealand</td>
<td>$6.41 \times 10^{-1}$</td>
<td>$6.35 \times 10^{-1}$</td>
<td>$7.04 \times 10^{-1}$</td>
</tr>
<tr>
<td>Sweden</td>
<td>$3.46 \times 10^{-9}$</td>
<td>$3.19 \times 10^{-8}$</td>
<td>$3.44 \times 10^{-1}$</td>
</tr>
<tr>
<td>Norway</td>
<td>$4.88 \times 10^{-6}$</td>
<td>$1.58 \times 10^{-5}$</td>
<td>$8.12 \times 10^{-1}$</td>
</tr>
<tr>
<td>Denmark</td>
<td>$5.70 \times 10^{-4}$</td>
<td>$3.42 \times 10^{-4}$</td>
<td>$3.89 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

Notes: Panel a): Figures denote p-values for the Ljung and Box (1978) $\chi^2$ test of no first-order serial correlation of the normal inverse probability transform series $x$. 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 $x$ are equal to 0. Panel c): Figures denote p-values from computing the LR test statistic of Berkowitz (2001), which is distributed as $\chi^2(3)$ under the null hypothesis.
### Table 3. Value-at-Risk calculations

#### Panel a) Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>RW</th>
<th>VECM</th>
<th>MSIAH-VECM</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRB</td>
<td>-0.0014</td>
<td>-0.0010</td>
<td>0.0024</td>
</tr>
<tr>
<td>RMSRB</td>
<td>2.43×10⁻⁶</td>
<td>1.45×10⁻⁶</td>
<td>6.65×10⁻⁶</td>
</tr>
<tr>
<td>AD</td>
<td>-</td>
<td>0.0132</td>
<td>0.1347</td>
</tr>
<tr>
<td>corr_{VaR,W}</td>
<td>-0.0121</td>
<td>0.0314</td>
<td>0.1553**</td>
</tr>
</tbody>
</table>

#### Panel b) VaR backtests

<table>
<thead>
<tr>
<th></th>
<th>RW</th>
<th>VECM</th>
<th>MSIAH-VECM</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>VR</td>
<td>0.23%*</td>
<td>0.23%*</td>
<td>0.72%</td>
</tr>
<tr>
<td>CD</td>
<td>0.998</td>
<td>0.996</td>
<td>0.994</td>
</tr>
</tbody>
</table>

**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 (8) and (9) 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 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 CD tests only p-values are reported.
References


