Available online at www.sciencedirect.com
ScienceDirect
Journal of Banking & Finance 32 (2008) 729–740
Journal of BANKING & FINANCE
www.elsevier.com/locate/jbf

Do Euro exchange rates follow a martingale? Some out-of-sample evidence

Jian Yang [a,*], Xiaojing Su [b], James W. Kolari [c]

[a] The Business School, University of Colorado Denver, Campus Box 165, P.O. Box 173364, Denver, CO 80217, USA
[b] Department of Economics, Texas A&M University, College Station, TX 77843, USA
[c] Department of Finance, Texas A&M University, College Station, TX 77843, USA

Received 9 March 2006; accepted 21 May 2007
Available online 22 August 2007

Abstract

Traditional autocorrelation and variance ratio tests are based on serial uncorrelatedness rather than martingale difference. As such, they do not capture potential nonlinearity-in-mean, which could lead to misleading inferences in favor of the martingale hypothesis. This paper employs various parametric and nonparametric nonlinear models as well as several model comparison criteria to examine the potential martingale behavior of Euro exchange rates in the context of out-of-sample forecasts. The overall evidence indicates that, while martingale behavior cannot be rejected for Euro exchange rates with major currencies such as the Japanese yen, British pound, and US dollar, there is nonlinear predictability in terms of economic criteria with respect to several smaller currencies.

JEL classification: C2; C5; F3

Keywords: Euro exchange rates; Martingale; Nonlinear models; Technical trading rule; Forecasting evaluation

1. Introduction

Since the breakdown of the Bretton Woods system in the early 1970s, the study of floating exchange rates has become a major research area in international finance. In a seminal paper Meese and Rogoff (1983) show that many structural models of exchange rate determination perform poorly against the benchmark of a naïve martingale hypothesis. Subsequent confirmation of this finding by numerous empirical studies has led to the conventional wisdom that nominal exchange rates (approximately) follow a martingale in the sense that future changes are unpredictable based on past exchange rate changes.[1] This martingale behavior is consistent with weak-form market efficiency in foreign exchange markets, which has been adopted by many researchers in modeling exchange rates.

Recently, several traditional methods for testing martingale behavior, including autocorrelation tests (Box–Pierce–Ljung's portmanteau) and variance ratio tests (Lo and Mackinlay, 1988), have been challenged by Hsieh (1993), Hong (1999) and Hong and Lee (2003), among others. Relevant to our purpose, the variance ratio test has been widely applied to efficiency tests of foreign exchange markets (Hsieh, 1989, 1993; Liu and He, 1991; Ajayi and Karemera, 1996; Anthony and MacDonald, 1999; Yilmaz, 2003; and Belairie-Franch and Opong, 2005). However, as pointed out by Hong and Lee (2003), these traditional methods only investigate serial uncorrelatedness rather than martingale difference. From a nonlinear time series perspective, there is a considerable difference between serial uncorrelatedness (or white noise) and a martingale difference sequence. A nonlinear time series can have zero autocorrelation but a non-zero mean conditional on its past history, which implies predictable nonlinearity-in-mean

* Corresponding author. Tel.: +1 303 556 5852; fax: +1 303 556 5899.
E-mail address: jian.yang@cudenver.edu (J. Yang).

[1] The terms "random walk" and "martingale" have been interchangeably used in the literature. However, strictly speaking, the innovations series is independent and identically distributed for "random walk", while it is a martingale difference sequence for "martingale".

0378-4266/$ - see front matter
doi:10.1016/j.jbankfin.2007.05.009

(Hsieh, 1989, 1993). Hence, when applied to nonlinear time series, these traditional methods could result in misleading inferences in favor of the martingale hypothesis. Theoretically, there is no reason to believe that economic systems must be intrinsically linear.

This paper attempts to empirically test whether Euro exchange rates introduced in 1999 follow a martingale. We seek to contribute to the exchange rate literature in the following ways. First, we apply the recent nonlinear modeling techniques of Hong and Lee (2003) to the problem of testing the martingale behavior of Euro exchange rates. The advent of the Euro marks one of the most dramatic developments in international finance since the collapse of the Bretton Woods system. Although it already has become a key international currency, the martingale behavior of Euro-based exchange rates has been little explored due to its short history.[2] A notable exception is Belairie-Franch and Opong (2005), who examine this issue using a nonparametric variance ratio test based on signs and ranks. Hence, the present study extends their work by applying Hong and Lee's (2003) methods that more thoroughly allow for potential nonlinear-in-mean behavior. Additionally, we use a nonparametric kernel regression technique to further explore potential predictable nonlinearity-in-mean.

Second, we provide out-of-sample rather than in-sample test results. Out-of-sample evidence bears directly on predictability and is particularly important to mitigate the concern of in-sample model overfitting for nonlinear models. Unlike previous studies cited above but similar to Hong and Lee (2003), our out-of-sample tests take into account both statistical and economic significance. Few studies have considered economic criteria as measured by the magnitude of trading returns and particularly the direction of exchange rate changes, which have practical value to investors and other decision-makers.

As noted by Lee and Mathur (1996) and Neely et al. (1997), significant trading rule profits in foreign exchange rates could imply that these markets are inefficient. Previous evidence on trading rule profits has primarily focused on dollar-based exchange rates; hence, we seek to contribute further evidence on trading rule profits by examining Euro-based exchange rates. Furthermore, the predictability of direction of changes as an alternative economic criterion has been little explored in the exchange rate literature, with the noticeable exception of Hong and Lee (2003). As discussed more thoroughly in Hong and Chung (2003), the direction of changes in economic variables may be a reasonable proxy for a utility-based measure of forecasting performance. Particularly from a perspective of decision-making under uncertainty, there exist important circumstances under which the direction-of-change criterion is exactly the right one for maximizing the economic welfare (e.g., profit) of the forecaster (Leitch and Tanner, 1991). Directional predictability in asset returns also has important implications for market timing and the resulting active asset allocation management, as asset managers may focus on the direction of excess returns rather than the magnitude to make buy/sell decisions (e.g., Merton, 1981; Abhyankar et al., 2005). Indeed, in practice most commonly-used technical trading rules in financial markets are based on the prediction of the direction of asset returns.

Third, and last, model comparisons in this study are improved relative to previous exchange rate studies by using White's (2000) novel test to address the concern of data snooping (i.e., spuriously superior predicative ability of some complex models due to chance).

2. Literature review

Numerous studies have examined the hypothesis that exchange rate changes follow a martingale (or the related concept of random walk). As observed by Liu and He (1991), this hypothesis has two important implications: a unit root and uncorrelated increments. Because unit root tests cannot detect some departures from a martingale and the possibility of autocorrelation has interesting implications with respect to alternative exchange rate models (Liu and He, 1991), most recent studies emphasize the analysis of uncorrelated increments using the variance ratio test. Among these studies, Liu and He (1991) find evidence against the random walk hypothesis for five weekly bilateral exchange rates. Similarly, Ajayi and Karemera (1996) find that the random walk hypothesis does not hold well for daily or weekly exchange rate changes in eight Asian-Pacific economies. Anthony and MacDonald (1999) report mixed evidence against the random walk hypothesis for many daily bilateral exchange rates in the European Monetary System. By contrast, based on 27 years of data, Yilmaz (2003) finds that daily exchange rates for seven major currencies follow a martingale most of time, with the exception of periods marked by coordinated central bank interventions. Most of these studies focus on US dollar-based exchange rates. Unlike these studies, Belairie-Franch and Opong (2005) report evidence in favor of the martingale hypothesis for 10 Euro exchange rates (excluding the Canadian dollar and Singapore dollar). Relevant to the present study, studies using the variance ratio test only provide in-sample evidence, which is not as meaningful to practitioners as out-of-sample evidence. Moreover, most

[2] Since its introduction, the Euro has been the second most widely held international reserve currency after the US dollar. More recently, due to the weak dollar and the strong Euro, many countries have held or planned to hold more Euros in the place of the US dollars in their currency reserves. As of December 2006, the Euro has surpassed the US dollar in terms of combined value of cash in circulation. The possibility for the Euro to become the first international reserve currency in the near future is now gaining wide attention, suggesting the importance of better understanding the behavior of Euro-based exchange rates. To some extent, Euro-based exchange rates could behave differently from dollar-based exchange rates, due to the differences in principal policy objectives, institutional constraints, and implementation practices between European Central Bank (or its predecessor the Bundesbank) and the US Federal Reserve (see, e.g., Wang et al., 2007).

Table 1
The summary of models

Name	Models for $E(Y_t|I_{t-1})$ and sign$[E(Y_t|I_{t-1})]$
Benchmark	$E(Y_t|I_{t-1}) = \mu$
AR(d)	$E(Y_t|I_{t-1}) = \beta_0 + \sum_{j=1}^{d}\beta_j Y_{t-j}$
PN(d,m)	$E(Y_t|I_{t-1}) = \alpha_0 + \sum_{j=1}^{d}\sum_{i=1}^{m}\alpha_{ij}Y_{t-j}^{i}$
NN(d,q)	$E(Y_t|I_{t-1}) = \beta_0 + \sum_{j=1}^{d}\beta_j Y_{t-j} + \sum_{i=1}^{q}\delta_i G(\gamma_{0i} + \sum_{j=1}^{d}\gamma_{ji}Y_{t-j}), G(z) = (1+\mathrm{e}^{-z})^{-1}$
FC(d,L)	$E(Y_t|I_{t-1}) = \alpha_0(U_t) + \sum_{j=1}^{d}\alpha_j(U_t)Y_{t-j}$ where $U_t = Y_{t-1} - L^{-1}\sum_{j=1}^{L}Y_{t-j}$
NP(k,m)	$E(Y_t|I_{t-1}) = g(Y_{t-1}, Y_{t-2})$
Combined (1–4)	Combined forecast of AR(2), PN(2,4), NN(2,5) and FC(2,50)
Combined (1–6)	Combined forecast of AR(2), PN(2,4), NN(2,5), FC(2,50), and two NP models
MATTR(L)	sign$[E(Y_t|I_{t-1})]$ = sign(U_t) where $U_t = Y_{t-1} - L^{-1}\sum_{j=1}^{L}Y_{t-j}$
Buy-and-Hold	sign$[E(Y_t|I_{t-1})] = 1$

Notes: The benchmark model is the martingale model. AR(d) is the autoregression model. PN(d,m) is the polynomial regression model. NN(d,q) is the neutral network model. FC is the functional coefficient model of Cai et al. (2000). NP is the nonparametric model estimated by the kernel estimation approach. For NP(k,m) models the smoothing parameter h is used in nonparametric estimation for minimizing k period out-of-sample forecast, and m is the regression length used to find h and make forecasts.

previous studies do not investigate potential nonlinearity-in-mean with respect to exchange rates.

The nonlinearity in the higher moments of exchange rates is well established in the literature. Hsieh (1989, 1993) argues that most nonlinearity of daily exchange rates arises due to time-varying volatility, which may not imply nonlinear predictability in mean (unless there are significant ARCH-in-mean effects). The possible existence of nonlinearity-in-mean has received increasing attention due to the growing interest in forecasting exchange rate changes using sophisticated nonlinear techniques. Artificial neural networks are used in Kuan and Liu (1995) and Gencay (1999), who report promising results regarding exchange rate forecasts. Kuan and Liu (1995) find substantially lower mean squared forecast errors (MSFEs) than the martingale model for several major exchange rates. Gencay (1999) also documents that simple technical trading rules generated from applying neural networks yield a lower mean squared forecast error (MSFE) than the martingale model for five major exchange rates. By contrast, a nonparametric regression technique (i.e., nearest neighbors regression) has been applied by Diebold and Nason (1990), Meese and Rose (1990), Hsieh (1993), and Gencay (1999) with mixed results. For example, while Gencay (1999) reports some positive evidence, Diebold and Nason (1990), Meese and Rose (1990), and Hsieh (1993) find little improvement in out-of-sample forecasts for many major dollar exchange rates. Finally, while no formal forecasting attempt is provided, a state-dependent nonlinear model such as threshold autoregression is considered in Hsieh (1989) and Clements and Smith (2001). They find little evidence for the nonlinearity-in-mean in terms of traditional statistical criteria. In the present study, in order to compare findings to previous studies, some variants of all three popular nonlinear models will be used in the analyses below.[3]

[3] Some studies (Neely et al., 1997; Neely, 2002) also employ the genetic algorithm to forecast dollar-based exchange rates and report positive evidence based on trading rule profits.

3. Econometric methodology

To forecast exchange rate changes we use various models for $E(Y_t|I_{t-1})$, where Y_t represents exchange rate changes (i.e., the first difference of exchange rates in logarithm), and $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \ldots\}$ is the information set available at time $t-1$. As Y_t might not be a martingale difference process, and its conditional mean $E(Y_t|I_{t-1})$ can be time-varying in a complicated form, we will apply various nonlinear parametric and nonparametric models (in addition to the basic linear parametric model).

We use the martingale model as a benchmark for comparison of other models. Table 1 lists the various models that we comparatively examine, including the autoregressive model (AR(d)), autoregressive polynomial model (PN(d,m)), feedforward artificial neural network (NN(d,q)), functional coefficient model (FC(d,L)), nonparametric regression model (NP(k,m)), and some combined models shown there. Additionally, we include two trading rules to compare the direction-of-change forecasts: (1) the moving average trading rule (MATTR) and (2) Buy-and-Hold. Since the estimation of AR(d) and PN(d,m) is relatively simple using the ordinary least squares method, we next briefly discuss how to implement more complicated econometric tools used in this study (i.e., neural network, functional coefficient and nonparametric models).

3.1. The feedforward artificial neural network

Artificial neural networks have proven to be useful in capturing nonlinearity-in-mean for forecasting financial time series. One of the greatest advantages of neural networks over other commonly-used nonlinear time series models is that a class of multilayer neural networks can well approximate a large class of functions. Instead of using a specific nonlinear function between inputs and output for common nonlinear models, the basic structure of neural networks combines many ‘basic’ nonlinear functions via a multilayer structure. Normally there is one intermediate, or hidden, layer between the inputs and output. The intuition is that the explanatory variables simultaneously activate the units in the intermediate layer through some function Ψ and, subsequently, output is produced through some function Φ from the units in the intermediate layer. The following equations summarize this approach:

$$h_{i,t} = \Psi\left(\gamma_{i0} + \sum_{j=1}^{m} \gamma_{ij} X_{j,t}\right) \quad i = 1, \ldots, q, \tag{1}$$

$$Y_t = \Phi\left(\beta_0 + \sum_{i=1}^{q} \beta_i h_{i,t}\right), \tag{2}$$

where $X_{j,t}$ is the input or an independent variable, $h_{i,t}$ is the node or hidden unit in the intermediate or hidden layer, and Y_t is the output or dependent variable. In this study the independent variable $X_{j,t}$ coincides with the lagged dependent variable Y_{t-j}. The functions Ψ and Φ can be arbitrarily chosen and still approximate a large class of functions given sufficiently large numbers of units in the intermediate layer.

In this study we use single layer feedforward neural networks, which is the most basic but perhaps most widely-used neural network in economic and financial applications. In this case the input variables are connected to multiple nodes (or hidden units), and at each node they are weighted (differently) and transformed by the same activation function Ψ. The output of each node is then weighted again by β_i and summed and transformed by a second activation function Φ.

Following Kuan and Liu (1995), Gencay (1999) and Campbell et al. (1997), we chose the logistic function for the function Ψ and the identity function for the function Φ, which is common practice in the literature. The use of the logistic function for the function Ψ is appealing due to its threshold behavior that characterizes many types of (nonlinear) economic responses, and its shape reflects a form of learning behavior. By contrast, the use of the popular Gaussian function for the function Ψ allows for little or no response when the input variables take extreme values, whereas the logistic function would show some response. Also, as discussed in Gencay (1999) and Campbell et al. (1997), the use of the identity function for the function Φ does not necessarily have any major impact on empirical results due to the universal approximation property of neural networks. Coefficients for the NN(d,q) model are estimated using nonlinear least squares via the Newton–Raphson algorithm. The final equation we will estimate is as follows:

$$E(Y_t|I_{t-1}) = \beta_0 + \sum_{j=1}^{d} \beta_j Y_{t-j} + \sum_{i=1}^{q} \delta_i G\left(\gamma_{0i} + \sum_{j=1}^{d} \gamma_{ji} Y_{t-j}\right), \tag{3}$$

where $G(z) = (1 + \mathrm{e}^{-z})^{-1}$ and is a function of Ψ.

3.2. The functional coefficient model

The functional coefficient model introduced by Cai et al. (2000) is a new nonlinear time series model with time-varying and state-dependent coefficients. It includes threshold autoregression models, smooth transition regression, and many other regime switching models as special cases. The basic model can be expressed as follows:

$$E(Y_t|I_{t-1}) = \alpha_0(U_t) + \sum_{j=1}^{d} \alpha_j(U_t) Y_{t-j}, \tag{4}$$

where $\{(Y_t, U_t)'\}$ is a stationary process, and Y_t and U_t are scalar variables. It is important to choose an appropriate smoothing variable U_t. U_t may be chosen as a function of explanatory variable vector Y_{t-j} or as a function of other variables. In our forecasts of exchange rates using past returns, U_t should be a certain combination of the lagged independent variables. There are several specific ways to choose U_t. Here we chose U_t as the difference between the exchange rate (p) at time $t-1$ and the moving average of the most recent L periods of exchange rates at time $t-1$, or

$$U_t = p_{t-1} - L^{-1} \sum_{j=1}^{L} p_{t-j}. \tag{5}$$

Following Gencay (1999), we chose $L = 50$ and 200. We only report the results for $L = 50$, as the results for $L = 200$ are similar. Traders often use U_t as a buy or sell signal based on its sign, which reveals information on changes in direction. Thus, this model might be well suited to forecasting the direction of foreign exchange rate movements.

Following Cai et al. (2000), we estimate the term $\{a_j(U_t)\}$ nonparametrically using a local linear estimator. We approximate $a_j(U_t)$ locally (when U_j is close to u) by $a_j(U_t) = a_j + b_j(U_t - u)$. The local linear estimator at point u is $\hat{a}_j(u) = \hat{a}_j$, and $\{(\hat{a}_j, \hat{b}_j)\}$ are chosen by minimizing the sum of locally weighted squares defined as

$$\sum_{t=1}^{N} [Y_t - a_j - b_j(U_t - u)]^2 K_h(U_t - u), \tag{6}$$

where $K_h(\cdot)$ is the kernel function used as weights for points that are included to estimate $\{(\hat{a}_j, \hat{b}_j)\}$. We use the normal distribution as the kernel function, and h is the smoothing parameter or the bandwidth of the window of the kernel function, which is determined by the modified leave-one-out least square cross-validation method proposed in Cai et al. (2000). Note that the choice of the kernel function has little effect, while h is the most important factor in a nonparametric regression model (and it will converge to zero as n approaches infinity).

3.3. The nonparametric kernel regression model

Because nonlinearities in the conditional mean may be very complicated and cannot be expressed explicitly, it is desirable to use the nonparametric regression to estimate the model without specifying the forms of functions. In general, a nonparametric regression model can be generally expressed as

$$E(Y_t|I_{t-1}) = g(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-j}). \tag{7}$$

As mentioned above with respect to the nonparametric estimator of $a_j(U_t)$ in the functional coefficient model, $g(\cdot)$ can be estimated by local linear regression. At each point $y_t = \{y_{t-1}, y_{t-2}, \ldots, y_{t-j}\}$, we can approximate $g(\cdot)$ locally by a linear function $g(Y) = a + (Y - y)'b$. We can also approximate $g(y)$ locally simply by a constant function $g(Y) = a$ (i.e., the local constant estimator), which is the approach taken here. The local constant estimator is relatively simple to implement and has been widely used in applied research. Compared to other estimators, it has also drawn the most theoretical attention and thus has clear theoretical properties for estimation and inference of nonparametric models. The local constant estimator at point y is given by $g(y) = \hat{a}$, where $\hat{a}$ minimizes the sum of local weighted squares:

$$\sum_{t=1}^{N}[Y_t - a]^2 \prod_{s=1}^{j} K_{hs}(Y_{t-s} - y_{t-s}), \tag{8}$$

where $\prod_{s=1}^{j} K_{hs}(Y_{t-s} - y_{t-s})$ is the product kernel, $K_{hs}(\cdot)$ is the univariate kernel function, and $h = (h_1, \ldots, h_j)$ is chosen by the leave-one-out cross-validation procedure. Note that the asymptotic result for cross-validation is the same for weakly dependent data in the time-series context as for i.i.d. data in the cross-sectional context (see e.g., Li and Racine, 2007). As already noted, the smoothing parameter h is the most important parameter in nonparametric estimation. An inappropriately chosen h will give poor in-sample and out-of-sample prediction. Traditional nonparametric forecasting uses the h that minimizes the in-sample sum of squared errors to forecast the next-period value based on previous in-sample data. However, while this$h$ is optimal for all in-sample data, it may not be the best h for out-of-sample forecasting. Consequently, we use a modified method to select the smoothing parameter.[4]

Our modified approach consists of finding the best h for out-of-sample forecasting and making forecasts based on this h^*. For example, suppose that we have data points of x_1–x_{100}, and that we want to forecast x_{101}. The traditional approach is to find the best h to minimize the 100 data points' in-sample sum of squared errors (based on x_1–x_{100}) and then use this h^* and these data points (i.e., x_1–x_{100}) to forecast x_{101}. We propose the following modified nonparametric forecasting methodology. We use h^* and data points of x_1–x_{80} to forecast x_{81}, data points from x_2 to x_{81} are used to forecast $x_{82}, \ldots$, and data points from x_{20} to x_{99} are used to forecast x_{100}. We find the h^* that minimizes the sum of squared errors of out-of-sample forecast of points x_{81}–x_{100} and then use this h^* and data points x_{21}–x_{100} to make our final forecast of x_{101}. In this procedure we have two parameters to establish: (1) the out-of-sample evaluation length k is set equal to 20 ($\hat{x}_{81}$–$\hat{x}_{100}$) in the example, and (2) the regression length m is set equal to 80 in the example. Hence, we denote the model as NP(k,m), where the parameters (k,m) are crucial to the forecasting performance of this modified nonparametric regression model. Different exchange rate changes series might have different time series properties. For a more (less) volatile time series, a shorter (longer) evaluation length (m) may be better and vise versa. We thus chose two different evaluation lengths. Although different combinations could affect the estimation results for a particular time series, it appears that its impact is not substantial in this study. Below we also report the results for different combinations of parameters (k,m) to serve different needs. As shown in the tables, different combinations NP(25, 500) and NP(125, 400) yield largely similar results across the currencies.

3.4. Forecast combination

It has also been argued that no single forecasting model performs well for all time periods and under all different criteria, as the pattern of exchange rate changes can vary over time and may not follow a simple data generating process. In order to improve the predictability, we follow Yang (2004) and pool forecasts from the AR, PN, NN, FC, and two NP models to forecast the conditional means of exchange rate changes. These six models are denoted as models $1, \ldots, 6$, respectively. The combined model is given by:

$$\widehat{Y}_t^* \equiv \sum_{k=1}^{6} \omega_{kt} \widehat{Y}_{kt}, \tag{9}$$

where the weight ω_{kt} is determined as follows:

$$\omega_{kt} \equiv \frac{\exp\left[-\lambda_t \sum_{s=1}^{t-1}(Y_s - \widehat{Y}_{ks})^2\right]}{\sum_{j=1}^{6} \exp\left[-\lambda_t \sum_{s=1}^{t-1}(Y_s - \widehat{Y}_{ks})^2\right]} \tag{10}$$

with $\lambda_t = 1/(2S_t^2)$, S_t^2 is the sample variance of $\{Y_s\}$, s runs from 1 to $t-1$, and Y_{ks} is the out-of-sample prediction by model k. Intuitively, ω_{ks} gives higher weight to the model k if the prediction for model k is better than other models in previous forecasting exercises as measured by the mean squared forecast error (MSFE) criterion.

Finally, because several forecasting models using the same data are compared, it is important to take into account potential dependence among the models. Otherwise, inferences could be misleading due to data snooping. In this study we use White's (2000) novel test for out-of-sample multiple model comparisons that takes into account data-snooping bias.

4. Data and empirical results

Our data consist of the daily Euro exchange rates for the Australian dollar, British pound, Canadian dollar,

[4] We thank Qi Li for making this suggestion.

Japanese yen, Singapore dollar, Swiss franc and US dollar, which include all the important Euro-based exchange rates that are classified as "independently floating" by the International Monetary Fund. The data span the seven-year period from January 4, 1999 to September 30, 2005, as compiled by the European Central Bank and obtained from Datastream. The use of daily data is consistent with most previous studies. The short history of the Euro prevents us from examining alternative frequencies (e.g., monthly) due to the lack of sufficient observations to conduct out-of-sample tests for most of the nonlinear models.

We use basic rolling techniques to make out-of-sample forecasts. Suppose that there are N observations in total, where $N = R + P$, and P is the number of out-of-sample forecasts. The basic rolling technique uses observations 1 to R (time period 1 to time period R) to forecast period $R + 1$, which we can denote P_1. Then we use observations 2 to $R + 1$ to forecast $R + 2$, or P_2. We eventually obtain$P$ out-of-sample forecasts. Based on these out-of-sample forecasts, we apply four forecasting evaluation criteria: (1) the mean squared forecast error (MSFE), (2) the mean absolute forecast error (MAFE), (3) the mean forecast trading return (MFTR) defined as $\mathrm{MFTR} \equiv P^{-1}\sum_{t=R}^{n-1}\mathrm{sign}(\widehat{Y}_{t+1})Y_{t+1}$, where $\mathrm{sign}(\cdot)$ denotes $\mathrm{sign}(\widehat{Y}_{t+1}) = 1$ if $\widehat{Y}_{t+1} \geqslant 0$ and $\mathrm{sign}(\widehat{Y}_{t+1}) = -1$ if $\widehat{Y}_{t+1} < 0$, and (4) the mean correct forecast direction (MCFD) defined as $\mathrm{MCFD} \equiv P^{-1}\sum_{t=R}^{n-1}1[\mathrm{sign}(\widehat{Y}_{t+1})\mathrm{sign}(Y_{t+1}) > 0]$, with $\mathrm{sign}(\cdot)$ defined as above. Since exchange rate changes can be large at times, forecast errors can be volatile from period to period. For this reason the statistical accuracy of forecasts (as measured by the smallness of MSFE and MSAE) does not necessarily imply economic accuracy in the sense of maximizing investor profits. Since investors are most interested in correctly forecasting large changes in exchange rates, it is quite possible that wrong forecasts of large changes could have smaller MSFEs than correct forecasts of large changes. Hence, it is desirable in this case to compute economic measures of forecast accuracy, such as MFTR and MCFD (Neely et al., 1997; Gencay, 1999; Hong and Lee, 2003). In this regard, MFTR and MCFD are particularly relevant to profit-maximizing investors. Also, closely following Fama (1991), we do not explicitly allow for transaction costs in the evaluation of trading rule performance of various models, as our main interest lies in predictability rather than market efficiency.[5]

Tables 2–8 report the out-of-sample forecast results for the seven currencies under consideration. We should mention that the results in these tables are generally robust with respect to alternative $d = 1$ or $L = 200$ for applicable models. The tables provide the forecast results with R:P (regression length: total out-of-sample forecasts length) = 1:1, while the results with R:P = 2:1are similar. The value of P_{RC}^{1} is the bootstrap p-value for comparing a single model with the martingale model (the benchmark model) using White's (2000) test. The value of P_{RC}^{2} is the bootstrap reality check p-value for comparing k models with the martingale model, where the null hypothesis is that the best of the first k models has no superior predictive power over the martingale model. For example, for the Combined (1–4) model in Table 2, the criteria MSFE has P_{RC}^{1} equal to 0.644 (i.e., the probability that the Combined (1–4) model is not superior to the benchmark) and P_{RC}^{2} equal to 0.973 (i.e., the probability that the best of these models above including the Combined (1–4) model are not superior to the benchmark model). Note that the difference between P_{RC}^{1} and the last P_{RC}^{2} provides some insight into potential data-snooping bias.

Some explanations on each empirical model are in order. The number of lags (2) in the model AR(2) is determined by the minimization of information criteria AIC and SIC. For the model PN(2, 4), the number of lags (2) and the highest order of the exponential term (4) are selected by having relatively more significant parameters and good forecasting performance among preliminary regressions. For model NN(2, 5) the number of lags (2) and the number of nodes in the hidden layer (5) are chosen by relatively good foresting performance based on preliminary estimation. By "preliminary" we mean to use this combination of parameters to run the estimations on several randomly selected countries for certain periods, and this combination yields at least moderate performance compared with other combinations. However, there would not be any combination that always does better than others in all the situations considered. For the model FC(2, 50) 2 is the number of lags, and 50 is the length of a moving average. Due to the use of daily data, following the literature (e.g., Gencay, 1999), we chose a 50-day moving average. For the NP models the second number (400 or 500) is the regression length. Because the nonparametric estimation requires that the number of observations used in the (in-sample) estimation should be large enough, we somewhat arbitrarily chose this number to be 400 and 500, which seem reasonably large. In this regard, we show that the choice of 400 and 500 does not make substantial difference on our empirical results. The first number for the NP models, or 25 and 125, represents the forecasting evaluation length.

Overall, Tables 2–8 indicate that (even without allowance for data-snooping bias) none of the nonlinear models and their combinations can significantly outperform the martingale (benchmark) model in terms of MSFE and MAFE at nominal significance levels (1%, 5%, or 10%). The inability of the models to beat the benchmark

[5] We also follow Lee and Mathur (1996) to allow for transaction costs (albeit in a very simplified manner) by assuming that transaction costs of 0.01% are paid each time a new position is established. As noted in Cheung and Chinn (2001), since the late 1990s electronically-brokered transactions have risen substantially, mostly at the expense of traditional brokers, which should lead to greatly reduced transaction costs. The statistically significant difference between the best forecasting models and the martingale model largely remains after allowance for transaction costs. More importantly, however, note that predictability in three-out-of-four smaller currencies, as measured by the direction of changes, should not be affected by this consideration of transaction costs.

Table 2
Forecast comparison results for the Australian dollar

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0534			0.174			−0.004			0.478		
1	AR(2)	0.0534	.439	.468	0.174	.476	.478	0.002	.286	.250	0.499	.142	.149
2	PN(2,4)	0.0540	.950	.666	0.176	.750	.666	−0.008	.609	.355	0.500	.118	.211
3	NN(2,5)	0.0543	.914	.800	0.176	.912	.811	0.011	.074	.143	0.509	.066	.148
4	FC(2,50)	0.0553	.988	.929	0.178	.999	.925	−0.008	.627	.199	0.470	.653	.197
5	NP(25,500)	0.0555	.960	.966	0.177	.982	.960	0.010	.081	.248	0.520	.025	.109
6	NP(125,400)	0.0537	.671	.972	0.173	.228	.783	0.021	.011	.039	0.552	.000	.001
7	Combined (1–4)	0.0535	.644	.973	0.174	.484	.784	0.006	.133	.042	0.515	.026	.001
8	Combined (1–6)	0.0537	.687	.973	0.174	.419	.786	0.012	.041	.043	0.529	.004	.001
9	MATTR(50)							−0.007	.600	.057	0.485	.374	.003
10	Buy-and-Hold							−0.003	.465	.058	0.460	.805	.003

Notes: (1) The data are daily data from January 4, 1999 to September 30, 2005. R and P observations ($(R,P) = (854,854)$) are used for regression and predication, respectively. (2) P^1_{RC} is the bootstrap p-value for comparing a single model with the martingale model (the benchmark model) using White's (2000) test with 1000 bootstrap replications and a bootstrap smoothing parameter $q = 0.75$. The value of P^2_{RC} is the bootstrap reality check p-value for comparing k models with the martingale model, where the null hypothesis is that the best of the first k models has no superior predictive power over the martingale model. (3) AR, PN, NN, FC, NP are various models under considerations. See the notes to Table 1 also. The smaller MSFE and MAFE or the larger MFTR and MCFD, the better predictive ability of a model.

Table 3
Forecast comparison results for the Canadian dollar

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0625			0.194			−0.012			0.459		
1	AR(2)	0.0629	.879	.865	0.194	.742	.702	−0.004	.221	.221	0.481	.118	.105
2	PN(2,4)	0.0631	.953	.941	0.195	.951	.817	−0.001	.148	.204	0.489	.049	.098
3	NN(2,5)	0.0633	.922	.972	0.195	.910	.923	−0.004	.227	.279	0.481	.135	.140
4	FC(2,50)	0.0657	.992	.989	0.198	.994	.978	−0.008	.358	.110	0.464	.406	.005
5	NP(25,500)	0.0654	.994	.995	0.198	.996	.990	−0.000	.143	.141	0.498	.028	.006
6	NP(125,400)	0.0633	.879	.998	0.195	.784	.992	0.005	.045	.147	0.498	.022	.007
7	Combined (1–4)	0.0628	.770	.997	0.195	.812	.993	−0.006	.315	.152	0.475	.200	.007
8	Combined (1–6)	0.0629	.765	.997	0.195	.811	.993	−0.010	.449	.154	0.474	.188	.007
9	MATTR(50)							−0.009	.434	.197	0.463	.460	.024
10	Buy-and-Hold							−0.003	.080	.197	0.468	.180	.024

Notes: See the notes to Table 2.

Table 4
Forecast comparison results for the Japanese yen

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0566			0.183			0.003			0.499		
1	AR(2)	0.0569	.892	.906	0.184	.934	.946	0.002	.579	.574	0.482	.821	.835
2	PN(2,4)	0.0567	.627	.818	0.183	.497	.743	0.007	.320	.424	0.517	.163	.213
3	NN(2,5)	0.0575	.919	.931	0.184	.647	.900	−0.010	.649	.558	0.499	.475	.294
4	FC(2,50)	0.0577	.924	.976	0.184	.793	.959	−0.002	.725	.666	0.473	.911	.400
5	NP(25,500)	0.0598	.979	.991	0.186	.930	.986	0.009	.274	.659	0.525	.119	.307
6	NP(125,400)	0.0584	.998	.994	0.186	.992	.992	−0.001	.634	.698	0.486	.726	.333
7	Combined (1–4)	0.0569	.810	.994	0.183	.565	.992	0.003	.499	.704	0.508	.297	.342
8	Combined (1–6)	0.0569	.759	.994	0.183	.343	.957	0.008	.257	.707	0.522	.093	.345
9	MATTR(50)							−0.014	.939	.757	0.471	.888	.396
10	Buy-and-Hold							0.007	.167	.761	0.519	.042	.396

Notes: See the notes to Table 2.

martingale model in terms of statistical criteria is consistent with many previous studies (e.g., Clements and Smith, 2001). However, our results might be somewhat weaker than those reported by Hong and Lee (2003), who document some evidence for significantly improving forecasting performance in terms of MSFE and MAFE

Table 5
Forecast comparison results for the British pound

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0276			0.130			0.003			0.486		
1	AR(2)	0.0276	.428	.405	0.130	.829	.814	0.000	.661	.644	0.468	.801	.802
2	PN(2,4)	0.0277	.656	.572	0.130	.632	.836	0.003	.481	.630	0.486	.483	.659
3	NN(2,5)	0.0281	.958	.806	0.131	.988	.970	−0.002	.758	.744	0.488	.433	.729
4	FC(2,50)	0.0280	.887	.945	0.131	.959	.984	0.005	.329	.674	0.484	.502	.796
5	NP(25,500)	0.0287	.997	.969	0.133	.999	.996	0.001	.576	.748	0.472	.710	.835
6	NP(125,400)	0.0281	.986	.974	0.132	.995	.998	−0.008	.919	.768	0.459	.869	.851
7	Combined (1–4)	0.0277	.782	.976	0.130	.848	.998	0.003	.512	.771	0.475	.716	.855
8	Combined (1–6)	0.0277	.751	.976	0.130	.846	.998	0.002	.570	.774	0.470	.790	.859
9	MATTR(50)							−0.006	.868	.821	0.466	.776	.857
10	Buy-and-Hold							0.002	.532	.842	0.481	.626	.879

Notes: See the notes to Table 2.

Table 6
Forecast comparison results for the US dollar

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0705			0.204			−0.001			0.498		
1	AR(2)	0.0709	.871	.889	0.204	.721	.712	−0.004	.648	.654	0.487	.747	.751
2	PN(2,4)	0.0714	.986	.928	0.205	.980	.817	−0.007	.769	.763	0.481	.865	.850
3	NN(2,5)	0.0720	.976	.973	0.205	.950	.922	0.007	.166	.333	0.511	.220	.383
4	FC(2,50)	0.0712	.857	.992	0.205	.831	.967	0.015	.054	.162	0.514	.185	.260
5	NP(25,500)	0.0724	.986	.996	0.208	.996	.993	0.009	.104	.202	0.504	.329	.319
6	NP(125,400)	0.0714	.935	.996	0.205	.904	.996	0.000	.437	.217	0.509	.234	.347
7	Combined (1–4)	0.0713	.987	.996	0.205	.883	.996	−0.005	.627	.221	0.486	.765	.352
8	Combined (1–6)	0.0713	.980	.996	0.205	.920	.996	0.001	.395	.222	0.593	.630	.355
9	MATTR(50)							0.010	.185	.316	0.487	.665	.445
10	Buy-and-Hold							0.011	.075	.317	0.519	.054	.451

Notes: See the notes to Table 2.

Table 7
Forecast comparison results for the Swiss franc

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0081			0.069			−0.005			0.466		
1	AR(2)	0.0081	.201	.199	0.069	.554	.558	0.006	.006	.005	0.500	.049	.050
2	PN(2,4)	0.0082	.753	.383	0.070	.918	.671	0.008	.000	.000	0.499	.057	.071
3	NN(2,5)	0.0085	.998	.617	0.071	.999	.886	−0.001	.178	.001	0.463	.563	.115
4	FC(2,50)	0.0087	.979	.798	0.071	.994	.941	0.004	.013	.001	0.491	.123	.161
5	NP(50,200)	0.0087	.973	.921	0.072	.998	.964	0.000	.088	.001	0.457	.654	.024
6	NP(100,800)	0.0082	.827	.926	0.069	.834	.976	0.002	.060	.003	0.500	.052	.026
7	Combined (1–4)	0.0081	.422	.926	0.069	.672	.976	0.006	.003	.003	0.496	.074	.029
8	Combined (1–6)	0.0081	.410	.926	0.069	.725	.976	0.006	.008	.003	0.493	.096	.030
9	MATTR(50)							−0.003	.382	.009	0.467	.471	.043
10	Buy-and-Hold							0.003	.048	.013	0.508	.046	.047

Notes: See the notes to Table 2.

over the benchmark for (at least) four-out-of-five major currencies (without allowance for data-snooping bias). This difference in results is likely attributable to other differences, including the use of weekly data rather than daily data, US dollar base currency rather than the Euro, and different sample periods. Another difference is that, unlike their study, we test combined models but find that they do not necessarily outperform all single models (AR, PN, NN, FC, and NP).

Interestingly, like Hong and Lee (2003), we find that many models can outperform the martingale model in terms of MFTR and MCFD. The evidence sheds more light on whether exchange rates follow a martingale. As shown in Table 2 for the Australian dollar, the NN and

Table 8
Forecast comparison results for the Singapore dollar

k	Model	MSFE			MAFE			MFTR			MCFD		
		MSFE	P^1_{RC}	P^2_{RC}	MAFE	P^1_{RC}	P^2_{RC}	MFTR	P^1_{RC}	P^2_{RC}	MCFD	P^1_{RC}	P^2_{RC}
0	Benchmark	0.0455			0.162			0.002			0.501		
1	AR(2)	0.0459	.816	.791	0.164	.953	.933	0.005	.331	.329	0.493	.669	.675
2	PN(2,4)	0.0459	.899	.907	0.164	.959	.966	0.006	.282	.386	0.493	.692	.789
3	NN(2,5)	0.0458	.759	.949	0.164	.909	.988	0.016	.022	.070	0.522	.112	.272
4	FC(2,50)	0.0476	.988	.971	0.167	.995	.994	0.002	.501	.138	0.480	.836	.349
5	NP(25,500)	0.0480	.999	.995	0.168	.999	.999	0.006	.282	.173	0.507	.357	.406
6	NP(125,400)	0.0463	.978	.997	0.164	.966	.999	0.004	.409	.183	0.485	.818	.439
7	Combined (1–4)	0.0458	.850	.997	0.163	.925	.999	0.005	.289	.186	0.507	.350	.446
8	Combined (1–6)	0.0458	.875	.997	0.163	.950	.999	0.007	.192	.186	0.514	.186	.447
9	MATTR(50)							0.004	.453	.258	0.483	.784	.538
10	Buy-and-Hold							0.008	.115	.258	0.507	.275	.541

Notes: See the notes to Table 2.

the two NP models (as a single model) outperform the benchmark in terms of both MFTR and MCFD at the 10% significance level, and one NP model (model 6) outperforms the benchmark at the 5% or 1% significance levels. The importance of testing the NP models is further demonstrated by the fact that the forecast combination without NP models, or Combined (1–4), does not improve the forecast performance significantly at any nominal significance level (i.e., p-value = 0.133), while the forecast combination with NP models, or Combined (1–6), does so at the 5% level. Further, the NP model (model 6) yields the highest trading return (i.e., 0.021% per trading day or equivalently 5.27% per year with 251 trading days during the out-of-sample period) and highest frequency to correctly predict the direction of exchange rate changes (i.e., 55%), which even surpasses the combined model with the NP models. As shown under the column P^2_{RC}, allowing potential data snooping bias does not alter this inference, which suggests that this finding is fairly robust. Also, it is noteworthy that the MATTR model does not outperform better than the Buy-and-Hold model.

For the Canadian dollar in Table 3, only one NP model (model 6) yields significantly improved forecast performance in terms of MFTR at the 5% level. Nevertheless, trading returns (i.e., 0.005% per trading day) are clearly smaller compared to the Australian dollar. By contrast, in terms of MCFD, only the two NP models at the 1% level, in addition to the PN model at the 5% level, can better predict the direction of changes than the benchmark. Again, it is the NP models that yield the highest frequency to correctly predict the direction of exchange rate changes (i.e., 50%). These findings are robust to potential data-snooping bias. Neither of the two combined models can improve the forecast performance significantly at any nominal significance level. Lastly, there is no evidence for superiority of the MATTR model.

For the Swiss franc in Table 7, all models except NN outperform the benchmark in terms of MFTR at the 10% significance level. Also, the AR and PN models, as well as one NP model, outperform the benchmark in terms of MCFD at the 10% significance level. Hence, the Swiss franc exchange rate is among the easiest to predict among the seven currencies under consideration. Here we see that the NP models do not perform the best, albeit significantly better than the benchmark. Instead, the PN model is the best model and yields the highest trading return (i.e., 0.008% per trading day) and a relatively high frequency to correctly predict the direction of exchange rate changes (i.e., 50%). In these respects the simple AR model performed almost as well as the PN model. Furthermore, the MATTR model again performs worse than the Buy-and-Hold model.

Lastly, for the Singapore dollar in Table 8, no model can outperform the benchmark in terms of either MFTR or MCFD, with the exception of the NN model with higher trading return than the benchmark at the 5% level (i.e., p-value = 0.022). Nevertheless, this result is generally not very robust to potential data snooping bias. Hence, the predictability might not be as strong as the other smaller currencies.

As for the three most important Euro exchange rates with the Japanese yen, British pound, and US dollar, Tables 4–6 show that no model can outperform the benchmark across different forecasting criteria. An apparent exception might be the FC model for the US dollar in Table 6 (with p-value = 0.05 for P^1_{RC}), but it does not hold after allowance for potential data-snooping bias. Hence, there is significant (daily) predictability in terms of either MFTR or MCFD (or both) for the three (and possibly four) smaller currencies but not for any of the three major currencies. It is noteworthy that, although the modified nonparametric regression models do not improve forecasting performance in terms of MSFE and MAFE, they do the best in terms of MFTR and MCFD for two-out-of-four currencies in which predictability is found to be significant. This result agrees with Clements and Smith (2001), who argue that forecast evaluation based on traditional measures such as MSFE may mask the superiority of the nonlinear models.

As a robustness check, we also conducted the analysis based on the alternative ratio R:P = 2:1. The results (available on request) indicate predictability for two smaller currencies (i.e., Australian dollar and Canadian dollar) but not for two other smaller currencies (i.e., Singaporean dollar and Swiss franc). Again, we find no evidence for predictability for the three large currencies (i.e., US dollar, Japanese yen, British pound). Overall, the results remain qualitatively similar.

We also repeated these analyses using weekly data as an additional robustness check. The results based on $R{:}P = 2{:}1$ (available on request) are discussed here. Due to much smaller sample size using weekly data, this $R{:}P$ ratio seems to be preferred over alternative ratios in balancing in-sample estimation efficiency of several nonparametric models and power of the out-of-sample forecast evaluation tests. As might be expected, significant predictability particularly as measured by an economic criterion (i.e., MCFD) remains for the two smaller currencies (i.e., Australia dollar, Canadian dollar) (and perhaps Singapore dollar) but not for the Swiss franc. Likewise, the results for two-out-of-three major currencies (British pound and Japanese yen) are qualitatively unchanged. Surprisingly, we also find new evidence for predictability of the US dollar as measured by MCFD, which is not detected earlier at daily frequency. Some differences between daily and weekly results suggest that different (fundamental or non-fundamental) factors (to be discussed below) could play important roles for exchange rate dynamics at different frequencies.[6]

It should also be noted that exchange rate forecasts are typically based on two types of models: technical trading models or fundamentals-based models. Here we employ technical trading models, which use the past history of exchange rates to predict future movements (e.g., Cheung and Chinn, 2001).[7] As a theoretically appealing alternative, fundamentals-based models are based on the notion that macroeconomic fundamentals (including real economic activity and monetary fundamentals) can explain and predict exchange rate movements.[8] Unfortunately, previous research studies have not been able to establish that exchange rate forecasts obtained using fundamentals are better than forecasts from a naïve martingale model, particularly at the short- and medium-term horizons (Meese and Rogoff, 1983; Andersen et al., 2003; Abhyankar et al., 2005; Ehrmann and Fratzscher, 2005).

By contrast, a few recent studies (e.g., Andersen et al., 2003) have documented some success in forecasting several US dollar-based exchange rates with news about macroeconomic fundamentals using high frequency data. Nevertheless, intraday effects are generally very short-lived and often disappear within minutes, casting some doubt on their ability to explain overall exchange rate movements at daily or lower frequencies. Ehrmann and Fratzscher (2005) further show that macroeconomic news in the US and the Euro area (and Germany) can have somewhat longer-lived effects and explain much of the monthly directional changes but not magnitude of the US dollar/Euro (US dollar/DM) exchange rate. Their evidence is consistent with the use of direction of changes in the present study. In this regard, both Andersen et al. (2003) and Ehrmann and Fratzscher (2005) find the presence of asymmetries in exchange rate responses to fundamentals, which favors the use of nonlinear models applied here. In sum, our finding of predictability based on technical trading models could be driven by macroeconomic fundamentals, which is not surprising if there exist longer-lived effects of macroeconomic news beyond the daily horizon (especially for smaller currencies).

There are several other explanations for predictability in exchange rates also. For example, Liu and He (1991) argue that the autocorrelations of several dollar-based exchange rates may be consistent with an exchange rate overshooting or undershooting phenomenon. Another explanation for (short-horizon) predictability in (daily or weekly) exchange rates is central bank intervention. However, the empirical evidence is mixed in the regard. Both Yilmaz (2003) and Szakmary and Mathur (1997) document that several major US dollar-based exchange rates deviate from the martingale property and produce profitable trading returns during times of coordinated central bank interventions. On the other hand, Neely (2002) provides evidence that central bank intervention does not generate technical trading profits in several major US dollar-based exchange rates. Lee and Mathur (1996) also show that central bank intervention does not produce different results on trading rule profits for European currency cross-rates.

Despite mixed evidence on other exchange rates, it remains possible that the detected predictability in some Euro-based exchange rates (particularly those on smaller currencies) in this study may be related to central bank intervention. It is important to note that some central banks may not be involved in all policy interventions, and central bank market interventions cannot be expected to affect all bilateral exchange rates (Yilmaz, 2003). Further, although the huge daily turnover in the world's foreign exchange markets would tend to dwarf any attempts by central banks to alter trends of market forces (thus

[6] For example, the effect of central bank intervention or the effect of macroeconomic news on exchange rates could last for more than a day but diminish within a week, which would explain daily predictability and weekly unpredictability (see, e.g., Yilmaz, 2003; Ehrmann and Fratzscher, 2005).

[7] As discussed in Evans and Lyons (2005), the existing literature typically focuses on longer-horizon forecasting. Micro-based forecasting focuses on horizons from one day to one month, with overlap between micro and macro analysis of exchange rates at the one-month horizon. Hence, our study encompasses micro-based forecasting to some extent.

[8] The exchange rate predictability is important in itself, regardless of its implication for market (in)efficiency. The predictability as a result of market inefficiency clearly shows the promise of arbitrage or speculation opportunities in foreign exchange markets. Even if such predictability results from predictable time variations in fundamental risk factors and does not imply any market inefficiency, it still carries significant economic value in asset allocation decisions. It can affect the choice between domestic and foreign assets and optimal weights to the foreign assets (see, e.g., Abhyankar et al., 2005).

rendering trading rules unprofitable) (Lee and Mathur, 1996), it may be argued that such an effect would be less pronounced for smaller currencies due to their much lower turnover, which could explain their more significant predictability and trading profits. In line with this conjecture, Brandner et al. (2006) conclude that, during the period from 1993 to 1998, European Monetary System interventions had some (though limited) effects on the conditional means and variances of exchange rates.

5. Conclusions

This study attempts to investigate the martingale behavior of seven major Euro exchange rates using out-of-sample forecasts. A number of parametric and nonparametric nonlinear models are employed to capture potential nonlinearity-in-mean in foreign exchange rates. Traditional statistical criteria fail to reject the martingale hypothesis for all seven Euro exchange rates. However, economic criteria suggest predictability in the direction of daily exchange rate changes as well as trading returns for the three (or possibly four) smaller currencies (i.e., Australian dollar, Canadian dollar, Swiss franc, and perhaps Singapore dollar) but not for the three major currencies (i.e., Japanese yen, British pound, and US dollar). Hence, we find little evidence against the martingale behavior for the three major Euro exchange rates but supportive evidence for predictability of the smaller currencies.[9] By contrast, based on a variance ratio test, Belairie-Franch and Opong (2005) reject the random walk hypothesis for Euro exchange rates with the Canadian dollar and Singapore dollar but not with respect to eight other currencies, including several smaller currencies.

We contribute further evidence relevant to earlier works (Liu and He, 1991; Ajayi and Karemera, 1996; Anthony and MacDonald, 1999) that document some (in-sample) predictability of US dollar exchange rates. More specifically, we provide new evidence based on Euro exchange rates, as any inference based on US dollar exchange rates in previous studies cannot be presumably generalizable to exchange rates based on another major currency. Additionally, the evidence in this study is provided from a different (and at least equally interesting) perspective from previous studies. Essentially, we implement a model selection approach (e.g., Swanson and White, 1997), rather than the more traditional hypothesis testing approach using the variance ratio test. The model selection approach has two major advantages in the present context. First, it allows us to focus directly on the issue at hand – namely, out-of-sample forecasting performance which bears more directly on predictability. Second, and perhaps more importantly, unlike the traditional hypothesis testing approach, it does not require the specification of a correct model for its valid application. By contrast, earlier empirical findings based on variance ratio tests are very sensitive to potential model misspecification (e.g., heteroscedasticity). Lastly, our evidence takes into account not only statistical but economic significance, which may be more revealing to investors but is not reported in these earlier works.

Finally, with the availability of longer data series in the future, it would be interesting to explore whether Euro-based exchange rates at other frequencies (e.g., monthly) exhibit nonlinear predictability. The nonlinear-in-mean behavior (or lack of it) could be different for exchange rates at different (i.e., high versus low) frequencies. Also, another possible venue for future research is to apply nonlinear models and explore asset return predictability in other financial markets.

[9] As pointed out by a referee, since martingale means the existence of neither linear nor nonlinear dependence, we have to test all possible nonlinear dependence to check the martingale property of exchange rate changes, which is practically impossible. Hence, we can only reasonably conclude to have evidence against martingale (given the models examined) but not the acceptance of martingale.

Acknowledgements

We would like to thank Yongmiao Hong and Tae-Hwy Lee for sharing their computer program, Qi Li and particularly two anonymous referees for numerous helpful comments and suggestions. Kolari has received support from JPMorgan Chase and the Center for International Business Studies at Texas A&M University. This paper was reviewed and accepted during the tenure of the past Editorial Board in which Professor Giorgio Szego acted as Managing Editor of the Journal of Banking and Finance.

References

Abhyankar, A., Sarno, L., Valente, G., 2005. Exchange rates and fundamentals: Evidence on the economic value of predictability. Journal of International Economics 66, 325–348.

Ajayi, R.A., Karemera, D., 1996. A variance ratio test of random walks in exchange rates: Evidence from Pacific Basin economies. Pacific Basin Finance Journal 4, 77–91.

Andersen, T., Bollerslev, T., Diebold, F.X., Vega, C., 2003. Micro effects of macro announcements: Real-time price discovery in foreign exchange. American Economic Review 93, 38–62.

Anthony, M., MacDonald, R., 1999. The width of the band and exchange rate mean-reversion: Some further ERM-based results. Journal of International Money and Finance 18, 411–428.

Belairie-Franch, J., Opong, K.K., 2005. Some evidence of random walk behavior of euro exchange rates using ranks and signs. Journal of Banking and Finance 29, 1631–1643.

Brandner, P., Grech, H., Stix, H., 2006. The effectiveness of central bank intervention in the EMS: The post 1993 experience. Journal of International Money and Finance 25, 580–597.

Cai, Z., Fan, J., Yao, Q., 2000. Functional-coefficient regression models for nonlinear time series. Journal of American Statistical Association 95, 941–956.

Campbell, J., Lo, A., MacKinlay, C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton, NJ.

Cheung, Y.W., Chinn, M.D., 2001. Currency traders and exchange rate dynamics: A survey of the US market. Journal of International Money and Finance 20, 439–471.

Clements, M.P., Smith, J., 2001. Evaluating forecasts from SETAR models of exchange rates. Journal of International Money and Finance 20, 133–148.

Diebold, F.X., Nason, J.A., 1990. Nonparametric exchange rate prediction? Journal of International Economics 28, 315–332.

Ehrmann, M., Fratzscher, M., 2005. Exchange rates and fundamentals: New evidence from real-time data. Journal of International Money and Finance 24, 317–341.

Evans, M.D., Lyons, R.K., 2005. Meese-Rogoff Redux: Micro-based exchange rate forecasting. American Economic Review 95, 405–414.

Fama, E.F., 1991. Efficient capital markets: II. Journal of Finance 46, 1575–1617.

Gencay, R., 1999. Linear, nonlinear and essential foreign exchange rate prediction with simple trading rules. Journal of International Economics 47, 91–107.

Hong, Y.M., 1999. Hypothesis testing in time series via the empirical characteristic function: A generalized spectral density approach. Journal of American Statistical Association 84, 1201–1220.

Hong, Y.M., Chung, J. 2003. Are the directions of stock price changes predictable? Statistical theory and evidence, Working Paper, Cornell University.

Hong, Y.M., Lee, T.H., 2003. Inference on predictability of foreign exchange rates via generalized spectrum and nonlinear time series models. Review of Economics and Statistics 85, 1048–1062.

Hsieh, D.A., 1989. Testing for nonlinear dependence in daily foreign exchange rates. Journal of Business 62, 339–368.

Hsieh, D.A., 1993. Implications of nonlinear dynamics for financial risk management. Journal of Financial and Quantitative Analysis 28, 41–64.

Kuan, C.M., Liu, T., 1995. Forecasting exchange rates using feedforward and recurrent neural networks. Journal of Applied Econometrics 10, 347–364.

Lee, C.I., Mathur, I., 1996. Trading rule profits in European currency spot cross-rates. Journal of Banking and Finance 20, 949–962.

Leitch, G., Tanner, E., 1991. Economic forecast evaluation: Profits versus conventional error measures. American Economic Review 81, 580–590.

Li, Q., Racine, J., 2007. Nonparametric Econometrics: Theory and Practice. Princeton University Press, Princeton, NJ.

Liu, C.Y., He, J., 1991. A variance ratio test of random walks in foreign exchange rates. Journal of Finance 46, 773–785.

Lo, A.W., Mackinlay, A.C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Meese, R.A., Rogoff, K., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of International Economics 14, 3–24.

Meese, R.A., Rose, A.K., 1990. Non-linear, non-parametric, non-essential exchange rate estimation. American Economic Review 80, 192–196.

Merton, R., 1981. On market timing and investment performance: An equilibrium theory of value for market forecasts. Journal of Business 54, 363–406.

Neely, C., 2002. The temporal pattern of trading rule returns and exchange rate intervention: Intervention does not generate technical trading rule profits. Journal of International Economics 58, 211–232.

Neely, C., Weller, P., Dittmar, R., 1997. Is technical analysis in the foreign exchange market profitable? A genetic programming approach. Journal of Financial and Quantitative Analysis 32, 405–426.

Swanson, N.R., White, H., 1997. A model selection approach to real time macroeconomic forecasting using linear models and artificial neural networks. Review of Economics and Statistics 79, 540–550.

Szakmary, A.C., Mathur, I., 1997. Central bank intervention and trading rule profits in foreign exchange markets. Journal of International Money and Finance 16, 513–535.

Wang, Z., Yang, J., Li, Q., 2007. Interest rate linkages in the Eurocurrency market: Contemporaneous and out-of-sample Granger causality tests. Journal of International Money and Finance 26, 86–103.

White, H., 2000. A reality check for data snooping. Econometrica 68, 1097–1126.

Yang, Y., 2004. Combining forecasting procedures: Some theoretical results. Econometric Theory 20, 176–190.

Yilmaz, K., 2003. Martingale property of exchange rates and central bank interventions. Journal of Business and Economics Statistics 21, 383–395.