



Comparing Linear and Non-linear Benchmarks of Exchange Rate Forecasting

SJ Retief, M Pretorius and I Botha

ERSA working paper 494

February 2015

Economic Research Southern Africa (ERSA) is a research programme funded by the National Treasury of South Africa.

The views expressed are those of the author(s) and do not necessarily represent those of the funder, ERSA or the author's affiliated institution(s). ERSA shall not be liable to any person for inaccurate information or opinions contained herein.

Comparing Linear and Non-linear Benchmarks of Exchange Rate Forecasting

SJ Retief, M Pretorius and I Botha*

February 4, 2015

Abstract

Throughout the past 3 decades, the random walk model served as exchange rate forecasting benchmark to verify that a model is able to outperform a random process. However, its application as forecasting benchmark is contradictory. Rather than serving as a benchmark that explains exchange rate behaviour, it serves as a benchmark of what we do not understand in exchange rate forecasting – the random component. In order to accommodate for the observed mean reverting and non-linear patterns in exchange rate information, this study considers various univariate models to serve as linear or non-linear benchmarks of exchange rate forecasting. The results of forecasting performance indicate that the random walk model is an insufficient benchmark to explain exchange rate movements for non-static models. As linear alternative, an autoregressive model performed best to explain the mean reverting patterns in exchange rate information for quarterly, monthly and weekly forecasts of the exchange rate. As non-linear alternative, a Kernel regression was best able to explain volatile exchange rate movements associated with daily forecasts of the exchange rate.

Key words: random walk, autoregressive (AR), moving average (MA), Kernel regression

1 Introduction

Initial models to forecast exchange rates were based on the behaviour of monetary macroeconomic fundamentals. These monetary class models assumed that the exchange rate is a function of relative measures of money supply, short-term interest rates, income and the price level. However, the validity of monetary class models soon came into question when Meese and Rogoff (1983) showed that monetary fundamental models performed no better than a simple random walk model to forecast exchange rates over multiple forecast periods. This was an important finding, as it provided empirical evidence to show that macroeconomic information is not a prerequisite for exchange rate forecasting. The

*University of Johannesburg, Johannesburg, South Africa

exchange rate itself can be used as input to forecast exchange rates. More importantly, the contributions of Meese and Rogoff (1983) led to the identification of the random walk model as benchmark to test the validity of other exchange rate models which lacked the theoretical support enjoyed by monetary models of exchange rate determination.

Given the findings of Meese and Rogoff (1983), the behaviour of exchange rates can be linked to the findings of Bachelier (1900), who proposed that financial assets follow a random walk which makes them unpredictable.

In contrast, more recent research on the performance of fundamental models by Bissoondeeal, Binner and Elger (2009), Alvarez, Atkeson and Kehoe (2007) and Wu and Chen (2001) rejects the notion that exchange rates follow a random walk. These studies provide empirical evidence to show that fundamental models outperform the established random walk alternative. Ironically, Meese and Rogoff emphasize that exchange rates do not follow an exact random walk (Meese and Rogoff, 1983:14). However, if it is known that exchange rates do not follow a random process explicitly, alternative exchange rate benchmark models should be considered. Yet, judging by the universal acceptance of the random walk model as benchmark model for exchange rate forecasting, this finding has not been explored to its full potential.

To accommodate for the recognised mean reverting and non-linear patterns observed in exchange rate information, this paper considers both linear and non-linear models to serve as additional benchmarks of exchange rate forecasting. The forecasting performance of benchmark models is compared against the forecasting performance of the random walk model for a multitude of currency and exchange rate frequency combinations. Viable benchmark alternatives are identified for models that are able to consistently outperform forecasts of the random walk model and other competing benchmark alternatives.

The paper is organised as follows. Section 2 provides a review of the existing exchange rate literature to support random, linear or non-linear patterns in exchange rate information. Section 3 considers additional exchange rate benchmark models and section 4 the information used as data input for these models. The lag selection and statistical tests to establish a good fit of data is discussed in section 5, followed by a review of the methodology to assess forecasting performance in section 6. Section 7 identifies exchange rate benchmarks which are able to consistently outperform the random walk benchmark and other competing alternatives for various currency and frequency combinations and section 8 concludes.

2 Literature review

The findings by Meese and Rogoff (1983) of consistent outperformance of the random walk (RW) model against both fundamental and non-fundamental models established the RW as benchmark model. As noted by Engel (1994), Isard (1995), Sarno (2003) and later Evans and Lyons (2005) it became an accepted norm in subsequent studies to test the forecasting results of models against the

performance of the RW model. However, not all researchers agree with the conclusions of Meese and Rogoff (1983). Schinasi and Swamy (1987) provide compelling evidence that fundamental models can outperform the RW model. They apply the monetary models used by Meese and Rogoff (1983) to show that when the assumption of fixed coefficients is relaxed, fundamental models consistently outperform the RW model for both multi-step-ahead and one-step-ahead forecasts of the exchange rate.

To investigate the effect of real disturbances, Meese and Rogoff (1988) used real adjusted versions of the monetary models applied in their 1983 study (referred to above) to determine the relationship between real exchange rates and real interest rates in a monetary framework. In this study they reaffirm the superior performance of the RW model finding only occasional instances of outperformance by fundamental models. Sarantis and Stewart (1995) however, argue that Meese and Rogoff (1983, 1988) only used variants of monetary exchange rate models for forecast comparisons against the RW model. They support the findings of Schinasi and Swamy (1987) who argue that the dynamic specifications of the structural models applied by Meese & Rogoff (1983) are weak as they are only presented in static form.

From a non-fundamental perspective, Bleaney and Mizen (1996) also reject the hypothesis that exchange rates follow a random walk. They hypothesize that the exchange rate moves within a certain range with limited ability to wander without bound. A possible explanation for the mean reverting behaviour in data comes from De Bondt and Thaler (1985). They hypothesize that extreme price movements in traded financial assets are followed by subsequent movements in the opposite direction. In addition, they support an overreaction hypothesis which suggests that the more extreme the initial price movement, the greater will be the offsetting reaction.

The debate on whether or not a data set is random or mean reverting is perhaps best summarised by Poterba and Summers (1988). They argue that although the random walk hypothesis cannot always be rejected for an individual data set forming part of a greater data set, multiple data sets together strengthen the validity of overall mean reversion patterns in data. A similar argument holds to strengthen the validity non-linear patterns in data.

Limited research is available to propose alternative benchmarks of exchange rate forecasting however, Engel and Hamilton (1990) make a valuable contribution in this regard by addressing a key principle that leads us to question the validity of the RW without drift model as forecasting benchmark. They suggest that when the drift term is significantly different from zero for in-sample forecasts, the RW with drift model is more reasonable benchmark of exchange rate forecasting. Engel (1994) emphasises this point by arguing that if a RW with drift model outperforms a RW without drift model for in-sample forecasts, but fails to continue superior performance for out-of-sample forecasts, one must avoid the misleading notion to choose a benchmark based on out-of-sample performance.

The findings of Engel and Hamilton (1990) emphasise the fact that the drift component may change as forecasts move from the in-sample period to the out-

of-sample period. In order to adequately track changes in models with a drift component (typically linear models), out-of-sample forecasts should ideally be based on the latest available information about the exchange rate. Failing to comply with this requirement provides the RW without drift model with an unfair advantage as forecasts of the RW without drift model is based on the latest direction and magnitude of change in the exchange rate. The challenge for exchange rate forecasters is to find models that accommodate the patterns in exchange rate information for various frequencies of exchange rate information as well as their associated volatility.

A significant change in the forecasting performance of models due to a change in data frequency is reported by Altavilla and De Grauwe (2010). They show that the forecasting performance of both linear and non-linear models change considerably across currencies, forecast horizons and sub-samples. As motivation for the fluctuation in model forecasting performance, they support the hypothesis that more than one regime may prevail as the exchange rate moves through a sample period for recurring forecasts of the exchange rate. The findings of Altavilla and De Grauwe (2010) as well as Engel and Hamilton (1990) suggest that models which perform well for in-sample forecasts, should serve as viable additional benchmarks to measure out-of-sample forecasting performance. This is because such additional benchmarks include information about various regimes as opposed to only one (the most recent) regime.

3 Selection of the appropriate benchmark models

One of the major advantages of the RW model is that it serves as a benchmark to beat a random process. If an exchange rate model outperforms the RW benchmark it therefore provides evidence that the exchange rate information in question is forecastable. However, the RW model's association with the concept of randomness makes it insufficient to explain the linear and non-linear components in exchange rates. For this reason, and in line with the existing exchange rate theory, we now consider benchmark models that can accommodate either the linear or non-linear components that may exist in exchange rate information.

The Single Moving Average Model

The moving average (MA) method is one of the most naïve methods applied in linear forecasting to address concerns in volatility and is often used as smoothing method to explain introductory linear forecasting principles (Newbold, 1995:696). Data inputs for the model is created by averaging the values from k^{th} number of lags for each observation in time period $k + j$ as represented by equation (1) below where r represents the change in the exchange rate and $Mk + j$ the moving average value of order k for values ranging between r_{j+1} and r_{k+j} (Moosa, 2000:67). By extending the single MA method to a forecasting model where the error term (ϵ_t) is identically distributed we obtain equation (2)

presented below in regression form. Note that a single moving average model to the order 1 (k^{th} lag is equal to 1) is equivalent to the existing RW without drift benchmark.

$$\hat{r}_{k+j+1} = M_{k+j} = \frac{1}{k} \sum_{i=j+1}^{k+j} r_i \quad (1)$$

$$r_t = \frac{1}{k} \sum_{i=1}^k r_{t-i} + \epsilon_t \quad \epsilon_t \sim IDD(0, \sigma^2) \quad (2)$$

The Smoothed Linear Regression

To extend the moving average method to include more than just smoothed recent values of the exchange rate a time series model is created from equation (2) to represent a linear combination as given by equation (3). If only one lag is selected the model is equivalent to an autoregressive model with a lag order of 1. However, extending the lag to values greater than one expands the model to include both smoothed information and an expected trend. This characteristic aligns the single moving average process with the mean reverting patterns observed in linear exchange rate models. The resulting autoregressive (AR) process allows the data to speak for itself based on the assumption that the lagged values of the exchange rate contain information about the explanatory determinants of the exchange rate.

$$r_t = \mu + \frac{1}{k} \sum_{i=1}^k \theta_i r_{t-i} + \epsilon_t \quad (3)$$

The models univariate nature combined with a recurring one-step-ahead forecasting approach allows for effortless and direct comparisons to the RW benchmark as data inputs can be sourced from the same data set and the value of theta is recalculated for each forecast in a rolling window framework. The relative ease with which recurring forecasts can be calculated when a rolling window approach is followed accommodates the findings of Engel and Hamilton (1990), as well as Altavilla and De Grauwe (2010), who argue that more than one regime prevail as the exchange rate moves through a sample period for recurring forecasts of the exchange rate. It also addresses the concern of Engel (1994) who warns against choosing a benchmark based solely on out-of-sample performance. Since the sample selection moves one step forward in a rolling window framework, potential differences between the in-sample and out-of-sample regime is decreased to eliminate the informational disadvantage that may exist for comparisons of the smoothed linear regression against the established RW without drift benchmark (as discussed in the literature review). As is the case for the RW benchmark, the calculation process of the smoothed linear regression benchmark is simplistic and the tests of forecasting performance are easy to construct. The short calculation time leads to cost benefits for forecasters and improves the likelihood of the widespread use of this or a similar linear benchmark by the forecasting community.

The Autoregressive Model

As the exchange rate is often exposed to extreme volatile movements with the arrival of new information to the market, an autoregressive model with

a lag order of 1 (AR(1)) is evaluated as forecasting benchmark to assess the performance of AR models when the smoothing component of equation (3) is removed. The resulting linear representation is given by equation (4) which is simply an extension of the existing RW benchmark to include drift and slope components. r_t represents the change in the exchange rate, μ the expected drift and θ the slope of the AR model (Brooks, 2008:215).

$$r_t = \mu + \theta r_{t-1} + \epsilon_t \quad (4)$$

Following the reasoning of Engel and Hamilton (1990) the addition of a drift term (μ) to a random walk represents a more reasonable standard of estimation when the drift term is significantly different from zero. The addition of a directional trend (θ) removes the resulting AR(1) model from the concept of randomness to include a linear representation that recognises a constant directional trend in the exchange rate. In general terms the drift component and directional trend can be interpreted as additional information about the regime of the exchange rate over the whole sample period. The model is therefore sensitive to sudden changes in the regime as the drift component and trend represents an averaged view for all the observations in the sample (as is the case for all linear representations of exchange rates). However, the inclusion of a random component (r_{t-1}) to represent latest magnitude and direction of change provides for a more balanced forecasting approach to create a representation of both long (μ and θ) and short (r_{t-1}) term exchange rate information (in varying degrees depending on the sample size and frequency of information).

The selection of the AR(1) model as potential linear forecasting benchmark is supported by the findings of Lothian and Taylor (1996). They show that for annual data spanning the last 200 years the AR(1) process adequately explains the dollar-sterling and franc-sterling exchange rates. They argue that if the random walk model sufficiently explains exchange rates for the floating period, it should outperform a simplistic AR(1) benchmark for recurring forecasts of the exchange rate. In addition they show that the AR(1) model outperforms a RW with drift model thereby emphasising the importance of directional trend in exchange rate forecasting (as the latter models only differ by a slope component).

In summary the AR (1) model, single moving average model presented in equation (2) and smoothed linear regression presented in equation (3) are evaluated as linear forecasting benchmarks as they are firstly, more representative of the mean reverting or linear patterns associated with fundamental exchange rate models, secondly use readily available exchange rate information, thirdly allow for relatively easy model construction and calculation in a rolling window framework, fourthly improve on the limitations of the existing benchmark by including drift or smoothing components and lastly serve as implementable forecasting alternatives for comparisons of exchange rate model performance. Since similar linear AR models (or extensions of the AR models such as ARMA or ARCH models) may outperform the univariate benchmarks considered in this paper for a given data set, lag selection or data frequency, the main intent of this study is not to establish the ideal or superior benchmark for linear exchange

rate forecasting, but rather to investigate if simplistic linear models can consistently outperform a random walk benchmark. Such evidence would imply that although a given exchange rate model is able to outperform a random walk benchmark it may be unable to outperform a simplistic linear benchmark.

The Kernel Regression

Non-fundamental information contained in high frequency exchange rate data creates a challenge for short-term forecasts of the exchange rate as forecasters need to select a model that sufficiently accommodate a change in the directional trend of the exchange rate. For this reason, non-linear models are often selected to forecast exchange rates in the short-run. Their forecasts are driven by immediate changes in a non-linear trend and their assumptions are less restrictive than that of linear models (Mwamba, 2011:418). It is therefore appropriate that a non-linear benchmark is established against which the performance of non-linear models can be compared. This study proposes the use of an autoregressive univariate Kernel regression (also known as a heteroskedastic non-linear autoregressive model) as non-linear benchmark alternative.

In a Kernel regression a weighted function $\omega_{t,T}(x)$ is formed from a probability density function $K(x)$ which is non-parametric in nature. Since non-parametric models do not require any assumptions about the distributional form (for example normality), $K(x)$ does not play a probabilistic role. Nor is it assumed that X is distributed according to $K(x)$. The Kernel is simply a smoothing technique where the averaging of variables within a smoothing window is scaled according to the bandwidth h , as represented by equations (5) to (9) where the Kernel $K(x)$ and the bandwidth h has the following properties (Campbell, Lo and MacKinlay, 1997:500).

$$K(x) \geq 0, \quad \int K(u)du = 1 \quad (5)$$

$$K_h(u) \equiv \frac{1}{h}K\left(\frac{u}{h}\right), \quad \int K_h(u)du = 1 \quad (6)$$

The weight function used in the weighted average calculation is defined by equation (7) where the neighbourhood around X_t is determined from the value of the bandwidth h . To create a smoothed and well approximated convergence for the value of $m(\cdot)$ in equation (9), the value of h is adjusted to find the appropriate neighbourhood for the values of X_t . If the values of X_t are irregular the value of h is increased to include a greater neighbourhood of values. If on the other hand the values of X_t are closely scattered, the value of h is decreased to include a smaller neighbourhood of values (Campbell, Lo and MacKinlay, 1997: 500).

$$\omega_{t,T}(x) \equiv K_h(x - X_t)/g_h(x) \quad (7)$$

$$g_h(x) \equiv \frac{1}{T} \sum_{t=1}^T K_h(x - X_t) \quad (8)$$

The Nadaraya-Watson density function estimator for a Kernel regression is given by equation (9). We obtain the value for $\hat{m}_h(x)$ by substituting equation

(8) into (7). The resulting weighted average procedure is drawn from the assumption that if $m(\cdot)$ is sufficiently smooth for time series observations X_t near a value at point in time x_0 , then Y_t will be close to $m(x_0)$ for the corresponding values of X_t (Campbell, Lo and MacKinlay, 1997:500).

$$\hat{m}_h(x) = \frac{1}{T} \sum_{t=1}^T \omega_{t,T}(x) Y_t = \frac{\sum_{t=1}^T K_h(x - X_t) Y_t}{\sum_{t=1}^T K_h(x - X_t)} \quad (9)$$

In a Kernel regression the value of $K(\cdot)$ is usually chosen to be symmetric about zero, ensuring that the Kernel estimator $\hat{m}_h(x)$ is a density function itself. The Kernel estimator $\hat{m}_h(x)$ is therefore constructed from a Kernel that is centred at each observation with the Kernel function representing a spread of data near a value at point in time x_0 . To represent the spread of data we consider a Gaussian Kernel function regarded as the most popular choice of Kernel, as given by equation (10) (Campbell, Lo and MacKinlay, 1997: 501).

$$K_h(x) = \frac{1}{\sqrt[2]{2\pi} \frac{-x^2}{2h^2}} \quad (10)$$

From an exchange rate forecasting perspective the Kernel smoothing process removes the irregularity associated with high frequency exchange rate information to include an averaged interpretation of “noise” in the market. Härdle (1990) shows that under regularity assumptions for the Kernel K and the bandwidth h , $\hat{m}_h(x)$ converge to $m(x)$ as the population size increases. This convergence is also noted by Härdle and Linton (1994) for a wide variety of nonparametric methods. The noted convergence validates the use of nonparametric techniques to forecast high frequency exchange rate information on the assumption that exchange rates are mean reverting in an efficient market, even in the short-run.

To prevent that the neighbourhood for smoothing is too large, various methods are available to select the optimal bandwidth h . The most popular method is a cross-validation process that minimizes the weighted average squared error of the Kernel estimator (Campbell, Lo and MacKinlay, 1997:502). However, to improve the model calculation time, this study considers a scalar plug-in bandwidth as applied in the JMulti software and shown by Tschernig and Yang (2000) to perform as well as bandwidths selected from a cross-validation process. The plug in bandwidth is given by equation (11) where T is the sample size, $\hat{\sigma}_x$ the standard deviation and IQR the interquartile range of x_t (Lütkepohl and Krätzig, 1996:7).

$$h = 0.9T^{-\frac{1}{5}} \min\left(\frac{\hat{\sigma}_x IQR}{1.34}\right) \quad (11)$$

In addition, this study considers the corrected Final Prediction Error (FPE) criterion as applied by Tschernig and Yang (2000) to determine the number of lags for the Kernel regression. The lag(s) that yields the lowest FPE is used to calculate the optimal model. A major advantage of this method is that it adequately assigns the contribution of noise in the market to determine the

future directional trend. An added advantage is that the selection of the number of lags is based on a mathematical process rather than a process that includes human intervention.

From a theoretical perspective the non-linear Kernel regression provides four major improvements on the RW benchmark of exchange rates. Firstly, it selects the most relevant information to drive exchange rates through the selection of the optimal lag, number of lags and bandwidth (recall that the RW without drift model simply selects the most recent interpretation of noise and direction in the market with no indication of trend). Secondly, the Kernel regression improves on the RW benchmark by smoothing data to reduce the serial correlation that may exist in exchange rate data. Thirdly, it improves on the RW by not assuming that data will be independently and identically distributed (i.i.d) and finally by allowing for a non-linear directional trend in exchange rate information.

4 Data selection

The currencies evaluated in this study include the Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), Euro (EUR), British pound (GBP), Japanese yen (JPY) and South African rand (ZAR). All of the currencies pairs are expressed using indirect quotation against the United States dollar (USD) with the return data calculated as the difference in the logarithm of the exchange rate. The EUR, GBP and JPY exchange rates are selected to represent major international trade currencies, and the CAD and CHF to evaluate the exchange rate movements of relatively stable financial systems during the Global Financial Crisis. The AUD and ZAR are selected to respectively analyse the commodity based exchange rates of developed and underdeveloped economies. Overall, the currency selection aims to represent the exchange rates of a diverse number of economies.

The data for the study was obtained from the I-Net Bridge database. Four sample periods are selected to evaluate 120 point-sample observations of quarterly, monthly, weekly and daily end-of-day exchange rates for the period ending December 2011. The longest sample selection is therefore representative of exchange rate movements for the 3 decades (measured by calendar year) following the seminal work of Meese and Rogoff (1983) - who used floating exchange rate data for the period ending June 1981 to motivate the outperformance of the RW model. The same ending period is selected for all the frequencies to ensure that no preferential selection period is applicable across frequencies and currencies. However, it must be noted that the results for daily observations are susceptible to the volatility of the exchange rate over a relatively small sample period (July – December 2011). The four sample period selections stretches over a period ranging from 6 months to 30 years with the importance of recent information increasing with an increase in data frequency. For example, daily observations of the exchange rate for the past 6 months (last 120 observations) is deemed sufficient to forecast tomorrow's exchange rate, while weekly observations for almost 3 years (likewise 120 observations) is required to forecast the exchange

rate one week ahead (as opposed to selecting merely the latest observation as is the case for the RW benchmark). In this way, the fundamental and non-fundamental information that drive exchange rates is matched in line with data frequency on the assumption that the contribution of fundamental information diminishes as the frequency of information increases. This assumption is in line with the view held by exchange rate traders and forecasters that fundamental information drives exchange rates in the long-run as concluded by Cheung, Chinn and Marsh (2004).

5 Lag selection and tests of statistical significance

The selection of the optimal number of lags for moving average calculations is determined according to data frequency by selecting the number of k th observations required to equal or fit into the duration of the *next* data frequency. For example, to determine the number of lags for *daily* data we determine the number of trading days in a *week*. A similar approach is recommended by Brooks (2008, 149) for the lag selection of statistical tests despite the availability of a wide variety of models to determine the optimal number of lags. To verify if the lag selection according to the chosen frequency approach is appropriate for forecasts of the moving average (MA) based benchmarks considered in this study, the forecasting performance of the moving average method is considered for each currency and data frequency up to 5 lags. The results show that there is an acceptable level of underperformance for the frequency approach with none of the frequency based lag selections leading to a decline in model performance of more than 10% when compared against the optimal lag performance from the first 5 lags.

In this study the application of statistical tests to analyse data is somewhat arbitrary since a multitude of sample selections are used as data input for various models. While some data selections may be sufficient to explain the linear or non-linear trends in exchange rate data, statistical significance to confirm a good fit of data will be denied for others. In general, the relevance of statistical tests to establish a forecasting benchmark is questioned since the random walk model is often applied as forecasting benchmark despite the identification of a clear predictable structure in the exchange rate (Lothian and Taylor, 1996:505). The challenge for exchange rate forecasters is to match a particular forecasting benchmark with the information contained in a particular sample selection.

To analyse the distribution of data the Jarque-Bera test of normality is applied to models with a directional trend. Normality is confirmed when the null hypothesis is accepted at the 5% significance level. Overall, the results show that data normality can be confirmed for approximately 50% of the models, with the normality least prevalent in monthly observations of the exchange rate. To test for serial correlation in the error term of the linear and non-linear models, the Breusch-Godfrey test is applied with autocorrelation confirmed by the rejection

of the null hypothesis at the 5% significance level. The results show that only 11% of smoothed linear models conform to the requirement of no autocorrelation. AR(1) and Kernel models on the other hand performed well, conforming to the requirement of no autocorrelation 89% and 98% of the time respectively. As far as heteroskedasticity is concerned, the Lagrange multiplier (ARCH-LM) test is applied to trend based models to investigate the persistence of volatility in the exchange rate. Heteroskedasticity is confirmed by the rejection of the null hypothesis at the 5% significance level. The results are generally favourable with 89% of the benchmark models showing no heteroskedastic effects.

Considering all the statistical tests applied, the Kernel model is the clear outperformer compared to the smoothed linear and AR(1) benchmark models. The non-linear nature of Kernel regressions do not assume that data is normally distributed and provide the best results for tests of serial correlation and persistence in volatility. The satisfactory statistical results for the AR(1) and Kernel regressions suggest that exchange rates follow a non-linear mean reverting process.

6 Methodology to assess forecasting performance

A thorough forecasting approach is adopted to analyse the forecasting performance of the proposed benchmark models for multiple currencies and frequencies. Each consecutive forecast is obtained from a sample selection that operates in a rolling window framework. As such, the number of observations selected for each recurring forecasts remains constant at 90 observations. Consecutive one-step-ahead forecasts are created for the last 30 out of a 120 observations of quarterly, monthly, weekly and daily return values of the exchange rate. The first forecast result is obtained for the 91st observation and compared against the actual change in return. This process is repeated until the forecast for the 120th observation is calculated.

The forecasting approach differs from the approach followed by Meese and Rogoff (1983) which established the RW benchmark in three distinct ways. Firstly, each forecast horizon is driven by the information contained in the frequency of the exchange rate on the assumption that lower frequency information contains more fundamental information and less market “noise” when compared to high frequency information. In this way, the appropriate mix of fundamental vs. non fundamental information is applied to calculate one-step-ahead exchange rate forecasts for each horizon (as is the norm for one-step-ahead forecasts of the RW model). In their influential study Meese and Rogoff (1983) estimated multiple-step-ahead forecasts for up to ten months using only monthly exchange rate observations. This approach might have provided the RW model with a comparative advantage for forecasts comparisons of more than one-step-ahead, since the noise information contained in monthly data could adversely affect the accuracy of forecasts of more than one month ahead. However, consideration for the information contained in the frequency of the exchange rate does not explain why the one-month-ahead univariate models applied by Meese

and Rogoff (1983) could not outperform the RW model.

This brings us to the second distinct difference in terms of data interpretation. While Meese and Rogoff (1983) continuously increased their sample selection for each forecast by adding the latest available information to the sample selection for each recurring forecast, this study applies a rolling window framework. In a rolling window framework the oldest variable is removed from the sample selection with each new observation added. This process ensures that the sample selection evolves to exclude older information and include newer information for each recurring forecast. The evolution process therefore ensures that older information is discarded in favour of newer information with the aim to include the most recent information about the regime of the exchange rate. If the sample selection includes more information about an exchange rate's immediate regime it should improve the one-step-ahead forecast result as argued to motivate the selection of additional benchmark models earlier.

The third distinct difference in forecasting approach relates to the confirmation of model outperformance at various frequencies. While Meese and Rogoff (1983) provided fairly robust calculations given the forecasting tools available at the time, they only calculated forecasts for one frequency of the exchange rate. Recent advancements in computer forecasting software however, allow for recurring forecasts to be done at various frequencies with relative ease. To strengthen a model's case for consistent superiority, this study measures the forecasting performance of each model for quarterly, monthly, weekly and daily observations of the exchange rate. The resulting forecasting approach considers 30 recurring forecasts for 7 currencies at 4 frequencies to create 840 individual forecasts for each benchmark model presented. The chosen recurring forecast approach in a rolling window framework, combined with increased recognition for the information contained at each frequency, aims to improve the comparative playing field for all the models considered in this study. Quoting from the data description provided by Meese and Rogoff (1983), each model is allowed to include the "most up to date information" of the exchange rate. However, additional consideration is given to the information contained in the frequency of the exchange rate.

As far as the measure of forecasting performance is concerned, the root mean square error (RMSE) is used to align forecast results with the findings of Meese and Rogoff (1983) and subsequent studies on the performance of the RW model by researchers such as Schinasi and Swamy (1987), Engel (1994) and Altavilla and De Grauwe (2010) - who emphasise the informational importance of a change in the regime of the exchange rate. The overall model outperformance for forecasts of the additional benchmarks are judged with the RMSE as primary criterion. The calculation of the RMSE values for each consecutive forecast is given by equation (12) where r_i is the exchange rate return in period i (Moosa, 2000:341).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{r_i - \hat{r}_i}{r_i} \right)^2} \quad (12)$$

The model with the lowest RMSE among the competing models is identi-

fied as the superior model for a given currency and frequency combination. To improve the comparability of the results, all the RMSE values are expressed as a percentage improvement relative to the random walk models RMSE performance as represented by equation (13). The competing model RMSE is selected as numerator and the RW model RMSE selected as the denominator in the ratio calculation to create a measure of outperformance that can either be positive or negative when deducted from the value 1. The resulting percentage value is in turn used to serve as *relative* measure of a model’s outperformance against the RW model and other competing additional benchmark models under consideration¹. By expressing the RMSE performance as a percentage improvement, the average improvement in RMSE performance can be calculated for each model across currency and frequency combinations.

$$RMSE\ Improvement = \left(1 - \frac{Competing\ Model\ RMSE}{Random\ Walk\ Without\ Drift\ RMSE}\right) \quad (13)$$

7 Results of forecasting performance

Table I represents the forecasting performance of each additional benchmark across currencies and frequencies utilising the RMSE improvement measure as defined in the previous section (actual RMSE values for the respective models are provided in the Appendix). For the single moving average model, the results of forecasting performance show a consistent improvement in RMSE values for comparisons against the RW benchmark. On average, a 25.4% improvement in forecasting performance is obtained across currencies and frequencies suggesting a mean reverting, as opposed to random, component in the immediate preceding values of the exchange rate. When the single moving average model is expanded to include a linear representation of moving average information, the model average forecasting performance decreases to 21.7%. Despite a decrease in overall forecasting performance, the smoothed linear regression continues to outperform the RW model for all forecasts across currency and frequency combinations. The chosen frequency approach for the selection of the number of lags in the moving average calculation, coupled with the recurring one step ahead forecast approach, as well as consideration for the mix of fundamental vs. non-fundamental information in the frequency of exchange rate information, therefore yield forecasting results that significantly improve on a random walk approach.

When the MA component is removed from the smoothed linear regression to create an AR(1) model, the averaged forecast performance of the linear model

¹A ratio-based approach is also followed by Evans and Lyons (2005) to assess the performance of micro-based models against the performance of the RW model. In their study, the mean square error ratio’s show a consistent change (improvement or decrease depending on the model under evaluation) in forecasting performance as the forecasting horizon increased for daily forecasts of the exchange rate. Their identification of a consistent pattern in forecasting performance provides support for the application of mean square error ratio’s to allow for interpretable comparisons of forecasting performance

increases significantly to 30.7%. The resulting improvement on forecasting performance against the RW model indicates that the creation of a linear process from simply the most recent exchange rate information (latest change in return) provide noteworthy forecasts for one-step-ahead exchange rate prediction. More importantly, the results provide strong empirical evidence that the inclusion of a drift and slope component significantly improve on the forecasts of the RW model. The consistent outperformance of the AR(1) model against the RW model is further in line with the findings of Lothian and Taylor (1996), who provides strong evidence that a simple AR(1) model outperforms both the RW with-and-without drift benchmark alternatives in a floating exchange rate environment.

To accommodate for the non-linear trends that may exist in exchange rate information, a Kernel model is estimated considering 2 out of the first 5 lags as model inputs. The results show that the kernel model improves on the *overall* forecasts of the RW benchmark by 29.18% (for weekly

forecasts of the CHF, the Kernel model considered up to 12 lags to accommodate for the extreme volatility in return information over the forecasting period). The Kernel model further provides consistent outperformance across currency and frequency combinations with at least 92% of forecasts improving on RW forecasts by 20% or more.

Considering all the benchmark alternatives presented thus far, the AR(1) model performs the best for forecast frequencies of one week or more, while the Kernel model provides superior forecasts for daily exchange rate forecasting. It is also important to note that all of the benchmark alternatives are able to consistently outperform a simple random walk model for various currency and frequency combinations. Possible reasons for the superior performance of the additional benchmarks include firstly, that each model is allowed to include the most recent exchange rate return information through the application of a rolling window framework which discards old return information, in favour of new return information. Secondly, each forecast horizon considers the combination of fundamental vs. non-fundamental information contained in the frequency of exchange rate (on the assumption that lower frequency information contains more fundamental information and less “noise” information when compared to high frequency information). Thirdly, the RMSE performance is considered for the last 25% of the combined sample selection to create an averaged interpretation of a model’s forecasting performance. Although the outperformance of the RW model cannot be denied for some forecasts, multiple forecast together strengthen the validity of a model’s forecasting power as argued by Poterba and Summers (1988). Fourthly, since only return information is used as input for the linear and non-linear benchmarks, exchange rate movements have limited ability to deviate from their intrinsic values (under the assumption that they operate in an efficient market).

A summarised view of the performance of the additional benchmark models is reported in Table II. For each currency and frequency combination the percentage of times that a particular model outperforms its peers is presented to provide an indication of a model’s potential to be applied as forecasting bench-

mark from the alternatives. The results show that both the AR(1) model and Kernel model, provide competitive forecasts to serve as additional benchmarks of exchange rate forecasting. These models outperform their peers about 50% of the time. From all the benchmark alternatives, either the AR(1) model or Kernel² model provides superior forecasting performance, more than 95% of the time. Conversely, the single MA model and smoothed linear regression provide dissatisfactory results, with only one occurrence of outperformance achieved by the single MA model for monthly forecasts of the CAD.

Insert Table II here.

The outperformance of the additional benchmarks presented emphasise that model superiority is only partly established when a model is able to outperform a RW model. To strengthen the validity of a model’s forecasting power, researchers often compare their results to the forecasting performance of other models. However, these benchmark models differ across studies and as shown by the forecasts of the smoothed linear regression, even the worst performing benchmark model is able to outperform a RW model. If simplistic benchmark models such as the AR(1) and Kernel models are applied across studies to access model forecasting performance, it will go a long way to compare the information contained in the frequency, sample selection or forecast horizon of various exchange rate studies.

8 Conclusion

Given the consistent outperformance of the RW benchmark by all of the linear and non-linear benchmarks, there is sufficient evidence to suggest that additional benchmarks are required to assist both researchers and exchange rate forecasters to establish exchange rate model superiority. As linear benchmark, the AR(1) model performed best to explain the mean reverting patterns in exchange rate information by consistently outperforming the RW, single MA and smoothed linear models across currency and frequency combinations. In addition, the AR(1) model also outperformed the Kernel regression for quarterly, monthly and weekly forecasts of the exchange rate. As non-linear benchmark,

²As check of robustness the last 120 daily observations of 2008 was selected as sample period to evaluate the performance of the proposed exchanged rate benchmarks during the Global Financial Crisis. Applying currency selection, lag selection and model calculation principles identical to those applied in this study, the results continue to support the outperformance of the Kernel regression for daily forecasts of the exchange rate. A Kernel regression (with one lag selected) provided overall outperformance of 25% against the established RW benchmark. Model outperformance for the single MA, smoothed linear and AR(1) models were 17%, 19% and 22% respectively.

The results suggest that high frequency exchange rate information is better captured by a non-linear trend of immediate exchange rate values. The Kernel smoothing process removes the irregularity associated with high frequency exchange rate information to include an averaged interpretation of “noise” in the market - guided by and automated lag selection process. While the AR(1) continued to perform well, linear benchmark models are ideally intended for forecasting comparisons of low frequency information where the exchange rate is driven by fundamentals in an efficient market - where noise in the market is expected to have minimal impact on short-term forecasts.

the Kernel model consistently outperformed the RW, single MA and smoothed linear models and is best able to explain the volatile exchange rate movements associated with daily exchange rate forecasting.

The improved forecasting performance of the additional benchmarks are most likely due to the application of a rolling window framework to calculate forecasts for each recurring step of the exchange rate, as well as increased consideration for the informational mix contained in the frequency of the exchange rate. Together, these changes to the forecast approach eliminate a considerable portion of the informational advantage that the RW model may enjoy for one-step-ahead forecasts of the exchange rate. This conclusion is derived from the single moving average models ability to comfortably outperform the RW model following similar informational principals.

The outperformance of the RW model by the benchmark alternatives serves to confirm that the RW model is an insufficient benchmark to explain exchange rate movements for non-static models. Although the random walk model may continue to serve as viable benchmark for comparisons against static models (models where the coefficients are not allowed to change), recent advancements in forecasting software allow for the calculation of non-static exchange rate models with relative ease. This study finds that additional benchmarks are required to strengthen the validity of an exchange rate models forecasting outperformance as simplistic non-static linear and non-linear models consistently outperform the established random walk benchmark. The application of additional benchmarks to establish model superiority will ensure that researchers are challenged to continue with the development of sophisticated exchange rate models that outperform simplistic alternatives.

9 Recommendations

Since high frequency information is generally more volatile due to market risk, news and noise trading, the evaluation of additional benchmarks can be expanded to confirm consistent outperformance for hourly or higher frequency exchange rate information. However, it is advisable that several sample selections are used to strengthen the validity of high frequency benchmark outperformance. Since this study only considered daily information for a period of six-months as high frequency component, the validity of additional benchmarks for daily forecasts can likewise be strengthened if multiple six-month sample selections are introduced to confirm daily forecast outperformance.

References

- [1] Altavilla, C. & De Grauwe, P. (2010). Forecasting and Combining Competing Models of Exchange Rate Determination. *Applied Economics*, 42(27): 3455-3480.

- [2] Alvarez, F., Atkeson, A., & Kehoe, P. J. (2007). If Exchange Rates are Random Walks, then Almost Everything we Say About Monetary Policy is Wrong. *The American Economic Review*, 97(2): 339-345.
- [3] Bachelier, L. (1900). "Theory of Speculation", in Cootner, P. (1964). *The Random Character of Stock Market Prices*. Massachusetts: Institute of Technology Press.
- [4] Bissoondeeal, R. K., Binner, J. M., & Elger, T. (2009). Monetary Models of Exchange Rates and Sweep Programs. *Applied Financial Economics*, 19: 1117-1129.
- [5] Bleany, M. & Mizen, P. (1996). Nonlinearities in Exchange-Rate Dynamics: Evidence from Five Currencies, 1973-94. *The Economic Record*, 78(216): 36-45.
- [6] Brooks, C. (2008). *Introductory Econometrics for Finance*. New York: Cambridge University Press.
- [7] Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. New Jersey: Princeton University Press.
- [8] Cheung, Y., Chinn, M. D. & Marsh, I. W. (2004). How do UK-Based Foreign Exchange Dealers Think their Market Operates? *International Journal of Finance and Economics*, 9: 294-296.
- [9] De Bondt, W. F. M. & Thaler, R. H. (1985). Does the Stock Market Overreact? *Journal of Finance*, 40: 793-905.
- [10] Engel, C. (1994). Can the Markov Switching Model Forecast Exchange Rates? *Journal of International Economics*, 36: 151-165.
- [11] Engel, C. & Hamilton, J. D. (1990). Long Swings in the Dollar: Are They in the Data and Do
- [12] Markets Know It? *American Economic Review*, 80(4): 689-713.
- [13] Evans, M. D. & Lyons, R. K. (2005). Meese-Rogoff Redux: Micro-Based Exchange Rate Forecasting. *American Economic Review*, 95(2): 405-414.
- [14] Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press: Cambridge, UK.
- [15] Härdle, W. & Linton, O. (1994). *Applied Nonparametric Methods*. Available from: <http://dido.wss.yale.edu/P/cd/d10b/d1069.pdf>. (Accessed 12 December 2012).
- [16] Isard, P. (1995). *Exchange Rate Economics*. Cambridge University Press.
- [17] Lothian, J. R. & Taylor, M. P. (1996). Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries. *Journal of Political Economy*, 104(3): 488-509.

- [18] Lütkepohl, H. & Krätzig, M. (1996). *Initial Analysis in JMulti*. Available from: <http://jmulti.net/download/help/initanal.pdf>. (Accessed 12 December 2012).
- [19] Meese, R. & Rogoff, K. (1983). Empirical Exchange Rate Models of the Seventies: Do they Fit the Out of Sample? *Journal of International Economics*, 14: 3-24.
- [20] Meese, R. & Rogoff, K. (1988). Was It Real? The Exchange Rate-Interest Differential Relation over the Modern Floating-Rate Period. *Journal of Finance*, 43(4): 933-948.
- [21] Moosa, I. A. (2000). *Exchange Rate Forecasting Techniques and Applications*. London: Macmillan Press Ltd.
- [22] Mwamba, J. M. (2011). Modelling Stock Price Behaviour: The Kernel Approach. *Journal of Economics and International Finance*, 3(7): 418-423.
- [23] Newbold, P. (1995). *Statistics for Business and Economics*. New Jersey: Prentice-Hall Inc.
- [24] Poterba, J. M. & Summers, L. H. (1988). Mean Reversion in Stock Prices: Evidence and Implications. *Journal of Financial Economics*, 22(1): 27-59.
- [25] Sarantis, N. & Stewart, C. (1995). Structural, VAR and BVAR Models of Exchange Rate Determination: A Comparison of Their Forecasting Performance. *Journal of Forecasting*, 14: 201-215.
- [26] Sarno, L. (2003). Nonlinear Exchange Rate Models: A Selective Overview. Vol.3. International Monetary Fund.
- [27] Schinasi, G. J. & Swamy, P. A. V. B. (1987). The Out-of-Sample Forecasting Performance of Exchange Rate Models when Coefficients are Allowed to Change. *Board of Governors of the Federal Reserve System (US) International Finance Discussion Papers No.301*.
- [28] Tschernig, R. & Yang, L. (2000). Nonparametric Lag Selection for Time Series. *Journal of Time Series Analysis*, 21(4): 457-487.
- [29] Wu, J. L. & Chen, S. L. (2001). Real Exchange-Rate Prediction over Short Horizons. *Review of International Economics*, 9(3): 401-413.

Table I: Comparative Forecasting Performance of Additional Benchmark Models

Model	Frequency	AUD	CAD	CHF	EUR	GBP	JPY	ZAR	Frequency Average
MA	Quarterly	13.2%	3.2%	24.4%	19.8%	9.9%	36.6%	25.7%	19.0%
	Monthly	30.9%	49.0%	28.8%	27.7%	25.2%	34.7%	31.7%	32.6%
	Weekly	16.2%	25.9%	27.2%	25.1%	12.9%	39.1%	27.0%	24.8%
	Daily	15.9%	25.6%	29.2%	25.0%	24.8%	31.0%	24.3%	25.1%
	Currency Average	19.1%	25.9%	27.4%	24.4%	18.2%	35.4%	27.2%	25.4%
Smoothed Linear	Quarterly	8.6%	5.2%	18.1%	13.4%	7.3%	30.6%	22.2%	15.1%
	Monthly	25.8%	36.8%	23.3%	25.8%	25.5%	26.0%	22.7%	26.6%
	Weekly	13.0%	20.7%	27.6%	23.3%	15.1%	31.8%	21.5%	21.9%
	Daily	17.7%	24.0%	25.1%	20.8%	24.3%	27.8%	23.3%	23.3%
	Currency Average	16.3%	21.7%	23.5%	20.8%	18.1%	29.1%	22.4%	21.7%
AR(1)	Quarterly	22.7%	19.0%	31.3%	28.0%	20.6%	36.4%	31.3%	27.0%
	Monthly	33.0%	43.8%	29.5%	30.0%	31.1%	35.1%	34.6%	33.9%
	Weekly	26.8%	32.5%	32.1%	31.7%	26.1%	41.0%	30.5%	31.5%
	Daily	23.0%	29.9%	34.4%	29.9%	31.5%	33.6%	29.0%	30.2%
	Currency Average	26.4%	31.3%	31.8%	29.9%	27.3%	36.5%	31.4%	30.7%
Kernel	Quarterly	21.8%	5.8%	29.5%	26.6%	20.6%	42.5%	29.4%	25.2%
	Monthly	35.1%	43.9%	24.6%	27.4%	25.4%	35.3%	29.6%	31.6%
	Weekly	27.1%	30.7%	9.4%	28.5%	28.0%	41.6%	26.9%	27.5%
	Daily	29.4%	34.0%	34.4%	31.6%	34.4%	33.7%	29.6%	32.4%
	Currency Average	28.4%	28.6%	24.5%	28.5%	27.1%	38.3%	28.9%	29.18%

Source: JMulti results

Table 2: Superior Forecasting Performance of Additional Benchmark Models

Frequency	Single MA	Smoothed Linear	AR(1)	Kernel
Quarterly	0.0%	0.0%	71.4%	28.6%
Monthly	14.3%	0.0%	57.1%	28.6%
Weekly	0.0%	0.0%	57.1%	42.9%
Daily	0.0%	0.0%	14.3%	85.7%
Model Average	3.6%	0.0%	50.0%	46.4%

Source: JMulti results

Appendix 1: RMSE Performance

RMSE Performance							
Model	AUD	CAD	CHF	EUR	GBP	JPY	ZAR
Quarterly							
RW Without Drift	0.09532	0.06135	0.07890	0.07567	0.07102	0.08288	0.12928
Single MA	0.08272	0.05940	0.05966	0.06069	0.06401	0.05254	0.09599
Smoothed Linear	0.08716	0.05815	0.06462	0.06551	0.06585	0.05752	0.10057
AR(1)	0.07364	0.04969	0.05421	0.05445	0.05640	0.05275	0.08879
Kernel	0.07454	0.05779	0.05565	0.05557	0.05639	0.04762	0.09121
Monthly							
RW Without Drift	0.06772	0.05105	0.05796	0.05349	0.04082	0.04136	0.06752
Single MA	0.04679	0.02603	0.04128	0.03868	0.03053	0.02699	0.04610
Smoothed Linear	0.05026	0.03226	0.04447	0.03970	0.03041	0.03059	0.05217
AR(1)	0.04536	0.02870	0.04084	0.03745	0.02811	0.02685	0.04415
Kernel	0.04398	0.02863	0.04368	0.03881	0.03046	0.02677	0.04751
Weekly							
RW Without Drift	0.03482	0.02639	0.04162	0.02522	0.01462	0.01908	0.04266
Single MA	0.02917	0.01954	0.03028	0.01889	0.01273	0.01162	0.03114
Smoothed Linear	0.03029	0.02094	0.03011	0.01934	0.01241	0.01300	0.03348
AR(1)	0.02548	0.01782	0.02827	0.01723	0.01080	0.01126	0.02964
Kernel	0.02537	0.01828	0.03771	0.01803	0.01054	0.01113	0.03117
Daily							
RW Without Drift	0.01208	0.00944	0.01004	0.00826	0.00807	0.00736	0.01764
Single MA	0.01016	0.00702	0.00711	0.00619	0.00606	0.00508	0.01335
Smoothed Linear	0.00994	0.00717	0.00752	0.00654	0.00611	0.00532	0.01354
AR(1)	0.00931	0.00662	0.00658	0.00579	0.00553	0.00489	0.01254
Kernel	0.00853	0.00623	0.00659	0.00565	0.00529	0.00489	0.01242

Source: JMulti results