A Reappraisal of the Meese-Rogoff Puzzle

Imad Moosa and Kelly Burns*

School of Economics, Finance and Marketing, RMIT

Abstract

Several explanations have been put forward for the Meese-Rogoff puzzle that exchange rate models cannot outperform the random walk in out-of-sample forecasting. We suggest that a simple explanation for the puzzle is the use of the root mean square error (RMSE) to measure forecasting accuracy, presenting a rationale as to why it is difficult to beat the random walk in terms of the RMSE. By using exactly the same exchange rates, time periods and estimation methods as those of Meese and Rogoff, we find that their results cannot be overturned even if the models are estimated with time-varying coefficients. However, we also find that the random walk can be outperformed by the same models if forecasting accuracy is measured in terms of the ability to predict direction, in terms of a measure that combines magnitude and direction, and in terms of profitability.

Keywords: Forecasting, Random Walk, Exchange Rate Models, Direction Accuracy

JEL Classification Numbers: F31, F37, C53

* The corresponding author. Address: School of Economics, Finance and Marketing, RMIT, 445 Swanston Street, Melbourne, Victoria 3000, Australia. E-mail: kelly.burns@rmit.edu.au.

Introduction

Since the publication of the highly-cited paper of Meese and Rogoff (1983), it has become something like an undisputable fact of life that exchange rate determination models cannot outperform the naïve random walk model in out-of-sample forecasting.[1] Frankel and Rose (1995) argue that the negative results have had a "pessimistic effect" on the field of exchange rate modelling in particular and international finance in general. Sarno and Taylor (2002) argue that although empirical exchange rate models occasionally generate apparently satisfactory explanatory power in-sample, they generally fail badly in out-of-sample forecasting tests in the sense that they fail to outperform the random walk.[2] Bacchetta and van Wincoop (2006) point out that the poor explanatory power of existing exchange rate models is most likely the major weakness of international macroeconomics. In general, a view is held within the profession that no one has overturned the Meese-Rogoff results.

Several reasons have been put forward for the failure of exchange rate models to outperform the random walk. In their original paper, Meese and Rogoff (1983) attributed the failure of the models to simultaneous equations bias, sampling errors, stochastic movements in the true underlying parameters, misspecification and non-linearities. Meese (1990) adds other explanations such as improper modelling of expectations and over-reliance on the representative agent paradigm. Faust et al.

[1] Engle et al. (2003) argue that the Meese-Rogoff paper "has been vey influential in the development of empirical exchange rate studies".
[2] This proposition came years after the announcement by MacDonald and Taylor (1994) of their ability to outperform the random walk in out-of-sample forecasting. No explanation was suggested for the apparent contradiction.

(2003) suggest that forecasting power can be improved by using original release data rather than fully revised data and by using real time forecasts of future fundamentals instead of actual future fundamentals.[3] They conclude that their results contradict a "cherished presumption dating back to Meese and Rogoff"—the presumption being the finding that the random walk cannot be outperformed.

The objective of this paper is to demonstrate that the random walk appears to be superior to exchange rate models only because forecasting accuracy is measured by metrics that rely entirely on the magnitude of the forecasting error while overlooking the ability of the models (and the random walk) to predict the direction of change. Our proposition is that if forecasting power is judged by metrics that take into account factors other than the magnitude of the error, the Meese-Rogoff results can be overturned. This proposition is demonstrated by re-working the Meese-Rogoff results, using the same models, time period and forecasting horizons.

The Meese-Rogoff Methodology

Meese and Rogoff (1983) used out-of-sample forecasting power, as judged by the root mean square error, to determine the predictive power of a variety of exchange rate models. Engle and West (2005) dismiss this procedure completely by arguing that an exchange rate determination model should not be evaluated simply by whether it can beat the random walk in out-of-sample forecasting because exchange rates themselves are random walk. Engel et al. (2007) argue that beating the random walk in forecasting is too strong a criterion for accepting a model and suggest that exchange

[3] It is not obvious how the use of forecast values, as opposed to actual values, leads to improved forecasting accuracy when it is likely to compound the forecasting error.

rate models are not that bad.[4] They approach the problem from several angles, emphasising the observation that short-run movements in exchange rates are primarily determined by changes in expectations, which is exactly what the standard models imply. They explain the failure of exchange rate models to beat the random walk in terms of the "near random walk behaviour" of exchange rates, which makes the models' power to beat the random walk in out-of-sample forecasting rather low.[5]

Most economists, however, cast doubt on the appropriateness of the Meese-Rogoff procedure for the evaluation of predictive accuracy without dismissing it entirely as Engel et al. suggest. The underlying arguments are based on three aspects of the procedure: (i) in-sample versus out-of-sample forecasting, (ii) the use of random walk with drift as opposed to random walk without drift, and (iii) the use of measures of magnitude as opposed to measures of direction accuracy and profitability.

On the first issue, it has been suggested that it is more appropriate to use out-of-sample forecasting.[6] The conventional wisdom that it is more likely to find significant evidence of in-sample predictability compared to out-of-sample predictability reflects the lack of reliability of in-sample tests under the null of no predictability. For example, Tashman (2000) makes the statement that "forecasters generally agree that

[4] There is some sort of contradiction here. If the exchange rate is a random walk, then by definition it is not forecastable. If this is the case then it is not plausible to claim that "exchange rate models are not that bad".

[5] What Engle et al. (2007) say is that because the exchange rate behaves like a random walk the random walk model is the best representation of this behaviour. If this is the case then the random walk (model) is better than any other model. Again, in what sense are exchange rate models "not that bad" when the implication is that the models are misspecified or at best they do not represent the true behaviour of exchange rates?

[6] Clements and Hendry (2003) dismiss out-of-sample forecasting power as a model evaluation criterion. However, their argument is not for the use of in-sample forecasting—rather, they refer to "dichotomies" that "intrude" on any forecasting evaluation exercise.

forecasting methods should be assessed for accuracy using out-of-sample tests. Fildes and Makridakis (1995) suggest that "the performance of a model on data outside that used in its construction remains the touchstone for its utility in all applications".

On the second issue Engel and Hamilton (1990) argue that random walk with drift is a more reasonable standard of comparison when the drift factor is significantly different from zero. In their original study, Meese and Rogoff (1983) used both random walk with drift and without drift, without stating why one is preferable to the other.[7] However, the evidence provided by Engel (1994) indicates that it does not make much difference whether the random walk with or without drift is used. While this proposition may be valid for forecasting accuracy in terms of the magnitude of the forecasting error, it is definitely not so with respect to forecasting the direction of change. The random walk with drift predicts consistently either a positive or negative change, capturing the direction of change on some occasions. Therefore, it does make a difference whether the random walk with drift or without drift is used. If the random walk with drift is used when the drift factor is insignificant, the random walk will appear misleadingly better in forecasting the direction of change than it actually is. Therefore it is plausible to suggest that the choice between the random walk without and with drift depends (or should depend) on the statistical significance of the drift factor.[8]

---

[7] They measured the drift factor as the average first log difference of the underlying exchange rate.
[8] A generally accepted proposition is that exchange rates move as a random walk with little or no drift, which means that the random walk without drift should be used. However, whether or not exchange rates move as random walk without or with drift is an empirical issue where the outcome depends on the underlying exchange rate, the time period and the forecasting horizon.

The third issue pertains to measures of forecasting accuracy. In their original study Meese and Rogoff (1983) used conventional measures (predominantly the root mean square error), which are calculated from the magnitude of the forecasting error (the deviation of the forecast from the actual value). It is plausible to suggest that this may not be entirely appropriate because a correct prediction of direction can be more important than predicting the magnitude of the error and that the ultimate test of forecasting power is the ability to make profit by trading on the basis of the forecasts. However, whether the prediction of the magnitude of change is more or less important than the prediction of the direction of change depends on the underlying situation and the purpose for which the forecasts are used.[9]

Cheung et al. (2005) reassess exchange rate prediction using as criteria the mean square error, direction accuracy and the consistency test of Cheung and Chinn (1998).[10] They argue that using criteria other than the mean square error does not boil down to "changing the rules of the game" and that minimising the mean square error may not be important from an economic standpoint. They present a reason for not relying on the mean square error, suggesting that it may miss out on important aspects of prediction, particularly at long horizons. Christofferson and Diebold (1998) point out that the mean square error indicates no improvement in predictions that take into account cointegrating relations *vis-a-vis* univariate prediction. Leitch and Tanner

---

[9] In simple currency trading, where the interest rate differential is ignored (for example, intra-day trading), the prediction of the direction of change is the only thing that matters. When interest rates are taken into account, both factors become important because the decision depends on the expected rate of return, which consists of the interest rate differential and the expected change in the exchange rate. When the underlying situation involves speculation on combined currency option positions, such as straddles and strangles, the only thing that matters is the magnitude of change.

[10] The consistency criterion focuses on the time series properties of the forecast. A forecast is consistent if it moves in tandem with the actual value in the long run.

(1991) argue that the direction of change may be more relevant for profitability and economic concerns, while Cumby and Modest (1987) point out that it is also related to tests for market timing ability.

Profitability, or in general utility, is another criterion that can be used to test predictive power. Abhyankar et al. (2005) propose a utility-based criterion pertaining to the portfolio allocation problem. They find that the relative performance of a structural model improves when this criterion is used. Likewise, West et al. (1993) suggest a utility-based evaluation of exchange rate predictability. Li (2011) evaluates the effectiveness of economic fundamentals in enhancing carry trade, concluding that the profitability of carry trade and risk-return measures can be enhanced by using forecasts. Likewise, Boothe and Glassman (1987) compare the rankings of alternative exchange rate forecasting models using two different evaluation criteria: accuracy and profitability. The results show that the random walk ranks highest in forecasting accuracy and in terms of profitability for one of the two currency pairs used by them (German mark/U.S. dollar).

Leitch and Tanner (1991) suggest that economists are puzzled by the observation that profit-maximising firms buy professional forecasts when measures of forecasting accuracy indicate that a naïve model forecasts about as well. The explanation they present is that these measures bear very weak relation to the profit generated by acting on the basis of the forecasts. They point out that the only substitute criterion for profits is a measure of direction accuracy, as they find the relation between direction accuracy and profit to be almost as close as the relation between the other measures.

They further suggest that if profits are not observable, direction accuracy of the forecasts may be used as the evaluation criterion.

Modelling and Forecast Generation

Following Meese and Rogoff (1983) three macroeconomic exchange rate models are estimated and used to generate forecasts. These are the Frenkel-Bilson model (the flexible-price monetary model), the Dornbusch-Frankel model (the sticky-price monetary model) and the Hooper-Morton model. These models are specified as follows:

$$s_t = \alpha_0 + \alpha_1 (m_t - m_t^*) + \alpha_2 (y_t - y_t^*) + \alpha_3 (i_t - i_t^*) + \varepsilon_t \quad (1)$$

$$s_t = \alpha_0 + \alpha_1 (m_t - m_t^*) + \alpha_2 (y_t - y_t^*) + \alpha_3 (i_t - i_t^*) + \alpha_4 (\pi_t^e - \pi_t^{e*}) + \varepsilon_t \quad (2)$$

$$s_t = \alpha_0 + \alpha_1 (m_t - m_t^*) + \alpha_2 (y_t - y_t^*) + \alpha_3 (i_t - i_t^*) + \alpha_4 (\pi_t^e - \pi_t^{e*}) + \alpha_5 B_t + \alpha_6 B_t^* + \varepsilon_t \quad (3)$$

where s is the log of the exchange rate, m is the log of the money supply, y is the log of industrial production, i is the short-term interest rate, π^e is the long-run expected inflation rate and B is the trade balance (an asterisk indicates the foreign variable from a U.S. perspective). In these specifications the exchange rate is measured as the number of dollars per unit of the other currency (Japanese yen, JPY; Deutsche mark, DEM; and British pound, GBP).

The models are estimated over part of the sample period, $t = 1,2,\ldots m$, then a one-period-ahead forecast is generated for the point in time $m+1$. The forecast log exchange rate derived from equation (1) is

$$\hat{s}_{m+1} = \hat{\alpha}_0 + \hat{\alpha}_1 (m_{m+1} - m^*_{m+1}) + \hat{\alpha}_2 (y_{m+1} - y^*_{m+1}) + \hat{\alpha}_3 (i_{m+1} - i^*_{m+1}) \quad (4)$$

where $\hat{\alpha}_i$ is the estimated value of α_i. The process is then repeated by estimating the model over the period $t = 1,2,\ldots m+1$ to generate a forecast for point in time *m*+2, $\hat{s}_{m+2}$, and so on until we get to $\hat{s}_n$, where *n* is the total sample size. In general, the *k*-period-ahead forecast when the model is estimated over the period $t = 1,2,\ldots m+1$ is

$$\hat{s}_{m+k} = \hat{\alpha}_0 + \hat{\alpha}_1 (m_{m+k} - m^*_{m+k}) + \hat{\alpha}_2 (y_{m+k} - y^*_{m+k}) + \hat{\alpha}_3 (i_{m+k} - i^*_{m+k}) \quad (5)$$

Meese and Rogoff (1983) suggested that a possible reason for the failure to outperform the random walk is the use of models with constant parameters rather than allowing the parameters to vary over time—that is, estimating models in a time-varying parametric (TVP) framework. Schinasi and Swamy (1989) re-worked the Meese-Rogoff results by re-estimating the models in a TVP framework. They concluded that "while our results on fixed coefficient models support most of Meese and Rogoff conclusions, we find that when coefficients are allowed to change, an important subset of conventional models……can outperform forecasts of a random walk model".

Our proposition is that even models estimated in a TVP framework may not outperform the random walk if judged by criteria such as the root mean square error. To consider this proposition we also use TVP estimation. To estimate equation (1) in a TVP framework, the equation is re-written as

$$s_t = \mu_t + \phi_t + \alpha_{1t} (m_t - m^*_t) + \alpha_{2t} (y_t - y^*_y) + \alpha_{3t} (i_{,t} - i^*_t) + \varepsilon_t \quad (6)$$

where μ_t, ϕ_t and ε_t are the time series components of s_t: μ_t is the trend, ϕ_t is the cyclical component and ε_t is the random component. The trend, which represents the long-term movement of the dependent variable, is represented by the general specification

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{7}$$

$$\beta_t = \beta_{t-1} + \zeta_t \tag{8}$$

where $\eta_t \sim NID(0, \sigma^2_\eta)$, and $\zeta_t \sim NID(0, \sigma^2_\zeta)$. The cyclical component is specified as

$$\phi_t = \rho(\phi_{t-1} \cos\theta + \phi^*_{t-1} \sin\theta) + \omega_t \tag{8}$$

$$\phi^*_t = \rho(-\phi_{t-1} \sin\theta + \phi^*_{t-1} \cos\theta) + \omega^*_t \tag{9}$$

where ϕ^*_t appears by construction such that ω_t and ω^*_t are uncorrelated white noise disturbances with variances σ^2_ω and $\sigma^2_{\omega^*}$, respectively. The parameters $0 \le \theta \le \pi$ and $0 \le \rho \le 1$ are the frequency of the cycle and the damping factor on the amplitude respectively. The period of the cycle, which is the time taken by the cycle to go through its complete sequence of values, is $2\pi / \theta$ (Koopman et al., 2006). The model is estimated by maximum likelihood, using the Kalman filter to update the state vector. Once the model has been estimated, the forecasts are generated as in the case of OLS estimation.

Measures of Forecasting Accuracy

Once we have corresponding time series for the actual, s_t, and forecast, $\hat{s}_t$, for $t = m + k, \ldots n$, where k (the forecast horizon) takes the values 1, 3, 6 and 12, the root mean square error is calculated as follows

$$RMSE = \sqrt{\frac{1}{n - m - k - 2} \sum_{t=m+k}^{n} (\hat{s}_t - s_t)^2} \tag{10}$$

The root mean square error of the random walk is calculated as

$$RMSE = \sqrt{\frac{1}{n - m - k - 2} \sum_{t=m+k}^{n} (s_{t-1} - s_t)^2} \quad (11)$$

We also calculate direction accuracy as

$$DA = \frac{1}{n - m - k - 2} \sum_{t=m+k}^{n} a_t \quad (12)$$

where

$$a = \begin{cases} 1 & \text{if} \quad (\hat{s}_{t+1} - s_t)(s_{t+1} - s_t) > 0 \\ 0 & \quad\quad (\hat{s}_{t+1} - s_t)(s_{t+1} - s_t) < 0 \end{cases} \quad (13)$$

The other measure is the confusion rate, which is defined as

$$CR = 1 - DA \quad (14)$$

Moosa and Burns (2012) propose a measure of forecasting accuracy, the adjusted root mean square error (*ARMSE*), which combines the magnitude of the error and the ability of the model to predict direction correctly. It is constructed by adjusting the conventional *RMSE* to take into account the ability (of the model) to predict the direction of change. If two models have equal *RMSE*s, the model with the higher *CR* should have a higher *ARMSE*. Thus a possible formula for the adjusted *RMSE* is the following:

$$ARMSE = \sqrt{\frac{CR}{n - m - k - 2} \sum_{t=m+k}^{n} (\hat{s}_t - s_t)^2} \quad (15)$$

A nice property of *ARMSE* as defined by equation (15) is that it is not biased towards measures of either magnitude (*RMSE*) or direction (*CR*).

Meese and Rogoff (1983) reached their conclusion that exchange rate models cannot outperform the random walk in out-of-sample forecasting by comparing the numerical values of the RMSEs.[11] We go further by testing for the significance of difference between the root mean square error of the model and that of the random walk. This is the AGS test suggested by Ashley, Granger and Schmalensee (1980), which requires the estimation of the linear regression

$$D_t = \alpha_0 + \alpha_1 (M_t - \overline{M}) + u_t \tag{16}$$

where $D_t = w_{1t} - w_{2t}$, $M_t = w_{1t} + w_{2t}$, $\overline{M}$ is the mean of *M*, w_{1t} is the forecasting error at time *t* of the model with the higher RMSE, w_{2t} is the forecasting error at time *t* of the model with the lower RMSE. The null hypothesis of the equality of the two root mean square errors is $H_0 : \alpha_0 = \alpha_1 = 0$, which requires a Wald test of coefficient restrictions.

We also test for the significance of direction accuracy. To test the null hypothesis $H_0 : DA = \theta$ against the alternative $H_1 : DA > \theta$, we calculate the test statistic

$$z = \frac{DA - \theta}{\sqrt{(DA(1 - DA)/(m - n + k - 2)}} \tag{17}$$

The null hypothesis is rejected if *z* is greater than the critical value of the normal distribution. When we set $\theta = 0$, the test becomes that of the model against the random walk without drift. We will also set $\theta = 0.5$.

Measures of Profitability

---

[11] Unlike Meese and Rogoff (1983), Schinasi and Swamy (1989) highlight this point, stating that "because one cannot derive the exact finite sample distribution of the RMSE statistics (for even much simpler models) one cannot make probability statements about how significant these differences are". It seems therefore that neither Meese and Rogoff (1983) nor Schinasi and Swamy (1989) were aware that such a test was available in 1980 (the AGS test used in this paper).

Measures of profitability are calculated from the retrun on trading operations based on the random walk and the forecasts generated from the model. Under the random walk (without drift), the forecast change in the exchange rate is always zero, which means that a profitable strategy would be to go short on the low interest currency and long on the high interest currency. This operation represents the common carry trade, which in effect is also a forecasting-based strategy except that the forecasts are provided by the random walk without drift. Under this trading strategy, the period-to-period return is calculated as

$$\pi = \begin{cases} (i^* - i) + \dot{S}_{t+1} & \text{if} \quad i^* > i \\ (i - i^*) - \dot{S}_{t+1} & \phantom{\text{if}} \quad i^* < i \end{cases} \tag{18}$$

where $\dot{S}_{t+1}$ is the percentage change in the exchange rate. On the other hand, if forecasts are used for trading, the decision rule will be based on whether the forecast return, $\hat{\pi}$, is positive or negative.[12] In this case the realised return is calculated as

$$\pi = \begin{cases} (i^* - i) + \dot{S}_{t+1} & \text{if} \quad \hat{\pi} > 0 \\ (i - i^*) - \dot{S}_{t+1} & \phantom{\text{if}} \quad \hat{\pi} < 0 \end{cases} \tag{19}$$

When the drift factor is significant, profitability is calculated for what we may call a “modified carry trade” operation where the assumption is that the exchange rate is expected to change by the drift factor. In this case, the expected return is calculated by replacing the forecast percentage change in the exchange rate with the drift factor.

For a sample size *m*-*n*, where $t = m + 1, \cdots, n$, the mean value of the return is given by

$$\bar{\pi} = \frac{1}{m - n} \sum_{t=m+1}^{n} \pi_t \tag{20}$$

[12] The forecast return is calculated as the interest rate differential plus the forecast percentage change in the exchange rate.

The standard deviation of the rate of return is calculated as

$$SD = \sqrt{\frac{1}{n-m-1}\sum_{t=m+1}^{n}(\pi_t - \bar{\pi})^2} \tag{21}$$

The Sharpe ratio is used to measure the risk-adjusted return on carry trade. Following Burnside et al. (2010) and Gyntelberg and Remolona (2007), the Sharpe ratio is calculated as the ratio of the mean to the standard deviation of the rate of return.

$$SR = \frac{\bar{\pi}_t}{SD} \tag{22}$$

We conduct a conventional test of the difference between two means (mean returns), which has a t distribution.

Data and Empirical Results

The data source is International Financial Statistics (CD-ROM)—the sample covers the period March 1973-June 1981.[13] The construction of the composite variables follows the procedure used by Meese and Rogoff (1983). For example, they suggested that the expected inflation differential can be proxied by the long-term interest differential, the preceding 12-month inflation rate, or it can be extracted from an inflation rate autoregression. Our choice fell on the preceding 12-month CPI inflation rate. As in Meese and Rogoff (1983), the first forecasting period in this exercise is November 1976.

[13] Needless to say, the data set used here is bound to be different from the original data set used by Meese and Rogoff but this should not be detrimental to this exercise. As we are going to see later, the first set of results are qualitatively similar to those of Meese and Rogoff, leading to the same conclusion that none of the estimated models outperforms the random walk in out-of-sample forecasting as judged by the root mean square error.

Table 1 reports the results of testing the significance of the drift factor in the random walk process. The t statistics are calculated from a regression of the first log difference of the exchange rate on a constant (it is actually the t statistic of the constant term in this regression). We can see that the drift factor is significant at long horizons. Whenever the drift factor is significant, comparison is between the model and the random walk with drift. This makes a difference for the ability of the random walk to predict the direction of change. The random walk without drift fails consistently to predict the direction of change because it predicts a zero-change in the exchange rate when this can hardly be the case. On the other hand the random walk with drift predicts that the exchange rate changes by a percentage that is equal to the drift factor, which may be positive or negative. Hence the random walk predicts the direction correctly if the drift factor is positive (negative) when the exchange rate rises (falls).

Table 2 displays the root mean square errors of the estimated models and the corresponding random walk. Irrespective of the estimation method (OLS or TVP) and the underlying model, the root mean square error of the random walk is always lower than that of any of the models.[14] This finding confirms the results of Meese and Rogoff and our proposition—that even TVP estimation cannot overturn the Meese-Rogoff results if forecasting accuracy is measured by the RMSE. When we consider the results of direction accuracy, which are reported in Table 3, we can readily see that the random walk without drift is outperformed by all of the models, but the

[14] The only exception is the TVP Frenkel-Bilson model for the USD/DEM rate using a six month horizon.

random walk with drift at longer horizons outperforms the models.[15] We can also see that the direction of accuracy of the models improves when they are estimated in a TVP framework.

In Table 4, we can see that the AGS rejects the null of the equality of the root mean square errors of the random walk and the model in most cases, implying the superiority of the random walk, but we can also see that the models consistently beat the random walk without drift in terms of direction accuracy. The results of testing direction accuracy for the random walk are presented only when the drift factor is significant. We can see that (with one exception) the random walk with drift predicts the direction correctly more than 50 per cent of the times.

Table 5 reports the adjusted root mean square errors of the random walk and models, which shows a completely different story than that told by using the conventional RMSE. None of the OLS models beats the random walk for any of the forecasting horizons.[16] However, some TVP models outperform the random walk at horizons of one and three months for the USD/JPY rate, over horizons of one, three and six months for the USD/DEM rate and over a horizon of one month for the USD/GBP rate. The strange result is that none of the models outperforms the random walk at the 12 month horizon—strange because a consensus view has emerged that the forecasting power of the models (hence their ability to beat the random walk) is greater at long horizons. The difference here lies in the use of the random walk with drift over long horizons. Since the random walk with drift can predict the direction of

[15] Except for two cases where the TVP models outperform the random walk at the 6 month horizon.
[16] The one exception being the Dornbusch-Frankel model for the USD/DEM rate at the 6 month horizon.

change, it outperforms the models in terms of the ARMSE. Had we used the random walk without drift in all cases, the models would have looked in a better shape. Hence, it makes a difference for the results whether the random walk with or without drift is used.

Table 6 reports the results of evaluating forecasting power on the basis of the profitability. The OLS models exhibit a mixed performance with respect to beating the random walk in terms of profitability. However, all models estimated in a TVP framework outperform the random walk without drift. Where the drift factor was significant, models estimated using TVP still outperform the random walk in several cases. If profitability is indeed the ultimate test of forecasting accuracy, these results show that the Meese-Rogoff results can be overturned.

Conclusion

Since the publication of the Meese-Rogoff (1983) paper, failure to outperform the random walk in out-of-sample forecasting has become some sort of an undisputed fact of life. Several explanations have been put forward for this “puzzle” by Meese and Rogoff and others, including the use of time-varying parameters. In this paper we re-worked the Meese-Rogoff results to demonstrate that it is possible to overturn these results if forecasting accuracy is measured by metrics other than those that depend entirely on the magnitude of the error such as the root mean square error. We used either the random walk with drift or the random walk without drift, depending on whether or not the drift factor was significant. This actually gives the random walk greater power in predicting the direction of change.

This work provides some interesting results. By following the Meese-Rogoff methodology exactly (that is by using the RMSE as a measure of forecasting accuracy), we found that their results cannot be overturned, even if the models are estimated in a TVP framework. In all cases the random walk produces a lower RMSE than the model. In this case, the use of time-varying coefficients cannot explain the puzzle. However, when forecasting power is measured on the basis of direction accuracy only, the models outperform the random walk without drift, but the random walk with drift still outperforms some of the models. When both magnitude and direction are used, as represented by the adjusted root mean square error, some TVP models (and one OLS model) outperform the random walk except at a horizon of 12 months. In terms of the ultimate test of forecasting accuracy (based on profitability), most of the TVP models and several of the OLS models outperform the random walk. In more than one sense, therefore, the Meese-Rogoff results can be overturned. The mere use of TVP estimation does not overturn the results as long as the RMSE is used to measure forecasting accuracy.

While the inability to outperform the random walk in out-of-sample forecasting and the RMSE may look like a puzzle, it is not a puzzle at all. A simple explanation for the puzzle can be found in the argument put forward by Engel et al. (2007) that "beating the random walk is too strong a criterion for accepting a model". The random walk forecast for a particular point in time is the previous period's rate, which means that the forecasting error is the period-to-period change. If the exchange rate is not volatile, the error will be small, in which case it will be difficult for any model to beat the random walk. If, on the other hand, the rate is volatile, the RMSE of the random walk will increase but so will the root mean square error of the model as it is more difficult for any model to forecast a volatile than a less volatile exchange rate. It will be interesting to find out, by conducting Monte Carlo simulations, if by increasing volatility the RMSE of the model rises faster than the RMSE of the random walk. If this is the case, then the failure of exchange rate models to outperform the random walk in terms of the RMSE becomes rather intuitive.

References

Abhyankar, A., Sarno, L. and Valente, G. (2005) Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability *Journal of International Economics*, 66, 325-348.

Ashley, R., Granger, C.W.J. and Schmalensee, R. (1980) Advertising and Aggregate Consumption: An Analysis of Causality, *Econometrica*, 48, 1149-1167.

Bacchetta, P. and van Wincoop, E. (2006) Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?, *American Economic Review*, 96, 552-576.

Boothe, P. and Glassman, D. (1987) Comparing Exchange Rate Forecasting Models: Accuracy versus Profitability, *International Journal of Forecasting*, 3, 65-79.

Burnside, C., Eichenbaum, M., Kleshcelski, I. and Rebelo, S. (2010) Do Peso Problems Explain the Returns to the Carry Trade, *Review of Financial Studies*, 24, 853-891.

Cheung, Y.-W. and Chinn, M.D. (1998) Integration, Cointegration, and the Forecast Consistency of Structural Exchange Rate Models, *Journal of International Money and Finance*, 17, 813-830.

Cheung, Y-W, Chinn, M.D. and Pascual, A.G. (2005) Empirical Exchange Rate Models of the Nineties: Are They Fit to Survive?, *Journal of International Money and Finance*, 24, 1150-1175.

Christoffersen, P.F. and Diebold, F.X. (1998) Cointegration and Long-Horizon Forecasting, *Journal of Business and Economic Statistics*, 16, 450-458.

Clements, M.P. and Hendry, D.F. (2003) Evaluating a Model by Forecast Performance, Working Paper, Department of Economics, University of Warwick.

Cumby, R.E. and Modest, D.M. (1987 Testing for Market Timing Ability: A Framework for Forecast Evaluation, *Journal of Financial Economics*, 19, 169-189.

Engel, C. (1994) Can the Markov Switching Model Forecast Exchange Rates? *Journal of International Economics*, 36, 151-165.

Engel, C. and Hamilton, J.D. (1990) Long Swings in the Dollar: Are They in the Data and Do markets Know it?, *American Economic Review*, 80, 689-713.

Engel, C. and West, K.D. (2005) Exchange Rates and Fundamentals, *Journal of Political Economy*, 113, 485-517.

Engel, C., Mark, N. and West, K.D. (2007) Exchange Rate Models are not as Bad as you Think, NBER Working Papers, No 13318.

Engle, C., Rogers, J.H. and Rose, A.K. (2003) Empirical Exchange Rate Models, *Journal of International Economics*, 60, 1-2.

Faust, J., Rogers, J.H. and Wright, J.H. (2003) Exchange Rate Forecasting: The Errors We've Recently Made, *Journal of International Economics*, 60, 35-59.

Fildes, R. and Makridakis, S. (1995) The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting, *International Statistical Review*, 63, 289-308.

Frankel, J.A. and Rose, A.K. (1995) Empirical Research on Nominal Exchange Rates, *Handbook of International Economics*, vol 3, Amsterdam: Elsevier.

Gynelberg, J. and Remolona, E.M. (2007) Risk in Carry Trades: A Look at Target Currencies in Asia and the Pacific, *BIS Quarterly Review, December*, 73-82.

Koopman, S.J., Harvey, A.C., Doornik, J.A. and Shephard, N. (2006) *Structural Time Series Analyser, Modeller and Predictor*, London: Timberlake Consultants Ltd.

Leitch, G. and Tanner, J.E. (1991) Economic Forecast Evaluation: Profits versus the Conventional Error Measures, *American Economic Review*, 81, 580-590.

Li, M. (2011) An Evaluation of Exchange Rate Models by Carry Trade, *Journal of Economics and International Finance*, 3, 72-87.

MacDonald, R. and Taylor, M. (1994) The Monetary Model of the Exchange Rate: Long-Run Relationships, Short-Run Dynamics and How to Beat the Random Walk, *Journal of International Money and Finance*, 13, 276-290.

Meese, R.A. (1990) Currency Fluctuations in the Post Bretton Woods Era, *Journal of Economic Perspectives*, 4, 117-134.

Meese, R.A. and Rogoff, K. (1983) Empirical Exchange Rate Models of the Seventies: Do They Fit out of Sample? *Journal of International Economics*, 14, 3–24.

Moosa, I.A. and Burns, K. (2012) Can Exchange Rate Models Outperform the Random Walk? Magnitude, Direction and Profitability as Criteria, *Economia Internazionale* (forthcoming).

Sarno, L. and Taylor, M.P. (2002) *The Economics of Exchange Rates*, Cambridge: Cambridge University Press.

Schinasi, G.J. and Swamy, P.A.V.B. (1989) The Out-of-Sample Forecasting Performance of Exchange Rate Models when Coefficients are Allowed to Change, *Journal of International Money and Finance*, 8, 375-390.

Tashman, L.J. (2000) Out-of-Sample Tests of Forecasting Accuracy: An Analysis and Review, *International Journal of Forecasting*, 16, 437-450.

West, K.D., Edison, H. and Cho, D. (1993) A Utility Based Comparison of Some Models of Exchange Rate Volatility, *Journal of International Economics*, 35, 23-45.

Table 1: Statistical Significance of the Drift Factor

Exchange Rate	Estimated Drift	t Statistic
USD/JPY		
One month	0.005	0.95
Three months	0.016	1.72
Six months	0.035	2.26
Twelve months	0.070	2.72
USD/DEM		
One month	0.000	0.00
Three months	0.002	0.21
Six months	0.009	0.74
Twelve months	0.041	2.16
USD/GBP		
One month	-0.003	-0.86
Three months	-0.012	-1.69
Six months	-0.030	-3.23
Twelve months	-0.076	-7.25

Table 2: Root Mean Square Error

Exchange Rate/Model	1	3	6	12
USD/JPY				
Random Walk	3.70	6.83	11.39	17.08
Frenkel-Bilson (OLS)	12.72	15.46	18.18	23.01
Dornbusch-Frankel (OLS)	11.45	14.12	17.03	21.69
Hooper-Morton (OLS)	9.89	12.22	15.15	20.82
Frenkel-Bilson (TVP)	4.71	9.34	18.33	36.48
Dornbusch-Frankel (TVP)	4.87	9.24	18.11	34.95
Hooper-Morton (TVP)	8.76	13.43	24.29	42.33
USD/DEM				
Random Walk	3.89	5.99	8.97	12.47
Frenkel-Bilson (OLS)	12.52	13.76	15.07	18.44
Dornbusch-Frankel (OLS)	7.44	8.61	10.50	15.02
Hooper-Morton (OLS)	7.56	9.09	11.45	16.90
Frenkel-Bilson (TVP)	4.61	7.84	8.09	12.87
Dornbusch-Frankel (TVP)	5.78	9.80	10.75	20.58
Hooper-Morton (TVP)	5.60	12.46	21.09	32.60
USD/GBP				
Random Walk	3.02	5.42	6.47	6.94
Frenkel-Bilson (OLS)	15.42	17.42	19.99	24.84
Dornbusch-Frankel (OLS)	14.77	17.04	20.00	25.87
Hooper-Morton (OLS)	13.86	15.99	18.58	22.72
Frenkel-Bilson (TVP)	3.66	7.89	12.52	19.70
Dornbusch-Frankel (TVP)	3.80	8.38	13.55	21.69
Hooper-Morton (TVP)	8.03	15.77	24.03	33.81

Table 3: Direction Accuracy

Exchange Rate/Model	1	3	6	12
USD/JPY				
Random Walk	0.00	0.00	0.63	0.64
Frenkel-Bilson (OLS)	0.29	0.13	0.20	0.49
Dornbusch-Frankel (OLS)	0.27	0.15	0.29	0.47
Hooper-Morton (OLS)	0.30	0.19	0.33	0.56
Frenkel-Bilson (TVP)	0.60	0.73	0.78	0.51
Dornbusch-Frankel (TVP)	0.58	0.75	0.82	0.56
Hooper-Morton (TVP)	0.58	0.71	0.60	0.64
USD/DEM				
Random Walk	0.00	0.00	0.00	0.73
Frenkel-Bilson (OLS)	0.32	0.37	0.35	0.36
Dornbusch-Frankel (OLS)	0.36	0.39	0.37	0.36
Hooper-Morton (OLS)	0.36	0.39	0.35	0.36
Frenkel-Bilson (TVP)	0.40	0.60	0.64	0.67
Dornbusch-Frankel (TVP)	0.42	0.58	0.62	0.64
Hooper-Morton (TVP)	0.42	0.46	0.44	0.64
USD/GBP				
Random Walk	0.00	0.00	0.71	0.89
Frenkel-Bilson (OLS)	0.32	0.37	0.33	0.24
Dornbusch-Frankel (OLS)	0.32	0.39	0.31	0.24
Hooper-Morton (OLS)	0.34	0.39	0.35	0.27
Frenkel-Bilson (TVP)	0.42	0.42	0.42	0.54
Dornbusch-Frankel (TVP)	0.40	0.35	0.38	0.51
Hooper-Morton (TVP)	0.52	0.40	0.40	0.59

Table 4: Hypothesis Testing of Forecasting Accuracy

Exchange Rate	RW	FB(O)	DF(O)	HM(O)	FB(T)	DF(T)	HM(T)
USD/JPY (1)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		191.71	160.64	102.23	3.38	4.64	51.01
$z(DA = 0)$		4.73	4.53	4.94	8.66	8.31	8.31
$z(DA = 0.5)$		-3.55	-3.92	-3.20	1.44	1.15	1.15
USD/JPY (3)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		148.74	127.21	70.56	6.13	5.96	36.49
$z(DA = 0)$		2.84	3.06	3.50	11.37	12.00	10.80
$z(DA = 0.5)$		-8.10	-7.28	-5.96	3.57	4.00	3.18
USD/JPY (6)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		63.61	59.69	24.91	18.63	18.67	41.26
$z(DA = 0)$	9.27	3.53	4.61	5.05	12.55	14.43	8.22
$z(DA = 0.5)$	1.88	-5.47	-3.23	-2.52	4.48	5.65	1.37
USD/JPY (12)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		14.72	14.66	4.15	28.70	26.54	75.11
$z(DA = 0)$	9.03	6.56	6.27	7.50	6.41	7.10	8.35
$z(DA = 0.5)$	2.02	-0.15	-0.45	0.75	0.16	0.81	1.84
USD/DEM (1)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		96.32	30.62	29.81	16.17	3.22	7.13
$z(DA = 0)$		5.15	5.58	5.58	5.77	6.02	6.02
$z(DA = 0.5)$		-2.86	-2.23	-2.23	-1.44	-1.15	-1.15
USD/DEM (3)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		40.10	32.49	16.98	5.38	18.90	36.09
$z(DA = 0)$		5.64	5.86	5.86	8.56	8.20	6.37
$z(DA = 0.5)$		-1.97	-1.67	-1.67	1.48	1.17	-0.58
USD/DEM (6)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		96.32	30.62	29.81	16.17	3.22	7.13
$z(DA = 0)$		5.27	5.50	5.27	9.03	8.61	6.00
$z(DA = 0.5)$		-2.20	-1.88	-2.20	2.02	1.69	-0.75
USD/DEM (12)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		27.42	115.55	57.11	0.68	13.99	54.25
$z(DA = 0)$	11.12	4.98	4.98	4.98	8.83	8.35	8.35
$z(DA = 0.5)$	3.54	-2.02	-2.02	-2.02	2.21	1.84	1.84

Table 4 (continued)

Exchange Rate	RW	FB(O)	DF(O)	HM(O)	FB(T)	DF(T)	HM(T)
USD/GBP (1)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		494.44	437.43	373.74	2.53	4.31	79.54
$z(DA = 0)$		5.15	5.15	5.36	6.02	5.77	7.36
$z(DA = 0.5)$		-2.86	-2.86	-2.54	-1.15	-1.44	0.28
USD/GBP (3)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		200.73	183.52	165.14	13.18	18.21	105.87
$z(DA = 0)$		5.64	5.86	5.86	5.86	5.13	5.61
$z(DA = 0.5)$		-1.97	-1.67	-1.67	-1.17	-2.11	-1.48
USD/GBP (6)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		221.93	222.92	176.34	45.77	70.86	174.87
$z(DA = 0)$	11.06	5.05	4.83	5.27	5.73	5.23	5.48
$z(DA = 0.5)$	3.23	-2.52	-2.87	-2.20	-1.06	-1.69	-1.37
USD/GBP (12)							
$\chi^2(\alpha_0 = \alpha_1 = 0)$		286.07	322.92	221.65	82.54	116.09	524.49
$z(DA = 0)$	18.97	3.82	3.82	4.05	6.75	6.41	7.49
$z(DA = 0.5)$	8.30	-3.99	-3.99	-3.54	0.48	0.16	1.14

The critical values for the χ^2 and z test statistics are 5.99 and 1.96, respectively

Table 5: Adjusted Root Mean Square Error

Exchange Rate/Model	1	3	6	12
USD/JPY				
Random Walk	3.70	6.83	6.62	10.18
Frenkel-Bilson (OLS)	10.75	14.42	16.30	16.45
Dornbusch-Frankel (OLS)	9.80	13.04	14.31	15.84
Hooper-Morton (OLS)	8.25	11.03	12.37	13.88
Frenkel-Bilson (TVP)	2.98	4.86	8.64	25.47
Dornbusch-Frankel (TVP)	3.16	4.62	7.64	23.07
Hooper-Morton (TVP)	5.67	7.25	15.36	25.36
USD/DEM				
Random Walk	3.89	5.99	8.97	6.44
Frenkel-Bilson (OLS)	10.31	10.91	12.12	14.81
Dornbusch-Frankel (OLS)	5.96	6.73	8.32	12.06
Hooper-Morton (OLS)	6.06	7.11	9.21	13.56
Frenkel-Bilson (TVP)	3.57	4.93	4.83	7.43
Dornbusch-Frankel (TVP)	4.40	6.33	6.61	12.33
Hooper-Morton (TVP)	4.26	9.17	15.72	19.53
USD/GBP				
Random Walk	3.02	5.42	3.51	2.31
Frenkel-Bilson (OLS)	12.71	13.82	16.32	21.59
Dornbusch-Frankel (OLS)	12.17	13.32	16.57	22.48
Hooper-Morton (OLS)	11.27	12.50	14.94	19.46
Frenkel-Bilson (TVP)	2.79	6.03	9.52	13.38
Dornbusch-Frankel (TVP)	2.94	6.73	10.69	15.14
Hooper-Morton (TVP)	5.56	12.26	18.62	21.65

Table 6: Measures of Profitability

Exchange Rate	RW	FB(O)	DF(O)	HM(O)	FB(T)	DF(T)	HM(T)
USD/JPY (1)							
$\bar{\pi}$	0.25	-16.45	-16.35	-15.67	12.35	11.82	12.68
SD	45.22	42.06	42.10	42.37	45.86	46.00	45.76
SR	0.01	-0.39	-0.39	-0.37	0.27	0.26	0.28
t		112.34	111.60	106.70	-71.74	-68.84	-73.56
USD/JPY (3)							
$\bar{\pi}$	0.61	-18.01	-18.71	-17.41	14.82	14.87	12.60
SD	28.64	22.15	21.55	22.63	25.83	25.81	27.01
SR	0.02	-0.81	-0.87	-0.77	0.57	0.58	0.47
t		202.21	211.92	194.13	-133.8	-134.2	-113.0
USD/JPY (6)							
$\bar{\pi}$	3.95	-11.32	-6.60	-6.45	11.91	14.38	0.24
SD	24.01	21.49	23.41	23.45	22.30	20.75	25.34
SR	0.16	-0.53	-0.28	-0.27	0.53	0.69	0.01
t		120.34	65.19	63.55	-127.8	-154.3	-22.93
USD/JPY (12)							
$\bar{\pi}$	4.39	-3.37	0.20	-2.96	1.87	5.97	8.05
SD	20.06	20.21	20.49	20.27	20.30	19.47	18.69
SR	0.22	-0.17	0.01	-0.15	0.09	0.31	0.43
t		17.86	-19.02	13.56	-40.64	-76.65	-96.31
USD/DEM (1)							
$\bar{\pi}$	1.17	-1.48	4.46	0.88	7.77	7.18	9.28
SD	47.03	47.02	46.83	47.04	48.88	48.97	48.61
SR	0.02	-0.03	0.10	0.02	0.16	0.15	0.19
t		16.59	-20.56	1.85	-27.01	-24.04	-34.70
USD/DEM (3)							
$\bar{\pi}$	0.09	1.45	1.45	4.00	8.24	4.08	0.91
SD	24.26	24.22	24.22	23.92	24.25	25.30	25.62
SR	0.00	0.06	0.06	0.17	0.34	0.16	0.04
t		-15.54	-15.54	-45.07	-74.68	-35.09	-6.16
USD/DEM (6)							
$\bar{\pi}$	-2.09	1.16	0.13	2.68	9.99	7.24	-1.71
SD	17.69	17.77	17.81	17.60	15.87	17.33	18.73
SR	-0.12	0.07	0.01	0.15	0.63	0.42	-0.09
t		-46.72	-31.94	-69.98	-146.1	-108.3	-3.47
USD/DEM (12)							
$\bar{\pi}$	5.01	-2.13	-1.26	-0.01	5.60	4.53	2.22
SD	12.03	12.54	12.64	12.72	11.34	11.82	12.48
SR	0.42	-0.17	-0.10	0.00	0.49	0.38	0.18
t		-33.11	-47.18	-69.00	-120.6	-103.2	-68.91
USD/GBP (1)							
$\bar{\pi}$	-8.46	-2.51	-2.51	-1.54	5.96	1.41	11.33
SD	34.83	35.78	35.78	35.83	36.92	37.38	35.62
SR	-0.24	-0.07	-0.07	-0.04	0.16	0.04	0.32
t		-49.47	-49.47	-57.45	-100.1	-68.94	-138.8

Table 6 (continued)

Exchange Rate	RW	FB(O)	DF(O)	HM(O)	FB(T)	DF(T)	HM(T)
USD/GBP (3)							
$\bar{\pi}$	-5.77	-1.02	-0.17	0.01	0.26	-2.85	-1.76
SD	20.14	20.94	20.96	20.97	22.08	21.89	22.01
SR	-0.29	-0.05	-0.01	0.00	0.01	-0.13	-0.08
t		-64.38	-75.85	-78.18	-71.74	-38.37	-50.02
USD/GBP (6)							
$\bar{\pi}$	3.37	-0.85	-2.66	-0.43	-2.29	-2.59	-0.32
SD	13.06	13.46	13.22	13.48	14.10	14.05	14.28
SR	0.26	-0.06	-0.20	-0.03	-0.16	-0.18	-0.02
t		-41.96	-7.25	-49.89	-21.23	-16.78	-50.72
USD/GBP (12)							
$\bar{\pi}$	5.06	-3.21	-3.52	-2.92	2.39	0.70	2.73
SD	7.48	8.52	8.39	8.63	9.26	9.54	9.16
SR	0.68	-0.38	-0.42	-0.34	0.26	0.07	0.30
t		32.10	40.09	25.06	-71.83	-40.39	-78.29