CAN EXCHANGE RATE MODELS OUTPERFORM THE RANDOM WALK? MAGNITUDE, DIRECTION AND PROFITABILITY AS CRITERIA

1. 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. Several explanations have been put forward for this empirical regularity. In their original paper, Meese and Rogoff (1983) attributed their findings to some econometric problems, including simultaneous equations bias, sampling errors, stochastic movements in the true underlying parameters, model misspecification and failure to account for nonlinearities. Meese (1990) adds other explanations such as improper modelling of expectations and over-reliance on the representative agent paradigm – hence it appears that Meese questions the theoretical pillars of exchange rate models. As a matter of fact the majority of economists attribute the apparent failure to outperform the random walk to model inadequacy, in the sense that exchange rate models do not provide a valid representation of exchange rate behaviour (for example, Cheung and Chinn, 1999).

An important reason for the failure of these models to beat the random walk is typically, but not universally, overlooked. It is the use of quantitative measures of forecasting accuracy that are calculated from the forecasting errors, such as the root mean square error (which was used by Meese and Rogoff). It has been suggested that relying entirely on these measures may not be appropriate because a correct prediction of the direction of change may be more important than the magnitude of the error (for example, Engel, 1994). Others have suggested that the ultimate test of forecasting power is the ability to make profit by trading on the basis of the forecasts – hence the appropriate criterion would be profitability (for example, Leitch and Tanner, 1991). Cheung *et al*. (2005) argue that using criteria other than the mean square error does not boil down to "changing

the rules of the game" while advocating the use of other criteria by suggesting that minimising the mean square error may not be important from an economic standpoint. They argue against the use of the mean square error on the grounds 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 relationships *vis-a-vis* univariate prediction. Leitch and Tanner (1991) argue that the direction of change may be more relevant for profitability and economic concerns, while Cumby and Modest (1987) suggest that direction accuracy is also related to tests for market timing ability. Engel (1994) advocates the use of direction accuracy, which he describes as "not a bad proxy for a utility-based measure of forecasting performance". He also argues that it is impossible to think of important circumstances under which the direction of change is exactly the right criterion for maximising the welfare of the forecaster, citing as an example the case of central banks under a fixed exchange rate system.

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, as they find that the relative performance of the 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. He finds 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 criteria: accuracy (measured by the root mean square error) and profitability.

Leitch and Tanner (1991) note that economists are puzzled by the observation that profit-maximising firms buy professional forecasts when measures of forecasting accuracy indicate that a naïve (random walk) model provides free forecasts that are just as good (or as bad) as those provided by fee-charging forecasters. The explanation they present is that these measures bear very weak relation to the profit generated by acting on the basis of the forecasts, suggesting that the only substitute criterion for profit is a measure of direction accuracy. They find the relation between direction accuracy and profit to be almost as close as the relation between other measures (for example, between the root mean

square error and the mean absolute error). They further suggest that if profits are not observable, direction accuracy of the forecasts may be used as the evaluation criterion.

The objective of this paper is to examine these propositions by using simulated data, generated specifically to cover the two extremes of being good at predicting magnitude but not direction, and vice versa. The use of simulated data to represent models with extreme abilities is also motivated by the desire to determine whether the ability to forecast direction is more or less strongly related to profitability than the ability to forecast the magnitude of the error. A contribution of this paper is the development of a new measure of forecasting accuracy that takes into account errors of magnitude and direction by adjusting the conventional root mean square error. It is demonstrated that random walk forecasts are not as good as they typically appear when forecasting accuracy is measured in terms of direction accuracy and profitability.

2. The simulated data and conventional measures of forecasting power

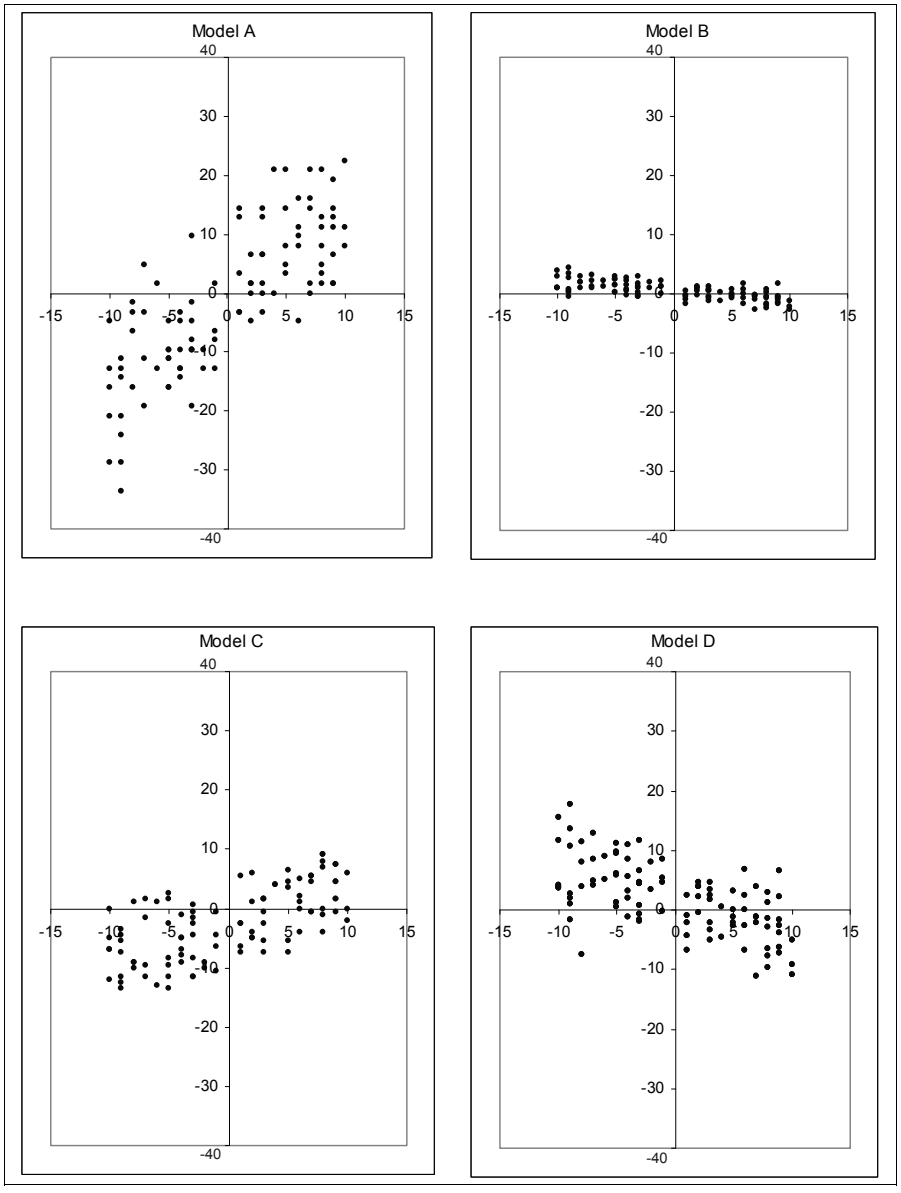
For the purpose of this exercise we generated 100 observations randomly on the actual percentage change in a hypothetical exchange rate. Then we generated eight sets of forecasts to represent a variety of models with various abilities to forecast magnitude and direction. The forecasting power of these hypothetical models (A, B, C, …, H) is represented by their prediction-realisation diagrams, which are plots of the actual on the predicted percentage changes in the exchange rate. The prediction-realisation diagrams, which are displayed in Figure 1, show that the best model in predicting the direction accurately is Model A because most of the points fall in the first and third quadrants. On the other hand, it seems that the best model in terms of the magnitude of the forecasting error is Model H, as the points cluster close to the 45 degree lines (not shown in the graphs).

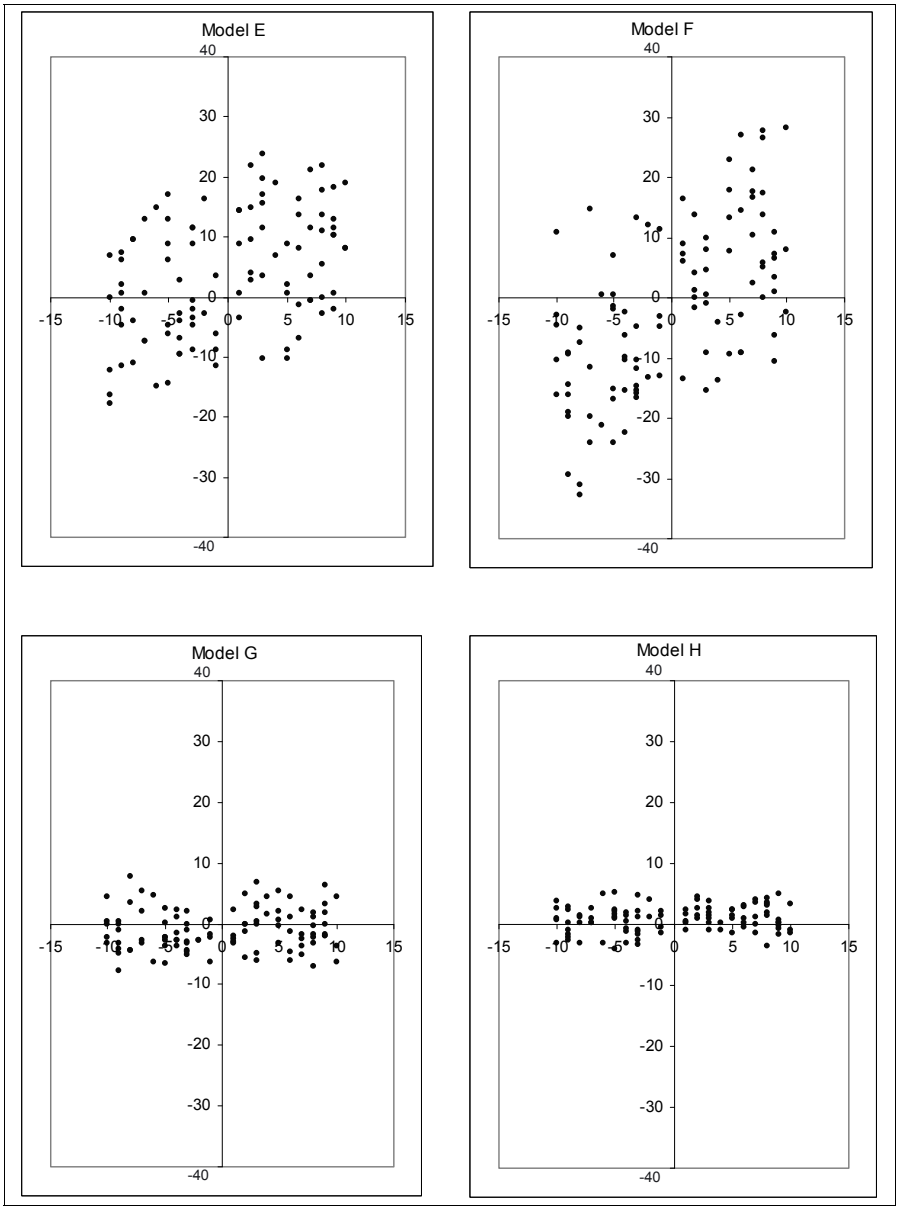
To confirm these observations we calculate the root mean square error (*RMSE*) and the confusion rate (*CR*). The root mean square error in percentage terms is calculated as

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} \tag{1}$$

where e is the percentage forecasting error, which is defined as

FIGURE 1 - *Prediction-realisation Diagrams of the Models*




Model E
Model F
Model G
Model H
40
30
20
10
0
-10
-20
-30
-40
-15
-10
-5
0
5
10
15

$$e_t = \frac{\hat{S}_t - S_t}{S_t} \tag{2}$$

where S_t and $\hat{S}_t$ are the actual and predicted values of the exchange rate, respectively.

The confusion rate is the percentage of occasions on which the model predicts the direction of change incorrectly. It is calculated as

$$CR = 1 - DA \tag{3}$$

where DA is direction accuracy, which is calculated as

$$DA = \frac{1}{n-1}\sum_{t=1}^{n} a^2 \tag{4}$$

where

$$a = \begin{cases} 1 & \text{if } \begin{cases} (\hat{S}_{t+1} - S_t)(S_{t+1} - S_t) > 0 \\ (\hat{S}_{t+1} - S_t)(S_{t+1} - S_t) < 0 \end{cases} \\ 0 \end{cases} \tag{5}$$

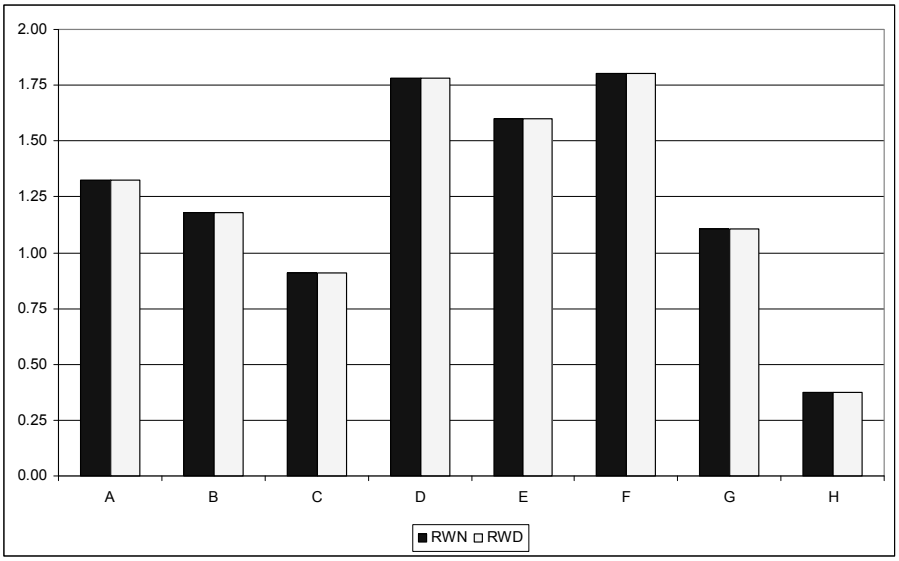
Table 1 shows the *RMSE* and *CR* of each of the eight models as well as their rankings according to the two criteria. According to the *RMSE* the best model is H (ranked 1) and the worst is F (ranked 8). By using *CR* as a criterion, the best model is A (ranked 1) and the worst is B (ranked 8). Hence, different criteria produce different rankings. Consider now the rank correlations resulting from the use of the two criteria. The rank correlation between *RMSE* and *CR* is -0.21. Negative correlation between *RMSE* and *CR* is not necessarily the case – it is due to the particular characteristics of the simulated data set. This result does not mean that a model that is good in predicting direction is necessarily bad in predicting magnitude, and vice versa. This result is obtained here by construction.

TABLE 1 - *Ranking of Models by Measures of Magnitude and Direction*

Model	*RMSE*	Ranking	*CR*	Ranking
A	8.30	5	0.09	1
B	7.38	4	0.77	8
C	5.70	2	0.26	3
D	11.16	7	0.74	7
E	10.04	6	0.30	4
F	11.31	8	0.22	2
G	6.95	3	0.41	5
H	2.33	1	0.66	6

Consider now the performance of the eight models against the random walk without drift, which always predicts zero change in the exchange rate, and the random walk with drift, which consistently predicts a positive or a negative change in the exchange rate (depending on the sign of the drift factor). Following Meese and Rogoff (1983), the drift factor is estimated as the average value of the percentage change in the exchange rate, which turns out to be 0.07 per cent. The results are presented graphically in Figure 2 and Figure 3, which display the ratios of the *RMSE*s and *CR*s to those of the random walk with drift (RWD) and without drift (RWN). In Figure 2 we can readily observe that only two models (C and H) outperform the random walk. The reason why the results are identical whether we use the random walk with or without drift is simple: the drift factor is numerically small and statistically insignificant, which makes the *RMSE* of the random walk with drift equal to that of the random walk without drift.

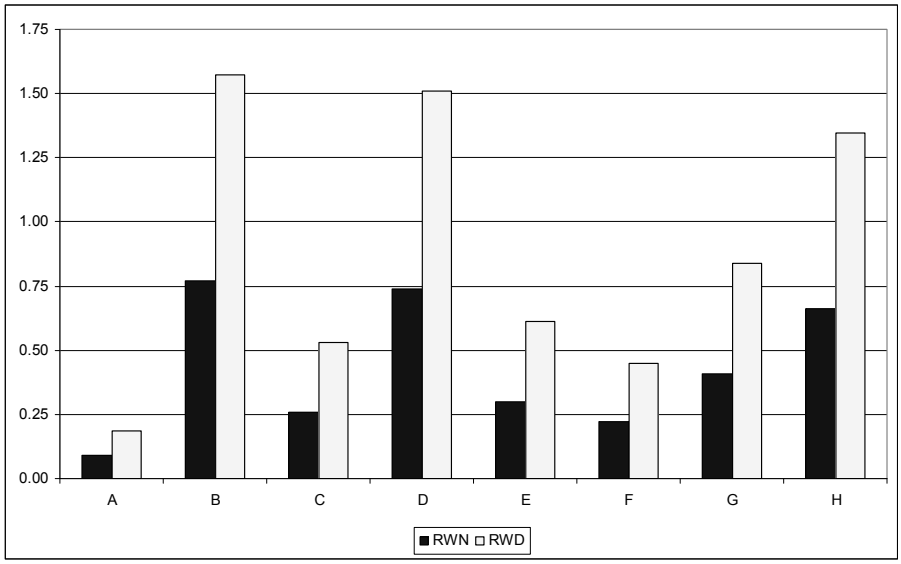
FIGURE 2 - *Ratio of RMSE Relative to Random Walk*



According to *CR*, on the other hand, all models outperform the random walk without drift and five of the eight models outperform the random walk with drift (only B, D and H fail to do so). The reason why the random walk without drift appears at the bottom of the league is simple: it always predicts no change, so it has a

CR of 1 – that is, it fails to predict the direction of change on each occasion. Since the drift factor has a positive sign, the random walk with drift always predicts a positive change – this is why it has a *CR* of 0.49 rather than 1. If, however, we consider the significance, rather than the value, of the drift factor we reach the conclusion that all models outperform the random walk with drift. Actually in this case there is no drift factor, which is quite consistent with the empirical observation that exchange rates move as random walk with little or no drift (see, for example, Moosa, 2000).

FIGURE 3 - *Ratio of CR Relative to Random Walk*



It seems therefore that it is justifiable to advocate the use of direction accuracy as a criterion. It also seems plausible to speculate that had Meese and Rogoff used direction accuracy as a criterion they might have been able to outperform the random walk.

3. THE ADJUSTED ROOT MEAN SQUARE ERROR (ARMSE)

As we argued earlier, the relative importance of magnitude and direction depends on the underlying decision making situation. For example, Moosa (2006) demonstrates that the notion of forecasting

accuracy is not unique and that it should be defined and measured depending on the underlying decision rule. Specifically he shows that in some situations what matters entirely is the direction of change (for example, intra-day trading where the interest rate factor is negligible). In other situations what matters is the magnitude of change (for example, betting on market volatility by using straddles and strangles). There are also situations where both the magnitude and direction do matter (for example, carry trade). Thus we may emphasise measures of forecasting accuracy that are based on magnitude or direction or both, depending on the underlying situation – using the *RMSE* in some cases, the *CR* in others and both when the underlying situation requires the prediction of magnitude and direction.

In an exercise (such as the Meese-Rogoff exercise) where the objective is to assess the forecasting power of various models without reference to the underlying situation, it may be useful to devise a measure of forecasting accuracy that takes into account both magnitude and direction without bias to either. Such a measure would also be useful for situations requiring the prediction of both magnitude and direction. The following is a proposal to come up with such a measure, which we call the adjusted root mean square error (*ARMSE*).

The *ARMSE* can be constructed simply by adjusting the conventional *RMSE* to take into account the ability or otherwise 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:

$$ARMSE = \sqrt{\frac{CR}{n}\sum_{t=1}^{n} e_t^2} \tag{6}$$

A nice property of *ARMSE* as defined by equation (6) is that it is not biased towards measures of either magnitude (*RMSE*) or direction (*CR*). The rank correlation between *ARMSE* and *RMSE* and between *ARMSE* and *CR* are close in value at 0.571 and 0.551, respectively[1]. The *ARMSE* can be modified to assign more weight to the prediction of direction or the magnitude of change (see appendix).

[1] The models are ranked in terms of *RMSE*, *CR* and *ARMSE*. Rank correlations are then calculated as the correlation coefficients between the *RMSE* and *CR*, on one hand, and the *ARMSE*, on the other.

Figure 4 shows the ranking of the eight models according to the three criteria (*RMSE*, *CR* and *ARMSE*). The best models (ranked 1) are H, A and H, respectively whereas the worst models (ranked 8) are F, B and D, respectively. The *RMSE* and *ARMSE* produce the same ranking once only, selecting model H as the best. Likewise, the *CR* and *ARMSE* produce the same ranking once, putting model D at number 7 in the ranking.

FIGURE 4 - *Ranking of Models by the three Criteria (1: the Best, 8: the Worst)*

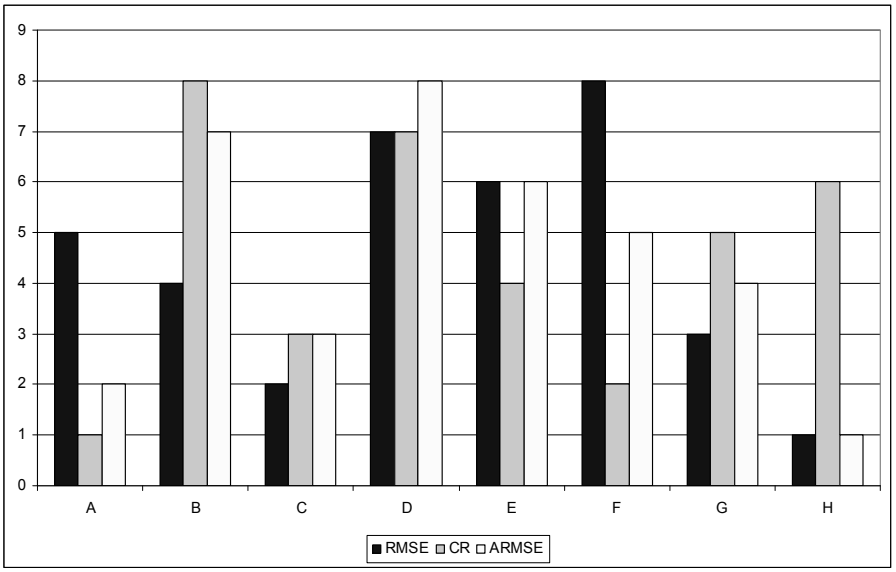
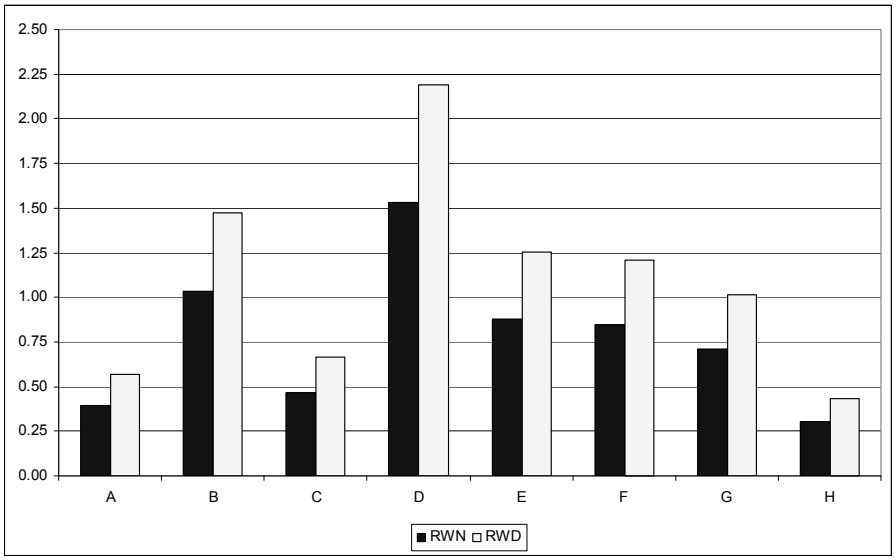


Figure 5 shows the ratio of the *ARMSE* of the model relative to that of the random walk with and without drift. Only models B and D fail to outperform the random walk without drift while five models fail to outperform the random walk with drift (B, D, E, F and G). If, however, we consider the statistical significance of the drift factor rather than its numerical value, the random walk with drift will turn out to be as inferior as the random walk without drift (in fact they become identical). These results suggest that Meese and Rogoff could have outperformed the random walk by evaluating forecasting power according a criterion that takes into account both magnitude and direction.

FIGURE 5 - *Ratio of ARMSE Relative to Random Walk*



4. PROFITABILITY AS A CRITERION

Leitch and Tanner (1991) use profitability to explain the apparent puzzle as to why profit-maximising firms pay for forecasts when the measures of forecasting accuracy based on the magnitude of the error are so poor and when these forecasters cannot outperform the random walk. They argue that measures of forecasting accuracy based on the magnitude of the error bear no predictable relation to the profitability of operations based on the forecasts, which means that these criteria are "unpredictable indicators of profits".

To use profitability as a criterion we simulate the interest rate differential such that it assumes values falling between two and four percentage points. Profitability is measured on the basis of the rate of return on an investment strategy that involves taking a short position on one currency and a long position on the other. The decision rule works as follows. First define the expected return, π^e, based on the forecast percentage change in the exchange rate as

$$\pi^e = (i_{y,t} - i_{x,t}) + \hat{S}^e_{t+1} \tag{7}$$

where $\hat{S}^e$ is the expected percentage change in the exchange rate, i_y is the interest rate on currency y and i_x is the interest rate on currency x, hence $i_y - i_x$ is the interest rate differential. The decision rule is to

take a short position on x and a long position on y if the expected return is positive, and vice versa. Hence the realised return is given by

$$\pi = \begin{cases} (i_{y,t} - i_{x,t}) + \hat{S}_{t+1} & \text{if} \quad \pi^e > 0 \\ (i_{x,t} - i_{y,t}) - \hat{S}_{t+1} & \quad\quad \pi^e < 0 \end{cases} \tag{8}$$

To evaluate the forecasting performance of the random walk (without drift) we basically conduct a carry trade operation, taking a long position on the high interest currency and a short position on the low intrest currency since the random walk implies that $\hat{S}^e_{t+1}=0$. A carry trade operation is implicitly based on the assumption of random walk without drift (see, for example, Moosa, 2004). The return on carry trade is therefore given by

$$\pi = \begin{cases} (i_{y,t} - i_{x,t}) + \hat{S}_{t+1} & \text{if} \quad i_{y,t} > i_{x,t} \\ (i_{x,t} - i_{y,t}) - \hat{S}_{t+1} & \quad\quad i_{y,t} < i_{x,t} \end{cases} \tag{9}$$

Once the period-to-period rates of return have been calculated, we can quantify the mean return, the standard deviation and the Sharpe ratio (the ratio of the mean to the standard deviation as a measure of risk-adjusted return).

Table 2 displays the ranking of models by profitability, using the Sharpe ratio as a measure of risk-adjusted return. The results show that four models (A, C, D and F) provide forecasts that enhance the profitability of carry trade, which means that the four models outperform the random walk in terms of profitability. The results also show that the rank correlation between *SR* and *CR* (0.76) is

TABLE 2 - *Ranking by Profitability*

Model	Mean	SD	*SR*	Ranking (*SR*)
A	4.89	5.12	0.96	1
B	2.72	6.52	0.42	6
C	3.79	5.95	0.64	2
D	0.18	7.14	0.02	8
E	3.73	6.02	0.62	3
F	3.54	6.03	0.59	4
G	2.02	6.82	0.30	7
H	-0.20	6.97	-0.03	9
RW	3.10	6.37	0.49	5

higher than that between *SR* and either *RMSE* (-0.19) or *ARMSE* (0.17). This means that the ability to predict the direction of change is more strongly related to profitability than the ability to predict the magnitude. This, however, is not a general result. It is dependent on the nature of the trading operation and the simulated data – in this case giving more weight to the percentage change in the exchange rate and hence to the ability to predict the direction of change. Still, by using profitability as a criterion, it can be demonstrated that the random walk can be outperformed.

5. Concluding remarks

Failure to outperform the random walk in out-of-sample forecasting has been accepted as something like an undisputed fact of life following the work of Meese and Rogoff (1983) who were the first to reveal this finding. While many reasons have been presented to explain this failure, a simple explanation is the use of measures of forecasting accuracy, such as the root mean square error, that depend entirely on the magnitude of the forecasting error. By using simulated data representing the forecasts of eight models with varying ability to forecast the magnitude and direction of change, it was demonstrated that it is possible to outperform the random walk (with and without drift). The results show that the random walk can be outperformed if the forecasting power is judged by measures of direction accuracy, by adjusting the root mean square error to take into account direction accuracy, and by using measures of the risk-adjusted return obtained from a trading strategy based on the forecasts.

Even the seemingly better performance of the random walk with drift (when *CR* and *ARMSE* are used as criteria) is due to the use of the numerical value of the estimated drift factor, which makes it better than the random walk without drift in predicting direction. Engel and Hamilton (1990) argue that (in theory) the random walk with drift is a more reasonable standard of comparison when the drift factor is significantly different from zero. If we follow this proposition, and given that the drift factor estimated in this study is not statistically significant, the results for the random walk with drift and without drift are identical. Because the drift factor invariably turns out to be insignificant, Engel and Hamilton suggest (based on their results) that it does not make much difference whether the random walk with or without drift is used as the standard for out-of-sample forecasting.

If this consideration is taken into account, our results are even stronger in refuting the proposition of the unbeatable random walk. It seems that Meese and Rogoff (1983) could have outperformed the random walk, had they judged forecasting accuracy by criteria other than the root mean square error.

Imad Moosa
RMIT University, School of Economics, Finance and Marketing, Melbourne, Australia

Kelly Burns
RMIT University, School of Economics, Finance and Marketing, Melbourne, Australia

APPENDIX

A General Formula for the ARMSE

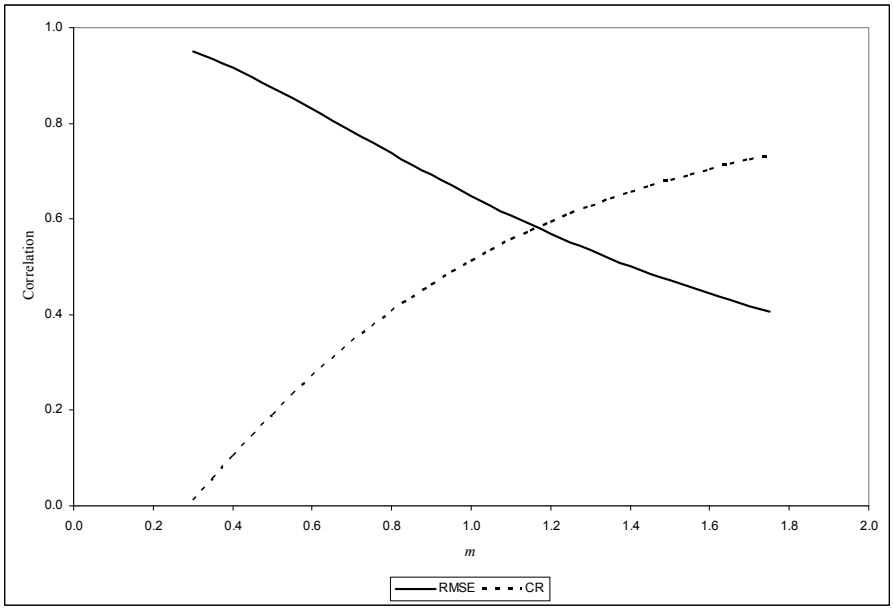
A general formula for the adjusted root men square error can be written as:

$$ARMSE = \sqrt{\frac{CR^m}{n}\sum_{t=1}^{n} e_t^2} \tag{10}$$

In our calculations we set $m=1$, which gives us a measure of forecasting accuracy that is not biased to either the ability to forecast the magnitude of the error or the direction of change. However, there are some circumstances under which the decision maker is more concerned with forecasting the direction rather than the magnitude or vice versa (see, for example, Moosa, 2006). It is possible to modify the measure in such a way as to favour either magnitude or the direction by changing the value of m. To emphasise the ability of the model to predict the direction of change (hence making the *ARMSE* more biased towards direction), we increase the value of m, and vice versa.

In the following exercise we try values for m ranging from 0.30 to 1.75 at intervals of 0.05. As we change the value of m, we

FIGURE 6 - *Rank Correlation as a Function of m*



calculate the corresponding rank correlation of the *ARMSE*, on the one hand, and the *RMSE* and *CR*, on the other. It turns out that rank correlations are related to m by the nonlinear relations:

$$\rho(ARMSE,\ RMSE) = 0.0804m^2 - 0.5563m + 1.1267 \tag{11}$$

and

$$\rho(ARMSE,\ CR) = -0.2926m^2 + 1.0831m - 0.279 \tag{12}$$

where $\rho(ARMSE,\ RMSE)$ is the rank correlation between *ARMSE* and *RMSE*, while $\rho(ARMSE,\ CR)$ is the rank correlation between *ARMSE* and *CR*. These curves are plotted in Figure 6, which shows that the rank correlation is a decreasing function of m for *RMSE* and an increasing function of m for *CR*. The point of intersection of the two curves represents the value of m that gives exactly no bias towards either magnitude or direction. If equations (10) and (11) are solved for m, we obtain $m = 1.164567$, which is the exact value that makes the *ARMSE* not biased to either magnitude or direction.

REFERENCES

Abhyankar, A., L. Sarno and G. Valente (2005), "Exchange Rates and Fundamentals: Evidence on the Economic Value of Predictability", *Journal of International Economics*, 66(2), 325-348.

Boothe, P. and D. Glassman (1987), "Comparing Exchange Rate Forecasting Models: Accuracy versus Profitability", *International Journal of Forecasting*, 3(1), 65-79.

Cheung, Y-W. and M.D. Chinn (1999), "Macroeconomic Implications of the Beliefs and Behavior of Foreign Exchange Traders", NBER Working Paper No. 7417.

Cheung, Y-W., M.D. Chinn and A.G. Pascual (2005), "Empirical Exchange Rate Models of the Nineties: Are any Fit to Survive?", *Journal of International Money and Finance*, 24(7), 1150-1175.

Christoffersen, P.F. and F.X. Diebold (1998), "Cointegration and Long-Horizon Forecasting", *Journal of Business and Economic Statistics*, 16(4), 450-458.

Cumby, R.E. and D.M. Modest (1987), "Testing for Market Timing Ability: A Framework for Forecast Evaluation", *Journal of Financial Economics*, 19(1), 169-189.

Engel, C. (1994), "Can the Markov Switching Model Forecast Exchange Rates?", *Journal of International Economics*, 36(1-2), 151-165.

Engel, C. and J.D. Hamilton (1990), "Long Swings in the Dollar: Are they in the Data and do Markets Know it?", *American Economic Review*, 80(4), 689-713.

Leitch, G. and J.E. Tanner (1991), "Economic Forecast Evaluation: Profits versus the Conventional Error Measures", *American Economic Review*, 81(3), 580-590.

Li, M. (2011), "An Evaluation of Exchange Rate Models by Carry Trade", *Journal of Economics and International Finance*, 3(2), 72-87.

Meese, R.A. (1990), "Currency Fluctuations in the post Bretton Woods Era", *Journal of Economic Perspectives*, 4(1), 117-134.

Meese, R.A. and K. Rogoff (1983), "Empirical Exchange Rate Models of the Seventies: Do they Fit out of Sample?", *Journal of International Economics*, 14(1-2), 3-24.

Moosa, I.A. (2000), Exchange Rate Forecasting: Techniques and Applications, Macmillan: London.

Moosa, I.A. (2004), "What is Wrong with Market-Based Forecasting of Exchange Rates?", *International Journal of Business and Economics*, 3(2), 107-121.

Moosa, I.A. (2006), "On the Notion of Forecasting Accuracy as Applied to Financial Decision Rules Involving Exchange Rates", *Forecasting Letters*, 1(2), 5-9.

West, K.D., H. Edison and D. Cho (1993), "A Utility Based Comparison of Some Models of Exchange Rate Volatility", *Journal of International Economics*, 35(1-2), 23-45.

ABSTRACT

While many explanations have been put forward for the failure of exchange rate models to outperform the random walk in out-of-sample forecasting, a simple explanation is the use of measures of forecasting accuracy that depend entirely on the magnitude of the forecasting error. By using simulated data representing the forecasts of eight models, it is demonstrated that the random walk can be outperformed if forecasting power is judged by measures of direction accuracy, by adjusting the root mean square error to take into account direction accuracy, and by using the risk-adjusted return obtained from a trading strategy based on the forecasts.

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

JEL Classification: C53, F31, F37

RIASSUNTO

I modelli di tasso cambio possono battere la "random walk"? Grandezza, direzione e profittabilità come criteri di comparazione

Sono state proposte diverse spiegazioni della incapacità dei modelli di previsione dei tassi di cambio di battere la *random walk*. Qui si avanza l'ipotesi di considerare come unica metodologia valida di comparazione la dimensione dell'errore di previsione. Utilizzando le previsioni di otto modelli si dimostra che la *random walk* può essere superata se la capacità di previsione è valutata attraverso misure di accuratezza direzionale, tramite lo scarto quadratico medio dell'errore e il *risk-adjusted return* ottenuto da una strategia di trading basata sulle previsioni.