Santamaria, D. (2012) Evaluating the predictiveness and profitability of foreign exchange forecasting models. The Journal of Prediction Markets, 6 (1). pp. 56-75. ISSN 1750-676X.

EVALUATING THE PREDICTIVENESS AND PROFITABILITY OF FOREIGN EXCHANGE RATE FORECASTING MODELS

Daniel SANTAMARIA

The Business School,
Faculty of Business and Management,
Canterbury Christ Church University
Canterbury, Kent,
Tel: 0044(0)1227 767700 Ext 3188
E-mail: daniel.santamaria@canterbury.ac.uk

ABSTRACT

This paper evaluates the performance of two competing currency models as a forecasting and trading tool in fund management. A dynamic vector error correction model is utilized to construct a currency forecasting and fair value forecasting model for the Euro-Dollar exchange rate. Emphasis is placed on robustness testing model performance by changing its specification and how they perform across different time periods. Based on the accuracy of the forecasts the fair value model outperforms the currency forecasting model; a finding that is not supported using directional forecasts. This is robust to changes in model specification and across different time spans that cover pre-and current financial crisis periods. It is also discovered that the evaluation criteria used and prevailing market conditions determines whether model performance translates into value added in a currency fund.

Keywords: exchange rate, model forecasts, fair value models, trading signals, model performance.

JEL Classification Code: F37 G15

INTRODUCTION

Much effort has gone into developing models that predict exchange rates. As a result, this poses a number of important generic questions that form the basis of this paper; to what extent can such efforts be expected to be successful? What is the value of these forecasts, and how can one evaluate it as a trading tool? This leads to the key issue amongst market participants of "getting the currency right". To achieve this on a consistent basis is a challenging task. Currencies are subject to short run speculative forces that can cause deviations from the long run path. Moreover, forecasts can go astray should one's interpretation of fundamental forces be flawed. Although much effort has gone into analyzing the performance of currency forecasting and fair value models against a random walk, evaluating both classes of models together has received little attention in the literature. The main paper on this issue by Evans and Lyons (2005) attempts to evaluate the performance of micro and macro models to find that fair value models, whilst useful in accounting for longer term trends have tended to perform poorly when explaining short and medium term currency movements. The poor performance of fair value models can be attributable to the fact that macroeconomic variables such as inflation, GDP differentials and relative money supply do not exhibit the same variability as exchange rates. This poses a challenge addressed in this study of developing a fair value model based on fundamentals that govern foreign exchange rates between major economic announcements.

This paper aims to provide a comprehensive evaluation of currency models and their usefulness as a forecasting and trading tool in the management of currency funds. This is investigated on the Euro-Dollar (EUR) exchange rate, although the same analysis can be applied to other currencies. To this end, a dynamic vector error correction model is utilized within the co-integration framework to develop two competing model types; a currency forecasting model and a fair value forecasting model. Both models are estimated in a rolling window that enables dynamic shifts in the relationships of the variables as it moves through time. In doing so, a number of important issues are addressed that make a significant contribution in this area: firstly, estimating both currency models in a rolling window will generate out-of-sample forecasts based on a model that allows dynamic shifts in the relationships of the variables as it moves through time. Secondly, a two step procedure is adopted to test how robust the performance of the currency forecasting model is in relation to the fair value model. Step one analyzes whether model performance is affected by changes in model specification of the error correction term. Step two involves performing a robustness test on model forecasts by splitting out-of-sample performance into pre and current financial crisis periods. Thirdly, this paper makes a further contribution by developing a comprehensive approach to model evaluation. Consistent with the literature, the model evaluation process will begin by focusing on forecasting accuracy as a measure of performance of both models. This is followed by computing the performance of model generated trading signals to determine whether it is consistent with the accuracy of the forecasts. The final stage investigates whether or not model performance translates into value added in an investment fund. Combining all three stages, the paper addresses the question of which model evaluation criteria is more important in determining the performance of a model driven fund; point forecasting accuracy or trading signals generated by directional forecasts. This leads to a fundamental question addressed in this paper on the robustness of the model evaluation results to changes in model specification and market conditions. In the light of the current financial market crisis, robustness analysis of this nature is of added importance.

This study is motivated by the ongoing debate amongst academics and practitioners on whether or not currency models can lead to a profit by forecasting short to medium term currency movements. If currency managers cannot profit from the use of currency models, then the foreign exchange market may be random and price efficient (Liu 2007 and Pukthuanthong-Le and Thomas 2008). The notion of applying and evaluating model performance beyond its ability to capture currency trends introduces the prospect of determining if both models add value as a trading tool in fund management. Owing to current financial market conditions, comparing the usefulness of both model types to previous studies is of added interest as they run their models in normal market conditions (Hamelink (2001)).

The remainder of the paper is as follows; next, the data and methodology used in this study. Section 3 presents the empirical results followed by an analysis of model performance is provided in section 4. Finally, section 5 summarizes and concludes the paper.

DATA AND METHODOLOGY

Data

The database comprises of two groups of independent variables that define the currency forecasting (CF hereafter) model and fair value forecasting (FVF) model. For the CF model, weekly spot rates for the Euro/Dollar (EUR), Dollar/Japanese Yen (JPY), Pound Sterling/Dollar (GBP) and Dollar/Canadian Dollar (CAD) are used from 10th February 1995 to 13th February 2009.[1,2] On the other hand, the variables used for the FVF model includes weekly interest rate sensitive two year government bonds for the euro-zone and US, longer term 10 year US government bond yields along with closing prices of Brent Crude per barrel.

---

[1] Prior to the introduction of the EURO in 1999, synthetic values calculated by Datastream International are used.

[2] The currencies chosen represent the most weight on the USD Index (5 7.6% for the EURUSD, 13.6% for the USDJPY, 11.9% for GBPUSD and 9.1% for the USDCAD) and are amongst the most important in terms of trade by the Federal Reserve Statistical Release. See www.federalresevre.gov/release/H10/Weights/ for information.

All series was downloaded from Datastream International and consist of the Friday's close of European trading with the start date determined by data availability.[3]

In developing the FVF model, the inclusion of such variables allows one to model the dynamics that govern foreign exchange rate movements between major macroeconomic news announcements. With interest rate data, two year interest rate differentials (IRD) was computed between the US and EU. The intuition behind using interest rate data lies in the dynamics of the currency market that is partially explained by IRD's when pricing forward exchange rates. Based on interest rate parity, traders formulate expectations of future currency rates and attach a premium on the current exchange rate futures contracts. The objective of using US 10 year government bond yields is to compute 2/10 year Treasury spreads as a proxy for market expectations of future economic conditions. It has been documented that there exists a long run relationship between exchange rates and long-short interest rate differentials (see for example Amoateng 1995). In testing the proxy for market expectations, Estrella and Mishkin (1996) finds that the yield spread between ten-year Treasuries and three-month bills outperforms other macroeconomic indicators in forecasting recessions two to six quarters ahead from 1971 to 1995. In a more recent paper, Wright (2006) uses a probit model to find more predictive power in the shape of the yield curve on the likelihood of a recession than the term structure. Finally, introducing Brent Crude oil prices is a useful variable in explaining the fair value of the EUR. For instance, in a recent paper, Coudert et al. (2008) use co-integration and causality analysis on the US effective exchange rate and crude oil prices between 1974 and 2004. They present evidence that the causality is running from oil prices to the exchange rate and that it is sensitive to the parameters used in the VAR estimation.

Methodology

In this study, a dynamic vector error correction model (henceforth VEC) is used within a co-integration framework to develop the CF and FVF model for the EUR. For the purpose of this paper, both models are used to generate out-of-sample forecasts. Previous studies have noted the usefulness of the co-integration approach. For instance, Tong (2001) provides evidence that using co-integrating models reduces the forecasting error associated with the random walk model and introduces the prospect of efficiency gains in forecasting. Robustness tests are performed on these models by changing the specification of the error correction term using a stepwise procedure.

Within this econometric framework, four steps are identified in the development of the forecasting models; the first step is to test each of the series for a unit root. To this effect, the Augmented Dickey Fuller (ADF) and Phillips Perron (PP) tests are utilized on the EUR rate and the independent variables used in the study. Using both methods serves the useful purpose of acting as a robustness test as both determine whether or not the null hypothesis that the series are integrated of order I(1) holds. Whist all series must be integrated of the same order, this condition does not necessarily imply that all series are co-integrated. A lack of co-integration of this nature implies no long run equilibrium of the variables with the implication that they can deviate arbitrary far from each other. Only stationary series determined by the ADF and PP statistic are applied in the forecasting process[4].

The next step involves testing for co-integration using the Johansen (1988) and Johansen and Juselius (1990) maximum likelihood approach. Given that several different variables are used to determine a number of co-integrating relations with the EUR, a system's based approach is more appropriate. Finding any equilibrium relationship with the EUR and the set of independent variables within both models implies that their stochastic trends are linked. However, it is affected by the choice of lag. As a result, the log likelihood ratio (LR) test is used. Chenug and Lai (1993) highlighted the importance of the LR test in determining the correct number of parameters within the Johansen co-integration approach. They present evidence that the Johansen test is sensitive to a model structure that has too few parameters

[3] U.S. Treasuries and Euro-zone bonds are traded across the globe 24 hours a day, 6 days a week and hence, there is no need to make any adjustments for the opening and close of markets at different times.

[4] In this paper, the lag length attached to the ADF test is determined by the Akaike Information Criterion and Schwarz Bayesian Information Criterion. The results are not presented in the paper, but are available upon request.

thus affecting the reliability of the model in estimating the long run relationship. Under these circumstances, the Alkaike Information Criterion (AIC) is employed to establish whether the optimal lag structure is consistent with the parameterization chosen by the Johansen approach.

Step three involves estimating the VEC specification for the CF and FVF models. The VEC model is a restricted vector autoregressive model for use with nonstationary series that are known to be co-integrated from the previous step. Hence, given the objective of generating out-of-sample forecasts for the EUR, the error (Z_t) in the co-integrating equation for both currency models are computed as

$$Z_{t-1} = \ln(\text{EUR})_{t-1} - \beta_0 - \beta_1 \ln(\text{GBP})_{t-1} - \ldots\ldots\ldots$$
$$\beta_2 \ln(\text{JPY})_{t-1} - \beta_3 \ln(\text{CAD})_{t-1}$$
$$Z_{t-1} = \ln(\text{EUR})_{t-1} - \beta_0 - \beta_1 \ln(\text{IRD})_{t-1} - \ldots\ldots\ldots\ldots$$
$$\beta_2 \ln(\text{Oil})_{t-1} - \beta_3 \ln(2/10\ \text{Spd})_{t-1} \quad (1)$$

and the following VEC specifications derived from equation (1) are estimated for both CF and FVF models respectively

$$\Delta \text{EUR}_t = \alpha_0 + \sum_{i=p}^{k+1} \beta_i \Delta \text{EUR}_{t-i} + \sum_{i=p}^{k+1} \gamma_i \Delta \text{JPY}_{t-i} + \sum_{i=p}^{k+1} \phi_i \Delta \text{GBP}_{t-i} + \ldots\ldots.$$
$$\sum_{i=p}^{k+1} \lambda_i \Delta \text{CAD}_{t-i} + \delta Z_{t-1} + \varepsilon_t$$
$$\Delta \text{EUR}_t = \alpha_0 + \sum_{i=p}^{k+1} \beta_i \Delta \text{EUR}_{t-i} + \sum_{i=p}^{k+1} \gamma_i \Delta \text{IRD}_{t-i} + \sum_{i=p}^{k+1} \phi_i \Delta \text{Oil}_{t-i} + \ldots\ldots.$$
$$\sum_{i=p}^{k+1} \lambda_i \Delta 2/10\ \text{Spd}_{t-i} + \delta Z_{t-1} + \varepsilon_t \quad (2)$$

where $\Delta(L)_{t-i}$ are independent variables that captures short run deviations from the long run path. The terms $\alpha, \beta, \gamma, \lambda, \delta$ are coefficients to be estimated of which δ measures the sensitivity of the EUR rate to the long run equilibrium. The i^{th} lags are determined using the LR test in the previous step that is sufficient to eliminate serial correlation in the model.

The final step involves deriving an error correction model that is not based on the co-integrating equation of (1) but determined by a stepwise procedure. Implementing this step enables robustness testing of model forecasting performance by changing the specification of the error correction term. First utilized by MacDonald and Taylor (1994), it treats the VEC model of equation (2) as a general model except that Z_{t-1} is excluded. Using probability values on the dependent variables, the stepwise procedure eliminates coefficients with p-values that are above a threshold value. With further simulations, the stepwise procedure develops a single equation error correction model that comprises of coefficients that are statistically significant. In such a model, the number of variables and lags in the equation will differ thereby changing the specification of both currency models. Hence;

If p-value > threshold value then eliminate variable from the general model
If p-value < threshold value then keep the variable in the model

The advantage of adopting this approach is that it avoids the scenario where the nature of the modeling process is bounded by economic theorems and hence, bias in the results towards a desired outcome consistent with theory. Furthermore, the derivation of a single equation error correction model is possible without altering the nature of the modeling process; that is, to generate information on the long run relationship and nature of the adjustment mechanism.

To estimate a dynamic VEC, this study takes advantage of changing relationships over a period of time by estimating both currency models in a fixed rolling window. With a rolling window, estimates are applied over a data sample that produces out-of-sample forecasts. For example, if one has data up to time t, the model is estimated to produce forecasts for time t+1. Then the starting observation is dropped and another one added at the end. The model is then re-estimated to produce forecasts for time t+2. The process continues until all out-of-sample observations are utilized. By estimating both econometric models in this way, rolling out-of-sample forecasts are generated based on a model that allows for dynamic shifts in the relationships of the variables as it moves through time.

EMPIRICAL RESULTS

The empirical analysis is performed in a ten year rolling window that starts with the initial in-sample period from 10th February 1995 to 4th February 2005.[5] This process is repeated 210 times throughout the in-sample period and ends when the last observation in the data set is reached. Hence, the results presented in this paper are based on the in-sample period from 19th February 1999 to 13th February 2009 and a four year out-of-sample forecast period 18th February 2005 to 20th February 2009; four years being sufficient to test the robustness of the models and their forecasts.

Owing to the volume of results generated the results of step one are not presented in the paper but are available upon request. With step one the results include the Augmented Dickey Fuller (ADF) and Phillips–Perron (PP) test statistics for a unit root in each of the series. All series are found to be integrated of order I(1) in the levels whereas the first differences are integrated of order zero I(0). These findings imply that all variables can be used in the forecasting process.

After establishing that the series are I(1), the next step involves testing for co-integration amongst the variables using the Johansen and the Juselius (1990) approach. Like the ADF test, the Johansen approach is sensitive to the model specification as this would impact on the test results. By using the LR test for the unrestricted VAR's, lag's 7 and 19 was found to be the optimal specification for the CF model using the 1% level. On the other hand, in applying the same test on the FVF model, the null hypothesis is accepted at all lags including lag 1 meaning that it selects lag zero. In view of the acceptance of the null hypothesis, the AIC test was performed 210 times in the rolling window to find lag 2 as the correct parameterization of the Johansen test. Table 1 presents the Johansen co-integration test results for both models.

[Insert Table 1 Here]

According to the trace statistic and maximum eigenvalues, the null hypothesis of no co-integrating relationship against the alternative hypothesis of at least one co-integrating relationship is rejected at the 1% and 5% significance level for both models. After testing for more co-integrating relationships in a stepwise fashion, the test results prove to be significant, especially for the FVF model. Finding multiple co-integrating relations for the FVF model is consistent with the results of Benassy-Quere et al. (2007), Coudert et al. (2008) and Breitenfellner and Cuaresma (2008).

Having identified at least one co-integrating relationship in both models, the next step is to estimate the unrestricted VEC models of equation (2). The model coefficients are presented in Table 2 for the CF model and FVF model respectively. According to the estimates for the CF model, 1.7% of the disequilibrium is corrected for each week by changes in the EUR rate as shown by the λ_0 coefficient. In addition, the EUR rate is sensitive to changes in JPY at shorter lags and GBP rate at longer lag periods. Robustness tests on the model diagnostics reveal that the CF model is well specified.

[Insert Table 2 Here]

The results for the FVF model reveal similar findings in that changes in the EUR rate corrects approximately 1.5% of the disequilibrium for each week. Interestingly, the EUR is very responsive to short run deviations from changes in 2/10 treasury yield spreads and two year interest rate differentials. This implies that changes in market perceptions on future economic

[5] The ten year rolling window ensures the number of data observations is sufficient to run the general models with many lags without risking model stability and statistical adequacy.

conditions as well as yield differentials play a statistically significant role in correcting deviations from the long run equilibrium. Once again, robustness tests on the model diagnostics suggest a well specified model.

In view of the VEC models estimated, the final stage involves running the stepwise procedure to develop single equation error correction models. Beginning with the general model estimated in the previous step (excluding Z_t), the stepwise procedure yields the following single equation models

$$\Delta EUR_t = \alpha_0 + \beta_i \Delta EUR_{t-11} + \gamma_i \Delta JPY_{t-6} + \gamma_i \Delta JPY_{t-13} + \ldots\ldots\ldots\ldots$$

$$\phi_i \Delta GBP_{t-2} + \phi_i \Delta GBP_{t-8} + \lambda_i \Delta CAD_{t-6} + \delta Z_{t-1} + \varepsilon_t$$

$$\Delta EUR_t = \alpha_0 + \gamma_i \Delta IRD_{t-2} + \lambda_i \Delta 2/10\ Spd_{t-1} + \ldots\ldots\ldots\ldots$$

$$\lambda_i \Delta 2/10\ Spd_{t-2} + \delta Z_{t-1} + \varepsilon_t \qquad (3)$$

The coefficients of the reduced single equation error correction specification for both models are presented in Table 3. Results from the single equation CF model show that only 0.6% of the disequilibrium is corrected by changes in the EUR, a reduction from 1.7% reported earlier. When comparing the variables simulated by the stepwise procedure with the significant coefficients reported in Table 2, this finding is not surprising. On the other hand, the change in the model specification of the FVF model reveals little difference in the magnitude of the error correction term. Whist the EUR rate remains sensitive to changes in 2/10 treasury yield spread, it is no longer responsive to changes in the two year interest rate differentials. Furthermore, in the reduced form model, the EUR appears to react to changes in oil prices. Although the statistical significance of changes in oil prices is consistent with the findings of Coudert et al. (2008), this is the result of re-defining the error correction term.

[Insert Table 3 Here]

MODEL PERFORMANCE

In establishing evidence of co-integration, it is logical to pose the question whether the existence of a long run relationship improves the forecasting ability of the model.[6] Previous studies provide encouraging results on the forecasting power of the error correction model (see MacDonald and Taylor 1994 and Tong 2001 on foreign exchange rates). To test this notion, this paper provides three stages to the model evaluation process. Consistent with the literature, stage one involves testing the forecasting accuracy of both models using weekly out-of-sample forecasts to compare the Sum of Squared Errors (SEE), the Root Mean Squared Errors (RMSE) and Mean Squared Errors (MSE). These statistics are computed on the basis of the following equations:

$$SSE = \left[F_t - A_t\right]^2$$

$$MSE = \sum_{j=0}^{T-1} \left[\frac{\left[F_t - A_t\right]^2}{T} \right]$$

$$RMSE = \sqrt{\sum_{j=0}^{T-1} \left[\frac{\left[F_t - A_t\right]^2}{T} \right]} \qquad (4)$$

where F_t is the one week ahead out-of-sample forecast, A_t is the actual EUR known for corresponding week and T is the total number of forecasts. Table 4 summarizes the forecast error statistics for both forecasting models throughout the out-of-sample period and the two sub periods that coincide with pre and current financial crisis. The results show that the FVF

[6] Before evaluating model performance, preliminary analysis was conducted using equality tests of the out of sample forecasts across the two sub periods; pre and current market stress conditions. Results not presented in this paper provide conclusive evidence that the forecasts are statistically significant from zero across the sub periods. The results also report a doubling of the standard deviation in the 2007 – 2009 period that correspond with a higher standard error of the forecasts.

model outperforms the CF model throughout the sample period, a finding that differs from the conclusions of Evan and Lyons (2005). Reducing the error correction specification into a single equation model improves the performance of the CF model. However when compared with previous studies such as Hogan (1986) and Tong (2001), only the FVF model compares favorably.

[Insert Table 4 Here]

Table 4 also breaks down the model prediction errors according to the directional forecasts. According to the results, much of the performance for both models arises when it forecasts a weaker Euro against the U.S Dollar. Once again, changing model specification of the VEC to a single equation model improves the forecasting performance of the CF model regardless of directional forecast. However, the FVF model performance deteriorates with the change in model specification.

The second stage of the model evaluation process involves testing whether the forecast accuracy of both models translates into directional forecasting performance from which buy and sell signals are generated. From a currency manager's point of view, this is of paramount importance as it determines the usefulness of the models as a trading tool. Table 5 provides analysis for both forecasting models under different error correction specifications.

[Insert Table 5 Here]

Interestingly, the CF model outperforms the FVF model irrespective of the trading signal. This is more so when the out-of-sample forecasts coincide with the 2007 – 2009 sub-sample. Changing the error correction specification makes little difference to this finding, although over the four year out-of-sample period the FVF model marginally outperforms the CF model. When compared to the statistics in Table 4, these findings lead to the proposition that forecasting accuracy does not necessarily translate into the performance of model generated trading signals. Of greater significance, when added together, is that the performance of both models is robust to changing market conditions.

The final stage addresses the issue of whether model performance translates into investment performance. This addresses the question of which model evaluation criteria is more important in determining the performance of a model driven fund; point forecasting accuracy or trading signals based on directional forecasts. To achieve this, the performance of all models is evaluated within a currency fund. A starting point is the assumption that the currency manager trades on the model signal at all times to determine medium term trading strategy. In addition, it is assumed there are no transaction costs when a trade is activated.[7] Based on these assumptions, Figure 1 presents the performance of a hypothetical model driven fund assuming an initial investment of EUR 20,000. Preliminary findings reveal that the CF model outperforms the FVF model using the unrestricted VEC specification. However, the CF model performance shows signs of deterioration during the height of the financial crisis that coincided with a marked improvement of the FVF model. Interestingly, the performance of the CF model is not robust to changes in model specification of the error term, a finding that contrasts with the FVF model.

[Insert Figure 1 Here]

To build on these findings, the Sharpe Ratio (SR) is used as the measure of performance. The SR is a statistic that sums up the desirability of a risky investment strategy by dividing the average period return μ in excess of the risk free rate (R_f) by the standard deviation of the return σ. Developed by Sharpe (1994), it is expressed as

[7] Owing to the competitiveness of transaction costs, its impact on the fund's profit and loss is small. Moreover, as transaction costs in foreign exchange trading comes from a variety of sources, ranging from activating the trade to keeping positions open overnight. Hence, for simplicity it is assumed to be zero.

$$SR \equiv \frac{\mu - R_f}{\sigma} \tag{5}$$

Given that μ and σ are the population moments of the distribution of returns R_t and hence, are not observable, both must be estimated using historical data.[8] Beginning with the mean and variance from a sample of historical returns (R_1, R_2, ...,R_T)

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t$$

and

$$\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T}\left(R_t - \hat{\mu}\right)^2$$

from which the estimate of the Sharpe ratio of equation (5) follows

$$\widehat{SR} = \frac{\hat{\mu} - R_f}{\hat{\sigma}} \tag{6}$$

Table 6 presents the SR analysis on the currency fund using both forecasting model types. The results reveal some interesting conclusions. Firstly, the results for both models compare favorably to the findings provided by Hamelink (2001). Although he uses non-linear forecasting models to generate three years out-of-sample forecasts for the Deutschmark, Swiss Franc and Japanese Yen, the SR values reported ranged from -1.24 to 2.75. Secondly, over the four year period, changes in specification of the error correction model add little or no value to investment performance of the FVF model, a finding that differs from the CF model. Thirdly, the CF model outperforms the FVF model over the out-of-sample period and the pre-crisis period. This follows from the findings in Table 9 that directional model performance translates into value added in an investment fund. However, this is not robust to changes in model specification. In addition, closer examination of the results reveals a marked improvement in the investment performance of the FVF model from the pre-crisis period to the 2007 – 2009 period, a finding that is robust to changes in model specification.[9] This implies that the superior forecasting accuracy of the FVF model (reported in Table 4) translates into performance of the investment fund over the past two years. This implies that the evaluation criteria used to determine forecasting performance and whether this translates into investment performance is subject to changes in model specification and market conditions.

[Insert Table 6 Here]

[8] The issue that arises with equation (5) is the reliability and precision of SR as determined by the statistical properties of the data. Lo (2002) examines this issue and illustrates how previous studies fail to test for serial correlation and heteroskedascity in returns invalidates the estimation of the SR if found to be present. Therefore, before evaluating the investment performance of the forecasting models, the statistical properties of the model generated return are tested for evidence of serial dependencies. This involves multiplying the actual return by the forecasting signal generated by the model. In the forecasting signal, the model produces 1 when it generates a buy signal and –1 for a sell signal. Hence, if actual weekly return is positive when it corresponds with a buy signal for that week then the currency manager will make a profit. Should actual return be negative and the model generates a buy signal, the currency manager will make a loss. The opposite is true for a sell signal. To determine the statistical properties of the data, the assumption that the model adjusted returns are Independently and Identically Distributed (i.i.d) is tested using the Breusch-Godfrey LM test for serial correlation and Autoregressive Conditional Heteroskedasticity (ARCH) effects. The test results indicate that the null hypothesis of no serial dependences in the first and second moments is accepted at the 1% level with the exception of the 2007 – 2009 sub sample. However, rejection of the null hypothesis of no serial correlation is reported at the 1% level for the 2007 – 2009 using model returns from the reduced form currency forecasting model. As a consequence, the model return was adjusted by running an ARIMA(2,2) model and repeating the serial correlation and ARCH tests on the adjusted series. This leads to acceptance of the null hypothesis. Although the results are not reported in the paper, it is available upon request.

[9] An avenue for future research is to extend the current analysis to previous crisis periods to evaluate model performance in relation to periods of stable market conditions. This is beyond the scope of this paper owing to restrictions on the availability of data.

CONCLUSION

The key issue of "getting the currency right" and what this implies for model evaluation and value added in an investment fund is the main theme of the paper. This study provides a comprehensive evaluation of currency models based on various error correction specifications and their usefulness as a trading tool with emphasis on robustness testing. In the light of the current market environment, robustness testing of trading models takes on greater importance and in doing so the results presented make a number of significant contributions in this area. Using a dynamic vector error correction model in a fixed rolling window, the existence of co-integrating relationships is exploited to develop forecasting models that successfully captures short to medium currency movements. On the basis of forecast errors, the FVF model outperforms the CF model throughout the four year out-of-sample period and that these findings compare favorably with the empirical literature. The results are robust after dividing the sample period into pre and current crisis periods and to changes in model specification in the error correction term. However, in using model generated trading signals as the evaluation criteria, the CF model outperforms the FVF model, a finding that is robust across the out-of-sample period. This implies that the superior forecasting accuracy of the FVF model does not necessarily translate into the performance of trading signals generated by directional forecasts. This paper also illustrates how forecasting performance translates into value added in a currency fund depends on the criteria used and prevailing market conditions. With the FVF model, forecast accuracy translates into investment performance especially during the current financial crisis. This contrasts with the directional forecast performance of the CF model that appears to translate into value added in the pre-crisis sample. Both findings are robust to changes in model specification of the error correction term, although investment performance of the CF model shows a marked deterioration using the more specific model.

The results presented in this study are of added significance given that the performance of the forecasting models coincides with the current financial crisis. The existence of co-integrating relations between the EUR exchange rate and all independent variables used warrants the use of non-linear co-integration analysis as a robustness test especially in the light of the financial crisis. Further investigation is warranted on the notion that the evaluation criteria used may determine whether this translates into value added in an investment fund. Another obvious avenue of research is to repeat the above analysis over much longer samples that cover previous financial crisis. If sufficient data is available, one would be able to perform robustness tests on model performance before, during and after a crisis. Therefore in view of the future avenues of research outlined, the results one can yield will be of great interest to both practitioners and academics alike.

REFERENCES

Amoateng, K.A.: The Relationship Between Real Long-Short Interest Spread Differentials and Real Exchange Rates, Applied Economic Letters, 2, pp.432-436 (1995)

Cheung Yin-Wong., Lai, K.S.: Finite-Sample of Jahansen's Likelihood Ratio Tests for Co-integration, Oxford Bulletin of Economics and Statistics, 55, 3, pp. 313-327 (1993)

Estrella, A., Mishkin, F.S.: The Yield Curve as a Predictor of U.S. Recessions, Current Issues in Economics and Finance, Federal Reserve Bank of New York, Vol.2, No.7, pp.1-6 http://courses.cit.cornell.edu/econ331/MishkinEstrella1996.pdf (1996)

Evans, M.D.D., Lyons, R.K.: Meese-Rogoff Redux: Micro-Based Exchange Rate Forecasting, Financial Economics, Macroeconomics, and Econometrics, Vol 95, No 2, pp. 405-414 (2005)

Federal Reserve Statistical Release, Currency Weights – Broad Index of the Foreign Exchange Value of the Dollar, www.federalresevre.gov/release/H10/Weights/

Hogan, L.I.: A Comparison of Alternative Exchange Rate Forecasting Models, The Economic Record, June, pp.215-223 (1986)

Johansen, S.: Statistical Analysis of Co-integration Vectors, Journal of Economic Dynamics and Control, 12, pp.231-254 (1988)

Johansen, S., Juselius, K.: Maximum Likelihood Estimation and Inference on Co-integration – With Applications to the demand for Money, BULLETIN, 52, pp.169-210 (1990)

Liu, S.: Currency Derivatives and Exchange Rate Forecasting, Financial Analysts Journal, Vol. 63, No 4, pp.72-78 (2007)

Lo, A.W.: The Statistics of Sharpe Ratios, Financial Analysts Journal, July/August, pp.36-52 (2002)

MacDonald, R., Taylor, M.: The Monetary Model of the Exchange Rate: Long Run Relationships, Short Run dynamics and How to Beat a Random Walk", Journal of International Money and Finance, 13 (3), pp.276-290 (1994)

Pukthuanthong-Le, K., Thomas III, L.R.: Weak-Form Efficiency in Currency Markets, Financial Analysts Journal, Vol. 64, No. 3, pp.31-52 (2008)

Sharpe, W.F.: The Sharpe Ratio, Journal of Portfolio Management, Vol. 21, No. 1, pp.49-58 (1994)

Tong, W.H.S.: Co-integration, Efficiency and Forecasting in the Currency Market, Journal of Business Finance and Accounting, 28 (1) and (2), January/March 2001, pp.127-150 (2001)

Wright, J.H.: The Yield Curve and Predicting Recessions, Working Paper – Finance and Economics Discussion Series, Federal Reserve Board, pp.1-21 http://www.federalreserve.gov/pubs/feds/2006/200607/200607pap.pdf (2006)

Coudert, V., Mignon, V., Penot, A.: Oil Price and the Dollar, Forthcoming in Energy Studies Review, Vol.15:2 (2008)

Benassy-Quere, A., Mignon, V., Penot, A.: China and the Relationship Between the Oil Price and the Dollar, Energy Policy, Elsevier, Vol. 35, No. 11, pp.5795 – 5805 (2007)

Breitenfellner, A., Cuaresma, J.: Crude Oil Prices and USD/EUR Exchange Rate, Monetary Policy and the Economy, No.4/08 (2008)

Table 1: Johansen Test for the Number of Co-integrating Relationships

	CF Model		**FVF Model**	
Number of Co-integrating Equations	Trace Statistic	Max-Eigen Statistic	Trace Statistic	Max-Eigen Statistic
None	60.07* (47.21)	32.09* (27.07)	73.84* (47.21)	33.71* (27.07)
At Most 1	27.99 (29.68)	14.84 (20.97)	40.13* (29.68)	21.83* (20.97)
At Most 2	13.15 (15.41)	8.075 (14.07)	18.30* (15.41)	17.29* (14.07)
At Most 3	5.07* (3.76)	5.07* (3.76)	1.02 (3.76)	1.02 (3.76)

Co-integrating Equation
(Standard Errors in parenthesis)

	Co-integrating coefficients			
CF Model	EUR	b_1	b_2	b_3
	1.0000	-1.70343	-3.25184	-1.10273
		(0.4230)	(0.5201)	(0.3759)
FVF Model	1.0000	0.27084	-0.36811	-0.42768
		(0.0846)	(0.0311)	(0.0927)

Note: The Johansen test is based on the optimal lags used; lag 19 for the currency forecasting model and lag 2 for the fair value forecasting model. The values in brackets are chi-squared critical values at the 5% level of significance. The asterisk * denotes rejection of null hypothesis of no co-integration at the 5% level. Trace and Max-Eigenvalue tests indicates 1 and 3 co-integrating relation(s) for the currency forecasting model and at most 2 co-integrating relationship(s) for the fair value forecasting model. The above results involve estimating the following co-integrating equations:

$$Z_{t-1} = \ln(EUR)_{t-1} - b_0 - b_1 \ln(GBP)_{t-1} - b_2 \ln(JPY)_{t-1} - b_3 \ln(CAD)_{t-1}$$

$$Z_{t-1} = \ln(EUR)_{t-1} - b_0 - b_1 \ln(IRD)_{t-1} - b_2 \ln(Oil)_{t-1} - b_3 \ln(2/10\,spd)_{t-1}$$

where Z_{t-1} is the co-integrating vector as defined in equation (1) that reflects deviations from the long run equilibrium rate and $b_0,\ldots\ldots\ldots b_3$ are coefficients to be estimated. Standard errors are in parenthesis.

Table 2: Unrestricted Vector Error Correction Model Estimations for the EUR (19th February 1999 – 13th February 2009)

Lag		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
CF Model																	
	δ_0	-0.0169															
		(-2.73)															
	β_i	0.1837	0.1532														
		(2.64)	(2.17)														
	γ	-0.2737	-0.1660				-0.2172		-0.1842								
		(-3.74)	(-2.21)				(-2.25)		(-2.25)								
	ϕ_i								-0.1442							-0.1123	-0.0993
									(-3.06)							(-2.32)	(-2.03)
	λ_i						-0.1365								-0.1471		
							(-2.18)								(-2.28)		
FVF Model																	
	δ_0	-0.0146															
		(-2.15)															
	β_i																
	γ	-0.0513	-0.0693														
		(-2.21)	(-3.01)														
	ϕ_i																
	λ_i		0.0564														
			(3.63)														

	Goodness of fit		Breusch – Godfrey Correlation Test				ARCH Test			
Diagnostic Tests	R^2	$R^{\wedge 2}$	F-Stat	p-value	$N*R^2$	p-value	F-Stat	p-value	$N*R^2$	p-value
CF	0.19	0.05	0.0920	(1.00)	1.9092	(1.00)	1.7178	(0.03)	33.463	(0.03)
FVF	0.05	0.03	0.7244	(0.80)	14.671	(0.79)	0.9851	(0.48)	19.753	(0.47)

Note: The above results involve estimating the following unrestricted vector error correction model of equation (2) for both models. The figures in parenthesis are t-statistics. Only coefficient values that are statistically significant at the 5% level are reported. The term $N*R^2$ represents Engle's LM test.

Table 3: Single Equation Error Correction Model Estimations for the EUR (19th February 1999 – 13th February 2009)

Lag			1	2	3	4	5	6	7	8	9	10	11	12	13
CF Model															
Co-integrating Coefficients		δ_0	0.0063												
			(2.85)												
b_1	-2.93947	β_i										0.1049			
	(1.3179)											(2.38)			
b_2	4.06518	γ_i		-0.0942						-0.1411					
	(1.0505)			(-2.31)						(-3.50)					
b_3	-1.73418	ϕ_i						-0.1428							0.1050
	(1.0246)							(-2.71)							(2.20)
		λ_i						-0.1068							
								(-1.98)							
FVF Model															
Co-integrating Coefficients		δ_0	-0.0139												
			(-2.21)												
b_1	-0.3030	β_i													
	(0.0989)														
b_2	-0.3625	γ_i													
	(0.0337)														
b_3	0.1265	ϕ_i											0.0539		
	(1.4006)												(3.60)		
		λ_i	-0.0320	-0.0717											
			(-2.01)	(-3.13)											

	Goodness of fit		Breusch – Godfrey Correlation Test				ARCH Test			
Diagnostic Tests	R^2	R^2	F-Stat	p-value	N^*R^2	p-value	F-Stat	p-value	N^*R^2	p-value
CF	0.07	0.06	0.712	(0.81)	14.72	(0.79)	0.930	(0.55)	18.82	(0.53)
FVF	0.04	0.03	0.735	(0.79)	14.88	(0.78)	1.061	(0.39)	21.22	(0.38)

Note: The above results involve estimating the following unrestricted vector error correction model of equation (3) for both models. The figures in parenthesis are standard errors for the co-integrating equation and t-statistics for the restricted error correction model. The term N^*R^2 represents Engle's LM test.

Table 4: Results on Model Forecasting Accuracy – CF and FVF EUR Models

	Vector Error Correction Model		Single Equation Error Correction	
	Forecasting Model – CF	Forecasting Model – FVF	Forecasting Model - CF	Forecasting Model - FVF
Out of Sample: Feb 2005 –Feb 2009				
SEE	0.0927	0.0697	0.0707	0.0698
MSE	0.0011	0.0000	0.0006	0.0001
RMSE	0.0330	0.0000	0.0235	0.0100
Out of Sample: Feb 2005–Feb2007				
SEE	0.0238	0.0217	0.0216	0.0213
MSE	0.0002	0.0000	0.0001	0.0000
RMSE	0.0153	0.0030	0.0121	0.0021
Out of Sample: Feb 2007 –Feb 2009				
SEE	0.0687	0.0479	0.0489	0.0484
MSE	0.0011	0.0001	0.0005	0.0003
RMSE	0.0334	0.0116	0.0230	0.0176
Breakdown of Forecasting Performance – Strong EURUSD				
$SEE_{(2005\text{-}2009)}$	0.0561	0.0299	0.0417	0.0337
$SEE_{(2005\text{-}2007)}$	0.0137	0.0063	0.0130	0.0076
$SEE_{(2007\text{-}2009)}$	0.0423	0.0235	0.0286	0.0260
$MSE_{(2005\text{-}2009)}$	0.0023	0.0000	0.0006	0.0000
$MSE_{(2005\text{-}2007)}$	0.0003	0.0000	0.0003	0.0000
$MSE_{(2007\text{-}2009)}$	0.0026	0.0000	0.0004	0.0000
$RMSE_{(2005\text{-}2009)}$	0.0482	0.0036	0.0239	0.0060
$RMSE_{(2005\text{-}2007)}$	0.0184	0.0022	0.0158	0.0031
$RMSE_{(2007\text{-}2009)}$	0.0510	0.0041	0.0191	0.0066
Breakdown of Forecasting Performance – Weak EURUSD				
$SEE_{(2005\text{-}2009)}$	0.0366	0.0398	0.0290	0.0361
$SEE_{(2005\text{-}2007)}$	0.0101	0.0154	0.0086	0.0137
$SEE_{(2007\text{-}2009)}$	0.0264	0.0244	0.0203	0.0224
$MSE_{(2005\text{-}2009)}$	0.0002	0.0000	0.0000	0.0000
$MSE_{(2005\text{-}2007)}$	0.0001	0.0000	0.0000	0.0000
$MSE_{(2007\text{-}2009)}$	0.0003	0.0001	0.0000	0.0001
$RMSE_{(2005\text{-}2009)}$	0.0152	0.0016	0.0004	0.0040
$RMSE_{(2005\text{-}2007)}$	0.0031	0.0052	0.0038	0.0052
$RMSE_{(2007\text{-}2009)}$	0.0176	0.0075	0.0039	0.0109

Note: This table provides summary statistics for the Sum of Estimation Errors (SEE), the Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE) These measures evaluates the forecasting performance of both models by comparing one period ahead forecast values of the EUR with the actual rate.

Table 9: Results on Model Directional Forecasts - Both Models

	Vector Error Correction Model		Single Equation Error Correction	
	Forecasting Model – CFM[1]	Forecasting Model – FVFM[2]	Forecasting Model – CFM	Forecasting Model - FVFM
Out of Sample: Feb 2005 –Feb 2009				
Overall	0.58	0.54	0.53	0.55
Buy Signal	0.58	0.55	0.54	0.55
Sell Signal	0.57	0.52	0.52	0.54
Out of Sample: Feb 2005–Feb2007				
Overall	0.56	0.55	0.50	0.55
Buy Signal	0.55	0.56	0.49	0.55
Sell Signal	0.57	0.55	0.50	0.55
Out of Sample: Feb 2007 –Feb 2009				
Overall	0.59	0.51	0.57	0.53
Buy Signal	0.59	0.53	0.58	0.54
Sell Signal	0.59	0.46	0.56	0.50

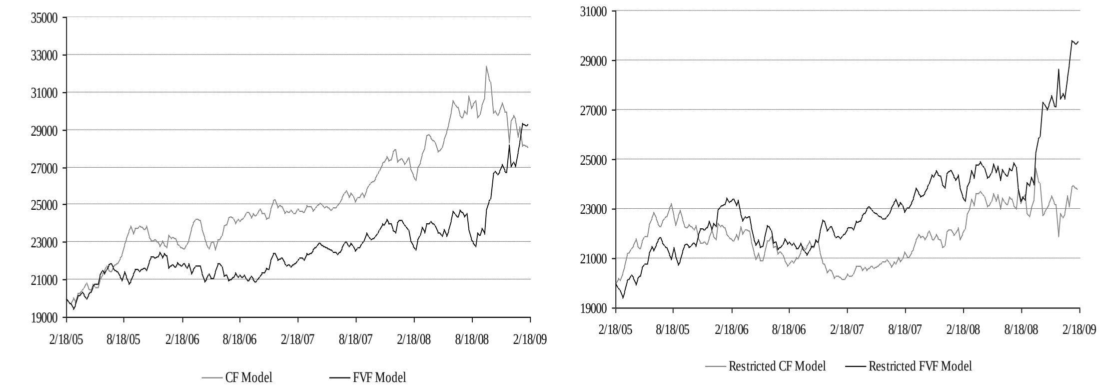
Note: These statistics summarizes the directional forecasting performance of both models. This is based on the model forecast in relation to the last observation that defines the horizon period; i.e. one week. It provides useful information for money market managers on the confidence levels of the forecasts and hence, model generated trading signals.

Table 10: Model Evaluation – Sharpe Ratio Analysis

	Vector Error Correction Model		Single Equation Error Correction	
Test Type	Forecasting Model – CFM[1]	Forecasting Model – FVFM[2]	Forecasting Model – CFM	Forecasting Model - FVFM
Sample Size: Feb 2005–Feb 2009				
(μ - R_f)	0.1122	0.1241	0.0712	0.1279
SR	1.1341	1.0215	0.4809	1.0603
SE(SR)	0.0891	0.0857	0.0734	0.0869
% SE(SR)	0.0837	0.0848	0.1059	0.0844
SE (μ)	0.6086	0.6571	0.8963	0.6401
SE (σ^2)	0.3914	0.3429	0.1037	0.3599
Sub Sample: Feb 2005–Feb2007				
(μ - R_f)	0.1334	0.0760	0.0358	0.0786
SR	1.1849	0.4792	-0.0038	0.6641
SE(SR)	0.1285	0.1040	0.0985	0.1089
% SE(SR)	0.1181	0.1503	1.6063	0.1336
SE (μ)	0.5876	0.8970	1.0000	0.8193
SE (σ^2)	0.4124	0.1030	0.0000	0.1807
Sub Sample: Feb 2007 –Feb 2009				
(μ - R_f)	0.0909	0.1489	0.0109	0.1539
SR	0.5517	1.3728	0.0567	1.3747
SE(SR)	0.1058	0.1373	0.0986	0.1374
% SE(SR)	0.1424	0.1172	0.4143	0.1172
SE (μ)	0.8679	0.5149	0.9984	0.5141
SE (σ^2)	0.1321	0.4851	0.0016	0.4859

Note: (μ - R_f) is the excess model return. SE(SR) is the standard error of the Sharpe Ratio, SE (μ) and SE (σ^2) shows whether the standard error is attributable to errors in the mean or the variance. For the risk free rate, the US 3 Month LIBOR plus 100 basis points is used.

Figure 1: Model Fund Performance – Initial Investment: EUR 20 000



Note: The fund performance is based on the assumption that the currency manager acts on the basis of the model trading signal at all times. Given the use of Friday's European close, the trading signal lasts until the same time the following Friday, hence positions are closed and then re-opened according to the model recommendations for the next week.