HONG KONG MONETARY AUTHORITY

Working Paper 08/2008

June 2008

COMPARING FORECAST PERFORMANCE OF EXCHANGE RATE MODELS

Prepared by Lillie Lam, Laurence Fung and Ip-wing Yu[1]

Research Department

Abstract

Exchange-rate movement is regularly monitored by central banks for macroeconomic-analysis and market-surveillance purposes. Notwithstanding the pioneering study of Meese and Rogoff (1983), which shows the superiority of the random-walk model in out-of-sample exchange-rate forecast, there is some evidence that exchange-rate movement may be predictable at longer time horizons. This study compares the forecast performance of the Purchasing Power Parity model, Uncovered Interest Rate Parity model, Sticky Price Monetary model, the model based on the Bayesian Model Averaging technique, and a combined forecast of all the above models with benchmarks given by the random-walk model and the historical average return. Empirical results suggest that the combined forecast outperforms the benchmarks and generally yields better results than relying on a single model.

JEL Classification Numbers: C11, C52, C53

Keywords: Bayesian Analysis, Model Evaluation and Selection, Forecasting and Other Model Application

Author's E-Mail Address:

llflam@hkma.gov.hk, Laurence_KP_Fung@hkma.gov.hk; Ip-wing_Yu@hkma.gov.hk

[1] The authors acknowledge the comments from Hans Genberg and Cho-hoi Hui.

Executive Summary:

- *Exchange-rate movement is regularly monitored by central banks for macroeconomic-analysis and market-surveillance purposes. Despite its importance, forecasting exchange rate has been a challenge since the collapse of the Bretton Woods System.*

- *Abundant studies in the literature show that exchange-rate models perform poorly in out-of-sample prediction analysis, even though some of them have good fit in-sample analysis.*

- *This paper studies exchange-rate predictability based on different theoretical and empirical models, including the Purchasing Power Parity model, Uncovered Interest Rate Parity model, Sticky Price Monetary model and the model based on the Bayesian Model Averaging technique, and a combination of these models' forecasts. It presents out-of-sample forecasts of the euro, British pound and Japanese yen against the US dollar in the horizons of one-quarter to eight-quarter ahead.*

- *Empirical results show that depending on the currencies and the forecast horizons, some of these models outperform common benchmarks given by the random-walk model and the historical average return.*

- *No single model consistently stands out as the best exchange-rate forecasting model when assessed by different criteria. The combined forecast is in general better than the forecast based on a single model when the root-mean-squared forecast error and the direction of change statistics are used as the criteria. Given the limitation of individual models, predictions based on them should be used with caution.*

I. INTRODUCTION

Exchange rate movement is an important subject of macroeconomic analysis and market surveillance. Despite its importance, forecasting the exchange rate level has been a challenge for academics and market practitioners.

The collapse of the Bretton Woods system of fixed exchange rates among major industrial countries marked the beginning of the floating exchange rate regime. Since then, there has been considerable interest in forecasting exchange rate movements. However, empirical results from many of the exchange rate forecasting models in the literature, no matter they are based on the economic fundamentals or sophisticated statistical construction, have not yielded satisfactory results. For example, Mussa (1979) concludes that the spot exchange rate approximately follows a random-walk process and most changes in exchange rates are unexpected. This was also supported by the seminal work of Meese and Rogoff (1983), which shows that none of the structural exchange rate models used in their paper could significantly outperform a simple random-walk model in both short- and medium-terms. The results in many follow-up empirical studies, though somewhat mixed, have broadly reached similar conclusions. Even when the in-sample forecasts of the exchange rate models perform well, the out-of-sample forecasts are disappointing when compared to those of a naïve random-walk model.[2]

Motivated by the recent work of Wright (2003), who argues that a model based on the Bayesian Model Averaging (BMA) technique gives "promising results for out-of-sample exchange rate prediction" compared to the random-walk model, this paper re-examines the exchange rate forecasting capability of some related models in a more systematic manner. In addition to the BMA approach, we also study the forecasting capability of three well-discussed models in the literature, namely the Purchasing Power Parity (PPP) model, Uncovered Interest Rate Parity (UIP) model, Sticky Price Monetary (SP) model of Dornbusch (1976) and Frankel (1979), and a combined forecast based on the above models.[3] In this paper, the models' forecasting performances at different time horizons (from one-quarter ahead to eight-quarter ahead) are assessed by different sets of criteria. This provides a more comprehensive and systematic way to evaluate the models.

---

2 For example, Meese and Rose (1991) use a variety of non-linear and non-parametric techniques to modelling exchange rates. However, negative results are obtained. Engle (1992) analyses eighteen exchange rates and found that the Markov-switching models could perform better only in-sample but could not have out-of-sample forecasts superior to a random-walk model. For details, see Frankel and Rose (1994).

3 These models are chosen because of different reasons. For instance, the PPP model is included because of its importance in the exchange rate literature while the SP model is a typical structural model that has been the subject of previous systemic analysis. Fore details, see the discussion in Cheng, Chinn and Pascual (2004).

The remainder of this paper is organised as follows: Section II presents the specification of the five models examined in this study. Section III discusses data involved in this study and the criteria to evaluate the forecasts. Comparison of forecasts obtained from empirical estimation is presented in Section IV. Section V summarises and concludes. Technical details of the BMA are discussed in the Appendix.

II. EXCHANGE RATE FORECASTING MODELS

Given that there are many candidates of empirical models available for exchange rate determination, the models used in the paper are selected according to at least one of the following criteria: (i) prominent in economic literature; (ii) not restrictive to only theoretical or empirical model; (iii) readily replicable and available for implementation; and (iv) not previously evaluated in a systematic manner.

Based on the above criteria, the models examined in this study are (i) the PPP model; (ii) the UIP model; (iii) the SP model; (iv) the model based on the BMA technique; and (v) the composite specification incorporating the above four models.

2.1 Purchasing Power Parity (PPP) Model

The PPP model is a theoretical exchange rate model. The model explains the movements of the exchange rate between two economies' currencies by the changes in the countries' price levels. The goods-market arbitrage mechanism will move the exchange rate to equalise prices in the two economies.[4] Mathematically, the exchange rate determination under the PPP model is expressed as

$$\ln e_t = \ln p_t - \ln p_t^* \tag{1}$$

where e_t is the nominal exchange rate, p_t and p_t^* are domestic and foreign prices respectively. Equation (1) is the relative version of the PPP model, as price indexes instead of actual price levels are considered in the estimations.

---

4 For example, if the US goods are more expensive than those in Japan, consumers in the US and Japan may tend to purchase more Japanese goods. The increased demand for Japanese goods will drive the Japanese yen to appreciate with respect to the US dollar until the dollar-denominated prices of the US goods and Japanese goods are equalised.

The PPP model in this study is specified as a restrictive error-correction form, that is an error-correction model without the short-run dynamics.[5] This specification follows that used in Cheung et al. (2004). The restrictive setup explicitly allows the variation of the exchange rate as a correction of its last-period deviation from a long-run equilibrium. For the case of the PPP model, the specification of the restrictive error-correction form is written as

$$\ln e_{t+h} - \ln e_t = \alpha_0 + \alpha_1 (\ln e_t - \beta_0 - \beta_1 \ln \tilde{p}_t) + \varepsilon_t \qquad (2)$$

where $\tilde{p}_t$ is the domestic price level relative to the foreign price level, ε_t is a zero mean error term, and h is the forecast horizon.

2.2 Uncovered Interest Rate Parity (UIP) Model

The UIP describes how the exchange rate moves according to the expected returns of holding assets in two different currencies. Ignoring transaction cost and liquidity constraints, the UIP gives an arbitrage mechanism that drives the exchange rate to a value that equalises the returns on holding both the domestic and foreign assets. Specifically, if the UIP holds, the arbitrage relationship will give the following expression

$$E_t(\ln e_{t+h} - \ln e_t) = i_t - i_t^* \qquad (3)$$

where $E_t(\ln e_{t+h} - \ln e_t)$ is the market expectation of the exchange rate return from time t to time $t+h$, and i_t and i_t^* are the interest rates of the domestic and foreign currencies respectively.

Similar to the specification of the model based on the PPP model, the UIP model is also tested in the restrictive error-correction form, i.e.,

$$\ln e_{t+h} - \ln e_t = \alpha_0 + \alpha_1 (\ln e_t - \beta_0 - \beta_1 \ln \tilde{i}_t) + \varepsilon_t \qquad (4)$$

where $\tilde{i}_t$ is the domestic long-term interest rate relative to that in the foreign country.

---

[5] A general specification of the error correction model is: $\Delta y_t = \beta_1 + \mathbf{B}'\Delta X_t + \alpha\tilde{e}_{t-1} + \varepsilon_t$ where Δ is the first-difference operator; y is the dependent variable; X is the independent variables including a constant; α is the adjustment coefficient for the correction of the last-period deviation, $\tilde{e}$ is the error correction term; ε is a zero mean error term. The error correction term $\tilde{e}$ is the residual of the estimation of $y_t = \tilde{\mathbf{B}}X_t + \tilde{\varepsilon}_t$ where $\tilde{\varepsilon}$ is a zero mean error term The restrictive error-correction model employed in this study is without the terms of short-run dynamics, i.e. ΔX_t. The exclusion of the short-run dynamics allows us to obtain *ex-ante* forecasts for an h-period horizon without forecasting the h-period ahead right-hand-side variables. This is comparable to the BMA model specified in Section 2.4. For more discussions, see Cheung et al. (2004), Mark (1995) and Chinn and Meese (1995).

The UIP model is included in this study because it has recently been found to have forecast capability at longer horizons (see Cheng et al., 2004; Alexius, 2001; Meredith and Chinn, 1998).

2.3 Sticky Price Monetary (SP) Model

The SP model of Dornbusch (1976) and Frankel (1979) is a well-known model in the international finance literature. It can be interpreted as an extended PPP model by replacing the price variables in equation (1) with macroeconomic variables that capture money demand and over-shooting effects. This study includes the SP model because it has been the subject of systematic analysis in the literature.

According to the expression in Frankel (1979), the SP model is in the following form

$$\ln e_t = \ln m_t - \ln m_t^* - \phi(\ln y_t - \ln y_t^*) + \alpha(\ln i_t - \ln i_t^*) + \beta(\pi_t - \pi_t^*) \qquad (5)$$

where m_t is the domestic money supply, y_t is the domestic output, i_t is the domestic interest rate, π_t is the domestic current rate of expected long-run inflation, and all variables in asterisk denote variables of the foreign country.

As discussed in Section 2.1, the model examined in this study is specified in a restrictive error-correction form as

$$\ln e_{t+h} - \ln e_t = \alpha_0 + \alpha_1(\ln e_t - \beta_0 - \beta_1 \ln \tilde{m}_t - \beta_2 \ln \tilde{y}_t - \beta_3 \ln \tilde{i}_t - \beta_4 \ln \tilde{\pi}_t) + \varepsilon_t \qquad (6)$$

where $\tilde{m}_t$ is the seasonally adjusted money supply relative to the foreign country, $\tilde{y}_t$ is the seasonally adjusted real GDP relative to the foreign country, and $\tilde{\pi}_t$ is the inflation rate relative to the foreign country.

2.4 Bayesian Model Averaging (BMA) Method

The idea of the BMA method was first proposed by Leamer (1978) and has gained attention in the statistical literature since the 1990s. Recently, the use of the BMA in out-of-sample forecast in output growth and stock returns has been found to have some success in improving the predictive performance.[6] In particular, Wright (2003) shows that the application of the BMA method to the exchange rate forecast gives comparably favourable out-of-sample results as compared to the benchmark random-walk model.

[6] See Min and Zellner (1993), Avramov (2002) and Cremers (2002).

The focus of the BMA method is on how to choose the correct model given the uncertainty in a number of underlying candidate models. Instead of choosing only one final single model for forecasting, the BMA method has a mechanism to select different combination of the underlying models based on its performance over time.[7] For example, given that there are a total of k potential explanatory variables, the total number of linear models arising from different combination of these k variables is 2^k $(\sum_{r=0}^{k} {}_kC_r)$. Starting with some prior probabilities giving to a set of these 2^k candidate models using the in-sample information, the BMA method then assigns each candidate model a posterior probability determined by the Bayesian method.[8]

In this paper, the underlying model for the BMA method is specified as

$$\ln e_{t+h} - \ln e_t = \mathrm{X}_\mathrm{t}\beta + \varepsilon_\mathrm{t} \tag{7}$$

where X_t is a $T \times (k+1)$ matrix of exchange rate determinants including the constant term, k is the number of exchange rate determinants, T is the number of observations, β is a $(k+1)\times 1$ matrix of parameters to be estimated. Following Wright (2003), sixteen variables including both the economic and financial variables are considered as the determinants in the model. These variables are:

(i) stock price
(ii) change in stock price
(iii) long-term interest rate
(iv) short-term interest rate
(v) term spread
(vi) oil price
(vii) change in oil price
(viii) exchange rate return of the previous period
(ix) sign of exchange rate return of the previous period
(x) seasonally adjusted real GDP
(xi) change in seasonally adjusted real GDP
(xii) seasonally adjusted money supply
(xiii) change in seasonally adjusted money supply
(xiv) consumer price level
(xv) inflation rate
(xvi) ratio of current account to GDP

---

[7] The advantages of using the BMA to account for model uncertainty have been assessed for several different classes of models, including survival models by Raftery, Madigan and Volinsky (1996), and linear regression by Raftery, Madigan and Hoeting (1997).

[8] The Appendix presents the details of how to obtain the posterior model probability.

These variables, except the oil price index, the exchange rate return and the sign of exchange rate return, are relative to those of the foreign country. With these sixteen variables, there are 2^{16} or 65,536 candidate models. The forecasts from these models will be pooled together based on their posterior probabilities. The BMA approach gives the results by averaging over the forecasts of all models with their posterior probabilities, that is,

$$E_t(e_{t+h} \mid I_t) = \sum_{r=1}^{R} E(e_{t+h,r} \mid M_r, I_t) p(M_r \mid I_t) \tag{8}$$

where $E_t(e_{t+h,r} \mid M_r, I_t)$ is the predicted value of the exchange rate at $t+h$ by model M_r with information set I_t, $p(M_r \mid I_t)$ is the posterior model probability for model M_r as derived in the Appendix, r is the index of model $1,\ldots, R$, I_t is the information set at time t including observations of the exchange rate, its determinants and estimated error terms from time 1, …, t.

2.5 Combined Forecast

As different models have their own merits, the idea that a combination of forecasts outperforms any individual forecast has attracted attention since it was introduced by Bates and Granger (1969).[9] The combined forecast used in this paper is constructed by assigning different weights to the forecasts obtained from the above models. The weighting scheme is similar to that in Stock and Watson (2001), in which the weights are assigned based on the relative mean squared forecast error (MSE) of the models' forecasts in the past, i.e. the smaller the forecast's mean squared error of a model is, the larger is the weight assigned to that particular forecast. Specifically, at time period T, the MSE for model $i=1,\ldots,n$ is

$$MSE_{T,i} = \frac{\sum_{t=1}^{T} (\ln e_t - \ln \hat{e}_{t,i})^2}{T} \tag{9}$$

[9] Bates and Granger (1969) document that a composite set of forecasts combined by two separate sets of forecasts is more accurate in terms of lower mean squared forecast error. Their justification of combining forecasts is that some useful information is lost if a single 'better' or 'the best' forecast is chosen whereas forecasts of other models are discarded. A more general case of combining more than two forecasts is shown in Newbold and Granger (1974), in which the forecast combination is found to be "frequently profitable".

where e_t is the realised exchange rate in the past (t = 1, …, T) and $\hat{e}_{t,i}$ is the predicted rate by model i. The weight for model $i=1,\dots,n$ is

$$w_{T+1,i} = \frac{1/MSE_{T,i}}{\sum_{j=1}^{n}(1/MSE_{T,j})} \tag{10}$$

In the analysis in Section IV, these relative performance weights are derived according to the forecasts obtained from the models in Section 2.1 - 2.4 (i.e. the PPP model, the UIP model, the SP model and the BMA model). The combined forecast at $T+1$ ($\hat{e}_{T+1,c}$) is constructed as

$$\ln \hat{e}_{T+1,c} = \sum_{i=1}^{n} w_{T+1,i} \ln \hat{e}_{T+1,i} \tag{11}$$

III. DATA AND OUT-OF SAMPLE FORECAST EVALUATION

3.1 The Data

Three currencies including the euro (EUR), British pound (GBP) and Japanese yen (YEN) against the US dollar (USD) are examined in this section. Quarterly, period-average data of the US, the UK, Germany and Japan from Q1 1973 to Q4 2007 are employed for the estimations. The data of the four countries used for the empirical estimations and their sources are given in Table 1.

Table 1. Data Sources

Variables	Sources and Details
Exchange Rates	Since the EUR/USD is not available until 1999 Q1, the series from 1973 Q1 to 1998 Q4 is traced back using the series of the Deutsche mark against the USD (DEM/USD) with a conversion factor stated in IFS World and Country Notes as 1.95583 DEM per USD. The series of EUR/USD, DEM/USD, GBP/USD and YEN/USD from 1973 Q1 to 2007 Q4 are from the International Financial Statistics (IFS) of the International Monetary Fund (IMF).
Stock Price	The series of share price index are from IFS.
Long-Term Interest Rate	The series of government bond yields are from IFS.
Short-Term Interest Rate	The financing bill rate from IFS is used for Japan, the Treasury Bill rate from IFS is used for the UK and the US. The series of Treasury Bill rate from IFS of Germany is not used since it is a rate on 12-month Federal debt register claims. Comparing to the rates of the US, the UK and Japan, the tenor of Germany's Treasury Bill rate is too long to be a proxy of short-term interest rate. Hence, the series of Frankfurt banks 3-month money market rate from CEIC is used for Germany.
Oil Price	The series of average crude price index is from IFS. This series is the average spot petroleum price index of Dubai Fateh, UK Brent and West Texas Intermediate.
Real GDP	The series of seasonally adjusted nominal GDP and GDP deflator (2000=100) are from IFS.
Money Supply	The M1 of the European Union is from CEIC, which is then seasonally adjusted by X12-ARIMA seasonally adjustment procedure of EViews. The seasonally adjusted M2 of the UK is from CEIC. The seasonally adjusted M1 of Japan and the US is from CEIC and IFS respectively.
Price Level	The consumer price indices (2000 = 100) are from IFS.
Current Account	The series of seasonally adjusted current account are from CEIC for Germany and the US, and from the Office of National Statistics for the UK. The series of Japan from 1973 Q1 to 1995 Q4 is from BIS database whilst the one from 1996 Q1 to 2007 Q3 is from CEIC.

Note: For the case of the UK, due to the discontinuation of the narrow money M1 series, the M2 series is used for the estimation, the study period of the UK starts from 1986 Q4, when the M2 series was released. The monthly M2 series was first released in September 1986 while the quarterly M2 was first released in 1986 Q4. Refer to the website of the Bank of England at http://www.bankofengland.co.uk/statistics/about/faq_all.htm for the reason of the ceased production of M1 of the UK.

3.2 Out-of-Sample Forecast Evaluation Criteria

For the purpose of comparison and assessment of the performance of the models, two benchmarks are employed. The first one is a simple random-walk model which is suggested by Meese and Rogoff (1983), who find that no estimated renowned theoretical model outperforms the random-walk model. The driftless random-walk model for an exchange rate in level is specified as

$$e_{t+h} = e_t + \varepsilon_t \qquad (12)$$

While the random-walk model is a conventional benchmark in the literature, another benchmark usually used in many other studies is the historical average return. The performance of the models in Section 2.1 - 2.5 will be compared with both the forecasts based on the random-walk model and the historical average return.

Four measures are used to assess the forecast accuracy of the models. The first measure is the ratio of the root mean squared forecast error (RMSE) of each model to that of the benchmark over time. If the ratio is less than one, the model in general predicts the actual value with a smaller error than the benchmark. The smaller the RMSE ratio is, the better is the forecast. In this measure, instead of giving the same weight in calculating the RMSE of forecasts at different time of the sample period, a factor suggested by Litterman and Winkelmann (1998) is applied. A factor of one is given to the latest forecast error at time *t*, 0.9 to the forecast error at time *t-1*, 0.9^2 to that at time *t-2*, and 0.9^n to that at time *t-n*.[10]

The second measure is the ratio of direction of change (DoC) given by the forecasts of the model. For each *ex-post* forecasting period, a 'one' will be assigned for the period when the model predicts correctly the direction of the actual exchange rate movement and zero otherwise. The DoC statistic is the proportion of "ones" over all the forecasting periods. If the statistic is greater than 0.5, this indicates that the model is better than the random-walk model which cannot predict the direction of the exchange rate movement. The higher the DoC is, the better is the forecast.

The third measure is the t-statistic which tests the null of forecast errors with zero mean. An acceptance of the null is desired for the forecast. The last one is a quantitative measure which counts the number of forecast errors with absolute values smaller than those of the random-walk model and the historical average return respectively. A 'one' is given to the period with a smaller forecast error and zero otherwise. Then a 'FE ratio' is calculated, which is the proportion of smaller forecast errors over all *ex-post* forecast horizons. The higher the ratio is, the better is the forecast.

[10] A time-decaying factor is applied to the calculation of the RMSE ratio because the performance in the more recent time should be more important in the model evaluation.

IV EMPIRICAL RESULTS

The *ex-post* forecasts from 1998 Q1 to 2007 Q4 of the five models specified in Sections 2.1 - 2.5 are examined for the EUR/USD, GBP/USD, and YEN/USD. Recursive estimation is adopted with the first sample as 1973 Q1 - 1997 Q4, that is, once a one-period ahead forecast is obtained, the sample is updated with one more period for the next one-period ahead forecast. The forecast horizons, *h*, examined in this study are one, two, three, four and eight quarters. The assessment of the forecast performance based on the four statistics as discussed in Section 3.2 are shown in Tables 2 to 4 for the EUR/USD, GBP/USD and YEN/USD respectively.

Table 2 shows the forecast accuracy statistics for the EUR/USD. There are a number of interesting observations from this Table. Based on the RMSE criteria, all models outperform the benchmarks of the random-walk model and historical average return, as the RMSE ratios are less than one. Among them, no single model consistently stands out as the best forecasting model in terms of RMSE ratios at all horizons. While the forecast from the PPP model and the combined forecast clearly outperform the others at longer horizons (eight-quarter ahead), they do not stand out when compared to the BMA model at shorter horizons (one- to two-quarter ahead).

Based on the RMSE measure, only the PPP model has better forecasting capability at longer horizons, whereas the performances of the forecasts from the other models are similar across different horizons.

Regarding the DoC statistics, the BMA forecast fares better as it has a higher proportion in line with the actual directional changes of the exchange rates than the others (except the SP model at longer time horizons). The BMA model is able to predict the direction of the exchange rate changes at about 60% or more of the time for all horizons.

The t-test shows that the forecasts from the BMA model perform poorly as their t-statistics at all horizons are all significantly different from zero. Among the three structural models (the PPP model, UIP model and SP model), the t-values of the PPP model are the smallest, indicating that the PPP is the best forecasting model based on this criterion.

As for the comparison of the *ex-post* forecast errors of the models with the benchmarks (either the random-walk or the historical average return), the ratios of the absolute forecast errors of the model smaller than the random-walk (FE_RW) and the historical average return (FE_HM) are used. Based on these ratios, the results are mixed. While the BMA model generally has more than half of the time with their forecast errors smaller than those of the benchmarks, they are only superior at short horizons. For long horizons (such as four- and eight-quarter ahead), the combined forecast is the best among all when these measures are used.

In summary, for the case of the EUR/USD, the forecasts from most of the models outperform the common benchmarks. However, among them, no single method clearly stands out. The PPP model is the best when the RMSE criterion is used, while the BMA and the SP model are superior to the others if the criterion is the DoC.

Table 3 shows the results of the case of the GBP/USD. In general, the model forecasts are not as good as in the case of the EUR/USD. For example, based on the RMSE measures, only the forecasts from the SP model consistently outperform the benchmarks. The forecasts from the BMA model and the combined method are better than the benchmarks only at longer horizons (eight-quarter ahead). Similar results are also observed for other criteria such as the DoC measure and the FE ratios. With the exception of the eight-quarter ahead forecasts from the BMA model and the combined method, the SP model is always superior to the others.

Table 4 presents the comparison results of the forecasts of the YEN/USD from different models. Similar to the EUR/USD, the forecasting models in general outperform the benchmarks with smaller RMSE ratios. Among them, the UIP and the combined forecasts have relatively better performance than the others. For other measures, the forecasts from the models are not much better than the benchmarks as their DoC statistics and their FE ratios are mostly around 0.5. All the models exhibit slightly weaker forecast ability at long horizons (e.g. four- to eight-quarter ahead) than the short ones (e.g. two-quarter ahead). In particular, the PPP model and the SP model underperform the benchmarks at long horizons.

Table 2. Evaluation of Forecasts of the EUR/USD

	Forecast Horizons (Quarters)				
	1	**2**	**3**	**4**	**8**
RMSE_RW					
PPP Model	0.78	0.79	0.77	0.74	0.60
UIP Model	0.83	0.87	0.88	0.88	0.82
SP Model	0.79	0.82	0.82	0.79	0.76
BMA Model	0.77	0.54	0.87	0.83	0.76
Composite Forecast	0.81	0.84	0.84	0.76	0.67
RMSE_HM					
PPP Model	0.83	0.84	0.82	0.85	0.69
UIP Model	0.87	0.93	0.95	1.02	0.96
SP Model	0.84	0.88	0.87	0.90	0.83
BMA Model	0.86	0.97	0.96	0.88	0.87
Composite Forecast	0.86	0.96	0.82	0.89	0.78
DoC					
PPP Model	0.56	0.58	0.54	0.50	0.59
UIP Model	0.41	0.47	0.46	0.53	0.69
SP Model	0.49	0.50	0.65	0.69	0.69
BMA Model	0.64	0.61	0.59	0.61	0.63
Composite Forecast	0.62	0.53	0.54	0.69	0.44
T-statistics					
PPP Model	-0.26	-0.19	-0.03	0.30	0.79
UIP Model	-1.07	-1.34	-1.47	-1.46	-1.83*
SP Model	-1.39	-1.59	-1.63	-1.46	-1.59
BMA Model	-2.31**	-3.08***	-3.27***	-2.69**	-1.80*
Composite Forecast	-1.24	-1.57	-1.70*	-2.02*	-1.45
FE_RW					
PPP Model	0.56	0.55	0.54	0.44	0.50
UIP Model	0.41	0.45	0.43	0.15	0.56
SP Model	0.49	0.50	0.62	0.69	0.66
BMA Model	0.59	0.66	0.48	0.53	0.53
Composite Forecast	0.26	0.32	0.32	0.67	0.69
FE_HM					
PPP Model	0.44	0.45	0.43	0.39	0.44
UIP Model	0.41	0.45	0.35	0.50	0.63
SP Model	0.44	0.42	0.46	0.47	0.38
BMA Model	0.54	0.50	0.49	0.50	0.50
Composite Forecast	0.44	0.37	0.46	0.61	0.75

Source: HKMA estimates

Notes:
1. RMSE_RW is the ratio of the root mean squared forecast error of the model to that of the random-walk whereas RMSE_HM is to that of the historical average return. A ratio less than one shows that the model performs better than the random-walk or the historical average return. The smaller the ratio is, the better is the models' forecast accuracy.
2. DoC measures the proportion of the ex-post forecasts in line with the actual change in the exchange rate. A number greater than 0.5 shows that the model is better than the random-walk in this measure. The larger the DoC is, the better is the forecast accuracy of the models.
3. T-statistics are from a t-test with a null of zero mean of the forecast errors. An acceptance of the null is desired. *, **, *** indicate significance at 10%, 5%, and 1% levels respectively.
4. FE_RW measures the proportion of the absolute forecast errors of the models' ex-post forecasts smaller than that of the random-walk while FE_HM compares with the absolute forecast errors of the historical mean. The larger the FE_RW or FE_HM is, the better is the forecast accuracy of the models.

Table 3. Evaluation of Forecasts of the GBP/USD

	Forecast Horizons (Quarters)				
	1	**2**	**3**	**4**	**8**
RMSE_RW					
PPP Model	1.12	1.27	1.32	1.31	1.13
UIP Model	1.13	1.28	1.34	1.35	1.11
SP Model	1.06	0.73	0.72	0.72	0.61
BMA Model	0.86	1.29	1.38	1.38	0.51
Composite Forecast	1.09	1.20	1.28	0.95	0.65
RMSE_HM					
PPP Model	1.40	1.53	1.56	1.48	1.14
UIP Model	1.42	1.55	1.60	1.52	1.12
SP Model	1.42	0.88	0.86	0.82	0.62
BMA Model	0.82	1.25	1.42	1.50	0.57
Composite Forecast	1.15	1.49	1.54	1.09	0.66
DoC					
PPP Model	0.46	0.50	0.41	0.47	0.66
UIP Model	0.44	0.50	0.41	0.42	0.50
SP Model	0.77	0.74	0.86	0.89	0.81
BMA Model	0.64	0.42	0.38	0.61	0.75
Composite Forecast	0.26	0.26	0.35	0.39	0.81
T-statistics					
PPP Model	-1.02	-1.17	-1.41	-1.61	-2.28**
UIP Model	-0.71	-0.80	-1.02	-1.18	-1.67
SP Model	-0.43	-0.81	-1.21	-1.39	-2.47**
BMA Model	-0.85	-3.85***	-3.07***	-2.99***	-0.30
Composite Forecast	-1.11	-1.92*	-2.02*	-2.27**	-2.72**
FE_RW					
PPP Model	0.38	0.34	0.32	0.39	0.63
UIP Model	0.38	0.32	0.35	0.33	0.47
SP Model	0.59	0.66	0.81	0.72	0.72
BMA	0.54	0.42	0.30	0.53	0.72
Composite Forecast	0.31	0.45	0.49	0.69	0.81
FE_HM					
PPP Model	0.36	0.29	0.38	0.44	0.53
UIP Model	0.36	0.29	0.30	0.33	0.41
SP Model	0.49	0.68	0.68	0.81	0.75
BMA Model	0.54	0.29	0.30	0.56	0.78
Composite Forecast	0.49	0.37	0.38	0.72	0.84

Source: HKMA estimates

Notes:

1. RMSE_RW is the ratio of the root mean squared forecast error of the model to that of the random-walk whereas RMSE_HM is to that of the historical average return. A ratio less than one shows that the model performs better than the random-walk or the historical average return. The smaller the ratio is, the better is the models' forecast accuracy.
2. DoC measures the proportion of the ex-post forecasts in line with the actual change in the exchange rate. A number greater than 0.5 shows that the model is better than the random-walk in this measure. The larger the DoC is, the better is the forecast accuracy of the models.
3. T-statistics are from a t-test with a null of zero mean of the forecast errors. An acceptance of the null is desired. *, **, *** indicate the null is rejected at 10%, 5%, and 1% significance levels.
4. FE_RW measures the proportion of the absolute forecast errors of the models' ex-post forecasts smaller than that of the random-walk while FE_HM compares with the absolute forecast errors of the historical mean. The larger the FE_RW or FE_HM is, the better is the forecast accuracy of the models.

Table 4. Evaluation of Forecasts of YEN/USD

	Forecast Horizons (Quarters)				
	1	**2**	**3**	**4**	**8**
RMSE_RW					
PPP Model	0.70	0.78	0.90	1.01	1.33
UIP Model	0.69	0.71	0.73	0.72	0.79
SP Model	0.70	0.73	0.77	0.78	0.91
BMA Model	0.68	0.75	0.85	0.83	1.03
Composite Forecast	0.67	0.72	0.76	0.79	0.85
RMSE_HM					
PPP Model	0.74	0.83	0.98	1.07	1.38
UIP Model	0.72	0.75	0.78	0.74	0.78
SP Model	0.73	0.77	0.83	0.81	0.91
BMA Model	0.64	0.76	0.87	0.86	1.12
Composite Forecast	0.65	0.70	0.84	0.82	0.86
DoC					
PPP Model	0.49	0.47	0.54	0.56	0.47
UIP Model	0.46	0.47	0.54	0.56	0.47
SP Model	0.41	0.47	0.54	0.56	0.47
BMA Model	0.44	0.50	0.41	0.39	0.54
Composite Forecast	0.44	0.47	0.51	0.56	0.50
T-statistics					
PPP Model	-0.89	-1.71*	-2.61**	-3.72***	7.98***
UIP Model	0.36	0.53	0.77	1.30	3.34***
SP Model	0.34	0.49	0.70	1.15	2.68**
BMA Model	0.21	-0.24	-0.44	-0.45	0.18
Composite Forecast	-0.02	0.41	0.76	1.70*	3.19***
FE_RW					
PPP Model	0.41	0.47	0.51	0.39	0.31
UIP Model	0.44	0.42	0.57	0.50	0.41
SP Model	0.38	0.50	0.57	0.50	0.44
BMA Model	0.33	0.50	0.32	0.53	0.50
Composite Forecast	0.36	0.42	0.46	0.47	0.50
FE_HM					
PPP Model	0.38	0.37	0.43	0.39	0.25
UIP Model	0.62	0.58	0.73	0.72	0.81
SP Model	0.51	0.50	0.43	0.53	0.66
BMA Model	0.49	0.42	0.41	0.50	0.44
Composite Forecast	0.44	0.42	0.51	0.50	0.47

Source: HKMA estimates

Notes:

1. RMSE_RW is the ratio of the root mean squared forecast error of the model to that of the random-walk whereas RMSE_HM is to that of the historical average return. A ratio less than one shows that the model performs better than the random-walk or the historical average return. The smaller the ratio is, the better is the models' forecast accuracy. The RMSE ratios of the eight-quarter BMA model of the Japanese Yen exclude an outlier estimate at 2007 Q2 with return rate at 0.80 at 2005 Q2. The RMSE ratios to RW and to HM are 1.2375 and 1.8271 respectively.
2. DoC measures the proportion of the ex-post forecasts in line with the actual change in the exchange rate. A number greater than 0.5 shows that the model is better than the random-walk in this measure. The larger the DoC the better is the forecast accuracy of the models.
3. T-statistics are from a t-test of null of zero mean forecast errors. An acceptance of the null is desired. *, **, *** indicate the null is rejected at 10%, 5%, and 1% significance levels.
4. FE_RW measures the proportion of the absolute forecast errors of the models' ex-post forecasts smaller than that of the random-walk while FE_HM compares with the absolute forecast errors of the historical mean. The larger the FE_RW or FE_HM is, the better is the forecast accuracy of the models.

V. CONCLUSIONS

For macroeconomic-analysis and market-surveillance purposes, it is essential for policymakers to regularly monitor the exchange-rate movements of the major currencies. This study examines and compares the predictability of the exchange rate forecasting models supported by (i) the PPP model; (ii) the UIP model; (iii) the SP model; (iv) the BMA model; and (v) a combined forecast.

In terms of the exchange-rate predictability, empirical results suggest that the PPP model, UIP model and SP model are in general able to outperform the random-walk model as well as the historical average return for the forecast of the exchanges rates of EUR/USD and YEN/USD, but not for that of GBP/USD.[11] The results are consistent with Cheung et al (2004) who concludes that a particular model may do well for one exchange rate but not for the others.

For the forecast-performance comparison of the five models, the results suggest that no single model in this study consistently outperforms the others. The forecast capability of a particular model depends on the exchange rate of interest and the forecast horizon. Given the uncertainty in selecting a model, the combined forecasts seem to have an edge over the others and in general have relatively smaller RMSE ratios and higher percentage in predicting the direction of changes correctly.[12] In spite of the attractiveness of combining forecasts, it should be stressed that forecasting exchange rate movement is still a daunting task. The forecasts of exchange rates from any of these models should be used with caution.

[11] The SP model is the best for GBP/USD.

[12] Sometimes, the results from the t-statistics and the FE ratios show that the combined forecast outperforms other models. However, these criteria are not as important as the RMSE and DoC in the literature.

Appendix: Determination of Posterior Model Probability in the Bayesian Moving Average Model

Given model uncertainty among the different models, the BMA weighs each model by a posterior probability. Based on the Bayes' theorem, the posterior probability is a probability of parameters in interest, ϕ, which for example can be the coefficients of parameters of the models or the predicted value of the dependent variable, conditional on the set of observed data *y*. The posterior probability $p(\phi \mid y)$ is given as

$$p(\phi \mid y) = \frac{p(y \mid \phi) p(\phi)}{p(y)} \tag{A1}$$

where $p(y \mid \phi)$ is the probability density function (p.d.f.) of *y* given ϕ and $p(\phi)$ is the prior density.

There are altogether *R* candidate models, the posterior of ϕ in the *r*th model, M_r where $r = 1, \ldots, R$, is

$$p(\phi_r \mid y, M_r) = \frac{p(y \mid \phi_r, M_r) p(\phi_r \mid M_r)}{p(y \mid M_r)} \tag{A2}$$

As $\int p(\phi_r \mid y, M_r) d\phi_r = 1$, integrating both sides of equation (A2) with respect to ϕ_r gives

$$p(y \mid M_r) = \int p(\phi_r \mid M_r) p(y \mid \phi_r, M_r) d\phi_r \tag{A3}$$

The left hand side is the marginal likelihood which is essential to calculate the posterior model probability in the BMA method. To derive the posterior model probability, we first note that based on the Bayes' theorem, this can be written as follows

$$p(M_r \mid y) = \frac{p(y \mid M_r) p(M_r)}{p(y)} \tag{A4}$$

where $p(M_r \mid y)$ is the posterior model probability for model M_r, $p(y \mid M_r)$ is the marginal likelihood from equation (A3), $p(M_r)$ is the prior model probability. In the Bayesian framework, as the denominator *p(y)* is treated as a constant, the posterior model probability is given by

$$p(M_r \mid y) = c p(y \mid M_r) p(M_r) \tag{A5}$$

where *c* is a constant.

Two pieces of prior information enter equation (A5): the prior model probability ($p(M_r)$) and the prior of ϕ ($p(\phi \mid M_r)$) which is a component of $p(y \mid M_r)$. Meaningful posterior probabilities require a proper prior. Standard choices of priors are used in this study. For the first piece of prior information, an equal prior probability is allocated to each model so that

$$p(M_r) = \frac{1}{R} \tag{A6}$$

As $\sum_{r=1}^{R} p(M_r \mid y) = 1$, integrating both sides of equation (A5) gives the constant c as

$$c = \frac{R}{\sum_{i=1}^{R} p(y \mid M_i)} \tag{A7}$$

and the posterior model probability can be written as

$$p(M_r \mid y) = \frac{p(y \mid M_r)}{\sum_{i=1}^{R} p(y \mid M_i)} \tag{A8}$$

Equation (A8) presents the posterior model probability in terms of the marginal likelihood defined in (A3).

It is noted from equation (A3) that the marginal likelihood has two components. The first one is the prior of ϕ, $p(\phi \mid M_r)$.[13] In this study, we employ a commonly used benchmark prior for $p(\phi \mid M_r)$: the natural conjugate g-prior. This prior which is introduced by Zellner (1986) requires only one scalar parameter, g, to be specified. Moreover, it is in the class of natural conjugate priors which is distributed as normal-gamma.[14]

---

[13] Treating model M_r, for $r = 1,\ldots,R$, as a linear regression model in this study, the parameters to be estimated, which are the focus of interest ϕ, are the coefficients of the variables, β, and error variance, σ^2.

[14] Priors are information that the researchers have in the estimation. Basically, priors can be in any form. The natural conjugate priors possess two advantages. First, when it multiplies the likelihood, a posterior in the same distribution is yielded. Second, the natural conjugate prior has the same functional form as the likelihood function, and hence the prior information has the same interpretation as the likelihood function. While the likelihood function means the probability of the observed data given the parameters in certain distribution(s), the prior can be described as the probability of unseen data given the parameters in the same distribution(s) as those in the likelihood function. Since the likelihood function can be written as a product of a normal density and a gamma density, so does the prior.

The second component in equation (A3) is the likelihood function. With the assumption of normality, the likelihood function $p(y \mid \phi, M_r)$ is

$$p(y \mid \beta_r, \sigma_r^2, M_r) = \frac{1}{(2\pi\sigma^2)^{N/2}} \{\exp[-\frac{1}{2\sigma^2}(y - X_r\beta_r)'(y - X_r\beta_r)] \qquad \text{(A9)}$$

where N is the sample size, y is a $N \times 1$ matrix of the dependent variable, X_r is a $N \times k_r$ matrix of the independent variables in model M_r, and k_r is the number of independent variables including the constant in model M_r. The likelihood can be further written as[15]

$$p(y \mid \beta_r, \sigma_r^2, M_r) = \frac{1}{(2\pi)^{N/2}} \{\frac{1}{\sigma_r} \exp[-\frac{1}{2\sigma_r^2}(\beta_r - \hat{\beta}_r)' X_r' X_r (\beta_r - \hat{\beta}_r)]\} \{\frac{1}{\sigma_r^{\upsilon}} \exp[-\frac{\upsilon s_r^2}{2\sigma_r^2}]\} \qquad \text{(A10)}$$

where $\hat{\beta}_r$ is the Ordinary Least Square (OLS) estimators for β_r, υ is the degree of freedom in the OLS estimation for β_r, and s_r is the OLS standard error. By equation (A10), the likelihood function is a product of a normal density (the first curly bracket) and a gamma density (the second curly bracket). This form suggests that the prior for β_r and σ_r^2 ($p(\beta_r, \sigma_r^2)$) is distributed as normal-gamma. Moreover, the joint probability $p(\beta_r, \sigma_r^2)$ can be expressed as a product of two probabilities as

$$p(\beta_r, \sigma_r^2) = p(\beta_r \mid \sigma_r^2) p(\sigma_r^2) \qquad \text{(A11)}$$

In this study, $\beta \mid \sigma_r^2$ is specified as in a normal distribution while σ_r^2 is as in a gamma distribution. With the adoption of the g-prior in Zellner (1986), $\beta_r \mid \sigma_r^2 \sim N(0, \sigma_r^2 g_r (X_r' X_r)^{-1})$ which has mean zero and covariance matrix $\sigma_r^2 g_r (X_r' X_r)^{-1}$. The value of g is chosen according to Fernandez, Ley and Steel (2001) as

$$g_r = \begin{cases} 1/k_r^2 & if \quad N \le k_r^2 \\ 1/N & if \quad N > k_r^2 \end{cases} \qquad \text{(A12)}$$

According to Raftery, Madigan and Hoeting (1997), the prior of σ_r^2 is chosen as a chi-squared distribution.[16]

---

[15] See Koop (2003) for computational details.
[16] The chi-squared distribution is a gamma distribution with the mean equal to the variance.

With the likelihood specified in equation (A10) and the priors described above, the marginal likelihood for each model is then computed by equation (A3). Equation (A8) is then used to compute the posterior model probability. Finally, the posterior of ϕ (β and $\hat{y}$ in this study), is derived based on the rules of probability as

$$p(\phi \mid y) = \sum_{r=1}^{R} p(\phi \mid y, M_r) p(M_r \mid y) \quad \text{(A13)}$$

In this study, sixteen independent variables are included in the estimation. The number of candidate models, R, to assess equals to 2^{16}. With this number, the computation of the posterior model probability is very intensive. Therefore, the Markov chain Monte Carlo model composition (MC^3) methodology developed by Madigan and York (1995) is employed to simulate the model space for an approximation of the denominator of the posterior model probability in equation (A8).[17] The BMA program used in this study is modified from the one available on the website of the language **R**.[18]

[17] See Raftery et al. (1997) for a brief introduction of the MC^3 methodology.

[18] The original program is available at http://cran.r-project.org/web/packages/BMA/index.html. For details of the language **R**, please refer to http://www.r-project.org.

REFERENCES

Alexius, A. (2001): "Uncovered Interest Parity Revisited," *Review of International Economics*, 9, 505-17.

Avramov, D. (2002): "Stock Return Predictability and Model Uncertainty," *Journal of Financial Economics*, 64, 423-58.

Bates, J.M., and C.W.J. Granger (1969): "The Combination of Forecasts," *Operations Research Quarterly*, 20, 451-68.

Cheung, Y.W., M. Chinn, and A.G. Pascual (2004): "Empirical Exchange Rate Models of the Nineties: Are Any Fit to Survive?," IMF Working Paper WP/04/73.

Chinn, M.D., and R.A. Meese (1995): "Banking on Currency Forecasts: How Predictable is Change in Money," *Journal of International Economics*, 38, 161-78.

Cremers, K.J.M. (2002): "Stock Return Predictability: A Bayesian Model Selection Perspective," *Review of Financial Studies,* 15, 1223-49.

Dornbusch, R. (1976): "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84, 1161-76.

Engle, C.M. (1992): "Can the Markov Switching Model Forecast Exchange Rates?" NBER Working Paper No.4210.

Fernandez, C., E. Ley and M.F.J. Steel (2001): "Model Uncertainty in Cross-Country Growth regressions," *Journal of International Economics*, 60, 35-59.

Frankel, J.A. (1979): "On the Mark: A Theory of Floating Exchange Rates Based on Real Interest Differentials," *American Economic Review*, 69, 610-22.

Frankel, J.A. and A.K. Rose (1994): "A Survey of Empirical Research on Nominal Exchange Rates," NEBR Working Paper No. 4865.

Koop, G. (2003): *Bayesian Econometrics*, West Sussex: Wiley.

Leamer, E.E. (1978): *Specification Searches*, New York: Wiley.

Litterman, R., and K. Winkelmann (1998): "Estimating Covariance Matrices," *Goldman Sachs Risk Management Series*, January.

Madigan, D. and J. York, (1995): "Bayesian Graphical Models for Discrete Data," *International Statistical Review*, 63, 215-32.

Mark, N. (1995): "Exchange Rates and Fundamentals: Evidence on Long Horizon Predictability," *American Economic Review*, 85, 201-18.

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

Messe, R.A., and A. K. Rose (1991): “An Empirical Assessment of Non-Linearities in Models of Exchange Rate Determination,” *Review of Economic Studies*, 58,603-2-19.

Meredith, G. and M. Chinn (1998): “Long Horizon Uncovered Interest Rate Parity,” NBER Working Paper 6769.

Min, C. and A. Zellner (1993): “Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates,” *Journal of Econometrics*, 56, 89-118.

Mussa, M. (1979): “Empirical Regularities in the Behaviour of Exchange Rates and Theories of the Foreign Exchange Market,” in Karl Brunner and Allan H. Meltzer, (eds.), *Policies for Employment, Prices and Exchange Rates*, Carnegie-Rochester Conference 11, North-Holland, Amsterdam.

Newbold, P., and C.W.J. Granger (1974): “Experience with Forecasting Univariate Time Series and the Combination of Forecasts,” *Journal of the Royal Statistical Society*, 137, 131-65.

Raftery, A., D. Madigan and C.T. Volinsky, (1996): “Accounting for Model Uncertainty in Survival Analysis Improves Predictive Performance (with Discussion),” in J. Bernardo, J. Berger, A. Dawid and A. Smith (eds.), *Bayesian Statistics 5*, Oxford: Oxford University Press.

Raftery, A.E., D. Madigan and J.A. Hoeting (1997): “Bayesian Model Averaging for Linear Regression Models,” *Journal of the American Statistical Association*, 92, 179-91.

Stock, J.H. and M. Watson (2001): “A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series,” in R.F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting: a Festschrift in Honour of Clive Granger,* Oxford: Oxford University Press.

Wright, J.H. (2003): “Bayesian Model Averaging and Exchange Rate Forecasts,” International Finance Discussion Papers No. 779, Board of Governors of the Federal Reserve System.

Zellner, A. (19860: “On Assessing Prior Distributions and Bayesian Regression Analysis with g-Prior Distributions,” in P.K. Goel and A. Zellner (eds.), *Bayesian Inference and Decision Techniques: Essays in Honour of Bruno de Finetti*, North-Holland, Amsterdam.