The Out-of-Sample failure of Empirical Exchange Rate Models – A Cointegration perspective.

Martina Porubcova
University of St. Gallen
Zurich, CH 8057
martina.porubcova@student.unisg.ch

Marco Hassan
University of St. Gallen
Milano, IT 20124
marco.hassan@student.unisg.ch

April 19, 2019

Abstract

This study reports an estimation of the out-of-sample performance of univariate and multivariate monetary models contributing to the analysis of the Meese Rogoff paradox for the empirical CHF/USD, GBP/USD and YEN/USD series. 20 years of monthly observations and five models are analyzed. In line with the academic literature we find a random walk model to be the best univariate model to capture the unit root behaviour of FX-rates and to forecast the latter out-of-sample. Nonetheless, in contrast with the general academic literature we do not find multivariate models to underperform a random walk model when forecasting FX-rates out of sample. Finally, we find that a vector error correction model representing the cointegration evidence among macroeconomics fundamentals and FX-rates outperforms a random walk model for all the country series and estimation lags.

Keywords FX-rates · Monetary Models · Meese- Rogoff Puzzle · Cointegration · VECM

1 Introduction

Many approaches tried to model FX-rates since the start of a free-floating system after Bretton Woods in the 1970s. Monetary models, linking macroeconomics variables to FX-rates were the first to emerge and due to their compelling theoretical intuition soon became the reference in the field. The enthusiasm for such elegant models was nonetheless diminished by two seminal works. Meese and Rogoff (1983a, [1]) and Meese and Rogoff (1983b [2]) demonstrated how no one of the different monetary models could beat a random walk without a drift in modeling FX-rates out of sample. After the evidence was presented the opinion divided. Some argued for a sample issue and confined the problem to the specific historical period, some argued for the limit of the linear models used by the authors to approach monetary models, some argued for a fail of monetary models and their underlying assumptions and, finally, some started to approach the issue in an innovative way turning towards micro-based models aiming to capture the complexity of market information asymmetries and investors heterogeneity [1].

While all such different reactions contributed to a further understanding of FX-rates properties we question whether the runaway from linear monetary models is justified given the methodological approach presented in the studies of Meese and Rogoff, the small sample of use and the lack of a cointegration analysis later formalized by Granger and Engle (1987, [4]). Given the gap in academic literature exploring this cointegration relationship our study will propose a study of it at monthly lag. This methodological study stands therefore as a revision of the out of sample failure of foreign exchange rates testing whether macroeconomics variables are indeed incapable of explaining FX-rates movements before exploring more nuanced fields such as nonlinear and micro-based modeling. The hope is therefore to avoid over engineering and over complication of the issue going back to the fundamentals.

In the specific, our study aims at exploring the above at a monthly lag frequency on a self developed data set for the recent period ranging from 1986 to 2006. Such period will allow a consistent estimation of all of the necessary parameters given the models of choice, moreover it will avoid to model the most recent financial crises where structural breaks are present in the series. As in Meese and Rogoff (1983a and 1983b, [1, 2]) we are going to explore the out of sample forecasting perfomance of foreign exchange rates for three major monetary models. Firstly, the monetary models first presented by Frenkel (1976, [5]) and Mussa (1976, [6]) claiming for a link among monetary mass differentials, interest rates differentials, output gap and FX-rates. Secondly the sticky price model where inflation rates differential plays a crucial role as prices adjusts sluggish to macroeconomics shocks (Dornbush 1976, [7]). Finally, a portfolio balance model claiming for the importance of monetary flows and net current account differentials as a key determinant for FX-rates changes as in Hooper and Morton (1982, [8]).
Based on such monetary models five different models will be analyzed in order to forecast FX-rates. Three of them are presented in the seminal work of Meese and Rogoff (1983a, [1]) and will pose the benchmark to check whether the result of the paper hold in the time frame of interest. These are a simple random walk forecast, an OLS estimation without inclusion of lagged terms and a vector autoregressive model. On the top of it we will analyze the performance of transfer models discussed in Montgomery and Weatherby (1980, [9]) and the vector error correction model representation of the cointegration relationship between macroeconomics variables and FX-rates proposed by Granger and Engle (1987, [4]).

If our results are in line with the seminal papers of Meese and Rogoff in the case of univariate models the picture resulting from a multivariate fit contradicts the results of the reference papers. We do not find enough evidence to support the claim of an underperformance of multivariate linear models in comparison to a simple random walk model. Moreover we find strong evidence for cointegration among macroeconomics series and FX-rates for all of the analyzed country series. This suggests a long term stable relation between the fundamentals and the FX-rates rejecting the claim of Meese and Rogoff of a weak relation between the two. Finally, we do find that a vector error correction model representation of the cointegration relation between macroeconomics series and FX-rates outperforms the out-of-sample forecast of FX-rates at all lags and for all of the country series further strengthening the hypothesis that a link between monetary models and FX-rates does indeed exists at monthly lag frequency when the models are well behavied and properly estimated.

[1] See Balliu and King (2005, [3]) for a general survey.

The paper continues as follows. Section 2 outlines the methodological approach used in the paper and explains the various models and techniques used for the empirical analysis and the out-of-sample forecast. Section 3 continues by briefly introducing the dataset used and the chosen proxies to capture the macroeconomics fundamentals of interest. Section 4 reports the main results of the empirical analysis and Section 5 concludes.

2 Methodology

2.1 Structural Model

The basic structural model encompassing all of the different monetary models discussed in the introductory session is of the following form:

$$s = \beta_0 + \beta_1(m - m^*) + \beta_2(y - y^*) + \beta_3(r - r^*) + \beta_4(\pi - \pi^*) + \beta_5(TB - TB^*) + \varepsilon \tag{1}$$

where s represents the logarithm of the indirect quote of FX-rates, i.e. the foreign exchange value of a dollar unit, $m - m^*$ represents the logarithm of the U.S. and foreign country money mass supply differential, $y - y^*$ represents the logarithm of the U.S. and foreign GDP level, $r - r^*$ represents the short term U.S. forign country interest rate differential, $\pi - \pi^*$ represents the U.S. and foreign country inflation rate differential, $TB - TB^*$ represents the U.S. foreign country current account differential normalized to be one at the beginning of the sample period and, finally, where ε represents the error term of the regression capturing all of the other factors not expressed in the model.
The model above encompasses all of the monetary models discussed. In the specific the most basics monetary model, the Frenkel-Bilson model logically inferred from the relative PPP proposition assumes $\beta_4 = \beta_5 = 0$. The Dornbusch model allowing a sluggish price adjustment behaviour and predicting FX-overshooting assumes $\beta_5 = 0$ and finally the Hooper-Morton model poses no restriction on the coeffcients of equation 1. Such a structural model will consequently be estimated on four different parametric models of interest, which will be discussed next.

2.2 Univariate Models

The most basic model that will be tested is an OLS model as in the benchmark papers of Meese and Rogoff (1983a, 1983b, [1, 2]). In this basic first model the coefficients of equation 1 will be computed without looking at any lagged effect. In contrast to it a second more flexible model will be applied to capture possible lagged effects of the macroeconomics fundamentals. In comparison to the Meese and Rogoff papers where the authors decided to capture the possibility of lagged terms by incorporating autoregressive models giving a higher importance on more recent observations we decided to apply the transfer function models widely spread in the fields of engineering such as control systems and electronic circuits. These models have been poorly discussed in the field of economics with the exception of Tustin (1957, [10]) that tried to make the point for applying the models into the economics modeling field.
Transfer function models relates a given set of inputs to an output variable through the following general formula

$$Y_t = \mu + \frac{(\omega_0 + \omega_1 B^1 + \cdots + \omega_s B^s)}{1 - \delta_1 B^1 - \cdots - \delta_r B^r} X_{t-b} + \varepsilon_t \tag{2}$$

where X represent a matrix of exogenous terms, μ the optional term modeling the mean of the series and ε the non-captured variation in the series. Moreover the above ratio represents the transform function of the regressors matrix and is especially caracterized by the order of the denominator and nominator terms. r, the order of the denominator term, expresses the rate of the decay pattern, where a higher term indicates a slower decay. s, the order of the nominator term, expresses the persistence of so called unpatterned spikes, that is the persistence of effects that are not captured in the decay pattern. Finally, the b teerm in the matrix of input represent the dead time, that is the time it takes for the dependent variable to react to some changes in the input matrix. This is of primary importance as it might very well be that some shift in macroeconomics fundamentals just start to display effects after a certain amount of time when the economics actors start to perceive the change.
Due to the flexible nature of the latter model, we believe that it is better suited and preferrable to capture the true nature of distributed lags present in the structural model of equation 1 in comparison with the smoothed autoregressive model applied by Meese and Rogoff with the arbitrarly chosen smoothing term of 0.95 (See Meese and Rogoff (1983a pp. 7, [1])).

2.3 Multivariate Models

The above discussed models rely on the exogeneity of the independent variables. If the assumption fails the model will be biased and would lead to misleading conclusion as the independence of the sample distribution would not be guaranteed. In practice this poses an issue for the estimation of the structural model of equation 1. While some variables as the monetary mass and the output gap are commonly treated as exogenous variables in the underlying monetary models the practice suggest that such macroeconomic variables might very well be influenced by the movements of FX-rates. On the top of it other variables such as short term interest rates differentials are treated as endogenous even in the underlying monetary models and therefore require a different estimation compared to the one outlined in the models of the previous section.
To obviate the above issues Meese and Rogoff (1983a and 1983b, [1, 2]) analyzed the structural model of interest through a vector autoregressive model firstly introduced by Sims (1980, [11]). In the general case the model consists of a system of equations of the form

$$y_t = \Phi_0 + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \cdots + \Phi_n y_{t-n} + \varepsilon_t \tag{3}$$

where y is a vector containing the endogenous variables of interest, Φ_0 is a vector of constants, $\Phi_1, \dots, \Phi_n$ are matrices describing the effect of lagged endogenous variables on the levels of the current variables and ε captures the equations specific error term. The resulting model will capture the endogeneity present in the monetary structural model allowing a consistent OLS estimation as far as the error terms of the equations will be uncorrelated.

Despite the described vector autoregressive model well manages to model the endogeneity of macroeconomic variables we question whether a restricted form of it could yield more efficient and reliable estimates capturing the FX-rates and the monetary models relation. This is especially motivated by the cointegration theory developed by Engle and Granger (1987, [4]) and the well known evidence of non stationary macroeconomics times series [2]. Moreover, the impportance of such mehtod is underlined by evidence of Phillips (1986, [13]) that set down the theoretical fundamentals showing how

[2] See for instance Gil-Alana and Robinson (1997, [12])

parameter estimates of cointegrated series will not converge in probability and will not converge to any non-degenerate distribution in the asymptotic case if the case of a misspecified OLS estimate as potentially is equation 3.

We propose therefore a test for cointegration among the macroeconomics and the FX-rates series based on Johansen (1991, [14]) and we consequently estimate a vector error correction model of the form

$$\begin{aligned}
\Delta y_t &= \Pi y_{t-1} + \sum_{i=1}^{p-1} \Phi^* \Delta y_{t-i} + \varepsilon_t \\
\Phi^* &= -\sum_{i=j+1}^{p} \Phi_i, \;\; j = 1, \ldots, p-1 \\
\Pi &= -\left(I - \Phi_1 - \cdots - \Phi_p\right)
\end{aligned} \tag{4}$$

where Πy_{t-1} of equation 4 represent the error correction term and y_i and Φ refer to the variables described in the unrestricted vector autoregressive model of 3.

2.4 Forecasting Approach

The four described models described and a random walk model without drift will be validated by looking at their ability to forecast FX-rates out of sample. In this sense the 234 observations sample is splitted in a training and validation sample. Three fourth of the total observations will be used for training the four models outlined above and the rest of the observations will be used for the out of sample validation.

In comparison to Meese and Rogoff (1983b, [2]) we will not try any restricted estimation based on the theoretical monetary models literature but will rather estimate unrestricted versions for all of the models outlined with the exception of the vector error correction model described in 4, given the by product VAR model restriction imposed by the latter. All of the four different models will be estimated according to the most general structural model described in 1 without restrictions. Based on the structural model results Wald tests on the restricted structural model will be computed in order to check whether there is in sample evidence to consider restricted monetary models.

Based on the first estimation we proceeded by calculating the out of sample performance of the different models applying a rolling forecast in analogy to the benchmark papers. This will consist of a reestimation of the four outlined models for each new forecasting point. Given the decision to estimate the out of sample model performance at one, three, six and twelve months lags the rolling forecast technique involves a model reestimation at such frequencies. Important is nonetheless to underline how the results of the Wald tests will be extended for each of the subsequent model estimation in the rolling forecast. This means that given statistical significant evidence for the Null of a restricted model in the first three fourth of the sample the same restricted model will be used for the subsequent model estimations. Explicitely this method consists to fit a structural model containing all of the macroeconomics series in the first 176 observations, that is from September 1986 to April 2001, and perform Wald tests to check whether a restricted monetary model is supported in sample. Based on such restults the rolling forecast method involves first to estimate the next point forecast at 1, 3, 6 and 12 months and in a second step to shift the data sample of one period such that a new model will be estimated for the October 1986 to May 2001 series and, according to the new parametric fit, point forecast at 1, 3, 6 and 12 months will be estimated. This method will then be iterated until the last out of sample observation for February 2006 is reached.

Finally, as in the benchmark papers, we will allow the univariate models described in 2.2 a richer set of information compared to the random walk. In the specific the random walk and the multivariate model will use the $\mathscr{F}_{t-1} = F_0, \ldots, F_{t-1}$ information set, where F_i represent a set containing all of the available information at timepoint i. By contrast the univariate models will dispose of $\mathscr{F}_t \setminus s_t$, where s_t represents the FX-rate at time point t. In simple terms this means that we are going to give the univariate models the advantage of using the actual realizations of macroeconomics variables without the need of estimating them.

Based on the obtained rolling forecasts three statistics will be computed to compare the model fit. These are the root mean squared error (RMSE), the mean absolute error (MAE) and the mean directional accuracy (MDA). A particular importance will be assigned to the MAE results given Westerfield (1977, [15]) that analyzed the empirical exchange rates distribution finding evidence for FX-rates non-normal stable-Paretian distributions with infinite variance.
In contrast with Meese and Rogoff (1983a and 1983b, [1, 2]) we decided to further explore the point estimators of the above mentioned statistics by computing a model confidence set (MCS) as described in Hansen et al. (2011, [16]). The idea of the latter consists of a procedure yielding a model set, $\mathcal{M}^*$, built to contain the best model with a chosen level of confidence. The exact procedure is based on an equivalence test δ_m and an elimination rule ϵ_m, consistent with the chosen confidence level. In a first step the competing models will be compared with each other. If δ_m does not support evidence for the equal performance of the models, ϵ_m is applied and the poorly performing models are discarded and the general problem reiterated until δ_m is accepted for all of the surviving models.
In this paper we decided to make use of superior predictive ability test of Hansen (2005, [17]) and to obtain p-values for the equal predictive hypothesis of models according to a bootstrap implementation outlined by Hansen et al. (2011, [16]).

3 Dataset

The dataset, comprehensively discussed in Appendix A 6 consists of monthly observations for a 21 years time frame ranging from January 1986 to January 2006. All the times series of use are selected in accordance with the underlying monetary structural model 1 and are consistent with the series first utilized by Meese and Rogoff (1983a, [1]). Three different FX-rates series will be analyzed, namely the JPN/USD rate, the GBP/USD rate and the CHF/USD rate. With respect to the independent variables of the structural model, we decided to work with the 3-months treasury bills to capture the short term interest rates differentials with the only exception of CH series where we worked with the 3-month US-CH LIBOR spread because of missing reported public available data for the seasonally unadjusted short term treasury rates at monthly lag.
For measuring money mass differentials we worked using the M1 and M3 measures testing the both given the important difference among the two and the lack of theoretical models to favour one measure above the other. We obtained stronger results for the M3 series and therefore decided to report the results obtained trough such measure in the paper.
For measuring the inflation rate differentials important for the Dorbusch model we worked using the consumer price indeces and for measuring the trade balance we worked using the net trade amount for goods as a proxy for the current account balance. The latter is of course just a rough approximation given the fraction of traded goods in comparison to

the total amount of trade including services and monetary transfers, nonetheless we preferred the latter measure to an interpolation of the quarterly published current account balances.
Finally, we worked with the unemployment rate to capture the output differential of the economies as we believe that such variable well manages to capture the fluctuations in output levels, the quantitiy of interest, leaving the constant term in equation 1 to capture the other information present in the output measure such as the size of the economies.

As in our reference papers we decided to work using not seasonally times series in order to avoid the possible bias introduced by different seasonal adjustments on the structural parameters documented in Sims (1973a and 1973b, [18, 19]). Henceforth we decided to approach our model validation in two different ways fitting the models on seasonal adjusted and detrended times series and on untreated times series in levels. This will of course have important consequences for the interpretation of the model results. While detrended and seasonally adjusted stationary times series will allow an interpretability of the results the latter is not guaranteed for the not treated times series where spurious regressions might arise distorting the model interpretation. Given the fact, that the goal of this work is nonetheless a reliable forecast of FX-rates rather than the identification of the causal relation between macroeconomic variables and FX-rates an analysis of the series in levels is ultimately interesting for the analysis and sometimes resulted in superior results in comparison to treated times series as in Meese and Rogoff (1983a, [1]).

With respect to the times series adjustments we proceeded by detrending the series at first exploring three different possibilities. Firstly detrending via differentiation, secondly detrending through a linear time trend and thirdly detreding through a moving average filter. Detrending through differencing provided to be the most effective due to the quadratic behaviour displayed by the series. This approach was therefore applied to all of the series. The obtained mean stationary series were consequently inspected for the presence of seasonality and appropriately differenciated to the point where no significance was found for seasonal units roots according to the methods proposed in Canova and Hansen (1995, [20]) and in Wang and Smith (2006, [21]).

4 Empirical Results

4.1 Univariate Models

The results for the univariate model fit are presented in table 1. As in Meese and Rogoff (1983a and 1983b, [1, 2]) we decided to report in the paper the results for the fit of the times series in levels and not for the detrended adjusted series as the latter performed generally worst compared to the first.

For all of the univariate models described in section 2.2 a Wald test assessing the empirical evidence for a restricted structural model was computed. While we do not find any evidence in the data supporting a restricted structural model for the detrended and differentiated times series we find that the jointly Null of the Frenkel-Bilson model cannot be rejected on a 1% confidence level for the OLS modeling FX-rates in levels. We decided therefore to estimate the out of sample performance for the OLS applying both the most general structural model and the Frenkel-Bilson model. For the transfer function model described in 2 the rate of decay, the persistence term and the dead lag where estimated at each model estimation in the rolling forecast. The optimal lags for the terms above where selected minimizing the Akaike information criteria (See Akaike (1998, [22]). We found a compelling evidence for modeling the differentiated series with one term in the nominator capturing the effect of unexpected shifts in the independent variables and two terms in the denominator capturing the general high persistency of FX-rates movements through time.

The results of table 1 confirm the one of the general literature where an OLS model vastly underperforms a random walk in the out of sample forecast of FX-rate at all tested lags. This especially in the case of a restricted model that suggests that despite the strong uncertainty surrounding the necessity of the inflation differentials and current account movements in the sample, such variables are beneficial for estimating the FX-rates movements out of sample.

In contrast, transfer function models performed much better in modeling FX-rates out of sample. While the model still underperforms a random walk in forecasting the FX-rates of the countries of interest at 1 month lag, the evidence at higher lags is rather mixed and it seems the two models equally performs in forecasting FX-rates out of sample. This is confirmed by the superior predictive test of Hansen (2005, [17]) and the bootstrap obtained p-value reported in parenthesis in table 1 obtained by applying Hansen (2011, [16]). For all of the models the equal predictability hypothesis of the unrestricted and restricted OLS model is rejected with 5% confidence.

Two possible causes for such observations might exist. While on the one hand the increased forecast performance at higher lags might be explained by the different information set allowed to the two models as described in 2.4, on the other hand, the fact might be caused by an increasing importance of macroeconomic variables at explaining the long run behaviour of FX-rates movements.

Finally, the Null of equal predictability of the random walk and the transfer function models in the superior predictability test could not be rejected with 10% confidence. This suggests the random walk model as the best univariate model to fit FX-rates out of sample given its parsimonious computation power combined with a restricted information set compared to the transfer function model as discussed in 2.4.

We observe no systematic difference when looking at RMSE. The picture for the latter looks similar at all lags for all the models to the results obtained looking at the MAE. No particular important outliers in forecast errors seem to be of particular importance in the analysis. A final word is give to the directional accuracy measure defined as the average number of times the sign of the realized FX-rate difference is matched by an equivalent sign in the FX-rate forecast difference. The values for such statistics lies between 0.5-0.7 for all the frequencies and there seems not to be systematic differences among the models. Such realizations for the random walk model suggest a certain degree of momentum for FX-rates important to include in the models.

4.2 Multivariate Models

The evidence of the previous section together with the general empirical literature suggests the presence of a unit root in the FX-rates. This was further confirmed by augmented Dickey-Fuller tests (Dickey and Fuller (1979, [23])) applied to the FX-rates and macroeconomics variables suggesting the statistical significant presence of unit roots. Dickey-Fuller tests were estimated once again after differencing the series. All of the series but the US-JP interest rates differential and the US-UK unemployment rate differential provided to be stationary after such adjustment, which supports evidence for an integration order of 1 for most of the series.

While in the previous section the FX-rate unit root was proved to be best modeled by a simple random walk model this section further expand the analysis by looking at the performance of multivariate models, which may capture and

Table 1: Mean Absolute Error – Univariate Models

Univariate Models				
Lag	Model	Switzerland	United Kingdom	Japan
MAE 1	Random Walk	0.0193 (0.503)	0.0168 (0.506)	0.0162 (0.506)
	Transfer Function Model	0.0205 (0.119)	0.0176 (0.128)	0.0173 (0.228)
	OLS - unrestricted	0.0234 (0.031)	0.0211 (0.053)	0.0194 (0.192)
	OLS - restricted	0.1435 (0.000)	0.1352 (0.002)	0.0534 (0.000)
MAE 3	Random Walk	0.0390 (0.307)	0.0330 (0.093)	0.0291 (0.534)
	Transfer Function Model	0.0384 (0.499)	0.0341 (0.094)	0.0280 (0.111)
	OLS - unrestricted	0.0603 (0.133)	0.0443 (0.143)	0.0497 (0.000)
	OLS - restricted	0.1551 (0.000)	0.1389 (0.007)	0.0583 (0.000)
MAE 6	Random Walk	0.0580 (0.483)	0.0495 (0.785)	0.0439 (0.484)
	Transfer Function Model	0.0581 (0.468)	0.0498 (0.526)	0.0442 (0.396)
	OLS - unrestricted	0.0789 (0.003)	0.0531 (0.486)	0.0669 (0.000)
	OLS - restricted	0.1716 (0.000)	0.1446 (0.041)	0.0669 (0.000)
MAE 12	Random Walk	0.0964 (0.171)	0.0584 (0.726)	0.0753 (0.505)
	Transfer Function Model	0.0954 (0.489)	0.0598 (0.678)	0.0752 (0.147)
	OLS - unrestricted	0.1210 (0.000)	0.0744 (0.523)	0.1001 (0.000)
	OLS - restricted	0.2100 (0.000)	0.1660 (0.459)	0.0776 (0.000)

efficiently estimate the relation among macroeconomics variables and FX-rates in the case of endogeneity among the series.

While the reference paper of Meese and Rogoff (1983a and 1983b, [1, 2]) attempted a multivariate model through a vector autocorrelation model we will analyze the extent of cointegration among the series given the results of the previous analysis showing a higher order integration for the macroeconomic and FX-rates series.

In order to do that we tested the hypothesis of cointegration for the macroeconomics variables and the FX-rates taking into account the trend displayed by the times series. The results of such an analysis are based on the trace statistic for the computed eigenvector discussed in Johansen (1991, [14]). Running the Johansen cointegration test we obtained the results presented in table 2. All of the series display a profound evidence for the presence of three cointegration vectors at 5% percent confidence level.

Given the five series of macroeconomics fundamentals it might be that the three cointegration relations might exist just among such variables and that a long term term relation between the macroeconomics fundamentals and the FX-rate for the three different series does not exists. In order to test such an hypothesis we decided to run Johansen tests checking for cointegration between the FX-rate and a single macroeconomic variable and to iterate the process adding a macroeconomic variable more until the estimation of a model comprehending all of the variables present in the structural model as in 1. The results in such a case confirm the hypothesis of cointegration between FX-rates and the macroeconomic variables. For the period analyzed we found strong evidence for cointegration among the FX-rate and macroeconomics fundamentals for the United Kingdom - USA series. In such a case the Null of no cointegration could be rejected with 5% confidence for the FX-M3, FX-CPI, FX-Current Account, FX-interest rate series. The picture looks more fragmented in the case of the other series. In the case of the Switzerland - USA series evidence for cointegration is displayed between {M3, M1, Interest Rate} and the FX-rate respectively, while for the Japan - USA series the Null of no cointegration can be rejected only for the FX-M3 series.

Table 2: Johansen Cointegration Test – Trace Statistics

			Quantiles Test Statistics		
Series		Trace Score	90%	95%	99%
Structural JP-US	Cointegrated Series <= 3	26.46	28.71	31.52	37.22
	Cointegrated Series <= 2	48.79	45.23	48.28	55.43
	Cointegrated Series <= 1	87.80	66.49	70.60	78.87
	Cointegrated Series = 0	138.18	85.18	90.39	104.20
Structural CH-US	Cointegrated Series <= 3	25.97	28.71	31.52	37.22
	Cointegrated Series <= 2	53.22	45.23	48.28	55.43
	Cointegrated Series <= 1	85.88	66.49	70.60	78.87
	Cointegrated Series = 0	145.12	85.18	90.39	104.20
Structural UK-US	Cointegrated Series <= 3	26.94	28.71	31.52	37.22
	Cointegrated Series <= 2	51.26	45.23	48.28	55.43
	Cointegrated Series <= 1	88.47	66.49	70.60	78.87
	Cointegrated Series = 0	178.10	85.18	90.39	104.20

Given the evidence of cointegration relations among FX-rates and macroeconomics fundamentals the Granger's theorem postulates the existence of a vector error correction model representation as in 4. We selected the optimal lagged terms of it by estimating three different information criteria for the model fit. In the specific, we computed the Akaike, Quinn-Hannan (See Hannan and Quinn (1979, [24]) and Schwarz (See Schwarz (1978, [25]) information criteria for the models with different lagged terms and selected for parsimony reason the minimum lag number identified by any of the three models above. An analogous approach was used for determining the optimal lagged terms of the benchmark multivariate model, a vector autoregressive model in difference. Based on such an approach multivariate models of term one, and with a smaller frequence of term two, resulted in the model estimates of the rolling forecast method. This is an important result given the restricted sample size and the exponential increase of parameters in the number of lagged terms.

Table 3: Mean Absolute Error – Multivariate Models

		Multivariate Models		
Lag	Model	Switzerland	United Kingdom	Japan
	Random Walk	0.0193 (0.125)	0.0168 (0.177)	0.0162 (0.625)
MAE 1	VAR	0.0183 (0.207)	0.0186 (0.178)	0.0160 (0.170)
	VECM	0.0176 (0.491)	0.0174 (0.476)	0.0153 (0.355)
	Random Walk	0.0390 (0.392)	0.0330 (0.495)	0.0291 (0.529)
MAE 3	VAR	0.0394 (0.359)	0.0333 (0.486)	0.0292 (0.488)
	VECM	0.0376 (0.624)	0.0328 (0.016)	0.0320 (0.770)
	Random Walk	0.0580 (0.675)	0.0495 (0.081)	0.0439 (0.449)
MAE 6	VAR	0.0578 (0.666)	0.0488 (0.504)	0.0418 (0.575)
	VECM	0.0580 (0.535)	0.0483 (0.355)	0.0434 (0.647)
	Random Walk	0.0964 (0.258)	0.0584 (0.232)	0.0753 (0.458)
MAE 12	VAR	0.0974 (0.165)	0.0603 (0.092)	0.0763 (0.267)
	VECM	0.0881 (0.507)	0.0542 (0.676)	0.0731 (0.573)

The results of the rolling forecast are presented in table 3. As in the case of the univariate fit and in line with the expectations the error increase in the estimation lag. In contrast to Meese and Rogoff (1983a and 1983b, [1, 2]) we do not find evidence for a general underperformance of the vector autoregressive models in predicting out of sample movements of foreign exchange rates in comparison to a simple random walk model. Especially in the short term when looking at the one month out of sample performance of vector autoregressive models we find that the latter marginally outperform a random walk model. For higher term lags the evidence is rather mixed and the superiority of random walk models cannot be claimed. Moreover, important is to underline that in comparison to univariate models of 2.2, the multivariate models make use of the same information set $\mathscr{F}_{t-1}$ as the random walk.

One possible explanation for the striking difference between the results reported in this paper and the one presented in Meese and Rogoff (1983a and 1983b, [1, 2]) may be found in the different sample size used for the reported study compared to the one of Meese and Rogoff papers. In this sense, Meese and Rogoff worked with a sample size of 87 observations, using as few as 37 observations for fitting their models (See Meese and Rogoff (1983a Section 3 and 1983b Section 3.2, [1, 2]). It is therefore highly likely that the multivariate results of their study suffer from overfitting issues leading the vector autoregressive model to poorly perform out of sample.

Turning to the vector error correction model of equation 4 that models the cointegration relation previously described we can see from table 3 that the rolling forecast of such models beats a random walk fit for all of the countries and at all lags when measured in terms of mean absolute error.

While the vector error correction model well manged to outperform the a simple random walk model when looking at the MAE statistics, no statistical significant evidence for the difference among the two models is found when modeling a model confidence set as in Hansen (2011, [16]). Looking at the p value of the Null Hypothesis of equal predictability resulting from a bootstrapped superior ability tests reported in parenthesis of table 3 the results are clear. All of the three models are indistinguishable in terms of their out of sample performance across all of the estimation lags and series.

5 Conclusion

The reported results generally confirmed the conclusions reported in Meese and Rogoff (1983a and 1983b, [1, 2]). No one of the four models analyzed could statistically significantly improve the out of sample performance of a simple random walk model without drift when forecasting of foreign exchange rates. Moreover, we found evidence for a statistically significant underperformance of OLS models compared to a random walk model.

When looking at the point estimation of out of sample error statistics we observed a mixed evidence for the vector autoregressive models and the transfer function models. Both models could sporadically yield more accurate results compared to the random walk model but generally underperformed the latter.

Given the evidence for cointegration among macroeconomics variables and FX-rates we estimated a vector error correction model which displayed promising results outperforming the performance of a random walk model when looking at out of sample forecasting error statistics. These results were consistent among all estimation lags and country series but nonetheless did not display enough evidence for rejecting the Null of equal predictability with the random walk model based on the interval estimation through the model confidence test of Hansen (2011, [16]).

It is possible to conclude therefore by underlying that our study does support the strength of a random walk in comparison to univariate models for FX-rates estimation and as a general sound model for capturing the unit root behaviour observed in the FX-rates movements. Nonetheless, our study vastly contrasted with the results of Meese and Rogoff papers when analyzing multivariate models. The latter models were able to well capture the movements of FX-rates through a parametric model based on macroeconomic fundamentals suggesting an estimation failure due to a small sample bias in the results reported by Meese and Rogoff (1983a and 1983b, [1, 2]). Moreover the cointegration between macroeconomics fundamentals and FX-rates suggests that a long run stable relation between the two exists.

Based on our results we reject the general claim of weak relation among macroeconomics variables and foreign exchange rates and we claim that the latter should be considered as a primary determining factor driving the foreign exchange rates movements at monthly lag. If the well performance of the random walk models is explained by the long memory displayed by foreign exchange series that is visibile by investing autocorrelation plots for the series in virtually all time periods, the macroeconomic models managed a comparable performance simply leveraging the movements, long run relation and endogeineity among foreign exchange and macroeconomic variables and require therefore a stronger attention from the general academic literature.

6 Appendix A

Series: Unemployment rate.

Description: Unemployment Rate: Aged 15-64: All Persons.

Source: Main Economic Indicators - complete database, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

Series: 3-month Treasury Bills USA, percent annum.

Description: 3-Month Treasury Bill - Secondary Market Rate

Source Board of Governors of the Federal Reserve System (US), 3-Month Treasury Bill: Secondary Market Rate [TB3MS], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/TB3MS`, April 13, 2019.

Series: Treasury Bills UK.

Description: Treasury Bills discount rate, percent per annum.

Source: Bank of England, Treasury Bill Discount Rate in the United Kingdom [TBDRUKM], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/TBDRUKM`, April 13, 2019.

Series: Treasury Bills JP.

Description: Interest Rates, Government Securities, Treasury Bills for Japan.

Source: International Monetary Fund, Interest Rates, Government Securities, Treasury Bills for Japan [INTGSTJPM193N], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/INTGSTJPM193N`, April 13, 2019.

Series: 3-month LIBOR.

Description: 3-Month or 90-day Rates and Yields: Interbank Rates

Source: Organization for Economic Co-operation and Development, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

Series: M3 USA & JP.

Description: M3 comprises notes and coins in circulation outside banking corporations; demand and savings deposits, fixed and installment savings deposits, time deposits, and certificates of deposit of households, nonfinancial corporations, local governments, securities companies, Tanshi companies, and some other financial corporations such as securities finance companies with depository corporations in national and foreign currency; and nonresident deposits with banking corporations in national currency.

Source: International Monetary Fund, M3 for Japan, and US retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/MYAGM3JPM189N` & `https://fred.stlouisfed.org/series/MYAGM3USM052N`, April 13, 2019.

Series: M3 CHF & UK.

Description: M3 comprises notes and coins in circulation outside banking corporations; demand and savings deposits, fixed and installment savings deposits, time deposits, and certificates of deposit of households, nonfinancial corporations, local governments, securities companies, Tanshi companies, and some other financial corporations such as securities finance companies with depository corporations in national and foreign currency; and nonresident deposits with banking corporations in national currency.

Source: Organization for Economic Co-operation and Development, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

Series: M1 UK & JP & CH.

Description: M1 consists of: (1) currency outside the U.S. Treasury, Federal Reserve Banks, and the vaults of depository institutions; (2) traveler's checks of nonbank issuers; (3) demand deposits; and (4) other checkable deposits (OCDs), which consist primarily of negotiable order of withdrawal (NOW) accounts at depository institutions and credit union share draft accounts.

Source: Organization for Economic Co-operation and Development, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

Series: CPI.

Description: Consumer Price Index: All Items

Source: Organization for Economic Co-operation and Development, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

Series: Trade.

Description: Net Trade. Value of Goods.

Source: Organization for Economic Co-operation and Development, Main Economic Indicators (database), `http://dx.doi.org/10.1787/data-00052-en`, April 13, 2019.

References

[1] R. A. Meese and K. Rogoff, "Empirical exchange rate models of the seventies: Do they fit out of sample?," *Journal of international economics*, vol. 14, no. 1-2, pp. 3–24, 1983.

[2] R. Meese and K. Rogoff, "The out-of-sample failure of empirical exchange rate models: sampling error or misspecification?," in *Exchange rates and international macroeconomics*, pp. 67–112, University of Chicago Press, 1983.

[3] J. Bailliu, M. R. King, *et al.*, "What drives movements in exchange rates?," *Bank of Canada Review*, vol. 2005, no. Autumn, pp. 27–39, 2005.

[4] R. F. Engle and C. W. Granger, "Co-integration and error correction: representation, estimation, and testing," *Econometrica: journal of the Econometric Society*, pp. 251–276, 1987.

[5] J. A. Frenkel, "A monetary approach to the exchange rate: doctrinal aspects and empirical evidence," *The scandinavian Journal of economics*, pp. 200–224, 1976.

[6] M. Mussa, "The exchange rate, the balance of payments and monetary and fiscal policy under a regime of controlled floating," in *Flexible Exchange Rates and Stabilization Policy*, pp. 97–116, Springer, 1977.

[7] R. Dornbusch, "Expectations and exchange rate dynamics," *Journal of political Economy*, vol. 84, no. 6, pp. 1161–1176, 1976.

[8] P. Hooper and J. Morton, "Fluctuations in the dollar: A model of nominal and real exchange rate determination," *Journal of international money and finance*, vol. 1, pp. 39–56, 1982.

[9] D. C. Montgomery and G. Weatherby, "Modeling and forecasting time series using transfer function and intervention methods," *AIIE Transactions*, vol. 12, no. 4, pp. 289–307, 1980.

[10] A. Tustin, "The mechanism of economic systems," 1957.

[11] C. A. Sims, "Macroeconomics and reality," *Econometrica: journal of the Econometric Society*, pp. 1–48, 1980.

[12] L. A. Gil-Alana and P. M. Robinson, "Testing of unit root and other nonstationary hypotheses in macroeconomic time series," *Journal of Econometrics*, vol. 80, no. 2, pp. 241–268, 1997.

[13] P. C. Phillips, "Understanding spurious regressions in econometrics," *Journal of econometrics*, vol. 33, no. 3, pp. 311–340, 1986.

[14] S. Johansen, "Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models," *Econometrica: journal of the Econometric Society*, pp. 1551–1580, 1991.

[15] J. M. Westerfield, "An examination of foreign exchange risk under fixed and floating rate regimes," *Journal of International Economics*, vol. 7, no. 2, pp. 181–200, 1977.

[16] P. R. Hansen, A. Lunde, and J. M. Nason, "The model confidence set," *Econometrica*, vol. 79, no. 2, pp. 453–497, 2011.

[17] P. R. Hansen, "A test for superior predictive ability," *Journal of Business & Economic Statistics*, vol. 23, no. 4, pp. 365–380, 2005.

[18] C. A. Sims, “Seasonality in regression,” *Journal of the American Statistical Association*, vol. 69, no. 347, pp. 618–626, 1974.

[19] C. A. Sims, “Distributed lags,” 1973.

[20] F. Canova and B. E. Hansen, “Are seasonal patterns constant over time? a test for seasonal stability,” *Journal of Business & Economic Statistics*, vol. 13, no. 3, pp. 237–252, 1995.

[21] X. Wang, K. Smith, and R. Hyndman, “Characteristic-based clustering for time series data,” *Data mining and knowledge Discovery*, vol. 13, no. 3, pp. 335–364, 2006.

[22] H. Akaike, “Information theory and an extension of the maximum likelihood principle,” in *Selected papers of hirotugu akaike*, pp. 199–213, Springer, 1998.

[23] D. A. Dickey and W. A. Fuller, “Distribution of the estimators for autoregressive time series with a unit root,” *Journal of the American statistical association*, vol. 74, no. 366a, pp. 427–431, 1979.

[24] E. J. Hannan and B. G. Quinn, “The determination of the order of an autoregression,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 41, no. 2, pp. 190–195, 1979.

[25] G. Schwarz *et al.*, “Estimating the dimension of a model,” *The annals of statistics*, vol. 6, no. 2, pp. 461–464, 1978.