LUISS

Dipartimento
di Economia Finanza

Cattedra Empirical Finance

Empirical analysis of Taylor rule models for Exchange Rate determination By state space methods

Prof.
Paolo Santucci
DeMagistris

RELATORE

Prof.
Stefano Grassi

CORRELATORE

Andrea Sergi
Matr. 698311

CANDIDATO

Anno Accademico 2020/2021

Index

1 Introduction

Exchange rates influences a wide range of economic agents for different reasons: both countries and individuals trading goods, financial traders, foreign workers and investors are affected by changes in exchange rates.

Different models have been proposed and empirically tested with the goal of linking exchange rate movements to macroeconomic fundamentals such as interest rates, inflation, outputs and money supplies. All the theories proposed state that exchange rates are determined by macroeconomic variables. However no model has managed to produce definitive answer when empirically tested. As first established by Meese and Rogoff floating exchange rates between countries are best approximated as random walks. Meese and Rogoff (M-R) (1983a, 1983b) in their seminal papers used different monetary models for exchange rate determination (Frenkel Bilson, Dornbusch Frenkel, Hooper Morton) with the goal of explaining nominal exchange rate movements in terms of contemporaneous macroeconomic variables, showing that they failed to outperform a naïve random walk (RW without drift) when comparing the out of sample forecasts by root mean forecast square error (RMSE), mean forecast error (ME), and mean absolute forecast error (MAE). The models parameters were kept fixed and were at first estimated employing various econometric techniques, including ordinary least squares (OLS), generalized least squares (GLS), instrumental variables (IV), and then were constrained to values based on the economic and empirical theory of money demand and purchasing power parity. The authors list as possible reasons of the poor forecasting performance the sensitivity of the models to the choice of proxy used to represent expected differences in inflation, temporary or permanent deviation from Purchasing Power Parity (PPP), misspecification of the money demand function, simultaneous equation bias, and changes in the parameters values over time.

The results were nonetheless interpreted as evidence of the inadequacy of the models to establish a meaningful relationship between exchange rates and macroeconomic fundamentals, starting a discussion regarding the usefulness of the models and the forecastability of exchange rates, commonly referred to as "Meese Rogoff puzzle".

Cheung and Chinn (1998) attribute the empirical failure to theoretical flaws in the models. Cheung, Chinn, and Pascual (2005) test the out of sample forecasts for different monetary models obtaining generally negative results and conclude that different specifications may work better for different exchange rates and at different time horizons.

By following M-R's suggestion that it may be fruitful to account for parameter instability, this thesis tries to improve on previous attempts by using a model based on Taylor rule fundamentals, allowing for parameter variation and heteroscedasticity by putting the model in state space form and estimating it using the Kalman filter. Indeed this approach sounds, at least intuitively, more promising. It seems reasonable to assume that the impact of shocks in the macroeconomic variables on the exchange rate is related to the relative health of the economies of the home and foreign country. For instance if one of the two countries is experiencing a recession with very low inflation, the exchange rate reaction to a further lowering in the price level of this country is expected to be greater than what would happen if his economy was thriving.

It must be noted however that this approach is not new. Different attempts to forecast exchange rates by state space methods have been made producing mixed results. Both Wolff (1987) and Schinasi and Swamy (1989) make use of time varying parameters to forecast exchange rates using the same models as in MR and still taking the RW model as a benchmark with not very encouraging results. Rossi (2006, 2013) in an extensive study uses the Kalman filter with a random walk specification for the parameters with mixed results, but concludes that in general evidence of predictability is scarce. Bacchetta et al. (2009) study wether acoounting instability is sufficient for solving the Meese Rogoff puzzle and conclude that this is hardly the case. Molodstova and Papell (2009) deviating from H-R use different Taylor rule based models of exchange rate determination (1993) for a multitude of currencies, allowing for stochastic parameters using rolling OLS regressions they manage to outperform the RW model in few cases. Haskamp (2017), with the main focus of out of sample forecasting uses the Kalman filter to allow for parameter variation in the Molodtsova Papell model. He reports that the Kalman filter manages to incorporate more abrupt adjustments in the coefficients when compared to the OLS rolling regression, however he still fails to outperform a random walk in terms of out of sample forecasts.

Although accounting for time varying coefficients has in general failed to beat a random walk, it is clear that a sensible improvement has been made over the first attempts with fixed coefficients of MR.

Even if the results are encouraging neither any model has taken into account the well known time varying nature of exchange rates volatility, treated as constant over time, nor any attempt has been made to deviate from the Gaussianity assumption, despite large evidence that exchange rates follow a leptokurtic distribution as shown by the pioneering work of Mandelbrot (1960).

In this thesis the efforts of linking exchange rates to macroeconomic fundamentals through the state space formulation of Taylor rule based macro model are furthered by taking into account the heteroscedastic nature of exchange rates, and allowing Autoregressive Conditional Heteroscedasticity (ARCH) effects to enter the model.

In the following sections Taylor rule fundamentals for exchange rate determination will be presented, and a reduced version proposed, then the general space time formulation of the model and the related Kalman Filter recursions are illustrated. Then ARCH effects will be included and dealt with using the approach proposed by Harvey Ruiz and Sentana (1992). An univariate time series analysis using the model will be carried out with the goal of investigating the relationship between macroeconomic fundamentals and exchange rates.

2 Models for exchange rate determination

In this section a short presentation of the different monetary models used in the literature for forecasting exchange rate models will be given. We first present the models used by HR (Frenkel – Bilson, Dornbusch – Frenkel, Hooper – Morton), then we introduce models based on Taylor rule fundamentals[1], with particular emphasis to the model of Molodtsova and Papell (2008). For the rest of the paper exchange rates are defined so that an increase in s_t means depreciation of the Home currency and viceversa.

2.1 Monetary models

The flexible price Frenkel Bilson model is specified as follows:

$$s_t = \alpha_0 + \alpha_1(m_t - m_t^*) + \alpha_2(y_t - y_t^*) + \alpha_3(i_t - i_t^*) + \epsilon_t \tag{1}$$

Where s_t is the logarithm of the exchange rate measured as unit of domestic currency needed to buy one unit of foreign currency, m is the log of money supply, y is the log of real income, i is the short term interest rate, ϵ is a stochastic error term, and the star represent the respective variable for the foreign country. PPP is assumed to hold both in the short and long run, and expectations do not play any role in driving exchange rate movements. The parameter α_1 is expected to be positive and equal to one, meaning that a monetary expansion of the home country will lead to higher prices causing a depreciation of the home currency. α_2 is expected to be negative as an increase in real income for the home country will increase the demand for real money balances appreciating the home currency.

[1] All the variables are considered in log. Where for interest rate the following approximation has been made: $log(1 + x) \approx x$ using a first order Taylor expansion around $x = 0$.

α_3 is expected to be positive, higher interest rates are expected to have a negative effect on the demand of real money balances producing a depreciation of the home currency.

The so called sticky price monetary model of Dornbusch and Frenkel is formulated as follows:

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

Where m, y, i, and the starred variables represents the same variables as in (1) and π_t^e is the expected long run inflation rate. Prices are assumed to be fixed in the short run but not in the long run. Inflation expectations are added in the model to take into account the fact that prices do not react rapidly to shocks in the money supply. In this model a domestic monetary expansion leads to a fall in interest rate and consequently to the depreciation of the home currency.

The sticky price monetary model incorporating current account effects of Hooper and Morton is:

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

Where B_t is the current trade account and all the other variables have the same interpretation as in (2). B_t is included to take into account the impact that the trade account balance has on exchange rate expectations. In particular α_5 and α_6 are expected to be positive and negative respectively, implying that a current account surplus of the home country will result in an appreciation of his currency.

2.2 Taylor rule based models for exchange rate determination

Monetary model argues that it is level of money supply in domestic and foreign country, which acts as main determinant of exchange rate. Each of the model in the previous paragraph imply that monetary policy controls directly the money supply in each country.
In practice most central banks set their monetary policy in terms of interest rates. In this case the equilibrium in the money market becomes residual in the determination of exchange rates.

Taylor (1993) argued that empirical evidence suggested that policy rules that targeted the short term interest rate worked better than the ones focused on money supply and/or exchange rates, in particular speculating that short term rates should be related to the price level and the economic output he proposed the following policy rule, commonly known as Taylor's rule.

$$i_t = \tilde{\imath} + \pi_t + \gamma(y_t - \bar{y}) + \phi(\pi_t - \bar{\pi}) \qquad (4)$$

Where i_t is the targeted short term interest rate, $\tilde{\imath}$ is the equilibrium real interest rate, y_t is the economy output, usually measured with GDP, π_t is the inflation usually measured with CPI, $\bar{\pi}$ is the targeted level of inflation, and the term $y_t - \bar{y}$ is the output gap, or the difference between output (measured in GDP) and the economy potential. In its original formulation Taylor assumed that for the US economy the Federal Reserve targeted an inflation rate of 2%, a real equilibrium interest rate of 2%, estimated the economy potential to be equal to the growth trend of GDP between 1984 and 1992 at 2.2%, and $\gamma, \phi = 0.5$.
The Taylor rule has worked remarkably well over the years, providing and accurate description of the policy decisions of the Federal Reserve, especially under the guidance of Paul Volcker and Alan Greenspan.
Assuming that the foreign central bank follows an interest rate policy that targets inflation and output gap as in (4):

$$i_t^* = \tilde{\imath}^* + \pi^*_t + \gamma(y_t^* - \bar{y}^*) + \phi(\pi_t^* - \bar{\pi}^*) \qquad (5)$$

And under the assumption of no arbitrage opportunities, so that Uncovered interest rate parity[2] (UIRP) holds:

$$E_t(s_{t+1} - s_t) = i_t - i_t^* \tag{6}$$

the Taylor rule can be used for exchange rate determination purposes, subtracting (4) from (5) and substituting the left hand side with (6) LHS we get:

$$E_t(s_{t+1} - s_t) = (\widetilde{\imath} - \tilde{\imath}^*) + \gamma\Delta y_t + \phi\Delta\pi_t \tag{7}$$

With $\Delta y_t = (y_t - \bar{y}) - (y_t^* - \overline{y^*})$ and $\Delta\pi_t = (\pi_t - \bar{\pi}) - (\pi_t^* - \overline{\pi^*})$ and making the simplifying assumption that γ and ϕ are equal for both central banks, implying that they react in the same manner to changes in the output gap and inflation.

Although as already mentioned this first formulation of the Taylor rule has performed well, it has several limitations. By Taylor own admission the unobservable nature of equilibrium real interest rate, inflation target, and economic potential posits a practical restriction to the estimation of the model. Instead of tackling the problem directly he proposes, with poor results, an alternative policy rule in which exchange rates are added as explanatory variables and the inflation target and the equilibrium real rate are set to zero. Also Clarida, Gali and Gertler (CGG for the rest of the paragraph) (1997, 1999) defined this baseline formulation backward-looking , in the sense that central bank should set the nominal interest rate according to expectations of future inflation and output instead of reacting to past realized values, proposing a "forward looking" version of (7):

$$i_t = \tilde{\imath} + \gamma[E_t(y_{t+n}) - \bar{y}] + \phi[E_t(\pi_{t+n}) - \bar{\pi}] \tag{8}$$

Moreover, they notice that it would be unfeasible for policy makers to respond to economic changes as aggressively as prescribed by the rule, such behavior could in fact results in loss of credibility and/or destabilization of the capital markets. In this respect they suggest that the central bank adjusts gradually to changes in economic conditions following:

$$r_t = (1 - \rho)i_t + \rho r_{t-1} + \epsilon_t \tag{9}$$

[2] E_t denotes expectations conditional on all information known at time t

Where i_t is as in (4) and ρ is the smoothing parameter.

Finally setting $\alpha = \tilde{\imath} - \phi\bar{\pi}$ in (7) and then plugging it into (9) we get:

$$r_t = (1-\rho)\left(\alpha + \gamma[E_t(y_{t+n}) - \bar{y}] + \phi E_t(\pi_{t+n})\right) + \rho r_{t-1} + \epsilon_t \qquad (10)$$

The model estimated by GMM was shown by CGG to fit remarkably well the interest rate decisions of G3 and E3 countries after 1979. Moreover according to CGG all the countries reaction were strong relative to inflation and moderate at best relative to the output gap and all the other variables eventually taken under consideration (exchange rates, foreign interest rate), concluding that indeed the central banks set their monetary policy holding inflation in higher regard over other economic fundamentals.

This specification of the Taylor rule was cleverly mixed and rearranged with its original formulation by Molodtsova and Papell (2008), with the goal of providing an estimable model for exchange rate forecasting.

Assuming that the central banks have a reaction function as specified in (4) with the additional inclusion of the real exchange rate, and that the monetary policy is smoothed according to (9), we have:

$$r_t = (1-\rho)(\alpha + (1+\phi)\pi_t + \gamma x_t + \delta q_t) + \rho i_{t-1} + \epsilon_t \qquad (11)$$

where q_t is the real exchange, $x_t = y_t - \bar{y}$, and all the other terms the same as before. Assuming that (11) also applies for the foreign country subtracting the foreign rule to (11) we get the model formulated by Molodtsova and Papell (MP for the rest of the paragraph):

$$E_t[r_t - r_t^*] = \omega + \omega_\pi \pi_t - \omega_{\pi^*}^* \pi_t^* + \omega_x x_t - \omega_{x^*}^* x_t^* - \omega_{q^*} q_t^* \qquad (12)$$
$$+ \omega_i\, i_{t-1} - \omega_{i^*}^* i_{t-1}^* + \eta_t$$

Where the intercept ω is constant, $\omega_{\pi,t} = (1-\phi_t)(1-\rho)$, $\omega_{x,t} = \gamma(1-\rho)$, and $\omega_{q,t} = \delta(1-\rho)$, the same applies for the foreign country.

In order to obtain an exchange rate forecasting equation they predict that in accordance with empirical evidence UIRP do not hold, in particular they assume that an increase in domestic inflation and output gap, and consequently in the domestic nominal interest rate, will result in appreciation of the domestic currency, as opposed to the UIRP that predicts the opposite. In agreement with the survey evidence of Gourincha and Tornell (2004), they expect investors to underestimate the persistence of interest rates shocks. This leads to a mechanism defined as *updating effect* in which investors constantly revise upwards their expectations of future currency appreciation. This effects is even more pronounced if the central bank follows an interest rate rule which includes smoothing, in which case the investors underestimation will be more severe. In accordance with this belief they expect that any change in economic conditions (inflation above target, increase in output gap), that makes the central bank set an higher interest rate will produce currency appreciation.

Molodtsova and Papell combine this predictions with equation (12) to get the following exchange rate forecasting equation

$$\Delta s_{t+1} = \alpha - \alpha_{\pi}\pi_t + \alpha_{\pi}^{*}\pi_t^{*} - \alpha_x x_t + \alpha_x^{*} x_t^{*} - \alpha_i i_{t-1} + \alpha_i^{*} i_{t-1}^{*} + \eta_t \qquad (13)$$

In the formulation of the model in the authors assumed smoothing as in (11),different time varying coefficients for the two central banks, and also the existence of difference between inflation targets and equilibrium interest rates, captured by the intercept.

As the authors point out the model in (13) can be formulated for a variety of different specifications. The model is said to be symmetric/asymmetric if the foreign central bank doesn't/do target the exchange rate (i.e. $\omega_{q^*} = 0 \ / \ \omega_{q^*} \neq 0$). If lagged interest rates are included the model is said to be with smoothing, otherwise with no smoothing. Finally if the coefficients are the same for the domestic and foreign countries are the same the model is said to be homogeneous. Otherwise, it is heterogenous.

The model is readily estimable by rolling OLS regressions as done by the authors. They proceeded to calibrate the model with data from March 1973 to February 1982 . The forecast exercise was carried out using a moving window, as done by MR and many others, meaning that the estimation is started using the first n in sample observation (in this case 120), and a one step ahead (1 month) forecast is generated. As soon as a new observation is available the procedure is repeated.

The robustness of the forecasts was then tested using the Clark West test against a Random Walk without drift with mixed results.

3 Methodology

In this section the model that will be used to forecast exchange rates is presented, in both its general and state space form. We explain the estimation techniques used in the empirical analysis of the next section, in particular the Kalman Filter (1960) and an extended formulation proposed by Harvey, Ruiz, and Sentana (1992) to account for ARCH effects in the residuals.

3.1 Model

Following in the footsteps of Molodtsova and Papell, We compare two models (heterogeneous and homogeneous) based on the Taylor rule accounting for smoothing as in (13) with time varying parameters while deviating from MP in two important respects.

- The output gap is not included in the central bank response function.
- Estimation and forecast are made using the Kalman filter considering ARCH effects in the residuals.

The models are formulated as follows:

$$\Delta s_{t+1} = \alpha_t + \alpha_{\pi,t}\pi_t + \alpha_{\pi^*,t}\pi_t^* + \alpha_{i,t}i_{t-1} + \alpha_{i^*,t}i_{t-1}^* + \epsilon_t \tag{14}$$

$$\Delta s_{t+1} = \alpha_{\pi,t}(\pi_t - \pi_t^*) + \alpha_{i,t}(i_{t-1} - i_{t-1}^*) + \epsilon_t \tag{15}$$

Where $\Delta s_{t+1} = E_t[s_{t+1} - s_t]$, all the other variables are as before. Equations 14 and 15 represent the heterogeneous and the homogeneous model respectively.

The intercept is meant to capture difference among countries of the smoothing coefficients, inflation targets and equilibrium interest rates, as time varying as opposed to H-P.

The reason for excluding the output gap is dual. First as empirical evidence produced by Clarida et al. suggests the response of monetary policy to the output gap is mild at best, being primarily focused on inflation. The second reason is that for estimation purposes including the output gap can be almost comparable to adding a latent explanatory variable to the model. This indicator defined as the level of output that can be achieved when the factors of production are used at non-inflationary levels, in theory is an important proxy for the amount of slack in the economy, in practice is unobservable. There is no general consensus on how to best estimate this indicator either; as Orphanides (2002) noted: “different methods would yield very different estimates of the output gap”. Compounding the problem is the fact that the data used to model it are often subject to revision, and there is a significant degree of parameter uncertainty. This features are evident in the work of Molodtsova and Papell: the forecasts accuracy of their model is affected by the method used for measuring the output gap.[3] They also report in agreement with the finding of Clarida et al. that the time varying parameters for the output gap are in general very close to zero.

For what concerns the different choice of estimation method, the Kalman filter is ideal to deal with the varying parameters problem. This algorithm has the advantage of being able to capture the dependence structure of the series both in terms of mean and variance. This type of modelling is in fact employed to analyze time series for which the data generating process is subject to regime shifts. This feature of the filter helps in dealing with asymmetric behavior of the series, such as the one exhibited in exchange rates, which alternate periods of relative stability with periods of high volatility.

[4]3.2 Linear Gaussian State space models and Kalman Filter

The general linear Gaussian state space model (LGSSM) is formulated as follows:

$$y_t = Z_t \alpha_t + \epsilon_t \tag{16}$$

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \tag{17}$$

With $\epsilon_t \sim N(0, H_t)$ $\eta_t \sim N(0, Q_t)$

Where (16) is called the observation equation, (17) the state equation, y_t is the $(n \times 1)$ observation vector, α_{t+1} is the $(m \times 1)$ <u>unobservable</u> state vector, ϵ_t is the iid $(n \times 1)$

[3] They try three different methods: linear trend, quadratic trend, and Hodrick–Prescott (H-P) filter. H-P provides the best forecast.

[4] For a more thorough treatment of the subject see Durbin and Koopman (DK) (2012)

normally distributed vector of observation disturbances (not observed), η_t is the iid $(r \times 1)$ normally distributed vector of state disturbances (not observable), Z_t is an $(n \times \mathrm{m})$ matrix, T_t is the $(m \times m)$ state transition matrix, R_t is an $(m \times r)$ matrix, H_t is the $(\mathrm{n} \times \mathrm{n})$ observation disturbances covariance matrix, Q_t is the $(r \times r)$ state disturbances covariance matrix, and ϵ_t and η_t are mutually independent from each other and from α_t.

The unobserved α_t's state represents the states which define the development over time of the system under consideration, together with the observations y_t's which are related to the states by (16).

The goal of filtering is to update our knowledge of the system every time a new observation is brought in, or more precisely to obtain the conditional distributions of α_t and α_{t+1} given the set of all available information at time t.

Since in the LGSSM the distributions of the disturbances are assumed to be normal we have that all the conditional joint distributions are normal:

$$p(\alpha_t|Y_{t-1}) = N(a_t, P_t) \tag{18}$$

$$p(\alpha_t|Y_t) = N(a_t^*, P_t^*) \tag{19}$$

$$p(\alpha_{t+1}|Y_t) = N(a_{t+1}^*, P_{t+1}^*) \tag{20}$$

Where Y_t denotes the set of all available information at time t. a_t, P_t are assumed to be known and are equal respectively to $E_{t-1}(\alpha_t)$ and $Var_{t-1}(\alpha_t)$[5]. With a_t the vector of unconditional stetes means of size $(m \times 1)$, and P_t the unconditional states covariance matrix of size $(m \times m)$.

$a_t^* = E_t(\alpha_t)$, $P_t^* = Var_t(\alpha_t)$ are respectively called filtered state, and covariance matrix estimates.

$a_{t+1}^* = E_t(\alpha_{t+1})$,$P_{t+1}^* = Var_t(\alpha_{t+1})$ are respectively called the state , and covariance matrix updates.

And additionally it is assumed that for the initial states the following holds $\alpha_1 \sim N(a_1, P_1)$, With a_1, P_1 known[6].

[5] The notation $E_{t-1}(\alpha_t)$ is equivalent to $E(\alpha_t|Y_{t-1})$

[6] This is hardly the case in practice, however it is convenient to explain the functioning of the algorithm. This assumption will be relaxed later.

Having established this quantities we move on by defining one step prediction errors, which will be fundamental in the so called prediction error decomposition estimation that will be employed later.

We define the one step ahead forecast error as:

$$v_t = y_t - E_{t-1}(y_t) = y_t - Z_t a_t \tag{21}$$

Where the last equality in (21) comes from substituting (16) and taking expectations.

We note also that:

$$E_{t-1}(v_t) = 0 \tag{22}$$

$$Var_{t-1}(v_t) = Var_{t-1}(Z_t \alpha_t + \epsilon_t - Z_t a_t) = Z_t P_t Z_t' + H_t = F_t \tag{23}$$

$$\begin{aligned} Cov_{t-1}(\alpha_t, v_t) = E_{t-1}\big[\big(\alpha_t - E_{t-1}(\alpha_t)\big)\big(v_t - E_{t-1}(v_t)\big)'\big] = \\ = E_{t-1}[\alpha_t(\alpha_t - a_t)' Z_t'] = P_t Z_t' \end{aligned} \tag{24}$$

Now we can proceed to obtain the filtered state and covariance estimate.

In order to do so, since the model is conditionally Gaussian we can use a well known result from multivariate analysis[7], the so called regression lemmas:

$$E(x|y) = E(x) + Cov(x,y)Var(y)^{-1}\big(y - E(y)\big) \tag{25}$$

$$Var(x|y) = Var(x) - Cov(x,y)Var(y)^{-1}Cov(x,y)' \tag{26}$$

Using α_t and v_t as respectively x and y in (24), (25) yields:

$$a_t^* = E_t(\alpha_t) = E(\alpha_t | Y_{t-1}, v_t) = a_t + P_t Z_t' F_t^{-1} v_t \tag{27}$$

$$P_t^* = Var_t(\alpha_t) = Var(\alpha_t | Y_{t-1}, v_t) = P_t - P_t Z_t' F_t^{-1} Z_t P_t \tag{28}$$

Which are the filtered state and covariance estimates.

Now we can proceed to update the state and covariance update:

$$a_{t+1}^* = E_t(T_t \alpha_t + R_t \eta_t) = T_t E_t(\alpha_t) = T_t a_t^* \tag{29}$$

$$P_{t+1}^* = Var_t(T_t \alpha_t + R_t \eta_t) = T_t P_t^* T_t' + R_t Q_t R_t' \tag{30}$$

Equations (21), (23), (27), (28), (30) represents the Kalman filter recursions.

[7] For proofs see Feller (1970).

The model parameter can be estimated by Maximum likelihood estimation. Here I present the so called prediction error decomposition of the likelihood that will be used later.
The Likelihood function is :

$$L(Y_n) = p(y_1, \dots, y_n) = p(y_1) \prod_{t=2}^{n} p(y_t | Y_{t-1}) \tag{31}$$

With $Y_t = (y_1', \dots, y_t')$. Taking the log of (31) we have:

$$logL(Y_n) = \sum_{t=1}^{n} log\, p(y_t | Y_{t-1}) \tag{32}$$

Since $E(y_t|Y_{t-1}) = Z_t a_t$,$v_t = y_t - Z_t a_t$, $Var(y_t|Y_{t-1}) = F_t$, and $p(y_t|Y_{t-1})$ we get:

$$logL(Y_n) = -\frac{np}{2} log 2\pi - \frac{1}{2} \sum_{t=1}^{n} (log|F_t| + v_t' F_t^{-1} v_t) \tag{33}$$

Where the variables in (33) are either constants or obtained from the filtering procedures. The likelihood can then be optimized by iterative numerical evaluations.

3.3 Including ARCH effects in the model

The estimation procedure just explained has been already tried although for a different model specification. Wolff (1987) formulated the Dornbusch Frenkel and the Frenkel Bilson model in state space form and used the Kalman Filter to estimate time varying parameters and generating forecast up to 24 months ahead. The comparison with the Random Walk was mixed and in general not very favorable. Wolff imputed the low forecasting power of the model to correlation in the residuals. He do not address this problem directly. He goes on by proposing a revised version of the Dornbusch Frenkel model allowing for changes in the equilibrium real exchange rate, and tries to take into account correlation in the residuals by modelling the observation equation as a first order autoregressive process.

Here the problem of correlation and heteroscedasticity is dealt with directly, the assumption of homoscedasticity for the residuals is in fact too restrictive. This presents a considerable modelling limitation if one believes that important information is embedded in the conditional volatility of certain variables in the model. Thus, potentially important information is lost in a model that assumes conditionally homoscedastic errors. This is especially true if a series exhibits episodes of low variance followed by episodes of high variance, in which case it is said to have ARCH errors.

Harvey, Ruiz, and Sentana (1992) (HRS) showed how ARCH effects could be handled theoretically in a state-space model where the conditional heteroskedasticity was present in either the measurement or the transition equation innovations, or in both. For the estimation, the authors proposed using an approximate (or a quasi-optimal) filter, which is a modification of the usual Kalman filter. We illustrate the general formulation of their approach

$$y_t = Z_t \alpha_t + \Lambda \epsilon_t + \epsilon_t^* \tag{34}$$

$$\alpha_t = T_t \alpha_{t-1} + \Phi \eta_t + \eta_t^* \tag{35}$$

Equation (43) and (44) are respectively the observation and the state equation. Λ, Φ are respectively of size $(n \times 1)$ and $(m \times 1)$. The sizes of the other matrices and vectors are the same as before. The notable difference is the inclusion of additional residual terms in both the state and the observation equations.

We have $\epsilon_t^* \sim N(0, H_t)$ and $\eta_t^* \sim N(0, Q_t)$ normally distributed and mutually independent. The ARCH effects enter the model by means of the two other scalar residuals terms:

$$\epsilon_t = \sqrt{h_t}\widetilde{\epsilon_t} \tag{36}$$

$$\eta_t = \sqrt{q_t}\widetilde{\eta_t} \tag{37}$$

With $\widetilde{\epsilon_t} \sim N(0,1)$, $\widetilde{\eta_t} \sim N(0,1)$, $\epsilon_t, \eta_t, \widetilde{\epsilon_t}, \widetilde{\eta_t}$ mutually independent. As for h_t, q_t, they represent either ARCH(p) or GARCH(p,q) models.
Here, for illustration purposes, we consider the model with ARCH (1)

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 \tag{38}$$

$$q_t = \gamma_0 + \gamma_1 \eta_{t-1}^2 \tag{39}$$

The problem with the inclusion of these additional terms is that the usual Kalman filter, as described in section 3.2, is not operable. For the states and innovations variances, instead of (30) and (23) respectively we would now have:

$$F_t = Z_t P_t Z_t' + H_t + h_t \tag{40}$$

$$P_t^* = T_t P_t^* T_t' + R_t Q_t R_t' + q_t \tag{41}$$

The problem is that both h_t and q_t are functions of the past unobserved shocks ϵ_{t-1}^2 and η_{t-1}^2, in fact making the estimation of the model by the usual Kalman filter unfeasible.

Harvey et al. show that the problem at hand can be dealt with by treating the ARCH disturbances (45) and (46) as both observations and states and then replacing the disturbances in (47) and (48) by their conditional expectation.

$$h_t = \alpha_0 + \alpha_1 E_{t-1}(\epsilon_{t-1}^2) \tag{42}$$

$$q_t = \gamma_0 + \gamma_1 E_{t-1}(\eta_{t-1}^2) \tag{43}$$

The state equation in the general LGSSM can be formulated with the inclusion of ARCH terms in the following way:

$$\alpha_t^* = \begin{bmatrix} \alpha_t \\ \eta_t \\ \epsilon_t \end{bmatrix} = \begin{bmatrix} T_t & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{t-1} \\ \eta_{t-1} \\ \epsilon_{t-1} \end{bmatrix} + \begin{bmatrix} I_m & \Phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_t^* \\ \eta_t \\ \epsilon_t \end{bmatrix} \tag{44}$$

Or more compactly.

$$\alpha^*{}_t = T_t^* \alpha^*{}_{t-1} + e_t \tag{45}$$

Where α_t^* is called the augmented state vector., η_t^* is of size $(m \times 1)$, η_t and ϵ_t are scalars and I_m is the identity matrix of size m.

With

$$E_{t-1}(e_t e_t') = \begin{bmatrix} Q & 0 & 0 \\ 0 & q_t & 0 \\ 0 & 0 & h_t \end{bmatrix} = Q_t^* \tag{46}$$

Analogously the measurement equation now becomes

$$y_t = [Z_t 0\ \Lambda]\alpha_t^* + \epsilon^*{}_t = Z_t^* \alpha_t^* + \epsilon_t^* \tag{47}$$

With

$$E(\epsilon_t^* \epsilon^*{}'_t) = H \tag{48}$$

The values of the disturbances ϵ_t and η_t is not directly observable, in fact in general knowledge of prior observations do not imply knowledge of past disturbances. This yields a model that in fact is not conditionally Gaussian anymore. The Kalman Filter in this case would not yield minimum mean square error estimates (MMSE's) and is said to be quasi optimal.
The distribution of η_t conditional on η_{t-1} is normal with mean zero and variance given by q_t. However the distribution of η_t conditional on past observations, which is the one needed for filtering is not known. Although this raises a further problem, as HRS shows since the mean and variance of α_{t-1}^* is known at time t, we can evaluate the first two moments of the conditional distribution of η_t and ϵ_t.
Under the model specification it follows that:

$$\epsilon_{t-1} = E_{t-1}(\epsilon_{t-1}) + [\epsilon_{t-1} - E_{t-1}(\epsilon_{t-1})] \tag{49}$$

$$\eta_{t-1} = E_{t-1}(\eta_{t-1}) + [\eta_{t-1} - E_{t-1}(\eta_{t-1})] \tag{50}$$

We can retrieve from the previous two equations the terms $E_{t-1}(\epsilon_{t-1}^2)$ and $E_{t-1}(\eta_{t-1}^2)$

$$E_{t-1}(\epsilon_{t-1}^2) = E_{t-1}(\epsilon_{t-1})^2 + E_{t-1}\left[\left(\epsilon_{t-1} - E_{t-1}(\epsilon_{t-1})\right)^2\right] \tag{51}$$

$$E_{t-1}(\eta_{t-1}^2) = E_{t-1}\left(\eta_{t-1}\right)^2 + E_{t-1}\left[\left(\eta_{t-1} - E_{t-1}\left(\eta_{t-1}\right)\right)^2\right] \tag{52}$$

The Kalman recursions for this latest specification will be[8]

[8] The notation $\alpha_{t|t-1}$ and $E_{t-1}(\alpha_t)$ are equivalent

$$\alpha^*_{t|t-1} = T_t\, \alpha^*_{t-1|t-1} \tag{53}$$

$$P^*_{t|t-1} = T_t P^*_{t|t-1} T_t' + Q^* \tag{54}$$

$$v^*{}_t = y_t - Z^*_t \alpha^*_{t|t-1} \tag{55}$$

$$F^*_t = Z^*_t \alpha^*_{t|t-1}\, Z^{*'}_t +\ R \tag{56}$$

$$\alpha^*_{t|t} = \alpha^*_{t|t-1} + P^*_{t|t-1} Z^{*'}_t F^{*-1}_t v_t \tag{57}$$

$$P^*_{t|t} = P^*_{t|t-1} - P^*_{t|t-1} Z^{*'}_t F^{*-1}_t Z^*_t P^*_{t|t-1} \tag{58}$$

From which we can get the terms on the RHS of (60) and (61). The terms $E_{t-1}(\epsilon_{t-1})$ and $E_{t-1}(\eta_{t-1})$ are the last two elements of the filtered state vector $\alpha^*{}_{t-1|t-1}$ and the terms $E_{t-1}\left[\left(\epsilon_{t-1} - E_{t-1}(\epsilon_{t-1})\right)^2\right]$ and $E_{t-1}\left[\left(\eta_{t-1} - E_{t-1}(\eta_{t-1})\right)^2\right]$ are the last two diagonal elements of the filtered covariance matrix $P^*_{t-1|t-1}$.

The estimation procedure can be carried out using the prediction error decomposition of the likelihood as for the general case.

$$Log\ (Y_n) = -\frac{n}{2} log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(log(F^*{}_t) + \frac{v^{*2}_t}{F^*{}_t}\right) \tag{59}$$

In this case the procedure is referred to as quasi-maximum likelihood estimation, to stress the fact that the filter is quasi-optimal under the current specifications.

4. Empirical Analysis

4.1 Model estimation in state space form

Having established the Kalman filter recursions for the general state space model we proceed by illustrating how the model proposed by Molodtsova and Papell (15) will be estimated.
We estimate both the homogeneous and heterogenous specifications, without intercept and output gap. We decided to exclude the intercept since in general is very close to zero.
The two model specification (homogenous and heterogeneous) are respectively

$$\Delta s_{t+1} = \alpha_{\pi,t}(\pi_t - \pi_t^*) + \alpha_{i,t}(i_{t-1} - i_{t-1}^*) + \epsilon_t \tag{60}$$

$$\Delta s_{t+1} = \alpha_{\pi,t}\pi_t + \alpha_{\pi^*,t}\pi_t^* + \alpha_{i,t}i_{t-1} + \alpha_{i^*,t}i_{t-1}^* + \epsilon_t \tag{61}$$

Where starred variables denotes foreign variables. We illustrate how the methods introduced in the preceding chapter are used for the specific estimation of the homogeneous model. Of course for the heterogeneous it is sufficient to adapt the observation matrix Z_t and the state vector α_t as appropriate.
In state space form the homogeneous model becomes[9]

$$\Delta s_{t+1} = Z_t\alpha_t + \epsilon_t \tag{62}$$

$$\alpha_t = \mathrm{T_t}\alpha_{t-1} + \eta_t \tag{63}$$

Where Δs_{t+1} is now a scalar Z_t is a (1×2) vector of explanatory variables :

$$Z_t = [(\pi_t - \pi_t^*), \quad (i_{t-1} - i_{t-1}^*)]$$

α_t is of (2×1) vector state variables, in this case the coefficients:

$$\alpha_t = [\alpha_{\pi,t}, \ \alpha_{i,t}]'$$

$\epsilon_t \sim N(0, \sigma_\epsilon^2)$ is the scalar observation disturbance and $\eta_t \sim N(0, Q_t)$ is the (2×1) vector of states disturbances. ϵ_t, η_t are also considered mutually independent. In this case the state transition matrix T_t is considered to be the (2×2) identity matrix: we assume that the states (i.e. the parameters) follow a random walk. Also we specify that the states covariance matrix Q_t is assumed to be diagonal. This assumptions of course is limiting as it will inhibit the ability

[9] $R_t$is assumed to be a (5×5) identity matrix.

of estimating eventual interactions between the coefficients, however the decision stems from the fact that the Kalman filter is notoriously difficult to calibrate and sensitive to initial condition, and the data set used is relatively small, thus I decided to keep the parameter space as small as possible. We must also notice that the states, (in this case the parameters) ought not to follow necessarily a random walk, indeed in theory there is no difficulty to assume other type of processes. Here we decided for the simplest possible formulation, since as just mentioned we want to keep the parameter space as small as possible, and also because using different specifications have not yielded good results. Schinasi and Swamy assumed the parameters followed an AR(1) process with worse results compared to the random walk specification used by Wolff.

The Kalman filter recursions following the reasoning of the previous paragraph now are.

$$\alpha_{t|t-1} = \alpha_{t-1|t-1} \tag{64}$$

$$P_{t|t-1} = P_{t|t-1} + Q \tag{65}$$

$$v_t = \Delta s_{t+1} - Z_t\alpha_{t-1|t-1} \tag{66}$$

$$F_t = Z_t P_{t|t-1} Z_t' + \sigma_\epsilon^2 \tag{67}$$

$$\alpha_{t|t} = \alpha_{t|t-1} + P_{t|t-1} Z_t' F_t^{-1} v_t \tag{68}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} Z_t' F_t^{-1} Z_t P_{t|t-1} \tag{69}$$

The parameters to be estimated are 3 corresponding to the 2 elements on the diagonal of Q_t and the observation disturbance variance σ_ϵ^2.

The parameters estimates can be obtained optimizing the analogue of (33):

$$Log\ (Y_n) = -\frac{n}{2} log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(log(F_t) + \frac{v_t^2}{F_t}\right) \tag{70}$$

In this section we proceed by estimating different specifications of eq. (15) comparing in sample fit of the model with and without the inclusion of GARCH terms. We analyze the relationship between macroeconomic fundamentals and exchange rates in the model and finally we carry out an out of sample forecasting exercise.

We update the model specified in section 3.2 including GARCH effects in the measurement equation.

$$\Delta s_{t+1} = X_t \alpha^*{}_t + \epsilon^*{}_t \tag{71}$$

$$\alpha_t^* = \alpha^*{}_{t-1} + \eta_t \tag{72}$$

With $\epsilon^*{}_t \sim N(0, h_t)$, $\eta_t \sim N(0, Q)$, $\quad h_t = \gamma_0 + \gamma_1 \epsilon_{t-1}^{*2} + \gamma_2 h_{t-1}$
To proceed with the application of the filtering procedure illustrated above the model is modified accordingly

$$\Delta s_{t+1} = [(i_{t-1} - i_{t-1}^*) \quad (\pi_t - \pi_t^*) \quad 1] \begin{bmatrix} \alpha_{i,t} \\ \alpha_{\pi,t} \\ \epsilon_t^* \end{bmatrix} = Z_t^* \alpha_t^* \tag{73}$$

$$\alpha_t^* = \begin{bmatrix} \alpha_{i,t} \\ \alpha_{\pi,t} \\ \epsilon_t^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_{i,t-1} \\ \alpha_{\pi,t-1} \\ \epsilon_{t-1}^* \end{bmatrix} + \begin{bmatrix} \eta_{t,i} \\ \eta_{t,\pi} \\ \epsilon_t^* \end{bmatrix} = T_t^* \alpha_{t-1}^* + \eta_t^* \tag{74}$$

Where $E_t(\eta_t^* \eta_t^{*'}) = \begin{bmatrix} Q & 0 \\ 0 & h_t \end{bmatrix} = Q_t^*$

As mentioned before, the model can be formulated in many different ways. We opted for this reduced formulation, which excludes the intercept and the output gap as in addition to reasons explained in section 3.1 it provides the best in sample fit for the data both in terms of AIC and BIC information criteria, as it will be shown below.

The model will be estimated by Maximum Likelihood adapting the filtering recursion from the previous section.

$$\alpha_{t|t-1}^* = T_t \alpha_{t-1|t-1}^* \tag{75}$$

$$P^*{}_{t|t-1} = T^* P_{t-1|t-1}^* T^{*'} + Q_t^* \tag{76}$$

$$v_t^* = \Delta s_{t+1} - Z_t^* \alpha_{t|t-1}^* \tag{77}$$

$$F_t^* = Z_t^* P^*{}_{t|t-1} Z_t^{*'} \tag{78}$$

$$\alpha_{t|t}^* = \alpha_{t|t-1}^* + P_{t|t-1}^* Z_t^{*'} F_t^{*^{-1}} Z_t^* v_t \tag{79}$$

$$P_{t|t-1}^* = P_{t|t-1}^* - P_{t|t-1}^* Z_t^{*'} F_t^{*^{-1}} Z_t^* P_{t|t-1}^* \tag{80}$$

The Log Likelihood function analogously becomes

$$Log\ (Y^*{}_n) = -\frac{n}{2} log 2\pi - \frac{1}{2}\sum_{t=1}^{n}\left(log(F^*{}_t) + \frac{v^{*2}_t}{F^*{}_t}\right) \tag{81}$$

4.2 Exploratory data analysis: why GARCH disturbances?

Financial returns time series are known to display changing variance and volatility clustering phenomena, meaning that periods of high variance tend to be followed by high variance periods and viceversa. To deal with this feature of the data Bollerslev (1986) proposed a generalization of the ARCH model of Engle (1982). The intuition behind the model is that shocks of asset returns series are serially uncorrelated, but dependent, and this dependence is a quadratic function of its lagged values. The so called Generalized Autoregressive Heteroscedasticity model (GARCH) is an extension of the ARCH model with autoregressive moving average (ARMA) formulation.

The general GARCH(p,q) is

$$r_t = \sigma_t \epsilon_t \qquad \sigma_t^2 = \gamma_0 + \sum_{i=1}^{p} \gamma_{1,i} r_{t-i}^2 + \sum_{j=1}^{q} \gamma_{2,j} \sigma_{t-j}^2 \tag{82}$$

With $\epsilon_t \sim N(0,1)$

Several more sophisticated alternatives have been proposed in the literature such as the EGARCH of Nelson (1991), and the GJR-GARCH of Glosten et al. (1993), to allow for asymmetric effects between positive and negative shocks.

In this paper the easiest possible formulation is used. As shown by a Hansen and Lunde(2005) in an out of-sample comparison between 330 ARCH type model and the GARCH(1,1) for the Deutsch Mark/US Dollar exchange rate, different types of GARCH models generally fail to outperform the GARCH(1,1).

$$r_t = \sigma_t \epsilon_t \qquad \sigma_t^2 = \gamma_0 + \gamma_1 r_{t-1}^2 + \gamma_2 \sigma_{t-1}^2 \tag{83}$$

We require $\gamma_0, \gamma_1, \gamma_2 > 0$, and $\gamma_1 + \gamma_2 < 1$. If these conditions are satisfied the process is covariance stationary and the unconditional variance of r_t is

$$Var(r_t) = \frac{\gamma_0}{1 - \gamma_1 - \gamma_2} \tag{84}$$

If additionally we have $1 - 2\gamma_1^2 - (\gamma_1 + \gamma_2)^2 > 0$, then the following is true[10]

$$\frac{E(r_t^4)}{E(r_t^2)^2} = \frac{3[1 - (\gamma_1 + \gamma_2)^2]}{1 - (\gamma_1 + \gamma_2)^2 - 2\gamma_1} > 3 \tag{85}$$

This means that under these regularity conditions a GARCH(1,1) model has the nice property of having thicker tails than a normal distribution, which is a preferable for dealing with exchange rate returns, which are notoriously leptokurtic.

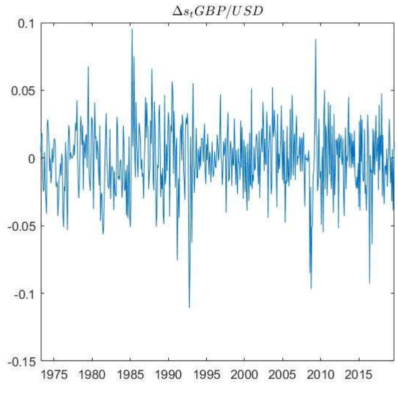
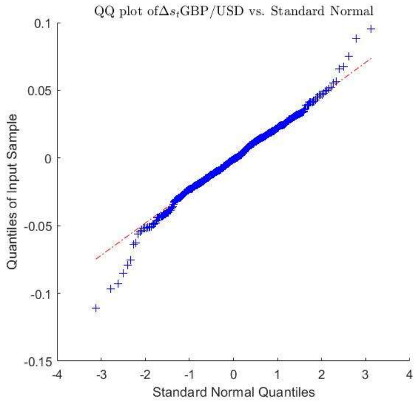
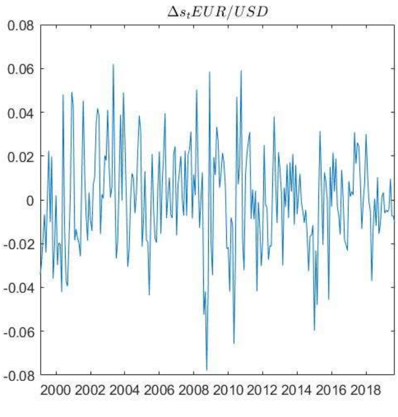
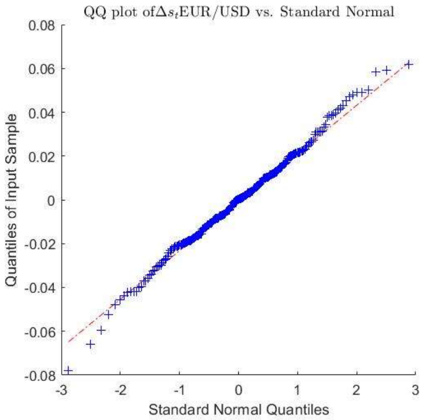
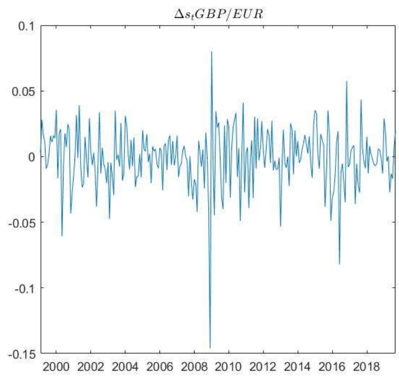
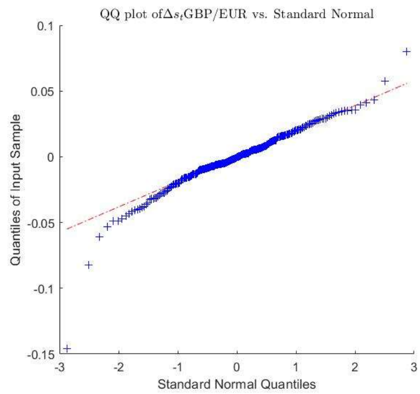
Although these attributes are more evident in high frequency data, we maintain that including GARCH disturbances in the model is preferable. As evident from an exploratory data analysis of returns for the three currencies reported in Fig 1, the returns show signs of being leptokurtic and even if the ARCH effects are not as strikingly visible by visual inspection as for higher frequencies series, a more accurate analysis using the ARCH test of Engle (1982) for lags of 1,5, and 10 months shows that we should reject the null hypothesis of no conditional heteroscedasticity for all three currency pairs. Also by observing the QQ plots we clearly see deviation from normality, although for the EUR/USD exchange rate deviation from gaussianity appears to be milder than for the others exchange rates.

Lags	**GBP/USD**	**GBP/EUR**	**EUR/USD**
1	**0.0001**	**0.011104392**	0.054600333
5	**0.0047**	**0.025009923**	**0.040012845**
10	**0.0449**	**7.88E-07**	**0.046737909**

Table 1 p-values for the ARCH test of Engle.

Figure 1 log returns plot and QQ-plots for the three currencies. Data from FRED Saint Louis Database.

[10] For proofs see Tsay (2010).


ΔstGBP/USD

QQ plot ofΔstGBP/USD vs. Standard Normal
Quantiles of Input Sample
Standard Normal Quantiles

ΔstEUR/USD

QQ plot ofΔstEUR/USD vs. Standard Normal
Quantiles of Input Sample
Standard Normal Quantiles

ΔstGBP/EUR

QQ plot ofΔstGBP/EUR vs. Standard Normal
Quantiles of Input Sample
Standard Normal Quantiles

4.3 Data

The data used to estimate the model are very similar to the ones used by Molodtsova and Papell, and are taken from the Federal Reserve Bank of Saint Louis's database and the International Monetary Fund's International Financial statistics database.

We use the Consumer Price Index (CPI) to measure the price level, the inflation rate is the annual rate, calculated as the 12 month difference of the CPI, the interest rate is the 3 month LIBOR rate based on the currency under consideration. All data are at monthly frequencies from March 1973 to September 2019 for GBP/USD, and from January 1999 to September 2019 for EUR/USD and GBP/EUR.

A point that must be emphasized is that the data used, as in Molodtsova and Papell, are the actual data, meaning the one we observe today. In the context of Taylor rule models this represents a problem. Since the central bank policymaking is forward looking an alternative suggested by Orphanides (2001,2004) would be to use historical real time forecast made by the central bank. However historical forecasts are available only for the Fed[11]. Another approach would be to use *ex post* data, however in this case producing out-of-sample forecasts would be impossible.

4.4 In sample estimation: evaluating model fit and diagnostic checking

We proceed by estimating the model specified in equation (59) with and without the inclusion of GARCH(1,1) disturbances for all three currency pairs. We will refer to the two homogeneous models as TVP and TVP GARCH, the two heterogeneous model are called TVP H and TVP GARCH H. We compare which model fits the data best by Akaike (AIC) and Bayesian (BIC) information criteria. For all the exchange rates under consideration the inclusion of GARCH disturbances produces an increase in the log likelihood function for both model specifications. The GARCH formulations are to be preferred by both AIC and BIC for GBP/USD and GBP/EUR. For EUR/USD the increase in the log likelihood is just marginal,

[11] The so called "Greenbook data".

AIC and BIC are lower for the standard TVP model, also the coefficients γ_0 and γ_1 are not statistically significant at the 5% level.

We report the results in the following tables.

GBP/USD TVP GARCH			
LogL			*1267.9663*
AIC			*-2.53E+03*
BIC			*-2.504E+03*
		std error	*p-values*
σ_i^2	**0.0907**	*0.0347*	*0.0132*
σ_π^2	0.0509	*0.013*	*0.1836*
γ_0	**0.0001**	*2.5E-05*	*0.0293*
γ_1	**0.085**	*0.0328*	*0.0085*
γ_2	**0.8118**	*0.0619*	*0.0000*

GBP/USD TVP			
LogL			1260.3577
AIC			-2.51E+03
BIC			-2.5017E+03
		std error	*p-values*
σ_i^2	**0.1376**	*0*	*0.0000*
σ_π^2	0.0033	*0.4237*	*0.4237*
σ_ϵ^2	**0.0005**	*3.00E-05*	*0.0000*

GBP/EUR TVP GARCH			
LogL			583.0307
AIC			-1.16E+03
BIC			-1.14E+03
		std error	*p-values*
σ_i^2	0.0236	0.0224	*0.2949*
σ_π^2	0.0000	0.0378	*0.9994*
γ_0	0.0001	4.76E-05	*0.1260*
γ_1	**0.1970**	0.0831	*0.0186*
γ_2	**0.6791**	0.1387	*0.0000*

GBP/EUR TVP			
LogL			572.84
AIC			-1.14E+03
BIC			-1.13E+03
		std error	*p-values*
σ_i^2	0.0463	*0.0421*	*0.2719*
σ_π^2	5.68E-07	*0.0014*	*0.9997*
σ_ϵ^2	**0.0005**	*3.63E-05*	*0.0000*

EUR/USD TVP GARCH			
LogL			588.1643
AIC			-1.166E+03
BIC			-1.15E+03
		std error	*p-values*
σ_i^2	**0.09430**	*0.04542*	*0.0389*
σ_π^2	0.07998	*0.07397*	*0.2806*
γ_0	0.00008	*0.00006*	*0.2233*
γ_1	0.06238	*0.04364*	*0.1542*
γ_2	**0.76030**	*0.16305*	*0.0000*

EUR/USD TVP			
LogL			586.5587
AIC			-1.1671E+03
BIC			-1.16E+03
		std error	*p-values*
σ_i^2	**0.0961**	*0.041927*	*0.0228*
σ_π^2	0.0059	*0.008649*	*0.4937*
σ_ϵ^2	**0.0005**	*0.000039*	*0.0000*

Table 2 parameter estimates and standard errors for the two homogeneous models. (Values in bold are statistically significant at the 5% level)

GBP/USD TVP GARCH H			
LogL			*1254.568*
AIC			*-2.50E+03*
BIC			*-2.465E+03*
		std error	*p-values*
σ_i^2	**0.0083**	*0.0113*	*0.4643*
σ_i^{2*}	0.0146	0.0109	0.1829
σ_π^2	**0.0247**	*0.0119*	*0.0389*
σ_π^{2*}	0.0169	0.0106	0.1131
γ_0	**9.69E-05**	*4.6E-05*	*0.0365*
γ_1	**0.1089**	*0.0398*	*0.0064*
γ_2	**0.7261**	*0.0989*	*0.0000*

GBP/USD TVP H			
LogL			1245.6559
AIC			-2.48E+03
BIC			-2.4597E+03
		std error	*p-values*
σ_i^2	0.0167	*0.0164*	*0.3088*
σ_i^{2*}	**0.0373**	0.0087	0
σ_π^2	**0.0274**	*0.0127*	*0.0314*
σ_π^{2*}	0.0002	0.0003	0.6177
σ_ϵ^2	**0.0006**	*0.00E+00*	*0.0000*

GBP/EUR TVP GARCH H			
LogL			*576.8925*
AIC			*-1.14E+03*
BIC			*-1.115E+03*
		std error	*p-values*
σ_i^2	0.0001	*0.03824*	*0.9989*
σ_i^{2*}	0.0262	0.0173	0.1307
σ_π^2	4.95E-05	*0.0350*	*0.9989*
σ_π^{2*}	3.27E-05	0.0232	0.9989
γ_0	6.17E-05	*3.2E-05*	*0.0519*
γ_1	**0.1624**	*0.071*	*0.0235*
γ_2	**0.7293**	*0.090*	*0.0000*

GBP/EUR TVP H			
LogL			567.8022
AIC			-1.13E+03
BIC			-1.1081E+03
		std error	*p-values*
σ_i^2	0.0005	*3.1084*	*0.9999*
σ_i^{2*}	0.0245	0.0354	0.4895
σ_π^2	0.0004	*8.366*	*1.0000*
σ_π^{2*}	2.07E-07	0.0032	0.9999
σ_ϵ^2	**0.0005**	*3.44E-05*	*0.0000*

EUR/USD TVP GARCH H			
LogL			*588.601*
AIC			*-1.16E+03*
BIC			*-1.139E+03*
		std error	*p-values*
σ_i^2	**3.56E-05**	*0.02542*	*0.9989*
σ_i^{2*}	0.00012	0.048673	0.99796
σ_π^2	0.00009	*0.049548*	*0.9986*
σ_π^{2*}	0.13729	0.058064	0.0189
γ_0	**8.09E-05**	*5.2E-05*	*0.1192*
γ_1	**0.1029**	*0.06177*	*0.0972*
γ_2	**0.7054**	*0.147893*	*0.0000*

EUR/USD TVP H			
LogL			586.215
AIC			-1.16E+03
BIC			-1.1449E+03
		std error	*p-values*
σ_i^2	**0.00086**	*2.1392426*	*0.9997*
σ_i^{2*}	0.0019	2.07261166	0.999
σ_π^2	0.0017	*0.9961426*	*0.9986*
σ_π^{2*}	0.0245	0.01714636	0.155
σ_ϵ^2	**0.00041**	*3.74E-05*	*0.0000*

Table 3 parameter estimates and standard errors for the two heterogeneouss models. (Values in bold are statistically significant at the 5% level)

We can clearly see that for the Euro exchange rates under consideration few parameters are statistically significant. Particularly concerning is the fact that the variances of the coefficients are not statistically different from zero This suggests that a time varying parameter model may not be too appropriate for the present application, as Kim and Nelson (1989) note, having coefficient variances equal to zero implies that coefficients are constant and not time varying. For what concerns the GARCH(1,1) coefficients, all three are significant only for GBP/USD data. Although this is not encouraging by looking at the standardized residuals we argue that it is still preferable to include GARCH effects in the model as residuals for the standard Gaussian TVP shows larger kurtosis and serial correlation. We define e_t and e_t^* the standardized residuals for the standard TVP model and the TVP GARCH respectively. They are computed as follows after the Kalman filter has been completed.

$$e_t = \frac{v_t}{\sqrt{F_t}} \qquad\qquad e_t^* = \frac{v_t^*}{\sqrt{F_t^*}} \tag{86}$$

We check if serial correlation and ARCH effects are present in the standardized residuals for both models by means of the Ljung Box test and the ARCH test of Engle. The results are reported in tables 3 and 5. We also check whether the residuals are conditionally gaussian by means of the Jarque Bera test, we also report the sample kurtosis. The results are reported in tables 4 and 6.

Looking at the above table we can see that for higher lags residuals of both models still shows sign of being autocorrelated, with the exception of the British Pound Euro exchange rate for which we cannot reject the null of no autocorrelation in the residuals. ARCH effects are displayed only for the TVP model making it inadequate for the present application. The inclusion of GARCH(1,1) disturbances manages to capture the ARCH feature present in the data. For the GBP/USD and GBP/EUR exchange rates deviation from normality in the residuals are still evident, however the inclusion of GARCH term produce an improvement of the sample Kurtosis. EUR/USD is the only currency pair for which the null hypothesis of normality cannot be rejected, the sample kurtosis for both models are very similar, indeed this confirms the argument made during exploratory data analysis that deviation from normality are milder for the Euro U.S Dollar exchange rate.

GBP/USD						
	Ljung Box e_t, e_t^*		*Ljung Box* e_t^2 e_t^{*2}		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*
1	1E-05	0.0094	0.001	0.9608	0.0001	0.9609
5	3E-05	0.0525	0.0005	0.9988	0.0002	0.999
10	3E-05	0.0088	0.0002	0.7409	0.001	0.7542

EUR/GBP						
	Ljung Box e_t, e_t^*		*Ljung Box* e_t^2 e_t^{*2}		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*
1	0.4663	0.9408	0	0.427	0	0.4306
5	0.1632	0.3922	0.0023	0.8504	0.0019	0.8484
10	0.2012	0.3388	0.0199	0.9643	0.0243	0.9640

EUR/USD						
	Ljung Box e_t, e_t^*		*Ljung Box* e_t^2 e_t^{*2}		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*
1	0	0.0015	0.0603	0.657	0.0621	0.6592
5	0.0023	0.0172	0.0003	0.421	0.0029	0.4388
10	0.0266	0.1038	0.0021	0.5932	0.013	0.6131

Table 3 p-values of diagnostic tests for correlation and ARCH effects for homogeneous model in standardized residuals

	GBP/USD		GBP/EUR		EUR/USD	
	TVP	*TVP GARCH*	*TVP*	*TVP GARCH*	*TVP*	*TVP GARCH*
JB test	1.00E-03	1.00E-03	1.00E-03	1.00E-03	0.2235	0.1916
Kurtosis	4.7268	3.81	9.0346	6.8075	3.4437	3.5278

Table 4 p-values of diagnostic tests for Jarque Bera test of standardized residuals, and standardized residuals sample kurtosis for homogeneous model.

	GBP/USD					
	Ljung Box e_t, e_t^*		*Ljung Box* $e_t^2\ e_t^{*2}$		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*
1	1.66E-05	0.0037	0.0005	0.6464	0.0006	0.6475
5	0.00028	0.0348	0.0004	0.9835	0.0035	0.9863
10	9.00E-05	0.0092	0.0005	0.8419	0.0057	0.8233

	GBP/EUR					
	Ljung Box e_t, e_t^*		*Ljung Box* $e_t^2\ e_t^{*2}$		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*
1	0.4309	0.7384	0	0	0	0.4120
5	0.1416	0.3244	0.0003	0.0002	0.0019	0.8857
10	0.1949	0.3353	0.0043	0.0043	0.0243	0.9908

	EUR/USD					
	Ljung Box e_t, e_t^*		*Ljung Box* $e_t^2\ e_t^{*2}$		*ARCH test*	
	p-value	*p-value*	*p-value*	*p-value*	*p-value*	*p-value*
Lags	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*
1	0	0.0092	0.0362	0.1524	0.0376	0.1551
5	0.0015	0.0521	0.0001	0.2004	0.001	0.2449
10	0.0181	0.1524	0.0011	0.4615	0.0077	0.4598

Table 5 p-values of diagnostic tests for correlation and ARCH effects for heterogeneous model in standardized residuals

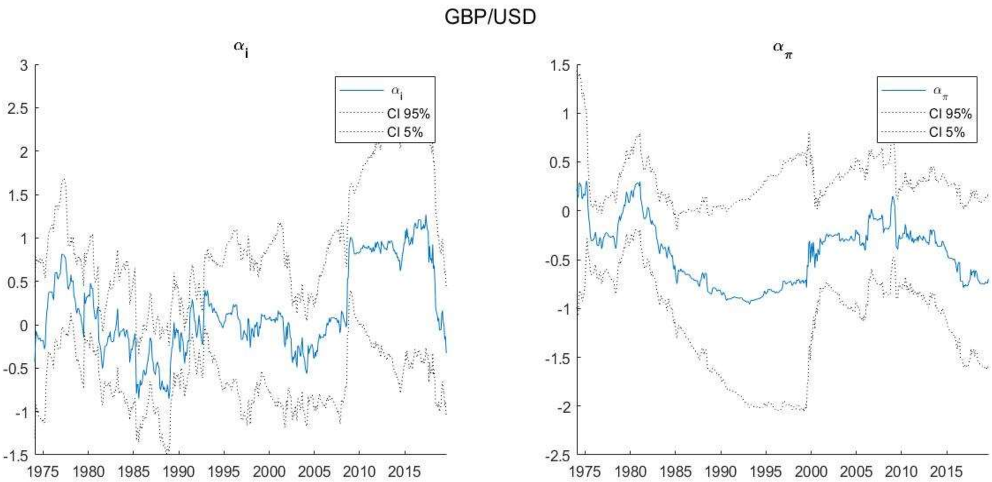
	GBP/USD		GBP/EUR		EUR/USD	
	TVP H	*TVP GARCH H*	*TVP H*	*TVP GARCH H*	*TVP H*	*TVP GARCH H*
JB test	1.00E-03	1.00E-03	1.00E-03	1.00E-03	0.2235	0.1916
Kurtosis	4.9217	4.2676	8.2198	7.445	3.6434	3.4503

Table 6 p-values of diagnostic tests for Jarque Bera test of standardized residuals, and standardized residuals sample kurtosis for heterogeneous model.

4.5 Time varying coefficients

We report the evolution of the parameters for the homogeneous TVP GARCH model in figure 2. For the British Pound Euro exchange rate the parameters are close to zero and shows little variation over time, as expected from the results in table 2.

For the two other currencies (GBP/USD, EUR/USD) α_i is in general estimated to be positive, meaning that a positive difference between domestic and foreign lagged interest rate results in depreciation of the home currency. The parameter for inflation differential α_π is in general negative meaning that higher inflation for the home country, which is associated with an increase in the interest rate target of the central bank, generates appreciation in the domestic currency. This would be in agreement with empirical evidence produced by Chinn, and Engle and West, of failure in the short of UIRP and rational expectations. However by looking at the confidence intervals we are not able to make any conclusive statement. It is evident that only the inflation differential for the Euro US Dollar exchange rate is barely significant at the 5% level from 2010 onwards.



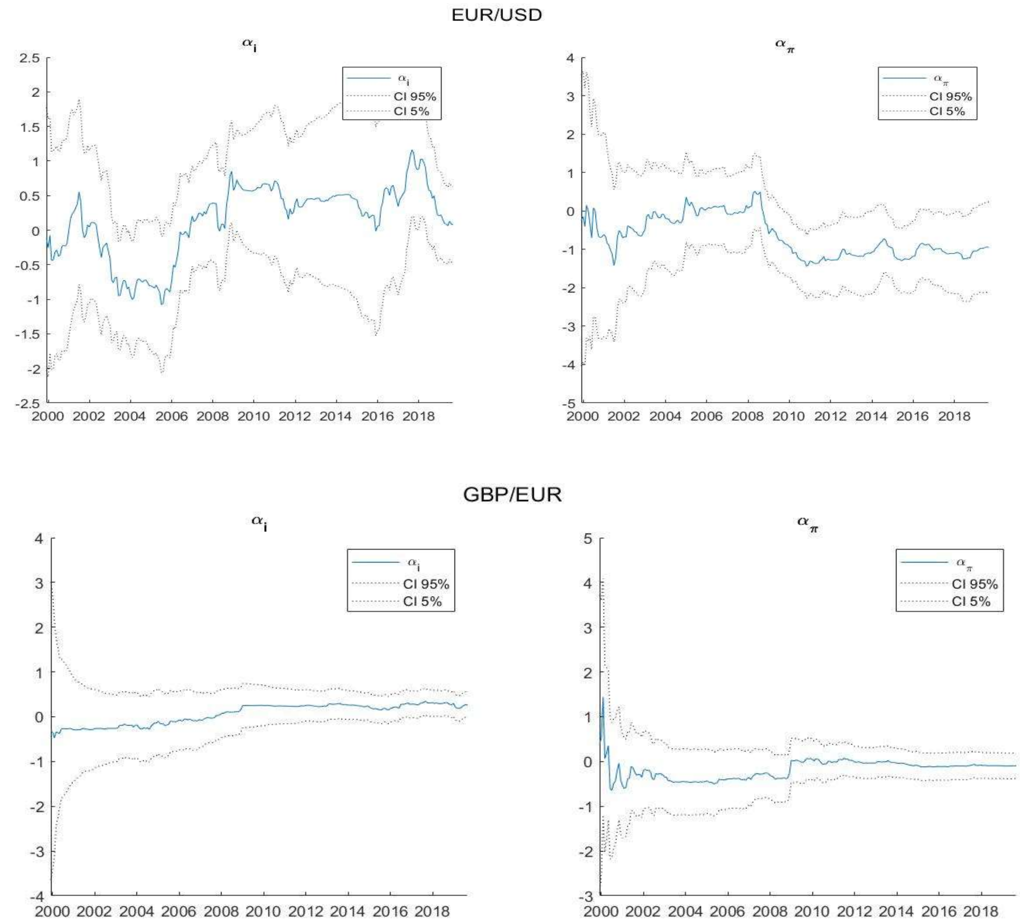


Figure 2 Evolution of the parameters for the three exchange rates TVP GARCH model.

We report the evolution of the parameters for the heterogeneous TVP GARCH H model in figure 3,4, and 5.

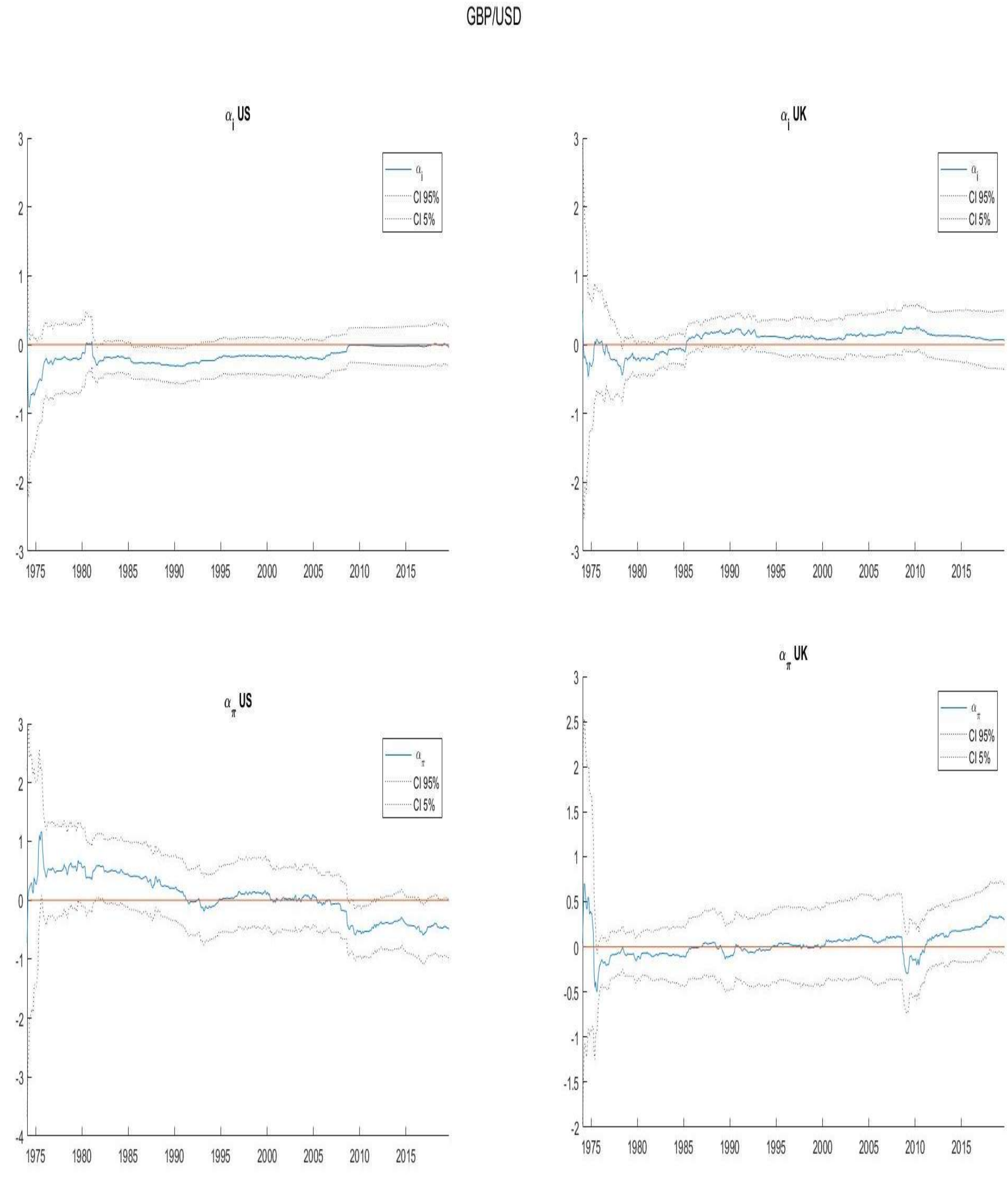


Figure 3 Evolution of the parameters for the GBP/USD exchange rate TVP GARCH H model.

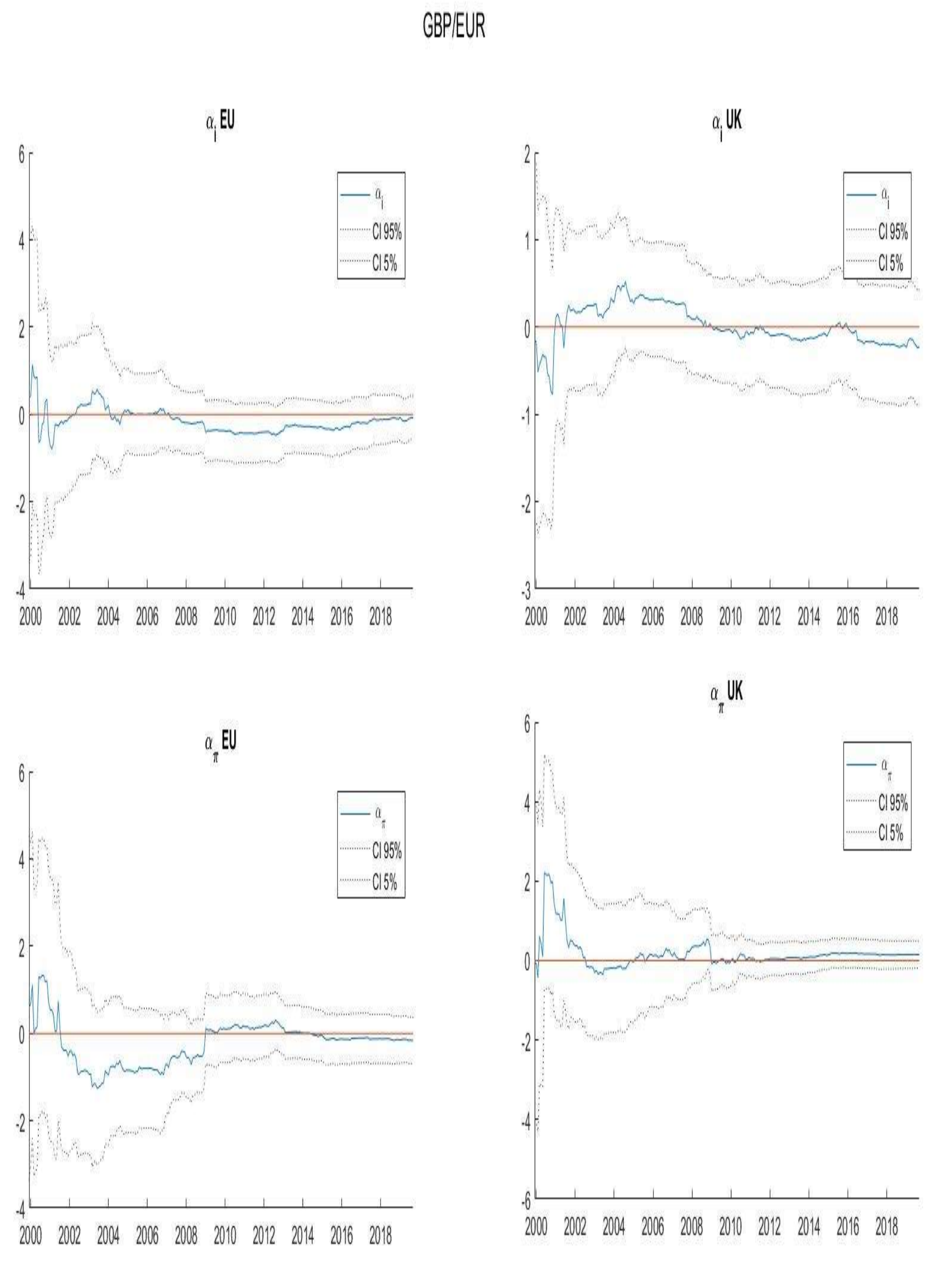


Figure 4 Evolution of the parameters for the GBP/EUR exchange rate TVP GARCH H model.

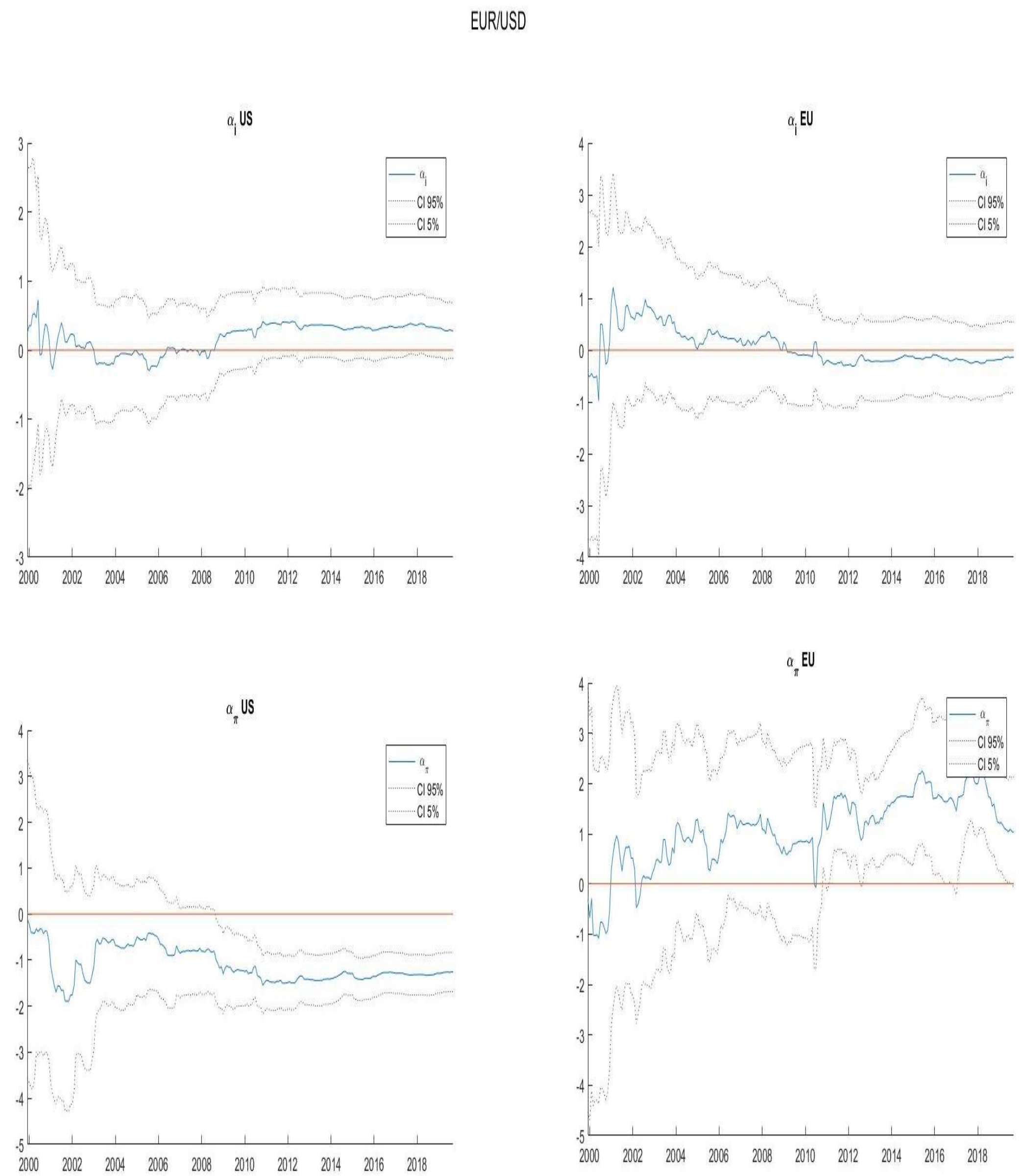


Figure 5 Evolution of the parameters for the EUR/USD exchange rate TVP GARCH H model.

The heterogenous model specification does not make any improvement in establishing a link between exchange rate and the explanatory variables. For British Pound US dollar and British Pound Euro none of the parameters can be said to be statistically different from 0 at the 5% level. For the Euro US Dollar exchange rate we confirm the results of the homogeneous model. Starting from 2010 the parameter for domestic inflation and foreign inflation are negative and positive respectively, while being statistically different from zero. We can't make any conclusive statements about the influence of lagged interest rates on EUR/USD movements. Athough after 2010 the parameter stays positive for US with the 95% confidence interval only narrowly including zero, suggesting that the exchange rate could agree with UIRP in the long run, the parameter for the Euro zone inflation despite being negative for the period under consideration is not statistically significant.

In figure 6 we report the level of uncertainty in exchange rate movements , measured by the conditional variance of the forecast errors. There are no clear differences between the conditional variances implied by the two different model specifications. The higher variance periods for all currency pairs coincide with the great financial crisis of the 2007-2009, for the Euro US Dollar exchange high variance is displayed during the sovereign debt crisis of the eurozone. Almost all the uncertainty in the exchange rates is explained by the GARCH disturbances with parameter uncertainty having residual explanatory importance at best.

Homogeneous *Heterogeneous*

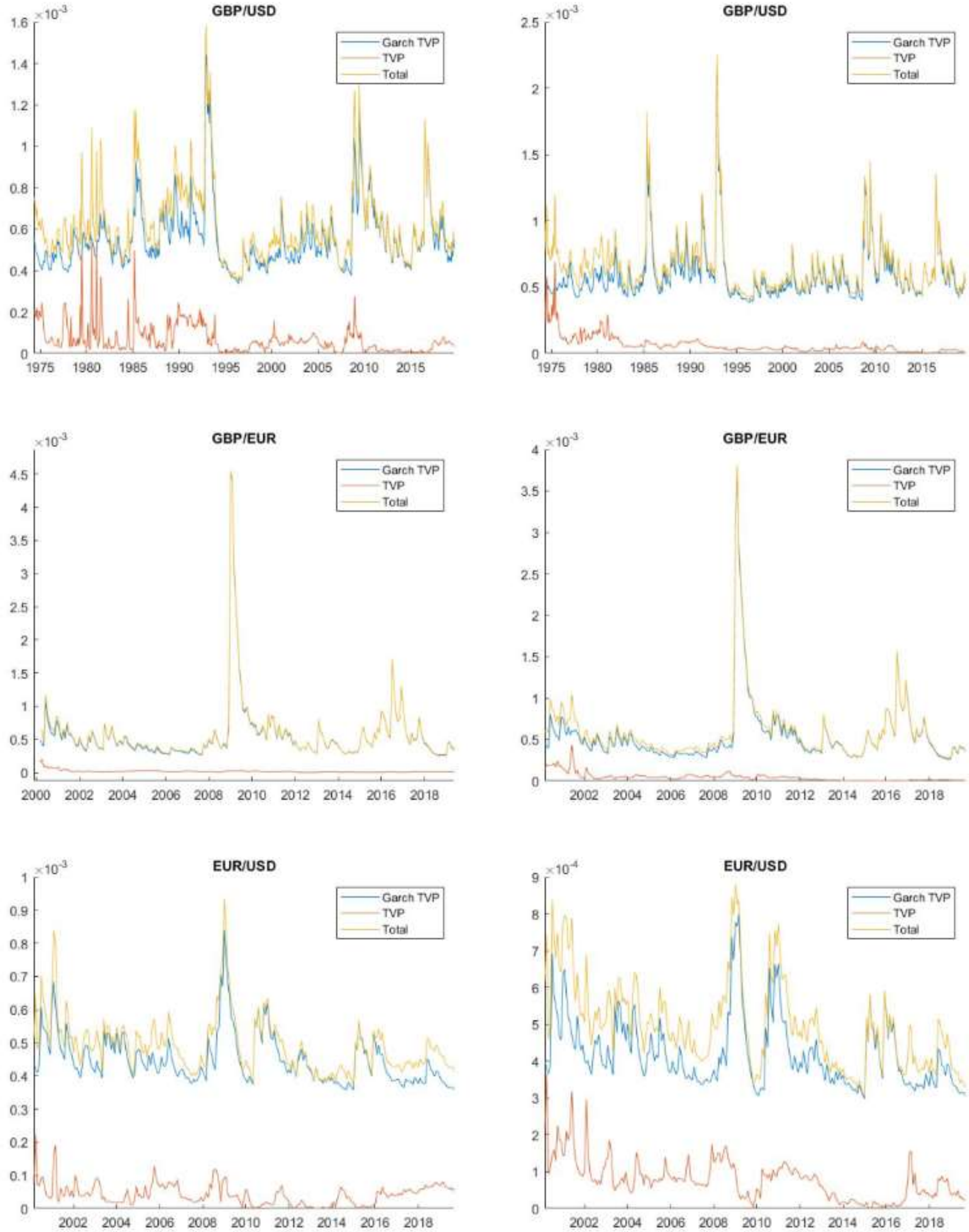


Figure 6 conditional variance implied by the homogenous TVP GARCH (left) and heterogeneous TVP GARCH H (right) models. The yellow line represents the total variance, the blue the part explained by the GARCH(1,1) disturbances, the orange line the part due to parameter uncertainty.

5. Conclusion

This paper deviates from existing empirical literature on exchange rate determination by Taylor rule fundamentals by focusing on in sample properties of the model. Previous attempts have their attention exclusively restricted towards the out of sample forecasting power of Taylor rule models. The most notable, and promising endeavour is the one made by Molodtsova and Papell, followed more recently by Haskamp, which uses the Kalman filter with negative out of sample performance. Both authors completely fixate on out of sample forecasts, even avoiding the inclusion of parameter estimates, standard errors, and the value of the likelihood function, and just selectively showing the evolution through time of the parameters (for which Haskamp does not even provide a confidence interval).

Starting from the encouraging results of Molodtsova and Papell, with the goal of establishing a relationship between macroeconomic fundamentals, we tried to improve by a state space formulation of a modified backward looking version of the Taylor rule of Clarida et al. that involves GARCH disturbances for three exchange rates (GBP/USD EUR/USD GBP/EUR). Even if this constitutes an improvement over the standard homoscedastic formulation, which fails to capture ARCH effects in the residuals, the results overall are disheartening.

For two exchange rates out of three (the exception being GBP/USD), we even reject the model specification of having parameters follow a random walk in favour of fixed parameters. Moreover we fail to give any definitive explanation of movements in the exchange rates in terms of changes in macroeconomic conditions, even for the British Pound US Dollar pair, as the parameters are never statistically different from zero. Also both model specifications shows little to no explanatory power in terms of exchange rate uncertainty, which is almost fully accounted for by the GARCH disturbances.

In light of the evidence presented, seems surprising that Molodtsova and Papell managed to outperform, (although still in few cases), a random walk in terms of out of sample forecasts. A further inspection reveals that indeed the few patterns of the parameters they reported are rarely statistically significant displaying similar behaviour to our results.
We can argue that the general outcome of their results can be sample dependent, indeed they consider the US Dollar as the domestic currency for all tries, and focus on data ranging from

1986 to 2006 which correspond to the FED chairmanships of Volcker and Greenspan, during which the Taylor rule fit remarkably well the monetary policy decisions of the United States. We conjecture that changes in macroeconomic conditions, and the great financial crisis have contributed to changes in the central banks monetary policy,

We conclude that the model based on the backward Taylor rule formulation of Clarida et al. in general can hardly be considered informative in terms of linking exchange rates to macroeconomic factors for the currency pairs taken under consideration. Although we cannot identify precisely any reason, considering the fact that the both model specification fits best for the British Pound US Dollar exchange rate (for which we use data starting from 1973); we conjecture that the inappropriateness of the Taylor rule for exchange rate determination can be due to changes in the monetary policy decisions of the central bank brought by changing macroeconomic conditions in the last 20 years.

Appendix: MATLAB

```
function [lik,a_hat,B,P,q,v,at,Pt,Q]=Kalman_GARCH(theta0,s,H,bsum)

%%% function for the Quasi ML procedure of Harvey et al. with GARCH(1,1)
%%% disturbances in the obsrevation eq. of a time varying paramter model
%%% with diagonal state variance covariance matrix and paramteres that
%%% folows a random walk

%%  Author: Andrea Sergi
%% input
% theta0: staring value of uknown paramaters
% s: dependent variable (exchange rate (t x 1))
% H: explanaatory variables (t x m) (CPI interest)

n=size(s,1);
t=size(s,1); % number of obs
m=size(H,2); % number of parameters
I=eye(n);

% preallocating the Q matric
q=zeros(n,1);
Q=zeros(m,m,n);
Q(:,:,1)=zeros(m,m);
Q(1:m-1,1:m-1,1)=diag(theta0(1:m-1));
q(1)=theta0(m)/(1-theta0(m+1)-theta0(m+2));
Q(m,m,1)=q(1);

% preallocating state transition matrix assuming paramters follows a random
% walk
A=zeros(m,m);
A(1:m-1,1:m-1)=eye(m-1,m-1);

% preallocating matrices
P=zeros(m,m,n);      % P_t|t-1
a_hat=zeros(m,n);    % a_t|t-1
F=zeros(n,1);        % variance of forecast errors
v=zeros(n,1);        % forecast error
at=zeros(m,n);        % a_t-1|t-1
Pt=zeros(m,m,n);      % P_t-1|t-1

P(:,:,1)=1e+8.*eye(m,m); %diffuse initialisation
P(m,m,1)=q(1);

v(1)=s(1)-H(1,:)*a(:,1);
a_hat(:,1)=A*a(:,1);
a_hat(:,1)=zeros(m,1);

F(1)=H(1,:)*P(:,:,1)*H(1,:)';

lik=zeros(n,1);
at=zeros(m,n);
Pt(:,:,1)=5.*eye(m,m);
Pt(m,m,1)=q(1);
F(1)=H(1,:)*Pt(:,:,1)*H(1,:)';

for i=2:n

    % garch term
    q(i)=theta0(m)+theta0(m+1)*(at(m,i-1)^2 + Pt(m,m,i-1))+theta0(m+2)*Q(m,m,i-1);
```

```
    Q(1:m-1,1:m-1,i)=diag(theta0(1:m-1));
    Q(m,m,i)=q(i);

    a_hat(:,i)=A*at(:,i-1);

    P(:,:,i)=A*Pt(:,:,i-1)*A'+Q(:,:,i);

    v(i)=s(i)-H(i,:)*a_hat(:,i);

    F(i)=H(i,:)*P(:,:,i)*H(i,:)'+R;

    at(:,i)=a_hat(:,i)+P(:,:,i)*H(i,:)'*inv(F(i))*v(i);
    Pt(:,:,i)=P(:,:,i)-P(:,:,i)*H(i,:)'*inv(F(i))*H(i,:)*P(:,:,i);

    lik(i)=-0.5*log(2*pi)-0.5*log(F(i))-0.5*(v(i)^2/F(i));
end

if bsum==1
    lik=-sum(lik(2:end));
else
    lik=lik(2:end);
end
end
```

The results reported were obtained by using MATLAB. Since there are no MATLAB built-in tools for the application of the Kalman Filter with GARCH disturbance of Harvey et al. the code used for the estimation of the model was written from zero. Here I provide the function used for the computation of the likelihood. Once the likelihood has been maximized filtered state and variance estimates are obtained by running the same function and using as inputs the parameters obtained from the likelihood maximization step.

References

Bacchetta, P., E. van Wincoop and T. Beutler (2010), Can Parameter Instability Explainthe Meese-Rogoff Puzzle? NBER International Seminaron Macroeconomics, University of Chicago Press.

Cheung, Y.-W., Chinn, M.D., Pascual, A.G., 2005. Empirical exchange rate models of thenineties: are anyfit to survive? Journal of International Money and Finance

Chinn, M.D., 2008. Nonlinearities, Business Cycles, and Exchange Rates. Manuscript,University of Wisconsin.

Durbin J., Koopman S.J. Time Series Analysis by State Space Methods, Oxford University Press,2012

Engle, R. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation." Econometrica. Vol. 96, 1988

Gourinchas, P.-O., Tornell, A., 2004. Exchange rate puzzles and distorted beliefs. Journalof International Economics

Harvey, A.C., 1989. Forecasting, Structural Time Series Models, and the Kalman Filter. Cambridge UniversityPress, Cambridge, UK.

Harvey, A.C., Ruiz, E., Sentana, E., 1992. Unobserved component time series models with ARCH disturbances.Journal of Econometrics

Haskamp, Ulrich (2017) : Forecasting exchange rates: The time-varying relationship between exchange rates and Taylor rule fundamentals, Ruhr Economic Papers

Kalman, R. (1960). A new approach to linear filtering and prediction problems.Journalof Basic Engineering

G. M. Ljung; G. E. P. Box (1978). "On a Measure of a Lack of Fit in Time Series Models". *Biometrika*.

Kim, C.-J., Nelson, C.R., 1989. The time-varying-parameter model for modeling changing conditional variance:the case of the Lucas hypothesis. Journal of Business and Economic Statistics

Meese, R., & Rogoff, K. (1983). Empirical exchange rate models of the Seventies: Do theyfit out of the sample?

Meese, R.A. and K.S. Rogoff (1983), The Out-of-Sample Failure of Empirical ExchangeRate Models: Sampling Error or Mis-speciÖcation?, in:Exchange Rates and InternationalMacroeconomics,

Meese, R.A. and A. Rose (1991), An Empirical Assessment of Non-linearities in Models of Exchange Rate Determination, Review of Economic Studies

Molodtsova, T., & Papell, D. (2009). Out-of-sample exchange rate predictability withTaylor rule fundamentals.

Orphanides, A., 2001. Monetary policy rules based on real-time data. AmericanEconomic Review

Orphanides, A., van Norden, S., 2002. The unreliability of output gap estimates in realtime. Review of Economics and Statistics

Rossi, B. (2006). Are exchange rates really random walks? Some evidence robust toparameter instability

Rossi, B. (2013). Exchange rate predictability.Journal of Economic Literature

Schinasi, G., & Swamy, P. (1987). The out-of-sample forecasting performance of exchangerate models when coefficients are allowed to change

Taylor, J. (1993). Discretion versus policy rules in practice. In Carnegie-Rochester Conference Series on Public Policy

Wolff, C.P (1987), Time-Varying Parameters and the Out-of-Sample Forecasting Performance of Structural Exchange Rate Models, Journal of Business and Economic Statistics

Summary

Exchange rates influences a wide range of economic agents for different reasons: both countries and individuals trading goods, financial traders, foreign workers and investors are affected by changes in exchange rates.

Different models have been proposed and empirically tested with the goal of linking exchange rate movements to macroeconomic fundamentals such as interest rates, inflation, outputs and money supplies. All the theories proposed state that exchange rates are determined by macroeconomic variables. However no model has managed to produce definitive answer when empirically tested. As first established by Meese and Rogoff floating exchange rates between countries are best approximated as random walks. Meese and Rogoff (M-R) (1983a, 1983b) in their seminal papers used different monetary models for exchange rate determination (Frenkel Bilson, Dornbusch Frenkel, Hooper Morton) with the goal of explaining nominal exchange rate movements in terms of contemporaneous macroeconomic variables, showing that they failed to outperform a naïve random walk (RW without drift) when comparing the out of sample forecasts by root mean forecast square error (RMSE), mean forecast error (ME), and mean absolute forecast error (MAE). The models parameters were kept fixed and were at first estimated employing various econometric techniques, including ordinary least squares (OLS), generalized least squares (GLS), instrumental variables (IV), and then were constrained to values based on the economic and empirical theory of money demand and purchasing power parity. The authors list as possible reasons of the poor forecasting performance the sensitivity of the models to the choice of proxy used to represent expected differences in inflation, temporary or permanent deviation from Purchasing Power Parity (PPP), misspecification of the money demand function, simultaneous equation bias, and changes in the parameters values over time.

The results were nonetheless interpreted as evidence of the inadequacy of the models to establish a meaningful relationship between exchange rates and macroeconomic fundamentals, starting a discussion regarding the usefulness of the models and the forecastability of exchange rates, commonly referred to as "Meese Rogoff puzzle".

Cheung and Chinn (1998) attribute the empirical failure to theoretical flaws in the models. Cheung, Chinn, and Pascual (2005) test the out of sample forecasts for different monetary

models obtaining generally negative results and conclude that different specifications may work better for different exchange rates and at different time horizons.

By following M-R's suggestion that it may be fruitful to account for parameter instability, this thesis tries to improve on previous attempts by using a model based on Taylor rule fundamentals, allowing for parameter variation and heteroscedasticity by putting the model in state space form and estimating it using the Kalman filter. Indeed this approach sounds, at least intuitively, more promising. It seems reasonable to assume that the impact of shocks in the macroeconomic variables on the exchange rate is related to the relative health of the economies of the home and foreign country. For instance if one of the two countries is experiencing a recession with very low inflation, the exchange rate reaction to a further lowering in the price level of this country is expected to be greater than what would happen if his economy was thriving.

It must be noted however that this approach is not new. Different attempts to forecast exchange rates by state space methods have been made producing mixed results. Both Wolff (1987) and Schinasi and Swamy (1989) make use of time varying parameters to forecast exchange rates using the same models as in MR and still taking the RW model as a benchmark with not very encouraging results. Rossi (2006, 2013) in an extensive study uses the Kalman filter with a random walk specification for the parameters with mixed results, but concludes that in general evidence of predictability is scarce. Bacchetta et al. (2009) study wether acoounting instability is sufficient for solving the Meese Rogoff puzzle and conclude that this is hardly the case. Molodstova and Papell (2009) deviating from H-R use different Taylor rule based models of exchange rate determination (1993) for a multitude of currencies, allowing for stochastic parameters using rolling OLS regressions they manage to outperform the RW model in few cases. Haskamp (2017), with the main focus of out of sample forecasting uses the Kalman filter to allow for parameter variation in the Molodtsova Papell model. He reports that the Kalman filter manages to incorporate more abrupt adjustments in the coefficients when compared to the OLS rolling regression, however he still fails to outperform a random walk in terms of out of sample forecasts.

Although accounting for time varying coefficients has in general failed to beat a random walk, it is clear that a sensible improvement has been made over the first attempts with fixed coefficients of MR.

Even if the results are encouraging neither any model has taken into account the well known time varying nature of exchange rates volatility, treated as constant over time, nor any attempt has been made to deviate from the Gaussianity assumption, despite large evidence that exchange rates follow a leptokurtic distribution as shown by the pioneering work of Mandelbrot (1960).

In this thesis the efforts of linking exchange rates to macroeconomic fundamentals through the state space formulation of Taylor rule based macro model are furthered by taking into account the heteroscedastic nature of exchange rates, and allowing Autoregressive Conditional Heteroscedasticity (ARCH) effects to enter the model by the methods of Harvey Ruiz and Sentana.

This paper deviates from existing empirical literature on exchange rate determination by Taylor rule fundamentals by focusing on in sample properties of the model. Previous attempts have their attention exclusively restricted towards the out of sample forecasting power of Taylor rule models. The most notable, and promising endeavour is the one made by Molodtsova and Papell, followed more recently by Haskamp, which uses the Kalman filter with negative out of sample performance. Both authors completely fixate on out of sample forecasts, even avoiding the inclusion of parameter estimates, standard errors, and the value of the likelihood function, and just selectively showing the evolution through time of the parameters (for which Haskamp does not even provide a confidence interval).

Starting from the encouraging results of Molodtsova and Papell, with the goal of establishing a relationship between macroeconomic fundamentals, we tried to improve by a state space formulation of a modified backward looking version of the Taylor rule of Clarida et al. that involves GARCH disturbances for three exchange rates (GBP/USD EUR/USD GBP/EUR). Even if this constitutes an improvement over the standard homoscedastic formulation, which fails to capture ARCH effects in the residuals, the results overall are disheartening.

For two exchange rates out of three (the exception being GBP/USD), we even reject the model specification of having parameters follow a random walk in favour of fixed parameters. Moreover we fail to give any definitive explanation of movements in the exchange rates in terms of changes in macroeconomic conditions, even for the British Pound US Dollar pair, as the parameters are never statistically different from zero. Also both model specifications shows little to no explanatory power in terms of exchange rate uncertainty, which is almost fully accounted for by the GARCH disturbances.

In light of the evidence presented, seems surprising that Molodtsova and Papell managed to outperform, (although still in few cases), a random walk in terms of out of sample forecasts. A further inspection reveals that indeed the few patterns of the parameters they reported are rarely statistically significant displaying similar behaviour to our results.
We can argue that the general outcome of their results can be sample dependent, indeed they consider the US Dollar as the domestic currency for all tries, and focus on data ranging from 1986 to 2006 which correspond to the FED chairmanships of Volcker and Greenspan, during which the Taylor rule fit remarkably well the monetary policy decisions of the United States. We conjecture that changes in macroeconomic conditions, and the great financial crisis have contributed to changes in the central banks monetary policy,

We conclude that the model based on the backward Taylor rule formulation of Clarida et al. in general can hardly be considered informative in terms of linking exchange rates to macroeconomic factors for the currency pairs taken under consideration. Although we cannot identify precisely any reason, considering the fact that the both model specification fits best for the British Pound US Dollar exchange rate (for which we use data starting from 1973); we conjecture that the inappropriateness of the Taylor rule for exchange rate determination can be due to changes in the monetary policy decisions of the central bank brought by changing macroeconomic conditions in the last 20 years.