Modelling and forecasting exchange rates with time-varying parameter models*

Angela Abbate[†] Massimiliano Marcellino[‡]

April 7, 2014

Abstract

We contribute to the literature on exchange rate modelling and forecasting in two distinct ways. First, we show that the interval and density forecasts of three major exchange rates vis-a-vis the US dollar can be improved by assuming time variation in the coefficients of the data generating process. Secondly, we show that the relationship between exchange rates and a set of macroeconomic and financial fundamentals can be unravelled through the modelling of parameter time variation. In particular, we find that controlling for macroeconomic predictors tends to deliver higher one-year ahead predictive likelihoods during economic recessions. A simple trading strategy further reveals that financial and monetary predictors are important from an economic perspective after the 2008 financial crisis.

J.E.L. Classification: C11, C53, F31, F37

Keywords: exchange rates, forecasting, density forecasts, BVAR, time-varying parameters, stochastic volatility

*We would like to thank Gino Cenedese, Pasquale Della Corte, and participants at an ECB - Bank of Italy workshop for useful comments on an earlier draft.

[†]European University Institute, Department of Economics, angela.abbate@eui.eu

[‡]Bocconi University, IGIER and CEPR, massimiliano.marcellino@unibocconi.it

1 Introduction

Exchange rates have an impact on the production decision of firms, on portfolio allocation, on a a country's prices, and more generally on its competitiveness. Hence, the need of having reliable models to track the current evolution of exchange rates and predict their future behaviour, especially in times of uncertainty and financial stress. In fact, exchange rate volatility has changed over the years. It has fallen after the price shocks and inflationary pressures of the 1970s, and it has increased again in the last decade, possibly as a consequence of the quantitative easing measures enacted by central banks around the world[1].

So far, a vast literature has been devoted to the construction and evaluation of the point forecasts of exchange rates. It has established, with few exceptions[2], that the best forecasting model is a simple random walk. This is surprising, given the relationship between exchange rates and a wide set of macroeconomic fundamentals posited by economic theory. This puzzle, originated by the seminal work of Meese and Rogoff [1983], has not yet been solved. A possible explanation lies in the instability over time of the link between exchange rates and fundamentals, as suggested by Bacchetta and van Wincoop [2009]. Alternatively, Engel and West [2005] have shown that an asset-pricing model where at least one of the fundamentals has a unit root, and the discount factor is close to unity, is able to generate exchange rate unpredictability. In addition, competing models have so far been evaluated mainly on the basis of their point forecasts. Though the latter are clearly of interest, for the decision making of economic agents and for the pricing of financial assets, interval and density forecasts of exchange rate are also relevant. On this the literature is more limited. Significant exceptions are Yongmiao et al. [2007] and Balke et al. [2013], who both show that the density forecasts of a random walk can be improved upon either with non-linear models, or with univariate Taylor-rule models with semiparametric confidence intervals.

Our contribution to the literature is twofold. First, we examine whether and to what extent the point, interval and density forecasts of three major exchange rates vis-a-vis the US dollar can be improved by assuming time variation in the coefficients of the data generating process. The exchange rates analysed are the monthly averages of the British Pound, the Japanese Yen, and the German Mark[3], used as a proxy for the Euro, over the period 1971m1 to 2013m6. As it can be seen in figure 1, the volatility of these three currencies has changed over time: a constant-volatility model could therefore lead to the incorrect estimation of forecast intervals, underestimating them in periods of high volatility and overestimating them otherwise. To model time variation, we experiment with two methods recently proposed in the literature: the time-varying parameter Bayesian vector autoregression with stochastic volatility developed by Cogley and Sargent [2005] and Primiceri [2005], and its approximation

[1] See Warwick-Ching, L. (2013, March 25). Currency wars: Volatility provides profit opportunity. *The Financial Times.* Retrieved from `www.ft.com`

[2] Exceptions include forecasts from error correction models (univariate, multivariate and panel), though these results are sensitive to the forecast horizon and to the sample used, as documented in Rossi [2013]. Carriero et al. [2009] show that a Bayesian vector autoregression with a large set of exchange rates beats the random walk in mean squared forecast errors. Moreover, linear models tend to perform better than non-linear ones, while the evidence on time-varying parameter models is mixed. For a comprehensive review see, e.g. Rossi [2013]. Recently, Dal Bianco et al. [2012] have used a mixed-frequency dynamic factor model with four weekly exchange rates and lower-frequency macroeconomic fundamentals. Their model delivers significantly smaller mean squared forecast errors than a random walk, and macroeconomic variables play a significant role.

[3] After the introduction of the euro, the currency refers to the contribution of the German mark to the euro, calculated using the conversion rate fixed on 1 January 1999.

proposed by Koop and Korobilis [2013], based on forgetting factors and on an exponentially weighted moving average estimator of the shocks' covariance matrix. The performance of these models is compared to that of two benchmarks, a Bayesian vector autoregression and a random walk (with and without GARCH innovations), by juxtaposing the respective point, interval and density forecasts. The analysis reveals that, though the point forecasts are similar, the time-varying models, and in particular the forgetting-factor one, deliver sharper and more accurately calibrated density forecasts, thus correctly estimating forecast uncertainty.

Our second contribution is to verify whether exchange rate predictability by fundamentals can be unravelled through a modelling of time variation. To answer this question we employ the forgetting factor methodology and consider the alternative inclusion of a wide set of macroeconomic and financial predictors. We find that models enriched with macroeconomic differentials tend to deliver higher predictive likelihoods at long horizons and in periods of economic recessions. In addition, simple trading strategies based on the competing forecast models reveal that controlling for monetary and financial fundamentals would have yielded positive returns to a US-based investor in the period of the 2008 financial crisis.

To our knowledge this comprehensive evaluation is the first of its kind in the empirical literature on exchange rate forecasting, not only for the methodology used but also for the emphasis on interval and density forecasts. Two works are closely related to this paper. Canova [1993] finds that a time-varying coefficient Bayesian model with exchange rates and short-term interest rates has a higher predictive ability than a random walk. More recently, the one-month ahead predictive ability of macroeconomic fundamentals in a time-varying setting has been evaluated by Della Corte et al. [2009]. Their findings support a time-varying treatment of volatility, in particular stochastic-volatility, as well as the use of forward premium models, which outperforms both the random walk and models with monetary fundamentals[4]. In contrast to our approach, Della Corte et al. [2009] do not allow for dynamic interrelationships across variables and countries, nor do they model time variation in the slope parameters. Though using the same currencies and frequencies, our sample size is longer and includes the 2008 financial crisis. Moreover, we focus on forecast horizons greater than one month, as well as on other statistical measure of predictive accuracy such as coverage rates. On the basis of these consideration, we can say that our approaches complement each other.

The paper is organised as follows. In the next Section we review the main exchange-rate determination models, as well as the empirical strategies that have been shown to improve on a simple random walk forecast model. Section 3 describes the two time-varying parameter models used in the modelling and forecasting exercises. Section 4 is dedicated to the evaluation and comparison of the results delivered by the different models. The assessment of the role of the macroeconomic and financial predictors is presented in Section 5. Section 6 discusses the results of a simple trading strategy based on the competing forecast models, and Section 7 concludes.

[4] Their evaluation criteria are both statistical, relative mean squared errors and log-likelihoods, as well as economical, based on the utility function of investors. In addition, the economic criterion supports an optimal combination of the model forecasts through Bayesian model averaging.

2 The role of macroeconomic and financial predictors

In this Section we briefly outline the theoretical and empirical relationships that link exchange rates to the set of macroeconomic and financial fundamentals used in this work. After introducing the main theoretical models of exchange-rate determination, we provide an overview of the methodologies used in the exchange-rate forecasting literature that have been shown to improve on the point forecasts of the random walk model, established as the benchmark forecast model since the seminal work of Meese and Rogoff [1983]. Lastly, we review two possible explanations for the random walk behaviour of exchange rates and the consequent low predictive ability of macroeconomic fundamentals that have been proposed in the literature.

2.1 Theoretical links between exchange rates and fundamentals

Several variables qualify as potential predictors of future exchange rates. The purchasing power parity theory (PPP), first developed by Cassel [1918], postulates that the nominal exchange rate (s_t) should be equal to the sum of the real exchange rate (q_t), and the difference in the general price level between the foreign and the home country ($p_t^* - p_t$):

$$s_t = p_t^* - p_t + q_t \ , \tag{1}$$

where, following the notation used in the empirical application, s_t is defined as the number of currency units per US dollar, while small case letters denote the logarithms of the variables, unless stated otherwise. Moreover, the uncovered interest rate parity (UIRP) condition suggests that exchange rate movements compensate differentials in the nominal interest rate levels ($i_t^* - i_t$):

$$E_t s_{t+1} - s_t = i_t^* - i_t + \rho_t \ . \tag{2}$$

This condition is based on rational expectation and risk neutrality, and ρ_t can be interpreted either as a forward premium or as an expectational error[5]. Empirical evidence on these models is mixed. Among others, Cheung et al. [2005] show that while the mean squared errors from PPP models are lower than those of a random walk for longer horizons, UIRP models do not significantly improve on the random walk at any horizon. On the contrary, both models are found to outperform the random walk by Della Corte and Tsiakas [2013], on the basis of statistical and economic criteria.

A richer relationship between exchange rates and fundamentals is posited by monetary models. By equating the money demand equations for the home and the foreign country, and assuming 1 and 2 to hold, the flexible price monetary model links the exchange rate to differentials in money, output and nominal interest rate through

[5]See, i.e. Engel and West [2005]. In-sample estimates of the UIRP model usually lead to opposite results from the theoretical relationship: i.e. that the currency of high-interest rate countries appreciates. See, for instance, the discussion in Della Corte and Tsiakas [2013].

the following reduced-form equation[6]:

$$s_t = \beta_0(m_t^* - m_t) + \beta_1(y_t^* - y_t) + \beta_3 q_t + \beta_4(v_t^* - v_t) + \beta_5 \rho_t + \beta_6 E_t s_{t+1} \ , \tag{3}$$

where v_t and v_t^* are shocks to the domestic and foreign money demand equations, and the parameters $\beta = (\beta_0, \ldots, \beta_6)$ are functions of the underlying structural parameters. Equation 3 is typically estimated using as regressors either just money or output differentials, or a combination of the two, $m_t^* - m_t - (y_t^* - y_t)$. The remaining variables are assumed to enter a generic error term. Alternatively, the sticky-price version of the monetary model 3, due to Dornbusch [1976], adds nominal interest rate differentials to the list of potential regressors. Empirical works have shown that the fundamentals proposed by the flexible monetary model have no predictive ability in the short run, see for instance Della Corte et al. [2009], but display comovements with the exchange rates at long horizons[7].

A relatively recent branch of exchange-rate prediction models is based on Taylor rules. These models build upon open economy frameworks, and assume that the policy rule followed by the central bank targets the country's exchange rate, as well as output and inflation. Equating the modified Taylor rules for the home and the foreign country yields a relationship between the exchange rate and differentials in output, inflation and interest rates. The good performance of Taylor rule models has been documented, among others, by Molodtsova and Papell [2009] and Inoue and Rossi [2012], while it has been questioned by Rogoff and Stavrakeva [2008].

Finally, increasing attention is being paid to financial predictors of exchange rates. Molodtsova and Papell [2012] find that the performance of their proposed Taylor rule models can be improved, in some cases, when they are augmented with credit spreads or measures of financial conditions. In addition, Shin et al. [2010] have shown how US credit aggregates, taken as proxies for the risk appetite of financial intermediaries, can help forecasting a wide set of exchange rates.

Other variables qualify as potential exchange rate predictors, including commodity and oil prices, trade balance differentials and productivity measures. We shall not however review the models behind them, as the focus of this paper lies on the most commonly used macroeconomic predictors, as well as on measures of financial market risk and liquidity.

2.2 A brief review of the empirical strategies used in the exchange-rate literature

A wide variety of methods has been used in the empirical literature on exchange-rate forecasting. The consensus has emerged that the toughest benchmark to beat, in terms of the accuracy of point forecasts, is the random walk model. Among the methodologies that have been shown to deliver lower mean squared forecast errors than a random walk[8] stand error-correction models, univariate, multivariate and panel. Carriero et al. [2009] reach a

[6] See, for instance, the derivations in Frankel [1984] and Engel and West [2005].

[7] On this, see Mark and Sul [2001], and the more recent results of Engel et al. [2008]. Both works find, using panel error-correction models, that monetary fundamentals have predictive ability at long horizons.

[8] These results appear however to be sensitive to the forecast horizon and to the sample used, as documented in Rossi [2013].

similar conclusion by relying on a Bayesian vector autoregression with a large set of exchange rates. Recently, Dal Bianco et al. [2012] have estimated a mixed-frequency dynamic factor model with four weekly exchange rates and lower-frequency macroeconomic fundamentals. Their model delivers significantly smaller mean squared forecast errors than a random walk, and macroeconomic variables play a significant role. Mumtaz and Sunder-Plassmann [2013] use a time-varying stochastic volatility vector autoregression to study the impact of asymmetric supply and demand shocks on the real exchange rate in four small open economies. The time-varying parameter model is found to outperform its constant-parameter counterpart on the basis of the mean squared forecast error and of the Bayesian deviance information criterion. A time-varying modelling of volatility is supported also by Della Corte et al. [2009], who show that univariate stochastic-volatility models based on forward premia deliver lower mean squared forecast errors than both the random walk and monetary models. For a recent and more comprehensive review of past works on exchange-rate forecasting, we refer to Rossi [2013].

2.3 Reconciling the exchange-rate disconnect puzzle

Contrarily to what exchange-rate determination models posit, macroeconomic fundamentals do not appear to be good predictors of future exchange rates. A possible explanation for this puzzle lies in the instability of the relationship that links exchange rates to their fundamentals. This instability has been documented, among others, by Rossi [2006] through a series of instability tests. Using survey data, Cheung and Chinn [2001] have explained that instability might result from the behaviour of foreign exchange-rate traders, who frequently change the weight they attach to fundamentals. In addition, Bacchetta and van Wincoop [2009] show that the unstable relationship between fundamentals and exchange rates can be generated within a model whose structural parameters are unknown to economic agents, and evolve gradually over time.

An alternative explanation for the low predictive ability of macroeconomic fundamentals has been suggested by Engel and West [2005], who argue that exchange-rate unpredictability is an implication of the structural models, rather than being evidence against them. Their contribution stems from the consideration that exchange rates are asset prices and are therefore influenced not only by current fundamentals, but also by the expectation on their future values. Engel and West [2005] show that all exchange-rate determination models can be rewritten so as to express foreign exchange rates as present discounted values of current and future fundamentals, as well as unobservable shocks. Intuitively, if fundamentals are very persistent and if agents are patient, implying that future fundamentals matter more than current ones, exchange rates will exhibit almost no correlation with current fundamentals. Analytical calculations provided in their paper show in fact that persistent processes for the macroeconomic variables, coupled with reasonable calibration values for the discount factor, generate very low correlations between exchange rates and fundamentals.

3 BVAR models with time-varying parameters and changing volatility

3.1 The time-varying parameter stochastic volatility BVAR

We start with a short description of the time-varying parameter stochastic volatility Bayesian vector autoregression (TVP-SV-BVAR), developed by Cogley and Sargent [2005] and Primiceri [2005], to whom we refer for additional details. The first component of the model is the measurement equation:

$$y_t = Z_t \beta_t + u_t \ , \tag{4}$$

where y_t is a $n \times 1$ vector of observed variables, Z_t is a $n \times k$ matrix of regressors, β_t is a $k \times 1$ vector of time-varying coefficients and u_t is a $n \times 1$ vector of innovations with covariance matrix Ω_t. Let Z_t contain a constant and p lags of each variable; it is then defined as $Z_t = I_n \otimes [1, y'_{t-1}, \dots y'_{t-p}]$ with dimension $n \times k = n \times n(1+np)$.

Following Primiceri [2005], the covariance matrix Ω_t can be decomposed as follows[9]:

$$A_t \Omega_t A'_t = \Sigma_t \Sigma'_t \ , \tag{5}$$

where Σ_t is a diagonal matrix, with the standard deviations of the structural innovations as its elements; while A_t is a lower triangular matrix with ones on its main diagonal, which summarises the contemporaneous relationships between the variables in y_t. Using the structural decomposition in 5, the measurement equation 4 can be rewritten in terms of the white-noise, homoskedastic and uncorrelated shocks ε_t:

$$y_t = Z_t \beta_t + A_t^{-1} \Sigma_t \varepsilon_t, \qquad \text{with } \ E[\varepsilon'_t, \varepsilon_t] = I_m \ . \tag{6}$$

To close the model, three transition equations are specified, describing the evolution of the parameters over time:

$$\begin{aligned} \beta_t &= \beta_{t-1} + \nu_t \ , \\ \alpha_t &= \alpha_{t-1} + \xi_t \ , \\ \log \sigma_t &= \log \sigma_{t-1} + \eta_t \ , \end{aligned} \tag{7}$$

where α_t is the vector of the non-zero, non-one elements of A_t stacked by rows, and σ_t is the vector of the diagonal elements in Σ_t. While the slope coefficients and those in the contemporaneous impact matrix are assumed to follow a random walk, the standard deviations of the structural innovations are modelled as geometric random walks. Finally, all the innovations of the model are posited to be distributed as a multivariate normal, with zero mean

[9] The decomposition in 5 emphasises the two drivers of the time variation in Ω_t: variation in the variance of the innovations and in their correlation structure.

and with the following block diagonal covariance matrix:

$$V = Var \begin{bmatrix} \varepsilon_t \\ \nu_t \\ \xi_t \\ \eta_t \end{bmatrix} = \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix} . \tag{8}$$

The objectives of the estimation are the unobserved paths of the parameters in 7, indicated by (B^T, A^T, Σ^T), and the hyperparameters in V. Sampling from the posterior density requires the specification of a prior distribution, as well as the use of a posterior simulator algorithm. A description of both is given below.

The covariance matrices (Q, W, S) are assumed to have an inverse-Wishart distribution and to be therefore characterised by a number of degrees of freedom and a scale matrix, set to a constant fraction of the covariance matrix's training sample estimate. The initial states of the three types of time-varying coefficients, $\beta_0, \alpha_0, \log \sigma_0$, are assumed to be normally distributed, with mean and variance calibrated through a training sample. The prior on the initial states and the transition equations in 7 imply that, conditional on Q, W, S, the prior distributions of the entire sequence of VAR coefficients, contemporaneous relationships, and log standard deviations are themselves normal. Further details on the prior specification, related to the empirical application of this paper, are provided in Section 4.

Sampling from the BVAR posterior density To generate a sample from the posterior of (B^T, A^T, Σ^T, V) we rely on a five-step Gibbs-sampler, following Primiceri [2005] and Del Negro and Primiceri [2013].

The first step is to sample the sequence of VAR coefficients β^T, given an initial guess of the parameters. For this task, a simulation smoother like the one proposed in Carter and Kohn [1994] can be used, exploiting the fact that the distribution of β^T, conditional on A^T and Σ^T, is linear and normal. The sequence of A^T can be drawn in a similar way, as its posterior distribution is normal, given B^T and Σ^T.

To draw the sequence of standard errors, the model needs to be transformed, given that it is neither linear nor Gaussian in Σ^T. More specifically, at this stage of the sampler the innovations to the measurement equation are distributed as a $\log \chi^2$. The transformation of the system can be achieved by using a mixture of normal approximations of the $\log \chi^2$ distributions, as described in Kim et al. [1998]. After sampling s^T, the matrix of indicator variables that rules the normal approximation[10], the system is approximately linear and Gaussian, conditional on A^T, B^T, V and s^T: a standard simulator smoother can then be applied, to recover the smoothened estimates of the volatility and of the variance of its innovations.

The last step is a draw from the inverse-Wishart distributions of the block components of V, yielding a sample of the model's covariance matrix.

[10]The parameters used for the approximation to the $\log \chi^2$ distribution are those of Primiceri [2005].

Sampling from the BVAR predictive density Let us denote with y^t and $\theta^t = (B^t,\, A^t,\, \Sigma^t,\, V)$, the history of the variables and of the coefficients from period 1 up to period t. We want to forecast up to h steps ahead in the future, that is, to make predictions on the vector $y^{t+h} = [y'_{t+1},\, \ldots,\, y'_{t+h}]$. For this, we need the predictive density of the BVAR model, which can be factored as follows, emphasising the different sources of forecast uncertainty:

$$p(y^{t+i},\, \theta^{t+i}|\, y^t) = p(y^{t+i}|\, \theta^{t+i},\, y^t) \cdot p(\theta^{t+i}|\, \theta^t,\, y^t) \cdot p(\theta^t|\, y^t), \;\; i = 1, \ldots, h\,. \tag{9}$$

To make the simulation from the predictive density $p(y^{t+h},\, \theta^{t+h}|\, y^t)$ less time consuming, we assume that the coefficients in θ^{t+i} are fixed out of sample[11]. Conditional on each Gibbs sampler draw from $p(\theta^t|\, y^t)$, we simulate a value for β^{t+1} by drawing the innovations ν_{t+1} in 7, and for the innovations u_{t+1}, drawn from a Normal distribution with variance Ω_t. A path for $\hat{y}_{t+i}$, $i = 1 \ldots h$ is then generated, conditioning on $\hat{y}_{t+i-1}$ and on $u_{t+i} \sim N(0, \Omega_t)$. We repeat this procedure a thousand times, and store the mean and the 68% and 95% percentiles of the simulated values $\{\hat{y}_{t+i,\,\kappa},\, i = 1 \ldots h\}_{\kappa=1}^{1000}$. After the Gibbs sampler is completed, we take the average of these values across the sampler draws.

3.2 The time-varying parameter forgetting factor VAR

Using the stochastic volatility Bayesian VAR for a recursive forecasting exercise generally demands a high computational time, as the number of iterations required for the convergence of the Gibbs sampler is large. To reduce the computational time, Koop and Korobilis [2013] have developed a procedure which approximates the model in 4, by replacing the posterior draws of the covariance matrices Q and Σ_t with empirical estimates.

The assumptions of the TVP-SV-BVAR model imply that, given information up to $t-1$, the slope coefficients in t are draws from a normal distribution:

$$\beta_t \,|\, \mathcal{F}_{t-1} \sim N(\beta_{t|t-1},\, P_{t|t-1})\,. \tag{10}$$

The Kalman filter routine, used in the first step of the Gibbs sampler, entails a prediction for the coefficients' covariance matrix, $P_{t|t-1} = P_{t-1|t-1} + Q$, which involves the posterior draw of Q. To circumvent this problem, the following approximation is used:

$$P_{t|t-1} = \frac{1}{\lambda_t} P_{t-1|t-1}, \quad \text{with } \lambda_t \in (0,\, 1]\,, \tag{11}$$

implying a possibly time-varying value for Q:

$$Q_t = \left(\frac{1}{\lambda_t} - 1\right) P_{t-1|t-1}\,. \tag{12}$$

The parameter λ_t is a forgetting factor, and discounts past information. In particular, a value of λ equal to

[11] This simplification is justified only if the time variation of the coefficients is moderate and, as is the case in our empirical application. Results obtained by letting θ^{t+i} drift for $h \geq 1$ are indeed very similar and available upon request.

0.99 implies, in the case of monthly data, that observations one year ago receive 89% as much weight as current observations. The case of a constant Q, like in the TVP-SV-BVAR model, is encompassed by choosing a constant $\lambda_t = \bar{\lambda}$, $\forall t$. On the other hand, time variation in Q can be accommodated by positing a law of motion for the forgetting factor, as specified by Park et al. [1991]:

$$\lambda_t = \lambda_{min} + (1 - \lambda_{min}) L^{f(\hat{\varepsilon}'_{t-1}\hat{\varepsilon}_{t-1})} \,, \tag{13}$$

where $\lambda_{min} = 0.96$, $L = 1.1$, f denotes a function that rounds its input to the nearest integer, and $\hat{\varepsilon}_{t-1}$ is the one-step ahead prediction error.

A similar approximation is used for the covariance matrix of the non-structural innovations, Ω_t. The latter is estimated as a weighted average of its past value, and of its current estimate[12]:

$$\hat{\Omega}_t = \kappa \hat{\Omega}_{t-1} + (1 - \kappa)\hat{\varepsilon}'_t \hat{\varepsilon}_t \,, \tag{14}$$

where the weight is represented by the decay factor κ. To summarise, the procedure developed by Koop and Korobilis [2013] is based on the Kalman filter and relies on the parametrisation of equations 13 and 14, as well as on the choice of initial conditions for the covariance matrix Ω_0, for the slope coefficients β_0 and their variance P_0. Further details on the paramerisation are provided in the next Section.

There are three main differences between the TVP-SV-BVAR model and its approximation based on the use of forgetting factors. Firstly, the latter delivers filtered, rather than smoothed, estimates and should hence be better suited for a forecasting exercise but less suited for a full sample evaluation. Secondly, and more importantly, equations 13 and 14 in the forgetting factor model do not provide any rule for the out-of-sample evolution of the covariance matrices Q_t and Ω_t. Lastly, while the TVP-SV-BVAR embeds a structural decomposition of the innovations covariance matrix, its approximation deals solely with non-structural innovations. However, this should not be a concern, so long as the objective of the researcher lies in forecasting, rather than in a structural analysis.

Sampling from the predictive density As it has been already noted, neither equation 13 nor equation 14 are proper laws of motion for the two covariance matrices of the model. Hence, in order to generate samples from the predictive density, we follow Koop and Korobilis [2013] and assume that both covariance matrices are fixed out of sample. That is, we assume that $\hat{Q}_{t+i} = \cdots = \hat{Q}_{t+1} = \hat{P}_{t|t}$ and that $\hat{\Omega}_{t+i} = \cdots = \hat{\Omega}_t$, $\forall i \geq 1$. In a similar way, the out-of-sample path for the slope coefficients β_{t+h} is assumed to be fixed out of sample and centred around the last estimated values for $\hat{\beta}_{t|t}$ and for $\hat{P}_{t|t}$. Given these assumptions, we simulate 5000 values for the vector $\hat{y}^{t+h} = [\hat{y}'_{t+1}, \ldots, \hat{y}'_{t+h}]$, and store the mean and the 68% and 95% percentiles of the values $\{\hat{y}_{t+i,\,\kappa},\ i = 1 \ldots h\}_{\kappa=1}^{5000}$.

[12] This is the Exponentially Weighted Moving Average estimator, commonly used in the finance literature.

4 Modelling and forecasting exchange rates

In this Section the two methodologies previously discussed are applied to jointly model and forecast three main exchange rates vis-a-vis the US dollar: the British Pound, the Japanese Yen[13], and the German Mark[14]. The currencies, defined such that an increase pertains to a depreciation, cover the period from 1971:m1 to 2013:m6 and have been downloaded from Datastream. Figure 1 shows that the volatility of the three exchange rates has changed over the years. In particular, note that after the 2008 financial crisis the volatility of the Mark and of the Pound has increased after a period of relative moderation in the previous decade.

After a description of the empirical exercise, we briefly discuss the in-sample estimates of the time-varying parameter models. We then assess their out-of-sample forecasting performance relative to four benchmarks: a random walk with or without GARCH innovations, and a constant-parameter BVAR estimated either recursively or with a rolling estimation window of eleven years.

4.1 Description of the estimation and forecasting exercises

We transform the exchange rates by taking either the logarithm of the levels, or their percentage change, approximated through the differences in the log levels[15]. All models but the random walk are estimated using a lag length of 12, to capture any seasonal component that might be present in the data. If a training sample is used, its length is set to be of four years. Depending on whether the model is estimated on the log levels or on the percentage changes, a driftless random walk prior or a white-noise one is used for the slope coefficients.

The parameterisation of all TVP-FF-VAR models, estimated on the log levels unless stated otherwise, follows that in Koop and Korobilis [2013] for the choice of the forgetting factor ($\lambda_{min} = 0.96$) but uses a smaller decay factor for the estimation of the covariance matrix ($\kappa = 0.90$ instead of $\kappa = 0.96$), on the account of the data being monthly, and not quarterly, and exhibiting a greater volatility. Interestingly, though the covariance matrix of the VAR coefficients is allowed to be time varying, the estimation reveals that the matrix is actually constant, thus matching the assumption of the TVP-SV-BVAR model.

The TVP-SV-BVAR is estimated on the percentage change of the variables, to ensure the stability of the estimates. Crucial for stability is also the prior specification of Q, the covariance matrix of the slope coefficients. Following Cogley and Sargent [2005], we impose a loose prior that pertains to a time-invariant model: the prior scale matrix is set close to zero, while the number of degrees of freedom is one plus the dimension of the matrix, the lowest possible[16]. All remaining details of the prior distribution follow closely those in Primiceri [2005] and are summarised in table 2. The estimation of the model follows the procedure described in the theoretical section

[13] The Yen has been standardised by a factor of a 100 to avoid computational problems due to the scale of the variable.

[14] After the introduction of the Euro, the exchange rate for the Mark is obtained by using the conversion rate fixed in January 1999. For further details on the calculation of this currency see `http://www.bankofengland.co.uk/statistics/Pages/iadb/notesiadb/effective_exc.aspx`

[15] Both these transformations are common in the literature and are chosen so that the forecasts need not be restricted on a subset of positive values.

[16] As pointed out by Cogley and Sargent [2005] and verified empirically in the case of our data, large values of Q make explosive draws of the VAR coefficients more likely. Note that in Primiceri [2005] the prior scale matrix is larger and the prior tighter (the number of degrees of freedom is higher).

Table 1: **Exchange-rate prediction models**

	MODELS	VARIABLES	DESCRIPTION
1	TVP-FF-VAR	$£$, DM, Yen	
2	TVP-SV-BVAR	$£$, DM, Yen	
3	BVAR	$£$, DM, Yen	
4	BVAR (ROLLING)	$£$, DM, Yen	rolling estimation window of 11 years
5	RW	$£ \vee DM \vee Yen$	
6	RW-GARCH	$£ \vee DM \vee Yen$	$s_{i,t} - s_{i,t-1} \sim$ GARCH(1,1), $i = \{$UK, DE, JP$\}$

	ADDITIONAL TVP-FF-VAR MODELS	VARIABLES	
7	UIRP	s_i, $R_i - R_{us}$	
8	PPP	s_i, $\Delta(p_i - p_{us})$	$i = \{$UK, DE, JP$\}$
9	M	s_i, $\Delta(m_i - m_{us})$	
10	Y	s_i, $\Delta(y_i - y_{us})$	
11	SP	s_i, $\Delta(a_i - a_{us})$	
12	GBY	s_i, $B_i - B_{us}$	
13	VIX	s_i, VIX	
14	IBR	e_i, $L_i - L_{us}$	

Note: Small case letters denote the logarithm of the variables, while Δ indicates that the transformation chosen is the monthly growth rate. s stands for the nominal exchange rate, R for the short-term interest rate, p for the CPI index, m for money, y for total industrial production, a for stock prices, B for the 10-year government bond yield, VIX stands for the log of the CBOE volatility index, and L for the 3-month interbank lending rate. The subscript i indexes the country and refers either to the UK, to Germany, or to Japan. See the text and references therein for a description and motivation of the different models.

with one major modification: explosive draws of the VAR coefficients are rejected in the first step of the Gibbs sampler. This is equivalent to sampling from a restricted posterior density, with the restricted law of motion for the VAR coefficients being a truncated and renormalised version of the unrestricted one[17].

The benchmark forecast models, estimated on the log-levels of the data, are the random walk, with or without GARCH residuals, and the Bayesian VAR. Point forecasts from a random walk are obtained by setting $\hat{y}^{rw}_{i,t+h|t} = y_{it}$, $\forall h$, while density forecasts are retrieved from a random walk model, whose residuals follow a GARCH(1,1) process[18]. The Bayesian VAR is estimated on the log-levels of the variables, using a Minnesota prior[19].

A summary of the competing models is given in the first block of table 1. Accounting for the training sample and for the initial observations required by the lag choice, the first estimation sample starts in 1976:m1 and ends in 2000:m1. The estimation sample is then progressively enlarged in a pseudo-real time exercise: at each step, models are re-estimated and forecasts up to one-year ahead are computed[20]. Finally, to gauge how the financial crisis might have influenced the forecasting ability of the competing models, we split the forecast sample in two: a pre-crisis sample that ends in August 2008, such that the last forecasted period is always one month before the Lehman bankruptcy filing, and a crisis sample that starts in September 2008 and ends in June 2013[21].

[17] For additional details, see Cogley and Sargent [2005].

[18] A random walk with GARCH innovations is better suited than a simple random walk to deliver density forecasts. Density forecasts from a simple random walk are available upon request, and are the worst among all models considered, a result consistent with the findings of Balke et al. [2013].

[19] We have tried to set the prior specification of the BVAR models as close as possible the initialisation of the forgetting-factor models.

[20] In the case of the time-varying stochastic volatility model of Primiceri [2005], the model is re-estimated every year as opposed to every month, due to the computational time required by the posterior simulation algorithm.

[21] The second forecast sample is inevitably shorter due to data availability.

4.2 In-sample evaluation: does time variation matter?

We preliminary assess whether the modelling of time variation suits the dynamics of the data by analysing in figure 2 the in-sample estimates of the forgetting-factor VAR model, fitted on the log-levels of the data over the full sample 1976:m1-2013:m6[22]. The first three rows of figure 2 plot the sum of the estimated VAR coefficients, by exchange-rate equation (rows) and regressor (columns), together with the relevant 68% confidence intervals. These graphs provide a partial justification for both the joint modelling of the three currencies, and the assumption of coefficient time-variation. The off diagonal panels show in fact that most of the cross-currency coefficients are different from zero and display a change in pattern, though moderate, after 2000. The sum of the non-autoregressive coefficients entering the Pound equation (first row of figure 2) are always significant but have decreased in magnitude after 2000. By contrast, while the Mark (second row of figure 2) seems to be affected solely by the Pound and only for the period comprised between 2000 ad 2010, the Yen is affected by both currencies, but the coefficients of the Mark are significant only after 2000. Further evidence of time variation is found in the standard deviation of the innovations, plotted in the last row of figure 2. These panels disclose a fall in volatility in the 1990s, and an increase after the 2008 financial crisis.

Lastly, we compare the in-sample likelihood[23] (in log scale) of the competing models. Figure 3 shows that the likelihood of the TVP-FF-VAR is generally higher than that of its Bayesian time-varying parameter counterpart (upper panel), and also of the non-time varying BVAR (second panel), especially in the first half and last part of the sample. Among the two Bayesian models, it is instead the non time-varying one that has the highest in-sample likelihood, as it can be inferred by the negative difference in the third panel of figure 3.

These preliminary results suggest that the assumption of time variation in the parameters and in the volatility of the innovations is supported by the data, and that the best in-sample modelling strategy seems to be the forgetting-factor VAR. However, as the in-sample and out-of sample performances are not always related, we proceed by assessing to what extent the modelling of time variation improves the point, interval and density forecasts of the three target variables.

4.3 Out-of-sample evaluation

Point forecasts: Point forecasts are compared through their relative mean squared forecast error:

$$RMSFE_{i,h}^{a,b} = \frac{1}{T_f}\left(\frac{\sum_{t=1}^{T_f}(\hat{y}_{i,t+h|t}^{a} - y_{i,t+h})^2}{\sum_{t=1}^{T_f}(\hat{y}_{i,t+h|t}^{b} - y_{i,t+h})^2}\right), \tag{15}$$

where T_f is the number of forecasts, while i and h index, respectively, the variable and the horizon. Throughout this work the numerator refers to a time-varying parameter model, while the denominator pertains to a constant-

[22]The in-sample estimates of the TVP-FF-VAR model fitted on the log-differences of the data, as well as of the TVP-SV-BVAR are available in a not-for-publication appendix companion to this paper.

[23]As no analytical formula is available for the likelihood of the TVP-SV-BVAR model, we have approximated it using the harmonic mean estimator suggested by Newton and Raftery [1994]. Though this estimator is in principle sensitive to outliers, we have verified that outliers are not a concern in our estimation procedure by checking the post burn-in draws of the parameters.

parameter benchmark.

The first two blocks of tables 3 and 4 report the mean squared forecast errors of the TVP-SV-BVAR and of the core forgetting factor model containing only exchange rates, relative to those of a random walk (table 3), or to those of the constant-parameter BVAR (table 4). Values in bold denote the horizon and variable for which the two forecast errors being compared are significantly different from each other, according to a Diebold and Mariano test at a 5% significance level, modified using the small sample size correction of Harvey et al. [1998]. Both time-varying models deliver lower mean squared forecast errors than a random walk in the pre-crisis forecast sample, and loose accuracy in the forecast sample that includes the financial crisis. However, while the one-step ahead forecast errors of the TVP-SV-BVAR are significantly smaller than those of a random walk across all forecast subsamples, the forgetting-factor model can never significantly beat the benchmark. In fact, the performance of the forgetting-factor model worsens considerably after 2008 and, over the entire forecast sample, it beats the random walk only at a one-month ahead horizon, and never significantly. The comparison with the constant-parameter BVAR, reported in the first two blocks of table 4, is instead generally in favour of the time-varying parameter models, and in particular of the TVP-SV-BVAR. At a one-year ahead forecast horizon the differences are however almost never statistically significant, though the time-varying models do on average better in the pre-crisis sample, and worse in the forecast sample that includes the financial crisis.

An explanation for the worsened performance of the forgetting factor model after 2008 can be traced back to equations 13 and 14, which describe the evolution over time of the coefficients' covariance matrices. The two transition equations adjust only partially to the current and past Kalman filter prediction errors. As a result, several periods are needed to fully incorporate a change in the model such as the sudden change in drift occurred at the time of the financial crisis, see e.g. first row of figure 4. A similar argument is made by Clements and Hendry [1996], where they show that vector error-correction models forecast worse than simple vector autoregressions when long-run equilibrium relationships alter over the forecast period. Moreover, the one currency for which the TVP-FF-VAR forecasting performance did not worsen considerably after the financial crisis is the Yen, for which the change in drift arguably did not occur, or did so in a less pronounced manner.

In summary, the time-varying stochastic volatility BVAR delivers one-month ahead point forecasts which are more accurate than both the random walk and the constant-parameter BVAR. On the other hand, the forgetting-factor model delivers accurate forecasts in the pre-crisis sample, though it never beats the random walk significantly, and its forecasting performance worsens considerably after the financial crisis.

Interval forecasts: We proceed to examine whether allowing for parameter time variation improves the calibration of the 68% and 95% forecast confidence intervals, the two most commonly used in empirical studies. The statistic we use is the coverage rate of each competing model, measured as the percentage of times in which the actual exchange rate is contained in the forecast confidence interval. As it has been previously discussed, an accurate assessment of the uncertainty surrounding point forecasts is likely to be of interest to a wide variety of forex market participants, from central banks to private investors. A model that delivers coverage rates which are significantly below their nominal counterparts underestimates forecast uncertainty. To the other extreme, coverage rates of a 100% imply that the estimated forecast confidence intervals always contain the actual values, but the confidence bands are so wide to be of little practical use. A model with correctly calibrated forecast intervals would have coverage rates which do not significantly differ from their nominal counterparts.

The empirical coverage rates of the different models, corresponding to 68% and 95% nominal coverages, are reported in table 6. Values in bold have not been found to be statistically different from their nominal counterparts, according to a likelihood-ratio test with 1 degree of freedom[24]. In the pre-crisis sample, the forecast confidence intervals of the TVP-FF-VAR model are correctly calibrated at all forecast horizons and for both values of nominal coverage, with few exceptions. By contrast, the random walk with GARCH innovations delivers accurately calibrated confidence intervals only for the Pound and the Yen at a one-month horizon. At longer horizons the confidence intervals are very large and always contain the actual outcome, thus overestimating forecast uncertainty. A similar problem affects the forecasts from the TVP-SV-BVAR and of the BVAR estimated using a rolling window, which systematically delivers excessively large confidence intervals. By contrast, the BVAR estimated without a rolling window tends to underestimate forecast uncertainty at long horizons, delivering one-standard deviation confidence bands which are too narrow.

In the forecast subsample that includes the financial crisis, the coverage rates of the TVP-FF-VAR model remain correctly calibrated. Also the performance of the remaining models remains substantially unaltered, with two relevant exceptions. The coverage rates of the TVP-SV-BVAR model improve significantly for all currencies but the Yen, making this model the second best performing one in this subsample, after the forgetting-factor model. Also the BVAR (without a rolling estimation window) delivers correctly calibrated forecast confidence intervals at a three-month horizon and, in the case of the Pound, at longer horizons as well.

To exemplify the results of table 6 we refer to figure 4 where the 68% forecast confidence intervals of the forgetting-factor model are plotted together with those of the BVAR model (darker area) and with the actual exchange rates (in red), across forecast horizons (rows) and currencies (columns)[25]. At short and medium forecast horizon, the forgetting-factor model provides the narrowest confidence bands, which we know from table 6 to be accurately calibrated. Hence, while the time-varying parameter model gives an efficient, as well as correct, estimation of uncertainty, the constant-parameter model overestimates forecast uncertainty. The latter model proves inaccurate also at long horizons, where forecast uncertainty is actually underestimated. As an example, the last row of figure 4 shows that the appreciation of the Pound and of the Mark between 2002 and 2005 is underrated by the BVAR, but it is instead contained in the confidence interval of the forgetting-factor model. Similarly, the appreciation of the Yen in 2013 (lower right panel of figure 4) is underestimated by the BVAR model but forecasted by the forgetting-factor model.

[24] For details of the test we refer to Clements [2005].

[25] The graphs of the remaining model's forecast confidence intervals are available upon request.

Density forecasts evaluation using probability integral transforms: Since the seminal work of by Dawid [1984] and Diebold et al. [1998], probability integral transforms have been extensively used to evaluate competing density forecasts. In the univariate case, the probability integral transform (p.i.t.) is the cumulative density function corresponding to the forecast density $p(\cdot)$, evaluated at the actual value of the series:

$$z_t = \int_{-\infty}^{y_t} p(y_t) d(y) = P(y_t) \ . \tag{16}$$

Diebold et al. [1998] have shown that, if the forecast model $p(y)$ coincides with the data generating process $f(y)$, the series $\{z_t\}_{t=1}^{T_f}$ of probability integral transforms is an i.i.d. sample from a $U(0,1)$ distribution. These conclusions are easily extended to a multivariate setting, such as ours[26]. Note that the joint predictive density can be factored as follows:

$$p(y_{1,t},\, y_{2,t},\, y_{3,t}) = p(y_{1,t}|\, y_{2,t},\, y_{3,t}) \cdot p(y_{2,t}|\, y_{3,t}) \cdot p(y_{3,t}) \ . \tag{17}$$

For 3 variables, there are 3! possible factorisations, but we shall use the one above for simplicity. Denote the probability integral transforms of the two conditional densities and of the marginal one with: $z^c_{1|2,3,t}$, $z^c_{2|3,t}$, $z^m_{3,t}$. If the predictive density is correct, the three sequences will each be i.i.d U(0,1) and independent of each other. As a result, the test for i.i.d. uniformity can be conducted on the following $3T_f \times 1$ stacked vector, as proposed by Diebold et al. [1998]:

$$S = [z^c_{1|2,3,1}, \ldots, z^c_{1|2,3,T_f},\, z^c_{2|3,1}, \ldots,\, z^c_{2|3,T_f},\, z^m_{3,1}, \ldots, z^m_{3,T_f}] \ . \tag{18}$$

We start with a visual inspection: figure 5 plots the histogram of the probability integral transforms of the one-month ahead density forecasts delivered by the competing models. The hump-shaped histograms of the constant-parameter BVAR model[27], and to a lesser extent of the time-varying BVAR, reveal that these models overestimate the variance of the variables[28], confirming the earlier findings of the coverage rates. On the contrary, the histogram for the TVP-FF-VAR and RW-GARCH models are essentially uniform.

A formal test of uniformity is achieved through the Kolmogorov-Smirnov test (KS), whose p-values are reported in table 5. The only model for which the null of uniformity is never rejected at any horizon is the forgetting-factor one. For the remaining models the conclusions are mixed. The p.i.t. sequences of the RW-GARCH model are found to be uniformly distributed only at a one-month and one-year ahead forecast horizons. In addition, the null of uniformity is rejected only for the first horizon, in the case of the TVP-SV-BVAR model, or for the first three months, in the case of the constant-parameter BVAR[29].

[26] On this, see Diebold et al. [1998] and Clements [2005]

[27] A similar hump-shaped histogram is obtained for the p.i.t. sequence of the BVAR estimated using a rolling estimation window, available upon request.

[28] See Mitchell and Wallis [2011].

[29] The plots of the empirical distribution and autocorrelation functions of the one-step ahead p.i.t. sequences are available upon request. This analysis, based on Diebold et al. [1998], confirms the results discussed so far and further reveals that using constant-parameter models leads to a misspecification of all conditional moments, and in particular of the conditional variance.

Density forecasts comparison: As a last step of the out-of sample analysis, we broaden our attention to the entire predictive densities of the competing models, and in particular to the evolution over time of the log-predictive likelihoods, i.e. the log likelihood of observing the actual realisation of the variable, given a forecast model:

$$\log p_{j,\,h,\,t}(y_{1,\,t+h} \,|\, \mathcal{F}_{j,\,t-1})\ , \tag{19}$$

where $p_{j,\,h,\,t}(\cdot)$ denotes the predictive likelihood of model j at horizon h (possibly time-varying and thus depending on time t), y_1 is a vector of target variables (one or all of the three exchange rates), $\mathcal{F}_{t-1,\,j}$ is the information set of model j available at t. Of interest is the cumulative difference between the log-predictive likelihood of the core TVP-FF-VAR model, $\log p_{1,\,h,\,t}$, and that of one of the alternative benchmarks, $\log g_{j,\,h\,t}$:

$$\mathcal{S}_{j,h} = \sum_{t=1}^{T_f-h} \Big[\log p_{1,\,h,\,t}(y_{1,\,t+h} \,|\, \mathcal{F}_{1,\,t-1}) - \log g_{j,\,h\,t}(y_{1,\,t+h} | \, \mathcal{F}_{g_j,\,t-1}) \Big]\ , \tag{20}$$

where $g_{j,\,h\,t}(\cdot)$ denotes, in turn, the predictive likelihood of the time-varying stochastic volatility BVAR, of the constant-parameter BVAR (with and without a rolling estimation window), and of a random walk (with or without GARCH innovations). This exercise is similar to what is undertaken in Amisano and Geweke [2010] and Amisano and Geweke [2013], and enables us to gauge the contribution of different observations over time in favour or against the core TVP-FF-VAR model. Moreover, the statistic in 20 can be interpreted as the summed difference in density forecast errors[30] and can be justified in terms of the Kullback-Leibler distance (KLIC). The latter can be expressed as the expected difference between the true log predictive density, $f_t(\cdot)$, and the predictive density of model j, $p_{j,t}(\cdot)$:

$$E[\log f_t(y_{1,\,t+1} \,|\, \mathcal{F}_{j,\,t-1}) - \log p_{j,\,t}(y_{1,\,t+1} \,|\, \mathcal{F}_{j,\,t-1})]\ , \tag{21}$$

where we consider the case $h = 1$ and drop the horizon subscript for expositional purposes. Under some regularity conditions, the average of the sample quantities of f_t and $p_{j,t}$ yields a consistent estimator of the KLIC distance. Hence, when two different predictive densities are being compared, $p_{1,\,t}$ and $p_{2,\,t}$, the average difference between their logarithms is inherently related to their KLIC distance:

$$\frac{1}{T_f}\sum_{t=1}^{T_f}\big(\log p_{2,\,t}(\cdot) - \log p_{1,\,t}(\cdot)\big) = \frac{1}{T_f}\sum_{t=1}^{T_f}\big(\log f_t(\cdot) - \log p_{2,\,t}(\cdot)\big) - \left[\frac{1}{T_f}\sum_{t=1}^{T_f}\big(\log f_t(\cdot) - \log p_{2,\,t}(\cdot)\big)\right]\ , \tag{22}$$

so that, among a class of alternative models, choosing the one with the highest average log-predictive likelihood entails selecting the model with the minimal KLIC distance.

Figure 6 plots the statistic $\mathcal{S}$ in equation 20, for different benchmark models and selected forecast horizons[31], providing further insights on the relative predictive ability of the competing models. At a one-month ahead horizon

[30] See the discussion in Hall and Mitchell [2007].

[31] The statistic obtained when the benchmark model is a simple random walk is not shown, but it is available upon request. It is always lower than that obtained when the benchmark model is a random walk with GARCH innovations.

the forgetting-factor model does significantly better than the constant-parameter BVAR, but performs similarly to both the random walk with GARCH innovations and the TVP-SV-BVAR. At medium and long forecast horizons the constant-parameter BVAR improves its performance relative to the forgetting-factor model and, at a one-year ahead horizon, always performs better. Finally note that a sharp drop in the likelihood of the forgetting-factor model can be observed in the financial crisis period, across all comparisons. Nevertheless, as time passes and the model discounts the Kalman filter prediction errors through equations 13 and 14, the predictive likelihood relative to those of the competing model increases again.

From this discussion it can be concluded that the hardest benchmarks to beat, in terms of predictive likelihoods are the random-walk with GARCH at a one-month horizon, and the constant-parameter BVAR at longer horizons. These two benchmarks are used when comparing the marginal log-predictive likelihoods in figure 7. The one-month ahead predictive likelihood of the forgetting-factor model is higher than the RW-GARCH model for all three currencies. At a three-months horizon, the time-varying model beats the constant parameter BVAR only in the case of the Yen, and of the Pound until 2008. The BVAR is instead a more accurate forecast model for the Mark, and for all three currencies when longer forecast horizons are considered.

To gauge whether the differences observed in figures 6 and 7 are statistically significant, we use the general test of equal predictive ability proposed by Amisano and Giacomini [2007]. The test statistics are reported in table 8, where positive values in bold denote combinations currency-forecast horizons at which the TVP-FF-VAR model performs better than the model indicated in the row header. The test confirms that the forgetting-factor model yields a significantly lower density forecast error than the constant-parameter BVAR at a one-month horizon, but is then beaten by the latter model at longer horizons. As far as the comparison with the RW-GARCH model is concerned, the predictive likelihood of the forgetting factor model is significantly higher at medium and long horizons, but it is not statistically different at a one-month horizon, despite being higher on average.

In this Section we have explored whether allowing for parameter time variation and stochastic volatility improves the in-sample fit, as well as the point, interval and density forecasts of the three target exchange rates. The in-sample estimates of the forgetting-factor VAR reveal time variation, albeit modest, in both the innovations' variance and in the slope coefficients, particularly in the cross-country ones. Turning our attention to the out-of-sample forecasting performance, we have found that the point forecasts of the TVP-SV-BVAR are significantly more accurate than those of the constant-parameter benchmarks at a one-month ahead horizon. Moreover, the time-varying parameter models, and in particular the forgetting-factor one, allow for a correct estimation of forecast uncertainty through an accurate calibration of the forecast confidence intervals. In addition, the joint modelling of the three exchange rates yields gains in terms of predictive accuracy over a simple random walk. The forgetting-factor model delivers in fact the lowest density forecast errors at short and medium horizons, though it is beaten by the constant parameter BVAR at longer horizons. Lastly, we have also noted that the forecasting performance of the forgetting-factor model worsens over the financial crisis period, and that this might be attributed to the nature of its transition equations. As the forecast errors made during the financial crisis are increasingly discounted, the likelihood of the forgetting-factor model increases relative to that of the competing models.

5 Do macroeconomic and financial variables matter for exchange-rate forecasting?

5.1 Empirical methodology and description of the competing models

In this Section, we assess how the performance of the core forgetting-factor model, containing only the three exchange rates, varies when the set of regressors is enlarged with different macroeconomic and financial predictors. The choice of the forgetting factor methodology is motivated not only by its computational advantages over its Bayesian counterpart, but also by the fact that it delivers similar point forecasts to those of the TVP-SV-BVAR, while improving on its interval and density forecasts.

The full set of exchange-rate prediction models analysed is summarised in the second block of table 1, and has been chosen on the basis of the theoretical suggestions in Section 2. In models 7-10 we add to the core model one group of macroeconomic predictors at a time: differentials in nominal interest rates (UIRP), inflation (PPP), money growth (M), and output growth (Y). The second group of models (11-14) adds differentials in stock market prices (SP), long-term government bond yields (GBY), interbank lending rates (IBR), or simply a measure of financial market stress like the CBOE volatility index (VIX). With these last four models we seek to capture the effects of the recent financial market developments, including the 2008 financial crisis and the liquidity injections that followed the credit easing programs of central banks around the world. For a more detailed description of the variables used we refer to Appendix A.

The dataset spans the period between 1982:m7 and 2013:m6, though the interbank lending rates and the VIX volatility index respectively start in 1985 and 1990. All variables are taken as differentials with respect to their counterpart for the US economy, with the exception of the VIX volatility index which pertains solely to the US. We transform all fundamentals by taking their monthly percentage changes. Exceptions include the VIX index, taken in log levels, and the interest rates (short-term, government bond yields, and interbank lending rates), kept in levels.

All models are estimated with the forgetting-factor methodology. The slope coefficients are initialised such that the exchange rates are shrunk to a random walk, and the fundamentals to white-noise processes. The initialisation of the variance, as well as the factor parametrisation follow those of the previous Section.

5.2 In-sample performance of the competing models

Figure 8 shows that the correlations[32] between exchange rates and fundamentals have evolved over the sample period considered, and that a clear change in pattern is particularly evident after the financial crisis. The correlation

[32]Standard errors are not available, due to the use of the EWMA estimator for the variance.

between exchange rates and interest rates is on average positive (with the exception of the Pound) but it is mostly negative after the financial crisis, when an increase in the interest rate with respect to the US is associated with a currency depreciation (an increase in the exchange rate). Hence, it is only in the years after the financial crisis (and for the Pound also between 1995 and 2000) that we observe the "forward bias puzzle" commonly found in empirical literature, i.e. the result that high-interest rate currencies tend to depreciate, rather than to appreciate as the UIRP theory would suggest[33]. Another interesting result is the estimated correlation between exchange rates and inflation differentials: higher inflation differentials are in fact associated with a currency depreciation (positive correlation in the second row of figure 8), as suggested by the purchasing power parity theory.

The estimated values of the reduced form VAR coefficients provide additional evidence of cross-country interdependencies and parameter time variation, though both are modest in entity and variable across currencies. Such time variation, mostly unaccounted for in empirical works, may be one of the causes of the apparent low predictive ability of macroeconomic fundamentals. A particularly important fundamental seems to be the differential in long-term government bond yields, as it can be inferred from the dynamics of the coefficients in figure 9.

5.3 Out-of-sample performance of the competing models

Point and interval forecasts: The mean squared forecast errors of the competing forgetting-factor VARs, relative to a random walk or to a constant-parameter BVAR, are reported in tables 3 and 4. Of all the models enriched with additional predictors, only the one that includes money growth differentials delivers lower mean-squared forecast errors for all currencies at a 6-month and 1-year ahead horizons, in the forecast sample that excludes the financial crisis. This difference is however significant only for 1-year ahead forecasts, and solely for the Pound. The forecast errors for the Yen delivered by the money growth model, as well as by the GBY and SP ones, are smaller than those of a random walk also in the forecast sample that includes the financial crisis, though the difference is never significantly significant.

Overall the inclusion of additional predictors to the core TVP-FF-VAR does not seem to improve on the point forecasts of a naive random walk forecast model. Similarly, controlling for additional predictors does not improve on the interval forecasts of the core forgetting-factor model: the forecast confidence intervals are correctly calibrated, as they are in the core model containing only exchange rates [34].

Density forecasts: As a last step of the out-of sample analysis, we turn once again to the entire predictive density and compare the cumulated differences in log-predictive likelihoods between the forgetting-factor models so far considered and the best-preforming benchmarks: the RW-GARCH at a one-month horizon, and the BVAR with constant parameters at longer horizons.

As far as the forecasts of the Pound and of the Mark are concerned, the fundamentals-based models generally perform worse than the core TVP-FF-VAR, and always worse than the BVAR at horizons greater than one month

[33] See, for instance, the discussion in Della Corte et al. [2009]
[34] See table 7.

(Mark) or three months (Pound).

By contrast, macroeconomic fundamentals perform better at forecasting the future dynamics of the Yen. As the upper panels of figure 11 show, inflation differentials forecast better at short horizons, and their relative importance has increased in the years preceding the financial crisis. Differentials in money growth and in government bond yields instead forecast better at medium and long horizons, with a further increase in their relative forecasting ability after the financial crisis, when they beat also the constant-parameter BVAR. Other important predictors are the differentials in short term interest rates and in industrial production. As shown in figure 12, the addition of these fundamentals improves the performance of the core forgetting-factor model at medium and long horizons, especially in the period between 2009 and 2011, though the predictive likelihoods are always lower than that of the constant-parameter BVAR.

Though a statistical test on the differences in log-predictive likelihoods generally favours the constant-parameter BVAR model[35], a suggestive pattern emerges when comparing the one-year ahead predictive likelihoods of the fundamentals-based models with those of the BVAR and of the core forgetting-factor model. As figure 13 shows, models enriched with macroeconomic or financial fundamentals tend to forecast better in recession periods[36]. This is especially true for the Pound in the 2008 recession, when the best performing models are those enriched with inflation and government bond yield differentials, for the Mark in the recessions between 2001 and 2005 and between 2011 and 2013, and for the Yen in the financial crisis years as well as in the current recession. These results provide evidence in favour of the hypothesis that in times of economic crises the expectation, and ultimately the determination, of exchange rates by forex market participants tend to be based on macroeconomic and financial fundamentals.

6 Economic evaluation: a simple trading strategy

So far, we have relied on purely statistical criteria to evaluate exchange-rate forecasts from competing models. However, an evaluation based on economic criteria might be of interest, particularly if the statistical models are to be used in real-world applications. In what follows, we asses the performance of the competing models through a simple trading strategy, described in Carriero et al. [2009]. We take the perspective of a US-based investor who bases her investment decisions on the predictions of a given forecast model, and has an investment horizon of one month[37]. The investor buys foreign currency only if she expects the latter to appreciate over the period of interest; no investment is made if the currency is instead expected to depreciate. At the end of the investment period, the investor liquidates the realised gain/loss (if the currency actually appreciated/depreciated) and reinvests the initial capital. We consider trading strategies based on the FF-VAR with only exchange rates, the FF-VARs augmented with fundamentals, the constant-parameter BVAR, and on a naive strategy that in each time attributes a 50% probability to a currency appreciation over the investment period. Trading strategies are evaluated on the basis of

[35] See table 9 for details.
[36] Recession periods are country specific and have been dated using the OECD series available on the FRED website.
[37] Results of the trading strategy with a longer investment horizon are available upon request.

their average return $\mu(\pi)$, on the returns' standard deviation $\sigma(\pi)$, as well as on the Sharpe ratio (SR)[38], over both the full forecast sample and the pre-crisis and crisis subsamples. The three statistics, together with the difference in the Sharpe ratio over the naive strategy (ΔSR) are reported in table 10.

As a preliminary assessment, note that across all forecast samples and currencies a trading strategy based on the core time-varying forgetting-factor model offers higher returns than both the naive strategy and the one based on the constant-parameter BVAR[39]. The profits of these three models are shown in figure 14, highlighting how the modelling of parameter time-variation proves useful at a one-month ahead horizon even when economic criteria, instead of statistical ones are used.

Next we assess whether controlling for macroeconomic and financial predictors would have offered any economic gain to an hypothetical investor. In the case of the Pound (first block of table 10), the only strategy that offers a positive return in the pre-crisis sample, despite having also the highest volatility, is the one based on the core forgetting-factor model. Interestingly, in the financial crisis sample almost all forgetting-factor models enriched with financial predictors (stock prices, government bond yields, and VIX index) offer a positive return. Of all these models however, only the one that includes the VIX indicator has a higher Sharpe ratio than the core forgetting-factor model. Similarly to what observed in the case of the Pound, a trading strategy for the Mark based on the core forgetting-factor model is the best performing one in the pre-crisis sample (middle block of table 10). If we instead limit our attention to the financial crisis sample, the strategies based on the models that incorporate macroeconomic and financial fundamentals offer higher returns and, in some cases, also a lower volatility. In particular, the Sharpe ratio of the strategy based on the output model is approximately 10 times higher than that based on the core model. Also the interbank lending rate model beats the core model in terms of Sharpe ratios, though its returns are much lower than those of the output model. Finally, the best models to form a strategy for the Yen (last block of table 10) are the M model in the pre-crisis sample, and the PPP in the crisis one, even though they are both beaten by the simple core model when the whole forecast sample is considered.

The results in this Section have highlighted two main points. First, a trading strategy based on the core forgetting-factor model with time-varying parameters yields a higher mean return, as well as a higher Sharpe ratio, than both a naive trading strategy and one based on a constant-parameter BVAR. Moreover, as it can be evinced from figure 15, trading strategies based on forgetting-factor models enriched with monetary and financial fundamentals would have offered higher returns to investors in the years between 2008 and 2010, one in which financial markets were characterised by high stress and uncertainty. This result is particularly evident in the case of the Mark, the currency for which the statistical criteria evaluated in the previous Section did not reveal any role for macroeconomic and financial fundamentals.

[38]The Sharpe ratio is defined as the ratio between the mean return and its standard deviation, and it is an effective way to summarise the mean-variance trade-off of a given investment strategy.

[39]The results of a trading strategy based on a BVAR estimated using a rolling window are very similar to those of the BVAR-based strategy, and are therefore not shown.

7 Conclusions

A big puzzle in the foreign exchange literature is the inability of macro and financial variables to predict the future behaviour of exchange rates. A few suggestions have been proposed in the literature to address this issue, and in this paper we have concentrated on the idea that the determinants of the exchange rate evolution can be time-varying. We have focused on the three major exchange rates vis-a-vis the US dollar and have assessed whether there are gains from a joint modelling of the three rates, as well as from the assumption of parameter time-variation. Next, we have evaluated whether adding a set of macro and financial variables to the model yields additional gains. Terms of comparison are the in-sample fit, the point, interval and density forecasts, as well as the results of a simple trading strategy. We have used two state-of-the-art time-varying parameter models: the stochastic volatility BVAR of Cogley and Sargent [2005] and Primiceri [2005], and its forgetting factor approximation recently proposed by Koop and Korobilis [2013]. These models have been compared with several benchmarks: the random walk, with or without GARCH innovations, and a constant-parameter BVAR.

Overall, the in-sample analysis is in favour the joint modelling of the three currencies and provides evidence of time variation, both in the VAR parameters and in the innovations' covariance matrix. Though time variation is modest and variable across currencies and models, this finding supports the conjecture by which the low predictive ability of fundamentals is driven by the instability of the relationship that links them to exchange rates. In particular, parameter time-variation is exhibited by the exchange-rate coefficients (both the autoregressive and the cross-country ones), and by the the reduced-form coefficients of the differentials in government bond yields.

The out-of-sample comparison of the competing models has revealed that accounting for parameter time variation, though improving the point forecasts of the target variables only at one-month ahead horizon, significantly refines the estimation of forecast uncertainty through an accurate calibration of the forecast confidence intervals. The analysis of the forecast probability integral transforms has further conveyed the result that it is the entire forecast density of the three exchange rates to be correctly calibrated, and not just the 68% and 95% confidence intervals.

A comparison based on log-predictive likelihoods has revealed that the forgetting-factor model yields the highest predictive densities at short and medium horizons, but a constant parameter BVAR beats the time-varying model at longer horizons. Thus, though on one hand we confirm that the joint modelling of exchange rates yields gains in terms of predictive accuracy over a simple random walk, the modelling of time-variation is best suited only at short horizons. We have also remarked that the error-correction nature of the forgetting-factor model might be the cause of its worsened forecasting performance over the financial crisis period. In fact, as the forecast errors made during the financial crisis are incorporated, the likelihood of the forgetting-factor model increases relative to that of the competing models.

We have then employed the TVP-FF-VAR, proven to deliver the best calibrated density forecasts, to gauge whether exchange rate predictability by fundamentals can be unravelled through the modelling of time variation. We have considered a wide set of macroeconomic predictors, as suggested by economic literature, as well as a number of financial variables and measures of market risk and liquidity, in an attempt to capture the effects of the recent financial crisis. Though not substantially improving point forecasts nor the calibration of forecast confidence intervals, we have found that controlling for macroeconomic and financial predictors can deliver predictive likelihood gains in times of economic recessions.

Modelling parameter time variation proves useful at short horizons even when economic evaluation criteria, instead of statistical ones, are employed. The results of a simple trading strategy have shown that the core time-varying forgetting-factor model offers higher returns than both a BVAR-based strategy and a naive rule that predicts a trading opportunity with 50% chance. Moreover, we have found that a strategy that controlled for monetary and financial fundamentals would have offered higher returns to investors in the turbulent years between 2008 and 2010.

Bibliography

Gianni Amisano and John Geweke. Comparing and evaluating bayesian predictive distributions of asset returns. *International Journal of Forecasting*, 26(2):216–230, April 2010.

Gianni Amisano and John Geweke. Prediction using several macroeconomic models. Working Paper Series 1537, European Central Bank, April 2013.

Gianni Amisano and Raffaella Giacomini. Comparing density forecasts via weighted likelihood ratio tests. *Journal of Business & Economic Statistics*, 25:177–190, April 2007.

Philippe Bacchetta and Eric van Wincoop. On the unstable relationship between exchange rates and macroeconomic fundamentals. NBER Working Papers 15008, National Bureau of Economic Research, Inc, May 2009.

Nathan S. Balke, Jun Ma, and Mark E. Wohar. The contribution of economic fundamentals to movements in exchange rates. *Journal of International Economics*, 90(1):1–16, 2013.

Fabio Canova. Modelling and forecasting exchange rates with a bayesian time-varying coefficient model. *Journal of Economic Dynamics and Control*, 17(1-2):233–261, 1993.

A. Carriero, G. Kapetanios, and M. Marcellino. Forecasting exchange rates with a large bayesian var. *International Journal of Forecasting*, 25(2):400–417, 2009.

C. K. Carter and R. Kohn. On gibbs sampling for state space models. *Biometrika*, 81:541–553, 1994.

G. Cassel. Abnormal deviations in international exchanges. *Economic Journal*, 28:413–415, 1918.

Yin-Wong Cheung and Menzie David Chinn. Currency traders and exchange rate dynamics: a survey of the us market. *Journal of International Money and Finance*, 20(4):439–471, August 2001.

Yin-Wong Cheung, Menzie D. Chinn, and Antonio Garcia Pascual. Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24(7):1150–1175, November 2005.

Michael P. Clements. *Evaluating Econometric Forecasts of Economic and Financial Variables*. Palgrave Macmillan, Basingstoke, Hampshire, 2005.

Michael P Clements and David F Hendry. Intercept corrections and structural change. *Journal of Applied Econometrics*, 11(5):475–94, Sept.-Oct 1996.

Timothy Cogley and Thomas J. Sargent. Drift and volatilities: Monetary policies and outcomes in the post wwii u.s. *Review of Economic Dynamics*, 8(2):262–302, April 2005.

Marcos Dal Bianco, Maximo Camacho, and Gabriel Perez Quiros. Short-run forecasting of the euro-dollar exchange rate with economic fundamentals. *Journal of International Money and Finance*, 31(2):377–396, 2012.

A Philip Dawid. Present position and potential developments: Some personal views: Statistical theory: The prequential approach. *Journal of the Royal Statistical Society. Series A (General)*, pages 278–292, 1984.

Marco Del Negro and Giorgio Primiceri. Time-varying structural vector autoregressions and monetary policy: a corrigendum. Staff Reports 619, Federal Reserve Bank of New York, 2013.

P. Della Corte and Tsiakas. Statistical and economic methods for evaluating exchange rate predictability. In L. Sarno James, J. and I.W. Marsh, editors, *Handbook of Exchange Rates*. Wiley, 2013.

Pasquale Della Corte, Lucio Sarno, and Ilias Tsiakas. An economic evaluation of empirical exchange rate models. *Review of Financial Studies*, 22(9):3491–3530, September 2009.

Francis X. Diebold, Todd A. Gunther, and Anthony S. Tay. Evaluating density forecasts with applications to financial risk management. *International Economic Review*, 39(4):863–883, Nov 1998.

Rudiger Dornbusch. Expectations and exchange rate dynamics. *Journal of Political Economy*, 84(6):1161–76, December 1976.

Charles Engel and Kenneth D. West. Exchange rates and fundamentals. *Journal of Political Economy*, 113(3): 485–517, June 2005.

Charles Engel, Nelson C. Mark, and Kenneth D. West. Exchange rate models are not as bad as you think. In *NBER Macroeconomics Annual 2007, Volume 22*, NBER Chapters, pages 381–441. National Bureau of Economic Research, Inc, October 2008.

Jeffrey A. Frankel. Tests of monetary and portfolio balance models of exchange rate determination. In *Exchange Rate Theory and Practice*, NBER Chapters, pages 239–260. National Bureau of Economic Research, Inc, November 1984.

Stephen G. Hall and James Mitchell. Combining density forecasts. *International Journal of Forecasting*, 23(1): 1–13, 2007.

David I Harvey, Stephen J Leybourne, and Paul Newbold. Tests for Forecast Encompassing. *Journal of Business & Economic Statistics*, 16(2):254–59, April 1998.

Atsushi Inoue and Barbara Rossi. Out-of-sample forecast tests robust to the choice of window size. *Journal of Business & Economic Statistics*, 30(3):432–453, April 2012.

Sangjoon Kim, Neil Shephard, and Siddhartha Chib. Stochastic volatility: Likelihood inference and comparison with arch models. *Review of Economic Studies*, 65(3):361–93, July 1998.

Gary Koop and Dimitris Korobilis. Large time-varying parameter vars. *Journal of Econometrics*, 177:185–198, February 2013.

Nelson C. Mark and Donggyu Sul. Nominal exchange rates and monetary fundamentals: Evidence from a small post-bretton woods panel. *Journal of International Economics*, 53(1):29–52, February 2001. URL `http://ideas.repec.org/a/eee/inecon/v53y2001i1p29-52.html`.

Richard Meese and Kenneth Rogoff. The out-of-sample failure of empirical exchange rate models: Sampling error or misspecification? In *Exchange Rates and International Macroeconomics*, NBER Chapters, pages 67–112. National Bureau of Economic Research, Inc, May 1983.

James Mitchell and Kenneth F. Wallis. Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness. *Journal of Applied Econometrics*, 26(6):1023–1040, 2011. ISSN 1099-1255.

Tanya Molodtsova and David Papell. Taylor rule exchange rate forecasting during the financial crisis. Working Paper 18330, National Bureau of Economic Research, August 2012.

Tanya Molodtsova and David H. Papell. Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of International Economics*, 77(2):167 – 180, 2009.

Haaron Mumtaz and Laura Sunder-Plassmann. Time-varying dynamics of the real exchange rate: an empirical analysis. *Journal of applied econometrics*, 28:498–525, 2013.

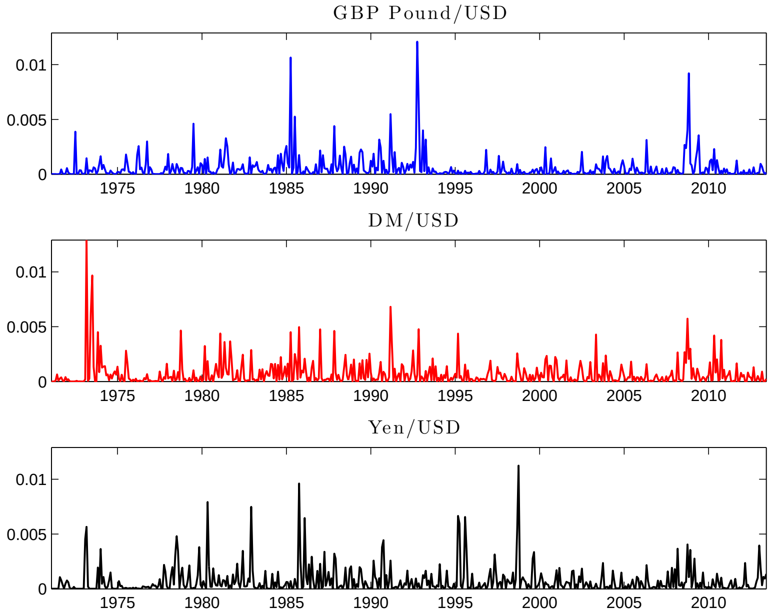
M. A Newton and A.E. Raftery. Approximate bayesian inference with the weighted likelihood bootstrap. *Journal of the Royal Statistical Society. Series B*, 56(1):3–48, 1994.

D. Park, B. Jun, and J. Kim. Fast tracking rls algorithm using novel variable forgetting factor with unity zone. *Electronics letters*, 27:2150–2151, 1991.

Giorgio E. Primiceri. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852, 2005.

Additional tables and figures

Figure 1: **Stylized facts**: Exchange rates volatility over the years.



Note: Volatility is measured as the square monthly percentage change in the exchange rate.

Kenneth S. Rogoff and Vania Stavrakeva. The continuing puzzle of short horizon exchange rate forecasting. NBER Working Papers 14071, National Bureau of Economic Research, Inc, June 2008.

Barbara Rossi. Are exchange rates really random walks? some evidence robust to parameter instability. *Macroeconomic Dynamics*, 10(01):20–38, February 2006.

Barbara Rossi. Exchange rate predictability. Working Papers 690, Barcelona Graduate School of Economics, February 2013.

Hyun Song Shin, Erkko Etula, and Tobias Adrian. Risk appetite and exchange rates. 2010 Meeting Papers 311, Society for Economic Dynamics, 2010.

Hong Yongmiao, Li Haitao, and Zhao Feng. Can the random walk be beaten in out-of-sample density forecasts? evidence from intraday foreign exchange rates. *Journal of Econometrics*, 141(2):736–776, December 2007.

Figure 2: **Reduced-form coefficients**: TVP-FF-VAR model with only exchange rates, in log levels. The first three panels plot the sum of the slope coefficients by each regressor (column) and exchange-rate equation (row), together with the pertinent 68% confidence band. The last row plots the standard deviations of the innovations to each exchange rate; confidence bands not available.

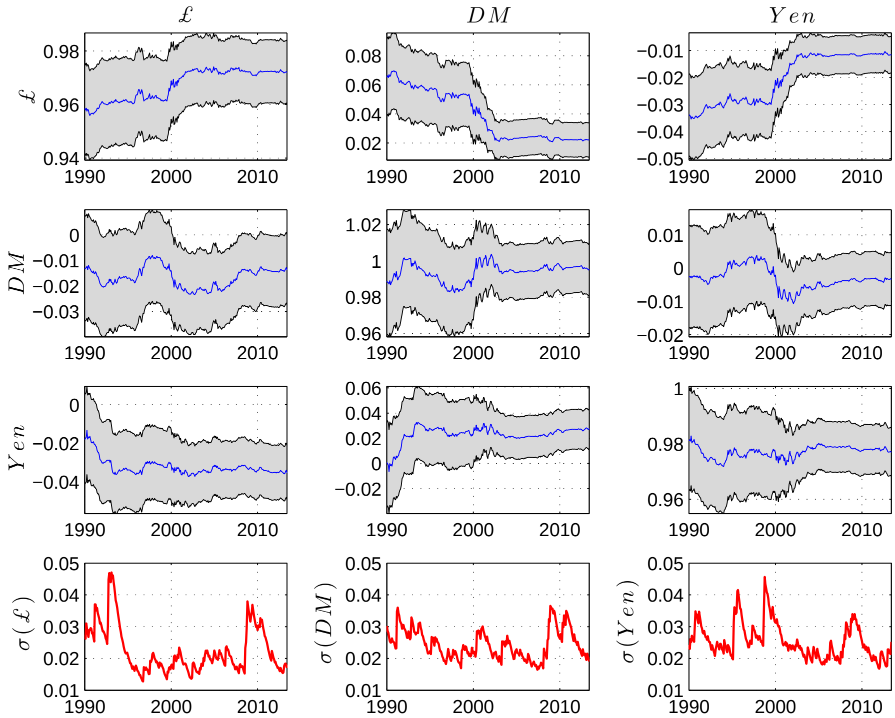
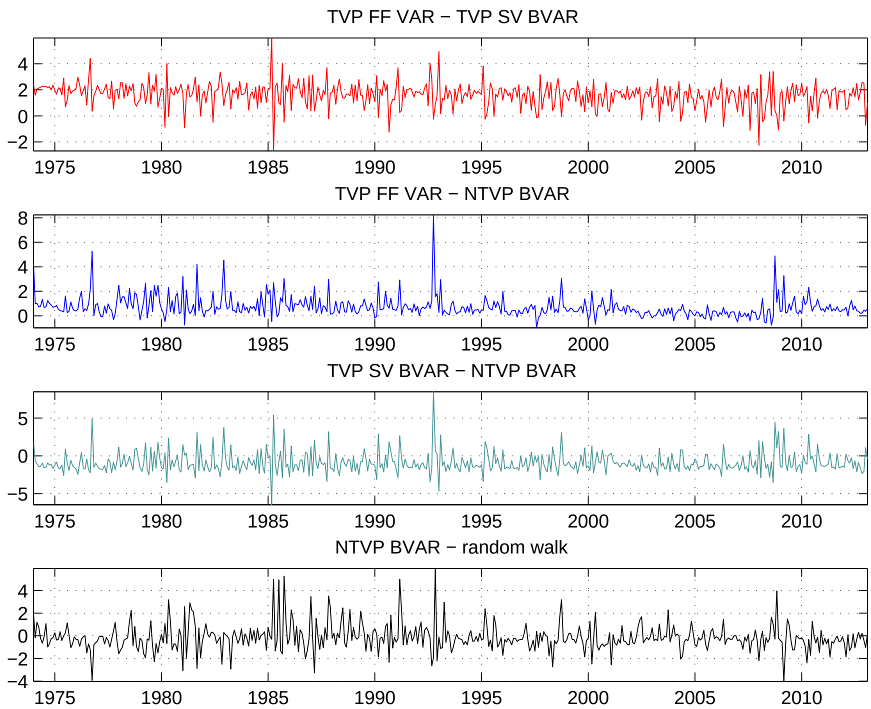


Figure 3: **Differences in the in-sample likelihood**: Each panel displays the difference between the in-sample likelihoods (in log scale) of the two models in the title.



Note: A positive difference in figure 3 indicates that the first of the two models in the title has a higher log likelihood. The likelihood of the TVP-SV-BVAR model has been approximated by using the harmonic mean estimator. The models in the first three panels have all been estimated on the log differences of the data, to allow comparison with the TVP-SV-BVAR. In the last panel, the BVAR is estimated on the log levels of the variables, to compare with a random walk.

Figure 4: **Forecast confidence intervals**: 68% forecast confidence intervals delivered by the TVP-FF-VAR, and by the constant-parameter BVAR (darker area), by exchange rate (columns) and horizon (rows). Actual exchange rates in red.

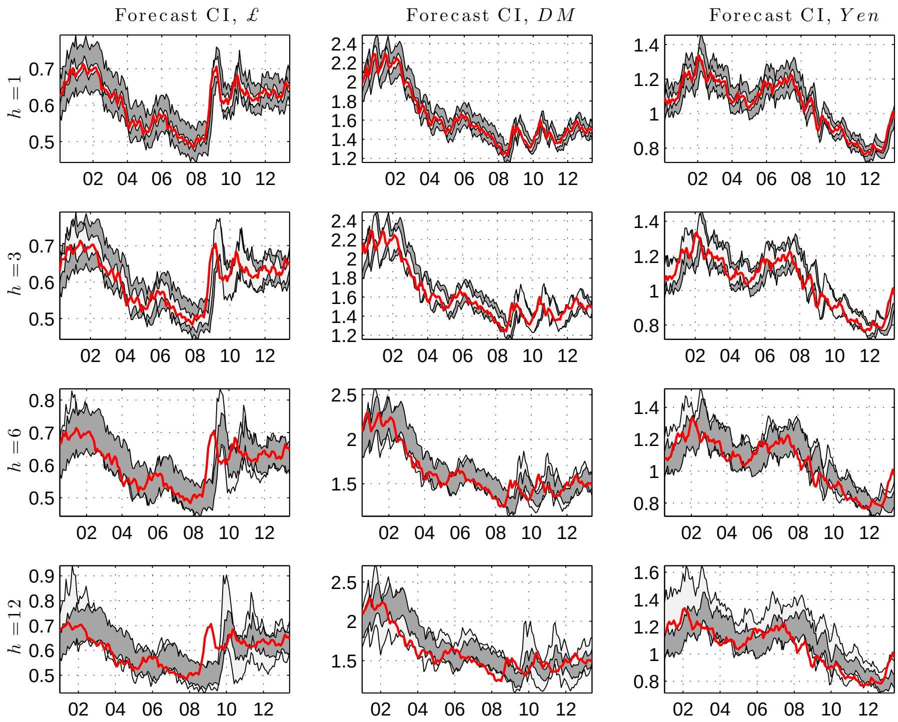


Figure 5: **Evaluating density forecasts**: Normalised histograms of the one-step ahead p.i.t. sequences, of 4 competing models.

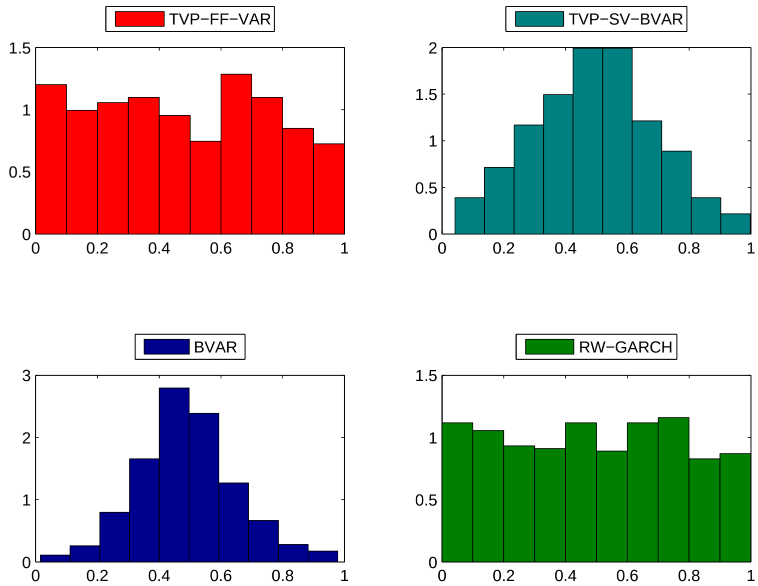


Figure 6: **Comparing log-predictive likelihoods**: Cumulative differences in log-predictive likelihoods between the core TVP-FF-VAR and various benchmarks: TVP-SV-BVAR (solid line), RW-GARCH (dashed-dotted line), BVAR with and without a rolling estimation window (respectively dashed and circled line), at selected forecast horizons.

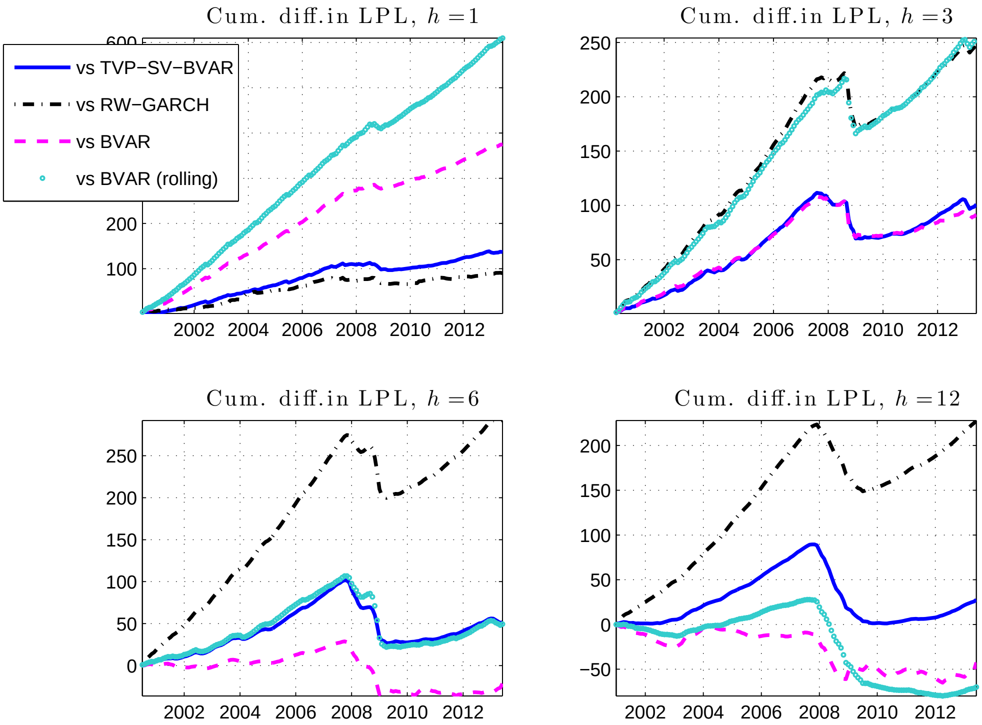
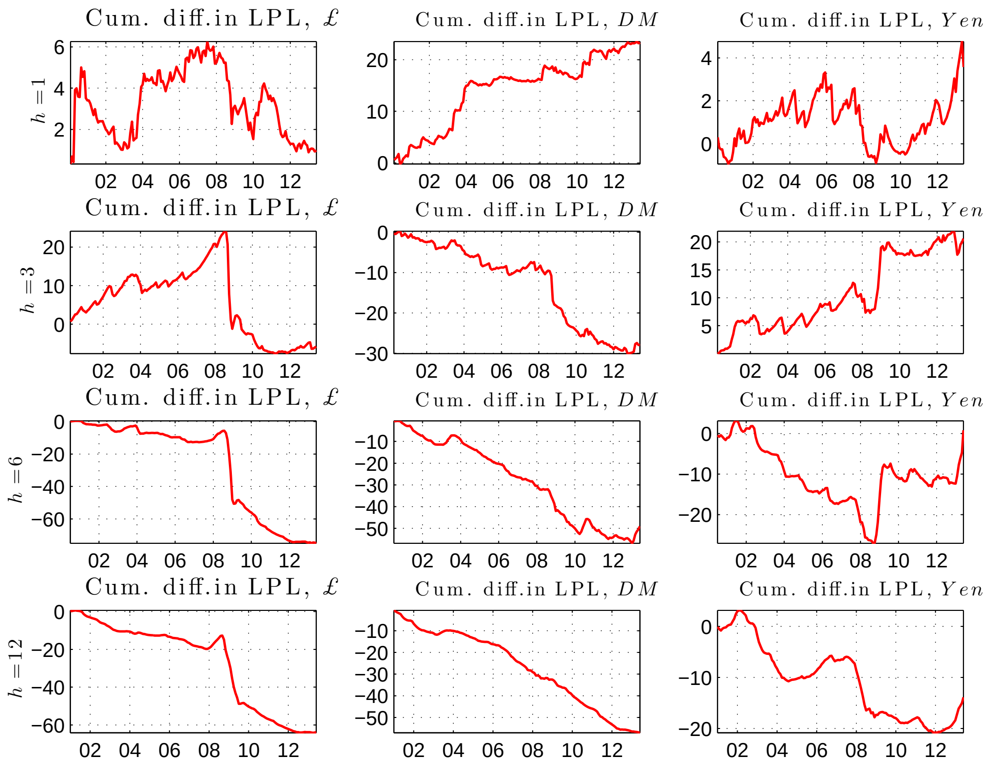
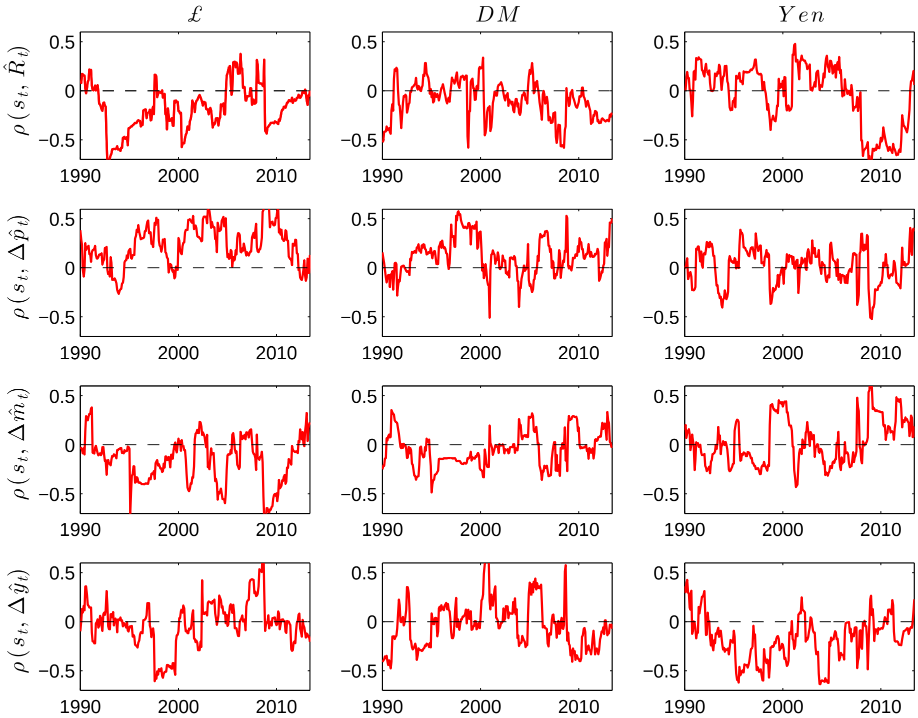


Figure 7: **Comparing log-predictive likelihoods**: Cumulative differences in log-predictive likelihoods between the core TVP-FF-VAR and the RW-GARCH (for $h = 1$) or the BVAR model (for longer horizons), by exchange rate (columns) and selected forecast horizons (rows).



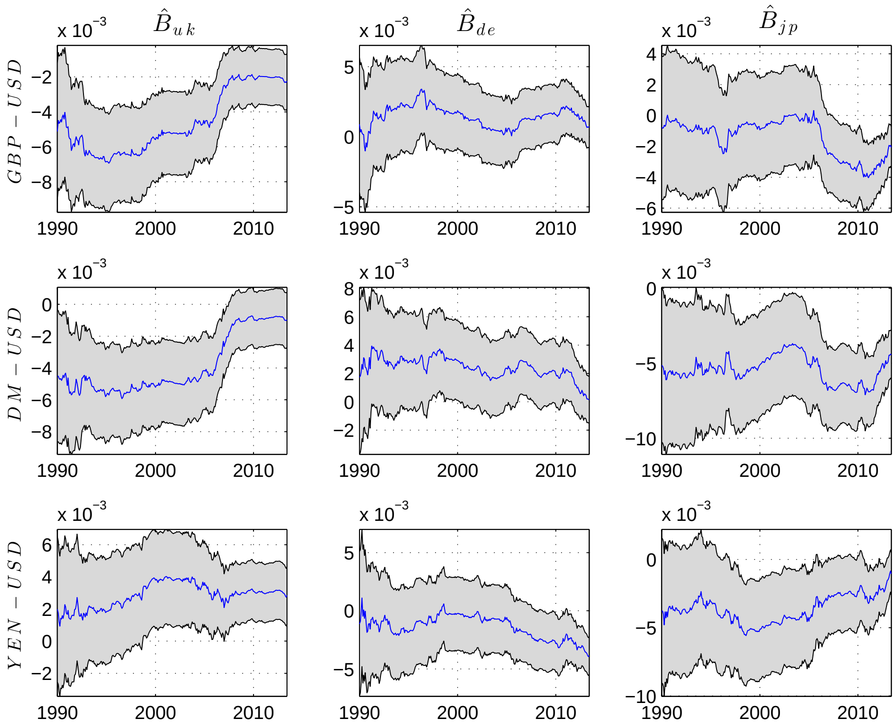
Note: Positive values of the plotted statistics indicate whenever the core TVP-FF-VAR has a higher predictive likelihood than the benchmark. Increases in the statistics denote dates in which the TVP-FF-VAR performs better than the benchmark.

Figure 8: **Time-varying correlations**: Correlations over time between exchange-rate and fundamentals, as estimated by the different TVP-FF-VAR models.



Note: The first three rows display the correlation between the exchange rate of each country (indicated in the column title) and its pertinent macroeconomic differential: nominal interest rate (first row, UIRP model estimates), money growth (second row, M model estimates), and industrial production growth (third row, Y model estimates).

Figure 9: **Reduced-form coefficients:** by exchange-rate equation (row); GBY model estimates.



Note: Shaded areas denote a 68% confidence interval for the sum of the differentials coefficients (indicated in the column title) entering each exchange-rate equation (rows). Off-diagonal elements mark the interdependencies between the differentials and the currency of different countries.

Figure 10: **Log-predictive densities for the Pound**: Cumulative differences in log-predictive likelihoods between competing TVP-FF-VAR models and the best performing benchmark, at selected forecast horizons.

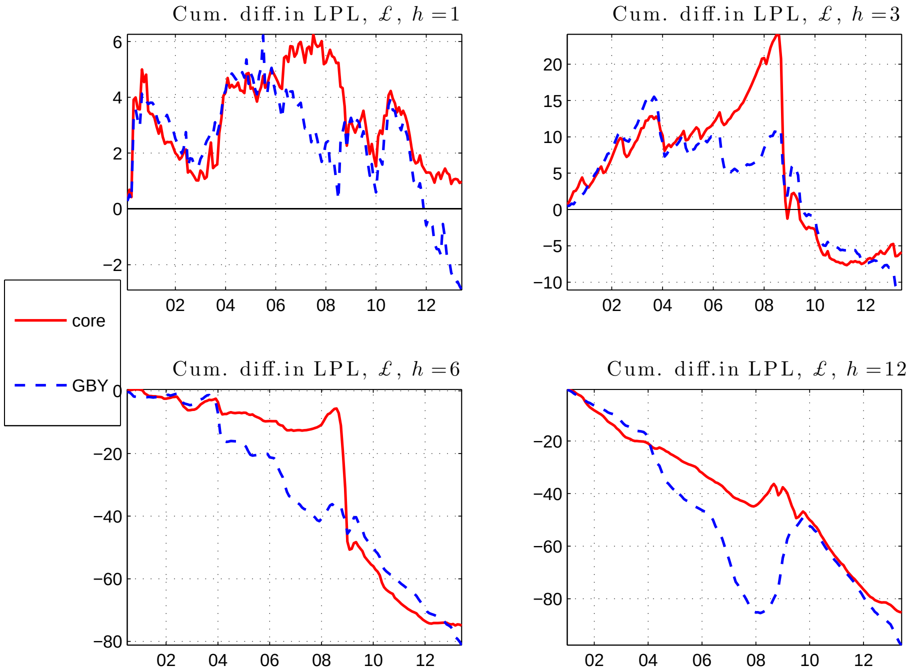
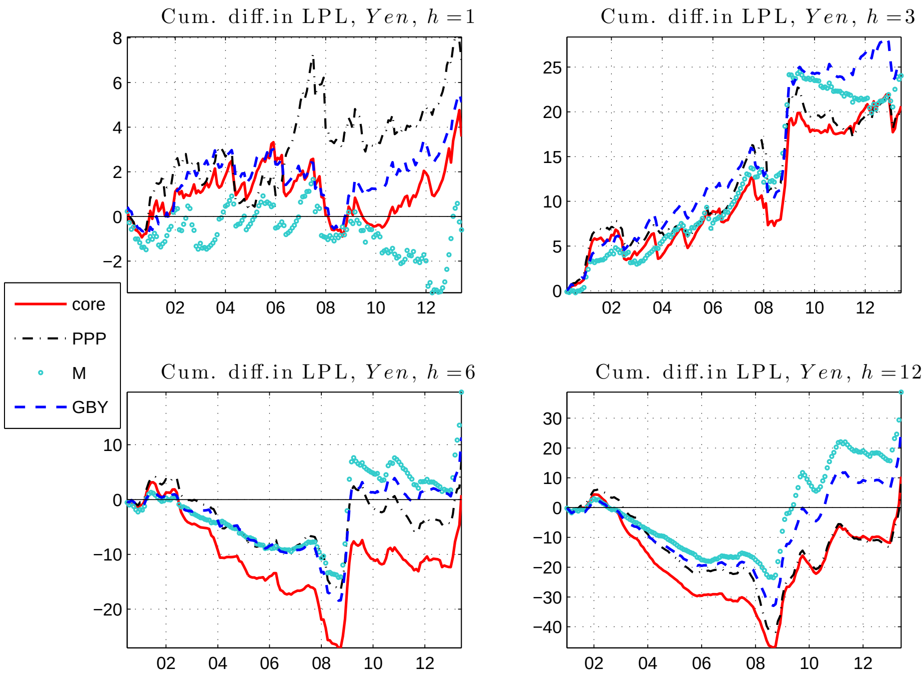
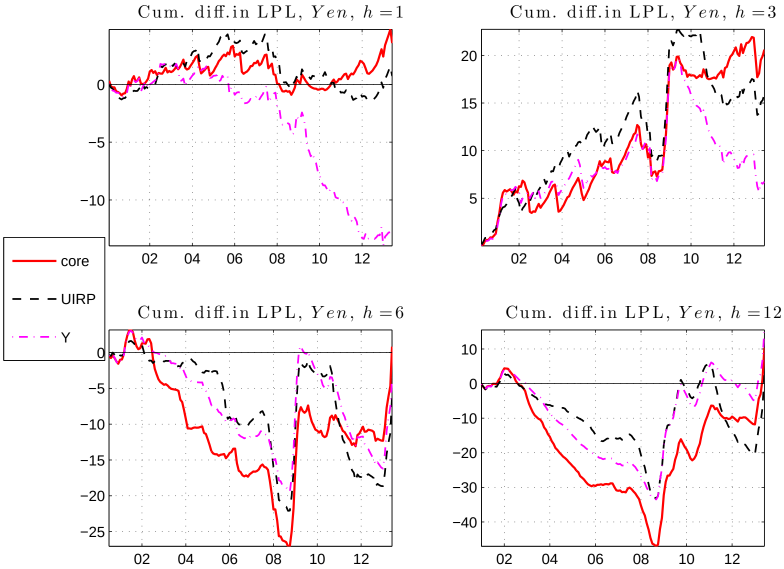


Figure 11: **Log-predictive densities for the Yen**: Cumulative differences in log-predictive likelihoods between competing TVP-FF-VAR models and the best performing benchmark, at selected forecast horizons.



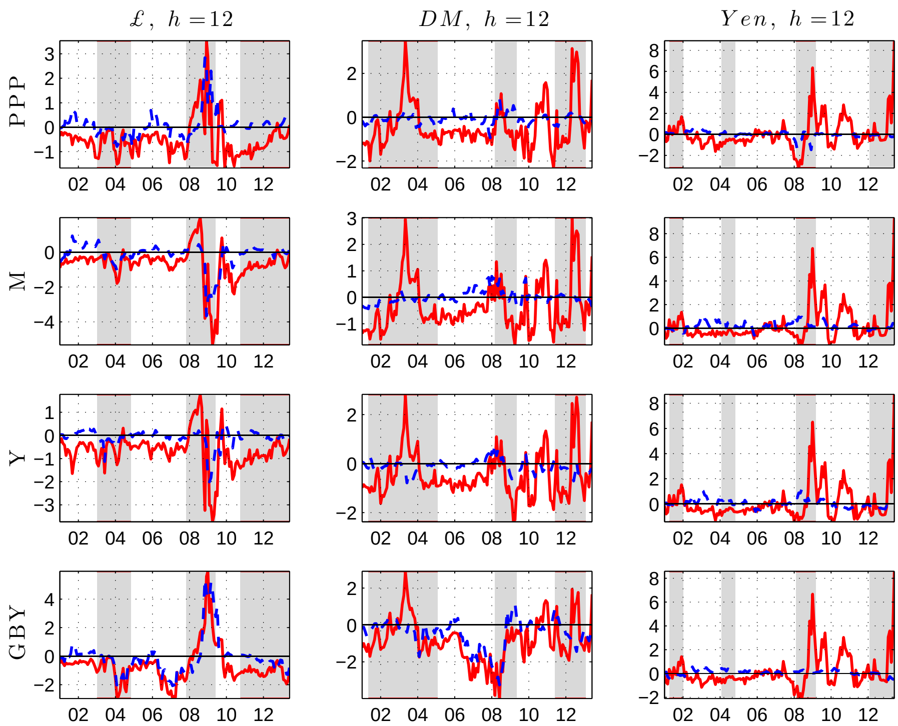
Note: The best performing forecast benchmarks are the RW-GARCH (at $h = 1$) and the constant-parameter BVAR (at higher horizons). A positive value of the statistic indicates a TVP-FF-VAR model with higher predictive likelihood than the benchmark. Increases in the statistic denote dates in which the TVP-FF-VAR performs better than the benchmark.

Figure 12: **Log-predictive densities for the Yen**: Cumulative differences in log-predictive likelihoods between competing TVP-FF-VAR models and the best performing benchmark, at selected forecast horizons.



Note: The best performing forecast benchmarks are the RW-GARCH (at $h = 1$) and the constant-parameter BVAR (at higher horizons). A positive value of the statistic indicates a TVP-FF-VAR model with higher predictive likelihood than the benchmark. Increases in the statistic denote dates in which the TVP-FF-VAR performs better than the benchmark.

Figure 13: **The role of fundamentals**: Differences in 1-year ahead log-predictive likelihoods between the TVP-FF-VAR model in the row header and the core TVP-FF-VAR with only exchange rates (dashed blue line), or the BVAR (solid red line).



Note: Positive values denote dates in which the forecasting performance of the fundamentals-based model is better. Shaded areas denote OECD-dated recession periods, respectively for the UK (first column), Germany (second column), and Japan (third column).

Figure 14: **Economic evaluation**: Returns of trading strategies based on different forecast models.

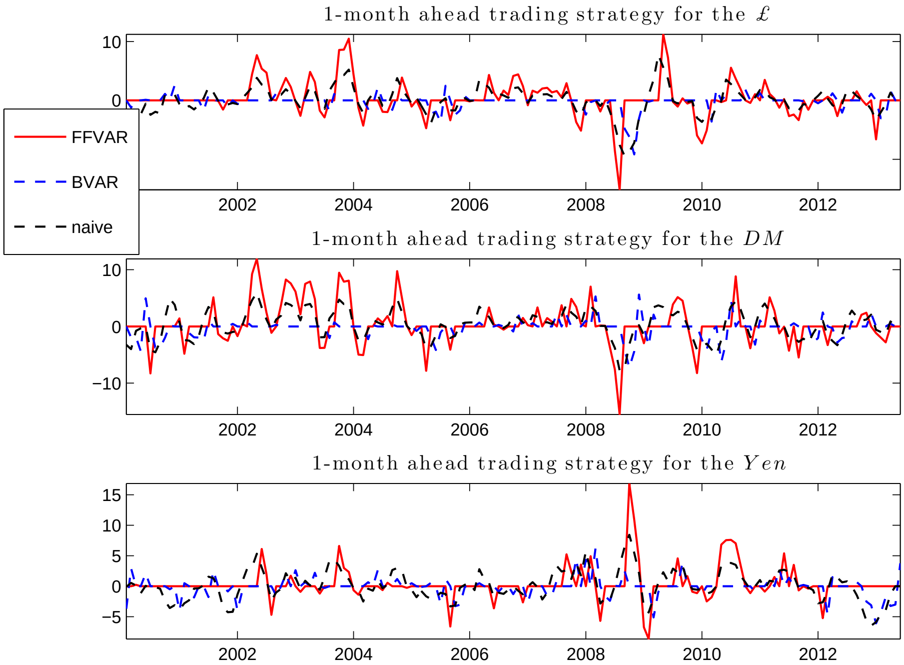


Figure 15: **Economic evaluation**: Returns of trading strategies based on macroeconomic and financial fundamentals, over the crisis forecast sample.

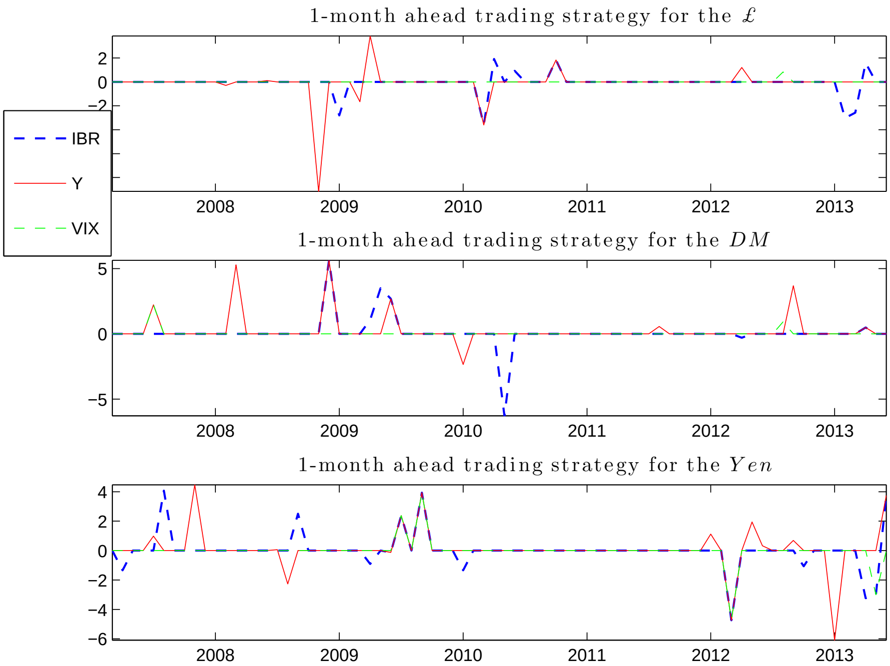


Table 2: **Priors of the TVP-SV-BVAR model**: The hat notation refers to OLS estimates on a 4-year training sample. M and K are the number of variables and of parameters. S_1 and S_2 pertain to the non-zero blocks of S, the covariance matrix of A. For the variables with an Inverse-Wishart prior distribution, the chosen number of degrees of freedom is the lowest admissible (one more than the size of the variable), to minimise the prior weight. The shrinkage parameters k_Q^2, k_S^2, k_S^2 are all set to 0.1.

VARIABLE	DISTRIBUTION	MEAN	VARIANCE	D.F.
B_0	N	$\hat{A}_{ols}$	$4 \cdot V(\hat{B}_{ols})$	$-$
A_0	N	$\hat{B}_{ols}$	$4 \cdot V(\hat{A}_{ols})$	$-$
$\log \sigma_0$	N	$\log \hat{\sigma}_{ols}$	I_M	$-$
Q	IW	$k_Q^2 \cdot V(\hat{B}_{ols})$	$-$	$K+1$
S_1	IW	$k_S^2 \cdot 2 \cdot V(\hat{A}_{1,ols})$	$-$	2
S_2	IW	$k_S^2 \cdot 3 \cdot V(\hat{A}_{2,ols})$	$-$	3
W	IW	$k_W^2 \cdot 4 \cdot I_M$	$-$	4

Table 3: **Mean squared forecast errors of the TVP-SV-BVAR and TVP-FF-VAR models relative to a random walk**: for different forecast samples and horizons. Values in bold denote significantly different RMSFE according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. [1998].

		PRE-CRISIS SAMPLE			CRISIS SAMPLE			FULL SAMPLE		
	h	£	DM	¥	£	DM	¥	£	DM	¥
TVP-SV-BVAR	$h=1$	0.94	**0.89**	1.02	**0.84**	0.94	**0.64**	**0.89**	**0.91**	**0.86**
	$h=3$	0.99	0.93	0.96	1.03	1.08	0.74	1.02	1.00	0.84
	$h=6$	**0.90**	0.96	0.91	1.13	1.16	0.89	1.05	1.05	0.90
	$h=12$	0.88	1.05	0.90	**1.32**	1.28	1.05	1.02	1.11	0.96
CORE FF-VAR	$h=1$	1.00	0.90	1.03	0.90	1.00	0.86	0.95	0.94	0.96
	$h=3$	1.07	0.97	1.08	1.09	1.18	0.89	1.09	1.06	0.98
	$h=6$	0.96	0.90	1.21	**1.33**	**1.35**	0.95	1.20	1.10	1.08
	$h=12$	0.97	0.86	1.34	2.07	**1.64**	1.02	1.26	1.05	1.24
UIRP FF-VAR	$h=1$	1.01	0.97	1.06	1.03	1.10	0.94	1.02	1.02	1.00
	$h=3$	1.05	1.04	1.06	1.24	1.47	1.01	1.18	1.22	1.03
	$h=6$	1.00	1.02	1.11	1.47	2.12	1.19	1.31	1.47	1.15
	$h=12$	1.02	1.10	1.07	2.49	**2.94**	2.16	1.52	1.54	1.51
PPP FF-VAR	$h=1$	**1.18**	1.08	0.95	0.90	1.18	0.89	1.04	1.12	0.92
	$h=3$	**1.34**	1.20	0.98	1.01	1.46	0.99	1.12	**1.32**	0.99
	$h=6$	1.19	1.19	0.95	**1.18**	1.85	0.98	1.14	1.44	0.98
	$h=12$	1.24	1.12	1.05	1.56	**2.34**	1.20	1.08	1.37	1.17
M FF-VAR	$h=1$	1.11	1.00	1.02	1.14	1.09	0.95	1.12	1.03	0.99
	$h=3$	**1.20**	1.01	0.94	1.26	1.17	0.89	1.24	1.09	0.91
	$h=6$	0.97	0.81	0.84	1.46	1.27	0.72	1.27	1.02	0.78
	$h=12$	**0.89**	0.77	0.59	2.93	**1.33**	0.66	1.34	0.91	0.68
Y FF-VAR	$h=1$	1.01	0.97	1.07	0.98	1.25	1.30	1.00	1.09	1.17
	$h=3$	1.08	1.06	1.05	1.09	**1.49**	1.20	1.09	1.25	1.13
	$h=6$	1.07	1.07	0.98	1.22	**1.30**	1.25	1.18	1.18	1.12
	$h=12$	1.07	0.96	0.80	2.10	**2.02**	1.24	1.36	1.26	1.02
GBY FF-VAR	$h=1$	1.06	1.07	1.04	0.91	1.07	**0.81**	0.98	1.07	0.94
	$h=3$	**1.36**	**1.35**	1.02	1.06	1.29	0.82	1.15	**1.31**	0.92
	$h=6$	1.84	**1.88**	1.04	1.33	1.76	0.87	1.37	**1.69**	0.95
	$h=12$	2.48	2.05	1.02	2.46	**2.96**	0.91	1.53	**2.13**	1.02
SP FF-VAR	$h=1$	**1.25**	1.16	1.17	1.04	1.05	0.93	**1.14**	1.11	1.06
	$h=3$	1.44	1.30	1.24	**1.20**	1.09	0.84	**1.28**	**1.20**	1.03
	$h=6$	1.26	1.17	1.47	1.42	1.13	0.89	**1.35**	1.13	1.16
	$h=12$	1.04	0.95	1.86	**2.53**	1.22	0.85	1.44	0.99	1.36

Note: The forecasting models are described in table 1. The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 4: **Mean squared forecast errors of the TVP-SV-BVAR and TVP-FF-VAR models models relative to a constant-parameter BVAR**: for different forecast samples and horizons. Values in bold denote significantly different RMSFE according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. [1998].

		PRE-CRISIS SAMPLE			CRISIS SAMPLE			FULL SAMPLE		
	h	£	DM	¥	£	DM	¥	£	DM	¥
TVP-SV-BVAR	$h = 1$	**0.47**	**0.42**	**0.52**	**0.32**	**0.38**	**0.41**	**0.39**	**0.40**	**0.48**
	$h = 3$	0.85	**0.76**	**0.74**	0.70	**0.76**	**0.67**	0.75	**0.76**	**0.70**
	$h = 6$	**0.79**	**0.85**	**0.81**	0.96	0.95	0.94	**0.92**	**0.91**	**0.88**
	$h = 12$	0.81	0.98	0.87	**1.23**	1.28	1.06	0.97	1.07	0.92
CORE FF-VAR	$h = 1$	**0.49**	**0.42**	**0.52**	**0.35**	**0.41**	**0.55**	**0.41**	**0.41**	**0.53**
	$h = 3$	0.92	**0.79**	0.83	0.74	0.82	0.79	0.80	**0.81**	**0.81**
	$h = 6$	0.85	0.80	1.08	1.13	1.11	1.01	1.05	0.96	1.06
	$h = 12$	0.89	0.80	1.30	1.93	1.63	1.03	1.20	1.02	1.19
UIRP FF-VAR	$h = 1$	**0.50**	**0.45**	**0.54**	**0.40**	**0.45**	**0.60**	**0.44**	**0.45**	**0.56**
	$h = 3$	0.90	0.85	0.82	0.83	1.02	0.90	0.86	0.93	0.86
	$h = 6$	0.88	0.91	0.99	**1.25**	1.74	1.25	1.15	1.28	1.13
	$h = 12$	0.94	1.02	1.04	2.32	**2.93**	2.17	1.45	1.49	1.45
PPP FF-VAR	$h = 1$	**0.58**	**0.50**	**0.48**	**0.35**	**0.48**	**0.57**	**0.45**	**0.49**	**0.51**
	$h = 3$	1.15	0.97	0.76	0.68	1.02	0.88	0.83	1.01	**0.82**
	$h = 6$	1.05	1.07	0.85	1.01	1.52	1.04	1.00	1.26	0.96
	$h = 12$	1.13	1.04	1.01	1.45	2.32	1.21	1.03	1.33	1.12
M FF-VAR	$h = 1$	**0.55**	**0.46**	**0.52**	**0.44**	**0.44**	**0.61**	**0.49**	**0.45**	**0.55**
	$h = 3$	1.03	0.82	**0.72**	0.85	0.82	0.80	0.91	**0.83**	**0.76**
	$h = 6$	0.86	0.72	**0.75**	1.24	1.05	0.76	1.11	0.89	**0.76**
	$h = 12$	**0.81**	0.72	0.57	2.73	**1.33**	0.66	1.28	0.89	**0.65**
Y FF-VAR	$h = 1$	**0.50**	**0.45**	**0.54**	**0.38**	**0.51**	0.83	**0.43**	**0.48**	**0.65**
	$h = 3$	0.93	0.86	0.81	0.73	1.04	1.07	0.80	0.95	0.94
	$h = 6$	0.94	0.96	0.87	1.04	1.07	1.32	1.03	1.03	1.10
	$h = 12$	0.98	0.90	0.77	**1.96**	**2.01**	1.25	1.30	1.22	0.98
GBY FF-VAR	$h = 1$	**0.52**	**0.50**	**0.53**	**0.35**	**0.44**	**0.52**	**0.43**	**0.47**	**0.52**
	$h = 3$	1.17	1.10	**0.79**	0.71	0.90	0.74	0.85	1.00	**0.76**
	$h = 6$	1.62	1.67	0.93	1.13	1.44	0.92	1.21	**1.47**	0.94
	$h = 12$	2.28	1.91	0.98	2.30	**2.94**	0.92	1.46	2.07	0.98
SP FF-VAR	$h = 1$	**0.62**	**0.54**	**0.59**	**0.40**	**0.43**	**0.59**	**0.50**	**0.49**	**0.59**
	$h = 3$	1.23	1.06	0.96	0.80	0.76	0.75	0.94	0.92	0.86
	$h = 6$	1.11	1.04	1.31	**1.21**	0.93	0.94	1.19	0.99	1.14
	$h = 12$	0.96	0.89	1.79	**2.36**	1.21	0.86	1.37	0.96	1.30

Note: The forecasting models are described in table 1. The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 5: **Testing the probability integral transforms**

	MODELS				
HORIZON	TVP-FF-VAR	TVP-SV-BVAR	BVAR	NTVP-BVAR (RW)	RW
1	0.39	0.00	0.00	0.00	0.65
3	0.70	0.05	0.02	0.00	0.00
6	0.73	0.28	0.43	0.12	0.01
12	0.66	0.42	0.72	0.43	0.08

Note: Main table values are p-values for the null hypothesis of the Kolmogorov-Smirnov test that the p.i.t. sequence of the model in the column is $U(0,1)$. The Bonferroni correction is used for horizons greater than 1; see text for details.

Table 6: **Coverage rates**

		PRE-CRISIS SAMPLE						CRISIS SAMPLE						FULL SAMPLE					
		68%			95%			68%			95%			68%			95%		
	h	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥
TVP-FF-VAR	1	**65.05**	**66.02**	**67.96**	**93.20**	**96.12**	**93.20**	**67.24**	**74.14**	**68.97**	**94.83**	**94.83**	**93.10**	**65.84**	**68.94**	**68.32**	**93.79**	**95.65**	**93.17**
	3	**73.27**	**64.36**	**70.30**	**95.05**	**96.04**	**93.07**	**68.97**	**62.07**	**58.62**	87.93	**96.55**	**91.38**	**71.70**	**63.52**	**66.04**	**92.45**	**96.23**	**92.45**
	6	79.59	**63.27**	**70.41**	**97.96**	98.98	**91.84**	**65.52**	**63.79**	**65.52**	87.93	**89.66**	87.93	**74.36**	**63.46**	**68.59**	**94.23**	**95.51**	90.38
	12	**75.00**	**67.39**	**68.48**	100.00	**97.83**	**93.48**	**77.59**	**77.59**	**60.34**	82.76	100.00	**96.55**	76.00	**71.33**	**65.33**	**93.33**	98.67	**94.67**
TVP-SV-BVAR	1	96.12	92.23	96.12	100.00	100.00	100.00	82.76	86.21	93.10	**98.28**	**98.28**	100.00	91.30	90.06	95.03	99.38	99.38	100.00
	3	95.05	85.15	94.06	100.00	100.00	100.00	**79.31**	**75.86**	84.48	**91.38**	**96.55**	100.00	89.31	81.76	90.57	**96.86**	98.74	100.00
	6	96.94	90.82	100.00	100.00	100.00	100.00	**72.41**	**72.41**	81.03	**91.38**	**96.55**	100.00	87.82	83.97	92.95	**96.79**	98.72	100.00
	12	100.00	86.96	100.00	100.00	100.00	100.00	**79.31**	89.66	93.10	**89.66**	100.00	100.00	92.00	88.00	97.33	**96.00**	100.00	100.00
BVAR	1	100.00	98.06	99.03	100.00	100.00	100.00	86.21	82.76	87.93	100.00	100.00	100.00	95.03	92.55	95.03	100.00	100.00	100.00
	3	98.02	86.14	85.15	100.00	100.00	100.00	**79.31**	**60.34**	**67.24**	**91.38**	**91.38**	**89.66**	91.19	76.73	78.62	**96.86**	**96.86**	**96.23**
	6	93.88	**61.22**	**71.43**	100.00	98.98	100.00	**68.97**	37.93	55.17	84.48	82.76	86.21	84.62	52.56	**65.38**	**94.23**	**92.95**	**94.87**
	12	**61.96**	40.22	52.17	100.00	81.52	**96.74**	**60.34**	41.38	29.31	79.31	79.31	72.41	**61.33**	40.67	43.33	**92.00**	80.67	87.33
BVAR (ROLLING)	1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	3	100.00	99.01	99.01	100.00	100.00	100.00	93.10	89.66	93.10	100.00	100.00	100.00	97.48	95.60	96.86	100.00	100.00	100.00
	6	97.96	97.96	97.96	100.00	100.00	100.00	84.48	81.03	86.21	**98.28**	100.00	**98.28**	92.95	91.67	93.59	99.36	100.00	99.36
	12	82.61	**64.13**	85.87	100.00	**97.83**	100.00	**79.31**	87.93	**74.14**	100.00	100.00	100.00	81.33	**73.33**	81.33	100.00	98.67	100.00
RW-GARCH	1	**66.02**	**65.05**	78.64	**91.26**	**90.29**	99.03	**65.52**	**62.07**	**67.24**	**94.83**	**93.10**	**93.10**	**65.84**	**63.98**	**74.53**	**92.55**	**91.30**	**96.89**
	3	100.00	100.00	100.00	100.00	100.00	100.00	93.10	98.28	96.55	100.00	100.00	100.00	97.48	99.37	98.74	100.00	100.00	100.00
	6	100.00	100.00	100.00	100.00	100.00	100.00	98.28	100.00	100.00	100.00	100.00	100.00	99.36	100.00	100.00	100.00	100.00	100.00
	12	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Note: Main table values denote actual coverage rates, corresponding to nominal coverages of 68% and 95%, for different models and forecast horizons h. Values in bold are not significantly different from the nominal counterpart. The forecasting models are described in the first block of table 1. The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 7: **Coverage rates**: Time-varying forgetting-factor models.

		PRE-CRISIS SAMPLE						CRISIS SAMPLE						FULL SAMPLE					
		68%			95%			68%			95%			68%			95%		
	h	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥	£	DM	¥
CORE	1	**65.05**	**64.08**	**68.93**	**93.20**	**96.12**	**93.20**	**67.24**	**74.14**	**68.97**	**94.83**	**93.10**	**93.10**	**65.84**	**67.70**	**68.94**	**93.79**	**95.03**	**93.17**
	3	**75.25**	**64.36**	**69.31**	**95.05**	**96.04**	**94.06**	**68.97**	**63.79**	**58.62**	**89.66**	**96.55**	**91.38**	**72.96**	**64.15**	**65.41**	**93.08**	**96.23**	**93.08**
	6	79.59	**67.35**	**70.41**	**97.96**	98.98	**90.82**	**65.52**	**63.79**	**65.52**	87.93	**89.66**	87.93	**74.36**	**66.03**	**68.59**	**94.23**	**95.51**	89.74
	12	**77.17**	**69.57**	**69.57**	100.00	**97.83**	**94.57**	**77.59**	**79.31**	**62.07**	82.76	100.00	**96.55**	77.33	**73.33**	**66.67**	**93.33**	98.67	**95.33**
UIRP	1	**64.08**	**64.08**	**68.93**	**93.20**	**96.12**	**93.20**	**68.97**	**72.41**	**68.97**	**94.83**	**93.10**	**93.10**	**65.84**	**67.08**	**68.94**	**93.79**	**95.03**	**93.17**
	3	**74.26**	**66.34**	**70.30**	**96.04**	**96.04**	**94.06**	**68.97**	**62.07**	**58.62**	87.93	**96.55**	**91.38**	**72.33**	**64.78**	**66.04**	**93.08**	**96.23**	**93.08**
	6	79.59	**65.31**	**70.41**	**97.96**	98.98	**91.84**	**65.52**	**65.52**	**63.79**	87.93	**89.66**	87.93	**74.36**	**65.38**	**67.95**	**94.23**	**95.51**	90.38
	12	**76.09**	**67.39**	**69.57**	100.00	**97.83**	**95.65**	**77.59**	**77.59**	**62.07**	82.76	100.00	**98.28**	76.67	**71.33**	**66.67**	**93.33**	98.67	**96.67**
PPP	1	**64.08**	**64.08**	**67.96**	**93.20**	**95.15**	**93.20**	**68.97**	**74.14**	**67.24**	**94.83**	**93.10**	**93.10**	**65.84**	**67.70**	**67.70**	**93.79**	**94.41**	**93.17**
	3	**73.27**	**64.36**	**69.31**	**95.05**	**96.04**	**94.06**	**67.24**	**62.07**	**58.62**	87.93	**94.83**	**91.38**	**71.07**	**63.52**	**65.41**	**92.45**	**95.60**	**93.08**
	6	78.57	**65.31**	**70.41**	**97.96**	98.98	**90.82**	**63.79**	**63.79**	**63.79**	87.93	**89.66**	87.93	**73.08**	**64.74**	**67.95**	**94.23**	**95.51**	89.74
	12	**76.09**	**68.48**	**68.48**	100.00	**97.83**	**94.57**	**77.59**	**77.59**	**60.34**	82.76	100.00	**98.28**	76.67	**72.00**	**65.33**	**93.33**	98.67	**96.00**
M	1	**70.87**	**66.02**	**68.93**	**93.20**	**94.17**	**93.20**	**70.69**	**77.59**	**65.52**	**94.83**	**93.10**	**93.10**	**70.81**	**70.19**	**67.70**	**93.79**	**93.79**	**93.17**
	3	**71.29**	**59.41**	**71.29**	**94.06**	**96.04**	**94.06**	**70.69**	**63.79**	55.17	**89.66**	**94.83**	**91.38**	**71.07**	**61.01**	**65.41**	**92.45**	**95.60**	**93.08**
	6	**69.39**	57.14	**68.37**	**96.94**	98.98	89.80	**68.97**	**62.07**	**56.90**	87.93	**91.38**	87.93	**69.23**	58.97	**64.10**	**93.59**	**96.15**	89.10
	12	**59.78**	56.52	**63.04**	**96.74**	**96.74**	**92.39**	**75.86**	**74.14**	50.00	82.76	100.00	**96.55**	**66.00**	**63.33**	58.00	**91.33**	**98.00**	**94.00**
Y	1	**63.11**	**64.08**	**68.93**	**93.20**	**96.12**	**92.23**	**67.24**	**74.14**	**65.52**	**94.83**	**93.10**	**93.10**	**64.60**	**67.70**	**67.70**	**93.79**	**95.03**	**92.55**
	3	**73.27**	**65.35**	**70.30**	**95.05**	**96.04**	**94.06**	**67.24**	**63.79**	**58.62**	87.93	**94.83**	**91.38**	**71.07**	**64.78**	**66.04**	**92.45**	**95.60**	**93.08**
	6	79.59	**64.29**	**71.43**	**97.96**	98.98	**92.86**	**65.52**	**63.79**	**63.79**	87.93	**89.66**	87.93	**74.36**	**64.10**	**68.59**	**94.23**	**95.51**	91.03
	12	**76.09**	**69.57**	**70.65**	100.00	**97.83**	**94.57**	**77.59**	**77.59**	**60.34**	82.76	100.00	**96.55**	76.67	**72.67**	**66.67**	**93.33**	98.67	**95.33**
GBY	1	**64.08**	**64.08**	**67.96**	**94.17**	**96.12**	**93.20**	**67.24**	**72.41**	**65.52**	**94.83**	**93.10**	**93.10**	**65.22**	**67.08**	**67.08**	**94.41**	**95.03**	**93.17**
	3	**73.27**	**64.36**	**70.30**	**95.05**	**96.04**	**93.07**	**67.24**	**62.07**	**58.62**	87.93	**94.83**	**91.38**	**71.07**	**63.52**	**66.04**	**92.45**	**95.60**	**92.45**
	6	77.55	**66.33**	**70.41**	**97.96**	98.98	**90.82**	**63.79**	**65.52**	**65.52**	87.93	**89.66**	87.93	**72.44**	**66.03**	**68.59**	**94.23**	**95.51**	89.74
	12	**76.09**	**68.48**	**69.57**	100.00	**97.83**	**94.57**	**77.59**	**77.59**	**62.07**	82.76	100.00	**98.28**	76.67	**72.00**	**66.67**	**93.33**	98.67	**96.00**
SP	1	**65.05**	**66.02**	**67.96**	**93.20**	**96.12**	**93.20**	**67.24**	**74.14**	**68.97**	**94.83**	**94.83**	**93.10**	**65.84**	**68.94**	**68.32**	**93.79**	**95.65**	**93.17**
	3	**73.27**	**64.36**	**70.30**	**95.05**	**96.04**	**93.07**	**68.97**	**62.07**	**58.62**	87.93	**96.55**	**91.38**	**71.70**	**63.52**	**66.04**	**92.45**	**96.23**	**92.45**
	6	79.59	**63.27**	**70.41**	**97.96**	98.98	**91.84**	**65.52**	**63.79**	**65.52**	87.93	**89.66**	87.93	**74.36**	**63.46**	**68.59**	**94.23**	**95.51**	90.38
	12	**75.00**	**67.39**	**68.48**	100.00	**97.83**	**93.48**	**77.59**	**77.59**	**60.34**	82.76	100.00	**96.55**	76.00	**71.33**	**65.33**	**93.33**	98.67	**94.67**

Note: Main table values denote actual coverage rates, corresponding to nominal coverages of 68% and 95%, for different models and forecast horizons h. Values in bold are not significantly different from the nominal counterpart. The forecasting models are described in the second block of table 1. The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 8: **Amisano-Giacomini test statistic**

		PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
	MODELS	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
£	BVAR	**8.68**	**3.40**	−0.68	**−3.31**	**5.10**	−1.55	**−2.03**	**−6.79**	**9.80**	−0.30	**−1.98**	**−5.27**
	BVAR (ROLLING)	**12.69**	**7.89**	**3.04**	−1.49	**12.21**	−0.18	−1.21	**−2.10**	**16.84**	**2.33**	−0.65	**−2.41**
	RW-GARCH	0.81	**10.40**	**16.24**	**19.88**	−1.19	−0.22	−0.44	−0.36	0.15	**2.83**	1.63	**1.92**
DM	BVAR	**6.01**	**−1.76**	**−4.59**	**−1.93**	**2.97**	**−2.26**	−1.30	−1.08	**6.55**	**−2.76**	**−3.98**	**−2.25**
	BVAR (ROLLING)	**11.51**	**4.64**	−0.96	**−4.65**	**9.44**	0.10	**−2.07**	**−5.97**	**14.63**	**2.30**	**−2.40**	**−6.95**
	RW-GARCH	**2.97**	**9.21**	**17.12**	**7.96**	1.06	1.10	**2.46**	**10.41**	**2.99**	**5.64**	**8.73**	**10.05**
Yen	BVAR	**9.83**	1.15	**−2.59**	**−3.04**	**5.27**	1.45	1.48	**2.22**	**11.35**	**1.75**	−0.37	−0.26
	BVAR (ROLLING)	**17.99**	**8.10**	1.61	−1.40	**12.50**	**3.10**	0.41	0.38	**22.44**	**7.61**	**2.19**	−1.37
	RW-GARCH	−0.15	**7.53**	**4.15**	1.06	1.55	**2.34**	1.12	0.51	1.05	**6.63**	**4.84**	0.93

Note: The table reports the t-statistics of the Amisano-Giacomini test for the null hypothesis that the TVP-FF-VAR model has the same predictive ability of the forecast model indicated in the row header. Positive values in bold denote the combination model-forecast horizon at which the TVP-FF-VAR model has a statistically higher log-predictive likelihood than the alternative model (for a 5% significance level). The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 9: **Amisano-Giacomini test statistic**: Forgetting-factor models vs best performing benchmarks.

		PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
	MODELS	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$	$h=1$	$h=3$	$h=6$	$h=12$
	CORE	0.81	**3.40**	−0.68	**−3.31**	−1.19	−1.55	**−2.03**	**−6.79**	0.16	−0.27	**−1.98**	**−5.41**
	UIRP	0.72	**3.32**	−0.82	**−3.94**	−1.55	−1.52	**−1.92**	**−5.59**	−0.38	−0.39	**−1.91**	**−5.74**
	PPP	−0.89	**1.74**	−1.51	**−4.27**	−0.94	−1.32	**−1.97**	**−1.97**	−1.24	−0.55	**−2.22**	**−3.58**
£	M	0.19	**3.06**	−0.49	**−2.65**	**−2.53**	−1.57	**−1.89**	**−3.69**	−1.13	−0.73	**−1.82**	**−3.91**
	Y	0.92	**3.29**	−1.16	**−3.47**	−1.50	**−1.80**	**−2.11**	**−5.87**	−0.23	−0.42	**−2.17**	**−5.47**
	GBY	0.42	1.30	**−2.34**	**−4.19**	−1.53	**−2.24**	**−5.89**	−0.47	−0.44	−0.80	**−4.46**	**−2.27**
	SP	−0.80	**1.75**	**−1.76**	**−3.59**	**−1.69**	**−1.68**	**−2.15**	**−7.89**	−1.56	−0.94	**−2.38**	**−6.00**
	CORE	**2.97**	**−1.76**	**−4.59**	**−1.93**	1.06	**−2.26**	−1.30	−1.08	**3.00**	**−2.81**	**−3.33**	**−2.18**
	UIRP	**2.29**	**−2.33**	**−4.23**	**−2.21**	0.40	**−2.55**	**−1.68**	−1.60	**1.98**	**−3.35**	**−3.52**	**−2.73**
	PPP	1.41	**−3.63**	**−5.84**	**−2.16**	0.06	**−2.32**	−1.50	−1.20	1.14	**−3.67**	**−3.96**	**−2.43**
DM	M	**1.90**	**−2.41**	**−5.06**	**−1.75**	0.27	**−2.05**	−1.41	−1.15	**1.73**	**−2.77**	**−3.37**	**−2.10**
	Y	**1.93**	**−2.60**	**−6.51**	**−2.36**	−0.41	**−3.06**	−1.58	**−2.14**	1.22	**−3.86**	**−3.96**	**−3.21**
	GBY	1.43	**−4.32**	**−5.56**	**−3.04**	0.35	**−2.87**	−1.57	**−2.97**	1.31	**−5.17**	**−5.06**	**−3.57**
	SP	1.15	**−3.05**	**−7.03**	**−2.49**	0.79	**−2.08**	−0.86	−0.78	1.40	**−3.64**	**−3.71**	**−2.34**
	CORE	−0.15	1.15	**−2.59**	**−3.04**	1.55	1.45	1.48	**2.22**	0.76	**1.84**	0.04	0.28
	UIRP	−0.06	1.56	**−2.13**	**−2.37**	0.38	0.67	0.71	0.99	0.17	1.40	−0.25	0.06
	PPP	0.66	**1.65**	**−1.68**	**−2.21**	1.39	1.07	1.28	**2.02**	1.25	**1.90**	0.33	0.22
Yen	M	−0.23	**2.86**	**−2.39**	**−2.63**	0.08	1.39	1.51	**2.24**	−0.13	**2.62**	0.80	1.09
	Y	−0.83	1.22	**−2.72**	**−3.19**	**−2.34**	−0.07	0.74	**1.78**	**−2.18**	0.61	−0.16	0.45
	GBY	−0.16	**2.02**	**−2.45**	**−2.69**	**2.62**	1.53	1.59	**2.36**	1.30	**2.33**	0.53	0.76
	SP	**−1.71**	0.42	**−3.77**	**−3.61**	0.30	1.64	1.45	**2.13**	−1.02	1.56	−0.14	0.12

Note: The table reports the t-statistics of the Amisano-Giacomini test for the null hypothesis that the TVP-FF-VARmodel indicated in the row header has the same predictive ability of the benchmark forecast model. Benchmark models are the RW-GARCH (for $h = 1$) and the constant-parameter TVP-FF-VAR(for longer horizons). Positive values in bold denote the combination model-forecast horizon at which the TVP-FF-VAR model has a statistically higher log-predictive likelihood than the benchmark (for a 5% significance level). The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Table 10: **Economic evaluation**: Key figures of the trading strategies based on the different forecast models.

		PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
	MODELS	$\mu(\pi)$	$\sigma(\pi)$	SR	ΔSR	$\mu(\pi)$	$\sigma(\pi)$	SR	ΔSR	$\mu(\pi)$	$\sigma(\pi)$	SR	ΔSR
£	CORE	0.49	1.51	0.33	0.21	0.07	1.70	0.04	0.15	0.29	1.61	0.18	0.19
	UIRP	−0.01	0.45	−0.01	−0.13	−0.16	0.89	−0.18	−0.06	−0.08	0.70	−0.11	−0.11
	PPP	−0.11	0.63	−0.17	−0.29	–	–	–	–	−0.06	0.46	−0.12	−0.12
	M	−0.08	0.36	−0.21	−0.33	−0.25	1.27	−0.20	−0.09	−0.16	0.92	−0.18	−0.17
	Y	−0.02	0.45	−0.04	−0.15	−0.10	1.25	−0.08	0.03	−0.06	0.92	−0.06	−0.06
	GBY	−0.06	0.63	−0.09	−0.21	0.01	0.66	0.01	0.12	−0.03	0.65	−0.04	−0.04
	SP	−0.21	1.00	−0.21	−0.33	0.02	0.59	0.03	0.14	−0.10	0.83	−0.12	−0.12
	IBR	−0.03	0.48	−0.06	−0.18	−0.07	0.78	−0.10	0.02	−0.05	0.64	−0.08	−0.07
	VIX	−0.17	0.82	−0.21	−0.33	0.01	0.09	0.11	0.23	−0.09	0.60	−0.14	−0.14
	BVAR	−0.11	0.91	−0.12	−0.24	−0.38	1.61	−0.23	−0.12	−0.24	1.30	−0.18	−0.18
	RW	0.12	1.00	0.12	–	−0.14	1.22	−0.11	–	−0.01	1.11	−0.01	–
DM	CORE	0.59	1.84	0.32	0.18	0.05	1.25	0.04	0.02	0.33	1.61	0.20	0.13
	UIRP	−0.10	0.65	−0.16	−0.29	0.00	0.00	0.00	0.00	−0.05	0.47	−0.12	−0.19
	PPP	−0.19	0.84	−0.23	−0.36	−0.03	1.18	−0.02	−0.04	−0.11	1.02	−0.11	−0.19
	M	−0.14	0.75	−0.18	−0.31	0.12	1.92	0.06	0.04	−0.01	1.43	−0.01	−0.09
	Y	−0.08	0.42	−0.18	−0.31	0.24	1.07	0.22	0.20	0.07	0.81	0.09	0.01
	GBY	−0.18	1.02	−0.17	−0.31	−0.06	1.31	−0.04	−0.06	−0.12	1.17	−0.10	−0.18
	SP	−0.08	1.04	−0.07	−0.21	0.03	1.50	0.02	0.00	−0.02	1.28	−0.02	−0.10
	IBR	−0.05	0.46	−0.11	−0.25	0.09	1.10	0.08	0.06	0.01	0.83	0.02	−0.06
	VIX	0.03	0.64	0.04	−0.09	0.05	0.28	0.17	0.15	0.04	0.50	0.07	−0.00
	BVAR	−0.35	1.28	−0.28	−0.41	−0.24	2.04	−0.12	−0.14	−0.30	1.68	−0.18	−0.26
	RW	0.16	1.22	0.13	–	0.03	1.32	0.02	–	0.10	1.27	0.08	–
Yen	CORE	0.11	0.87	0.13	0.19	0.25	1.47	0.17	0.05	0.18	1.20	0.15	0.12
	UIRP	0.02	1.33	0.02	0.08	0.03	0.62	0.05	−0.07	0.03	1.05	0.02	−0.01
	PPP	0.07	1.12	0.06	0.12	0.26	1.23	0.21	0.09	0.16	1.17	0.13	0.10
	M	0.17	1.09	0.16	0.22	0.14	1.26	0.11	−0.01	0.16	1.17	0.13	0.10
	Y	−0.07	1.02	−0.07	−0.01	0.08	1.29	0.07	−0.05	0.00	1.16	0.00	−0.03
	GBY	−0.21	1.08	−0.19	−0.13	0.06	0.67	0.09	−0.03	−0.08	0.91	−0.09	−0.12
	SP	−0.21	1.19	−0.18	−0.12	0.14	0.88	0.16	0.04	−0.04	1.06	−0.04	−0.07
	IBR	0.00	1.32	0.00	0.06	0.01	1.18	0.01	−0.11	0.01	1.25	0.00	−0.03
	VIX	−0.02	1.43	−0.01	0.05	−0.02	0.83	−0.02	−0.14	−0.02	1.18	−0.01	−0.05
	BVAR	−0.13	1.30	−0.10	−0.03	−0.14	1.94	−0.07	−0.19	−0.13	1.63	−0.08	−0.11
	RW	−0.07	1.08	−0.06	–	0.16	1.30	0.12	–	0.04	1.19	0.03	–

Note: Rows 1 to 9 in each table block display the results of trading strategies based on the forgetting-factor models, with or without macroeconomic and financial fundamentals. Rows 10 to 11 pertain instead to the two benchmarks: a constant-parameter BVAR and a naive strategy that predicts trading with 50% probability. ΔSR is defined as the difference in Sharpe ratios between the model in the row header and the naive strategy. The pre-crisis sample goes from 2000:m2 to 2008:m8. The crisis sample spans the period from 2008:m9 to 2013:m6.

Appendix A: Database description

Table 11: **Database description:** All the series below are monthly data. The yen is scaled down by 100.

SYMBOL	SERIES	SPAN	UNIT	SOURCE
£	POUND STERLING	1969:m1 - 2013:m6	currency units per USD	DATASTREAM
¥	YEN	1969:m1 - 2013:m6	currency units per USD	DATASTREAM
DM	DEUTSCHE MARK	1969:m1 - 2013:m6	currency units per USD	DATASTREAM
Y_{us}	INDUSTRIAL PRODUCTION US	1970:m1 - 2013:m6	index, 2005=100	OECD
Y_{uk}	INDUSTRIAL PRODUCTION UK	1970:m1 - 2013:m6	index, 2005=100	OECD
Y_{de}	INDUSTRIAL PRODUCTION DE	1970:m1 - 2013:m6	index, 2005=100	OECD
Y_{jp}	INDUSTRIAL PRODUCTION JP	1970:m1 - 2013:m6	index, 2005=100	OECD
P_{us}	CORE CPI US	1970:m1 - 2013:m6	index, 2005=100	OECD
P_{uk}	CORE CPI UK	1970:m1 - 2013:m6	index, 2005=100	OECD
P_{de}	CORE CPI DE	1970:m1 - 2013:m6	index, 2005=100	OECD
P_{jp}	CORE CPI JP	1970:m1 - 2013:m6	index, 2005=100	OECD
R_{us}	MONEY MKT INTEREST RATE US	1970:m1 - 2013:m6	%	IFS IFM
R_{uk}	MONEY MKT INTEREST RATE UK	1972:m1 - 2012:m6	%	IFS IFM
R_{de}	MONEY MKT INTEREST RATE DE	1970:m1 - 2013:m6	%	IFS IFM
R_{jp}	MONEY MKT INTEREST RATE JP	1970:m1 - 2013:m6	%	IFS IFM
M_{us}	MONEY SUPPLY M2 US	1959:m1 - 2013:m6	current prices	FED
M_{uk}	MONEY SUPPLY M2 UK	1986:m9 - 2013:m6	current prices	BANK OF ENGLAND
M_{de}	MONEY SUPPLY M2 DE	1973:m1 - 2013:m6	current prices	BUNDESBANK
M_{jp}	MONEY SUPPLY M2 JP	1960:m1 - 2013:m6	current prices	BANK OF JAPAN
B_{us}	10Y GOVERNMENT BOND YIELD US	1969:m1 - 2013:m6	%	OECD (MEI)
B_{uk}	10Y GOVERNMENT BOND YIELD UK	1969:m1 - 2013:m6	%	OECD (MEI)
B_{de}	10Y GOVERNMENT BOND YIELD DE	1969:m1 - 2013:m6	%	OECD (MEI)
B_{jp}	10Y GOVERNMENT BOND YIELD JP	1969:m1 - 2013:m6	%	OECD (MEI)
S_{us}	STOCK PRICE INDEX US	1969:m1 - 2013:m6	price index	REUTERS
S_{uk}	STOCK PRICE INDEX UK	1969:m1 - 2013:m6	price index	REUTERS
S_{de}	STOCK PRICE INDEX DE	1969:m1 - 2013:m6	price index	REUTERS
S_{jp}	STOCK PRICE INDEX JP	1969:m1 - 2013:m6	price index	REUTERS
L_{us}	3-MONTH INTERBANK LENDING RATE US	1986:m1 - 2013:m6	%	BBA
L_{uk}	3-MONTH INTERBANK LENDING RATE UK	1975:m1 - 2013:m6	%	BBA
L_{de}	3-MONTH INTERBANK LENDING RATE DE	1986:m1 - 2013:m6	%	BBA
L_{jp}	3-MONTH INTERBANK LENDING RATE JP	1986:m5 - 2013:m6	%	BBA
VIX	VIX VOLATILITY INDEX	1990:m1 - 2013:m6	%	CBOE