DEUTSCHE BUNDESBANK
EUROSYSTEM

Discussion Paper

Deutsche Bundesbank
No 19/2016

Point, interval and density forecasts of exchange rates with time-varying parameter models

Angela Abbate
(Deutsche Bundesbank)

Massimiliano Marcellino
(Bocconi University, IGIER and CEPR)

Discussion Papers represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank or its staff.

Editorial Board: Daniel Foos
Thomas Kick
Jochen Mankart
Christoph Memmel
Panagiota Tzamourani

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet http://www.bundesbank.de

Reproduction permitted only if source is stated.

ISBN 978–3–95729–263–6 (Printversion)
ISBN 978–3–95729–264–3 (Internetversion)

Non-technical summary

Research question

Exchange rates are relevant for a variety of economic agents: they influence production decisions, prices and portfolio allocations, and more generally they affect a country's competitiveness. There is hence a clear need for reliable models that track the future behavior of exchange rates, especially in times of uncertainty and financial stress. In this paper we investigate whether the forecasts of nine major currencies vis-a-vis the US dollar can be improved by assuming time variation in the coefficients of the data generating process.

Contribution

We explore whether modelling parameter time variation, and controlling for macroeconomic fundamentals in a time-varying manner, improves the point, interval and density forecasts of nine major exchange rates vis-a-vis the US dollar over the period 1976-2015. In addition, we perform an economic evaluation by building optimized trading strategies based on the competing forecast models.

Results

Three main results emerge from the empirical exercise. First, modelling parameter time-variation significantly refines the estimation of forecast uncertainty, though it does not improve point forecasts. In particular, time-varying parameter models deliver 68% and 95% forecast confidence intervals which are on average accurately calibrated. Second, using two scoring rules commonly used in the literature, we find that modelling parameter time variation is better suited at long horizons and in periods of high volatility, especially in the decade between 2000 and 2010. Third, the economic evaluation shows that trading strategies based on time-varying parameter models lead to higher portfolios returns, and to higher utility values for investors.

Nichttechnische Zusammenfassung

Fragestellung

Wechselkurse haben für eine Vielzahl von Wirtschaftsakteuren eine hohe Relevanz: Sie beeinflussen Produktionsentscheidungen, Preise sowie die Portfolioallokation und ganz allgemein die Wettbewerbsfähigkeit eines Landes. Dies macht deutlich, dass zuverlässige Modelle benötigt werden, die das künftige Verhalten von Wechselkursen, insbesondere in Phasen von Unsicherheit und finanziellen Spannungen, abbilden. In der vorliegenden Studie wird untersucht, ob die Prognosen zur Entwicklung von neun wichtigen Währungen gegenüber dem US-Dollar dadurch verbessert werden können, dass bei den Koeffizienten des Datengenerierungsprozesses eine Zeitvariation unterstellt wird.

Forschungsbeitrag

Wir prüfen, ob die Punkt-, Intervall- und Dichteprognosen neun wichtiger Wechselkurse im Verhältnis zum US-Dollar von 1976 bis 2015 durch die Modellierung der Zeitvarianz von Parametern und die zeitabhängige Berücksichtigung makroökonomischer Fundamentaldaten verbessert werden können. Überdies führen wir eine ökonomische Evaluierung durch, indem wir optimierte Handelsstrategien auf Basis der konkurrierenden Prognosemodelle konstruieren.

Ergebnisse

Aus der empirischen Studie lassen sich drei wesentliche Ergebnisse ableiten: Erstens verfeinern Modelle mit zeitvariablen Parametern die Schätzung der Prognoseunsicherheit erheblich, verbessern jedoch nicht die Punktprognosen. So liefern solche Modelle 68 %- und 95 %-Prognose-Konfidenzintervalle, die im Schnitt korrekt kalibriert sind. Zweitens kommen wir unter Verwendung zweier Bewertungsregeln, die in der Fachliteratur üblicherweise herangezogen werden, zu dem Schluss, dass die Modellierung der Zeitvarianz von Parametern für längere Zeiträume und in Phasen hoher Volatilität (insbesondere in der Dekade von 2000 bis 2010) besser geeignet ist. Drittens zeigt die ökonomische Evaluierung, dass Handelsstrategien, die auf Modellen mit zeitvariablen Parametern beruhen, zu einer höheren Portfoliorendite führen und einen größeren Nutzwert für Anleger haben.

Bundesbank Discussion Paper No 19/2016

Point, interval and density forecasts of exchange rates with time-varying parameter models*

Angela Abbate† Massimiliano Marcellino‡

June 7, 2016

Abstract

We explore whether modelling parameter time variation improves the point, interval and density forecasts of nine major exchange rates vis-a-vis the US dollar over the period 1976-2015. We find that modelling parameter time variation is needed for an accurate calibration of forecast confidence intervals, and is better suited at long horizons and in high-volatility periods. The biggest forecast improvements are obtained by modelling time variation in the volatilities of the innovations, rather than in the slope parameters. Moreover, we do not find evidence that parameter time variation helps to unravel exchange rate predictability by macroeconomic fundamentals. Finally, an economic evaluation of the different forecast models reveals that controlling for parameter time variation leads to higher portfolios returns, and to higher utility values for investors.

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

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

*We would like to thank Pasquale Della Corte for his help on the economic evaluation. We are also grateful to the Editor, two anonymous Referees, Gino Cenedese, and participants at the 2005 IAAE conference, the 8th ECB workshop on Forecasting Techniques and at an ECB - Bank of Italy workshop for useful comments on an earlier draft.

†Deutsche Bundesbank, angela.abbate@bundesbank.de. The views expresses in this paper do not reflect the views of the Bundesbank nor of its staff. *Corresponding author.*

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

1 Introduction

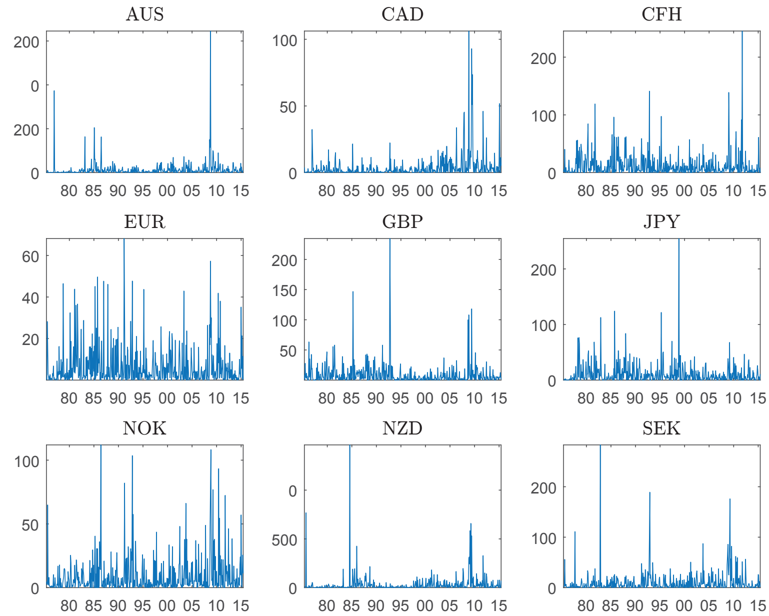
Exchange rates have an impact on the production decisions of firms, on portfolio allocations, and on a country's prices and competitiveness. Hence, there is a clear need for reliable models that track the current evolution of exchange rates and predict their future behaviour, especially in times of uncertainty. Since the seminal work of Meese and Rogoff (1983), a vast literature has been devoted to the construction and evaluation of exchange-rate point forecasts, and has established, with few exceptions, that the best forecast model is a simple random walk. One explanation for this puzzling exchange rate unpredictability is that it might arise from the instability of the underlying stochastic process.[1] In addition, though point forecasts are clearly of interest, interval and density forecasts of exchange rates are also relevant for the decision making of economic agents and for the pricing of financial assets. On this last point the literature is more limited.

In this paper we assess to what extent modelling parameter time variation improves the predictability of nine major currencies vis-a-vis the US dollar over the period 1976-2015. As it can be seen in figure 1, the volatility of these exchange rates has varied greatly over time: it has fallen after the price shocks and inflationary pressures of the 1970s, and it has increased again in the last decade, as a consequences of the crises in 2000-2001 and 2007-2008. Motivated by such considerations, we evaluate whether the point, interval and density forecasts of each of these exchange rates can be improved by controlling for other currencies and for macroeconomic fundamentals, as well as by allowing for parameter time-variation in the slope and volatility parameters. The latter is accounted for through a forgetting factor VAR model, recently proposed by Koop and Korobilis (2013), which is computationally tractable for medium-sized models, and exploits the information from current forecast errors to determine the extent of time variation.

Several results emerge from the statistical evaluation. First, modelling parameter time variation refines the estimation of forecast uncertainty, though it does not improve point forecasts. The point forecasts from time-varying autoregressive models are in fact more accurate than those from a random walk only at a one-month horizon, for roughly half the currencies in the sample. On the other hand, time-varying parameter models deliver 68% and 95% forecast confidence intervals which are on average accurately calibrated, according to unconditional coverage tests. Conversely, constant-parameter models generate forecast confidence intervals which are

[1] See for example Rossi (2006) and survey evidence documented by Cheung and Chinn (2001).

Figure 1: **Exchange rate volatility over the years**: measured as the square monthly percentage changes in the exchange rate reported in the title. For a description of the series, see Table 6 in the Appendix.



excessively large. Moreover, using two scoring rules commonly used in the literature, we find that modelling parameter time variation is better suited at long horizons and in periods of high volatility, especially in the decade between 2000 and 2010. The largest forecast improvements come in fact from modelling time variation in the volatilities of the innovations, rather than in the slope parameters of the VAR.

Second, modelling parameter time variation is not needed for macroeconomic fundamentals to show predictive content. A constant parameter model with macroeconomic fundamentals yields the highest predictive likelihoods at a one and two-year ahead horizons for one third of the currencies. Modelling variation in the slope parameters worsens the forecast performance of the constant-parameter model, while allowing for time-varying volatility is beneficial only up to a one-year ahead horizon.

The different performances of the competing models across periods suggest potential advantages from forecast combination. An optimal forecast combination based on continuous ranked probability scores reveals that controlling for macroeconomic differentials lowers point forecast errors for the Swiss Franc, the Euro, the Yen, and the New Zealand dollar, though this improvement is seldom statistically significant.

As a final step, we perform an economic evaluation of the different forecast models. In particular, we build trading strategies based on the forecasts of the competing models, taking the perspective of a US investor with a one-month trading horizon. The analysis reveals that strategies based on time-varying parameter models yield optimised portfolios with higher mean returns, but also higher volatility, than a static random walk strategy. Moreover, controlling for both macroeconomic fundamentals and time-varying parameters leads to the highest performance fee an investor would be willing to pay to switch from a benchmark random walk strategy, though the amount is not economically large.

Our comprehensive statistical and economic evaluation complements the empirical literature on exchange rate forecasting, not only for the methodology used, but also for the emphasis on interval and density forecasts. A wide variety of methods has been used in the empirical literature on exchange-rate forecasting, and the consensus is that the most difficult benchmark to beat, in terms of point forecast accuracy, is the random walk.[2] A small but increasing number of papers has instead focused on density forecasts. Relevant examples are Yongmiao et al. (2007) and Balke et al. (2013), who both show that the density forecasts of a random walk can be improved either with non-linear models, or with univariate Taylor-rule models with semiparametric confidence intervals. In addition, Mumtaz and Sunder-Plassmann (2013) show that a structural time-varying stochastic volatility vector autoregression outperforms its constant-parameter counterpart on the basis of the mean squared forecast error and of the Bayesian deviance information criterion. A paper closely related to ours is Della Corte et al. (2009), who establish that modelling time-varying volatility is important for the one-month ahead predictive ability of macroeconomic fundamentals. In contrast to the aforementioned articles, our forecast evaluation exercise comprises forecast horizons larger than one month, a wider set of evaluation criteria, and is based on a larger forecast sample that includes the 2008 financial crisis.

The paper is organised as follows. In the next Section we briefly outline the theoretical and empirical literature behind exchange-rate forecasting models. Section 3 describes our empirical strategy, while Sections 4 and 5 respectively present the statistical and economic evaluation of the competing models. Section 6 concludes.

[2]For instance, Carriero et al. (2009) and Dal Bianco et al. (2012) have improved upon the point forecasts of a random walk by relying respectively on a Bayesian vector autoregression with a large set of exchange rates, and on a mixed-frequency dynamic factor model with four weekly exchange rates and lower-frequency macroeconomic fundamentals. However, Rossi (2013) documents that in general the methodologies that deliver lower mean squared forecast errors than a random walk are typically sensitive to the forecast horizon and to the sample used.

2 Exchange-rate prediction models: a discussion

Since the seminal work of Meese and Rogoff (1983), empirical research has established that a simple random walk suffices to deliver the most accurate point forecasts of exchange rates. Models augmented with macroeconomic and financial fundamentals appear to improve random walk forecasts only at long horizons, though not systematically across exchange rates and samples.[3] There are two prominent explanations for this puzzling result. A first one, suggested by Engel and West (2005), argues that exchange-rate unpredictability is an implication of the structural models, rather than being evidence against them. The authors start from the consideration that all exchange-rate determination models can be rewritten so as to express exchange rates as present discounted values of current and future fundamentals, as well as unobservable shocks. Then, exchange rates will exhibit almost no correlation with current fundamentals so long as fundamentals are persistent and agents are patient, implying that future fundamentals matter more than current ones. A second explanation lies instead in the instability of the relationship between exchange rates and their fundamentals, which has been documented, among others, by Rossi (2006). Instability may arise from trading strategies that involve frequent changes to the weight attached to fundamentals, as documented through survey evidence by Cheung and Chinn (2001), or from a gradual time variation in the relationship between exchange rates and fundamentals, if the structural parameters are unknown to economic agents, as shown in Bacchetta and Wincoop (2009).

In this work we focus on the instability hypothesis and verify whether allowing for gradual parameter time variation improves the point and density forecasts relative to a random walk, and enables fundamentals to show predictive content. In particular, we model time variation in the relationships among nine major exchange rates vis-a-vis the US dollar, and between the latter and their macroeconomic fundamentals, as well as in the volatilities of these variables.

But which fundamentals should be considered? Several variables qualify in fact 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 ($p_t^* - p_t$):

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

[3] For a comprehensive review, see Rossi (2013).

where $e_t = \log s_t$ is defined as the number of currency units per US dollar in logs, following the notation used in the empirical application, and small case letters denote the logarithms of the variables.

Moreover, the uncovered interest rate parity (UIRP) condition suggests that exchange rate movements compensate differences 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 the rational expectation and risk neutrality hypotheses, and ρ_t can be interpreted either as a forward premium or as an expectational error.[4] Empirical evidence on both 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 et al. (2009), on the basis of statistical and economic criteria.

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 foreign central bank targets the country's exchange rate, as well as output and inflation. This assumption is not valid for the US, where it is assumed that the interest rate is set solely as a function of output and inflation fluctuations. 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).

Several other variables qualify as potential exchange rate predictors, including commodity and oil prices, trade balance differentials and productivity measures. We focus however on the most commonly used ones, given that our aim is to show whether the relationship between exchange rates and predictors varies over time, rather than to find the best predictor. Therefore, on the basis of the theoretical models outlined above, we allow exchange rates to be affected by differentials in inflation, in the short-term interest rate and in industrial production growth.

[4]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 et al. (2009).

3 Empirical strategy

In this Section we first introduce a flexible model of parameter time variation, and later outline the empirical strategy followed in the paper.

3.1 The time-varying parameter forgetting factor VAR

Time variation is modelled through a forgetting factor VAR (henceforth TVP FFVAR), recently introduced by Koop and Korobilis (2013). This model can be considered as an approximation of the Bayesian VAR with time-varying parameters and stochastic volatility developed by Cogley and Sargent (2005) and Primiceri (2005). Instead of using a sampler to draw the covariance matrices of the time-varying parameters and volatilities, the TVP FFVAR estimates them as a weighted average of previous estimates and of the current forecast error variance. It is a more parsimonious set up that considerably speeds up computational time, enables the analysis with a larger number of variables, and has been shown to perform well in forecasting analyses.[5]

The assumptions of the TVP FFVAR model imply a state-space model, with the following measurement equation:

$$y_t = Z_t \beta_t + \varepsilon_t \ , \tag{3}$$

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 drawn from a Normal distribution with mean 0 and covariance matrix Ω_t. Let Z_t contain a constant and p lags of each variable; it is then defined as $Z_t = I_n \otimes [1, y'_{t-1}, \ldots y'_{t-p}]$ with dimension $n \times k = n \times n(1 + np)$.

Given information up to $t-1$, the slope coefficients in t are draws from a normal distribution: $\beta_t |\, y^{t-1} \sim N(\beta_{t|t-1},\, P_{t|t-1})$. The Kalman filter routine, used in the first step of the Gibbs sampler, entails a prediction for the coefficients' covariance matrix, $P_{t|t-1} = P_{t-1|t-1} + Q$. To avoid drawing from the posterior distribution of Q, the following approximation is used:

$$P_{t|t-1} = \frac{1}{\lambda} P_{t-1|t-1}, \quad \text{with } \lambda \in (0,\, 1]\,, \tag{4}$$

where the parameter λ is a forgetting factor which discounts past information. A value of λ equal to 0.99 implies, in the case of monthly data, that observations one year ago receive 89%

[5] See, for instance, Koop and Korobilis (2013). Moreover, in a previous version of this paper we show that the TVP FFVAR generally outperforms the Bayesian counterpart in a comparison based on a fewer number of exchange rates. These results are available upon request.

as much weight as current observations.

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:[6]

$$\hat{\Omega}_t = \kappa \hat{\Omega}_{t-1} + (1-\kappa)\hat{\varepsilon}_t' \hat{\varepsilon}_t \,, \tag{5}$$

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 4 and 5, 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.

Note that equations (4) and (5) in the forgetting factor model do not provide any rule for the out-of-sample evolution for the two covariance matrices. In order to generate samples from the predictive density, we follow Koop and Korobilis (2013) and assume Ω_t to be fixed out of sample. 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 $\hat{P}_{t|t}$.[7] 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 relevant percentiles of the values $\{\hat{y}_{t+i,\,\kappa},\, i = 1 \ldots h\}_{\kappa=1}^{5000}$.

3.2 Exchange-rate prediction models

The goal of the paper is to assess whether the point, interval and density forecasts of the levels of nine major exchange rates vis-a-vis the US dollar can be improved relative to a random walk by (1) jointly modelling the nine exchange rates, (2) modelling the nine exchange rates jointly with macroeconomic differentials, and (3) allowing gradual time variation in the parameters of the forecast models.

To this purpose, we consider the relative performance of five models, described in Table 1: a time-varying univariate forgetting factor model (TVP AR), a time-varying forgetting factor VAR with only exchange rates (TVP FFVAR), a constant parameter Bayesian VAR with only exchange rates (BVAR), a time-varying forgetting factor VAR with exchange rates and fundamentals (TVP FFVAR MACRO), and a constant parameter Bayesian VAR with exchange rates and fundamentals (BVAR MACRO). The constant parameter models are estimated recursively

[6]This is the Exponentially Weighted Moving Average estimator, commonly used in the finance literature.
[7]Allowing for the coefficients to drift out of sample does not improve the predictive ability of the model.

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

	MODELS	DESCRIPTION
1	RW	Random walk
2	TVP AR	Single-equation forgetting factor model, time-varying slope and volatility, $\lambda = 0.99$ and $\kappa = 0.96$
3	TVP FFVAR	Forgetting factor model, time-varying slope and volatility, $\lambda = 0.99$ and $\kappa = 0.96$
4	BVAR	Bayesian model with constant parameters, estimated recursively
5	TVP FFVAR MACRO	Forgetting factor model, time-varying slope and volatility, $\lambda = 0.99$ and $\kappa = 0.96$ augmented with macroeconomic fundamentals
6	BVAR MACRO	Bayesian model with constant parameters, estimated recursively augmented with macroeconomic fundamentals

every month. The point and density forecasts are then compared to those of a random walk, with standard deviation estimated recursively every month.

As additional exchange rate predictors, we select differentials in the short term interest rate, in inflation and in output growth, respectively measured as the month-on-month percentage changes in the CPI and industrial production indexes. A full list and description of the series is provided in Appendix B. To reduce the dimensionality of the VAR models, we consider only the first principal component of the nine differentials in each macro category.[8]

The nine exchange rates are measured in log levels, in order not to lose relevant information. To correctly evaluate all forecast models, we ensure that their priors, or initial values, are comparable. We shrink the initial values of the exchange-rates coefficients towards random walk processes. The same is done for the macro fundamentals, as they exhibit a strong degree of persistence. Only three lags are used, to reduce dimensionality; however, results do not differ if we allow for more lags and a larger degree of shrinkage. Following Koop and Korobilis (2013), we initialise the priors (or initial values) for the covariance matrix of the innovations to the variance of the data in the first estimation sample.

Accounting for the lag length, the estimation sample runs from 1976m4 to 1986m1, while the forecast sample runs from 1986m2 to 2015m6. The estimation sample is then progressively enlarged in a pseudo-real time exercise: in each month we re-estimate the models and compute forecasts from one-month to two-years ahead.

[8]We have also explored alternative specifications such as modelling all exchange rates jointly with all differentials, and having 9 single country VARs that model an exchange rate with its country-specific differentials. Since none of these models improves on those shown in the paper, we do not report these results but make them available upon request.

4 Empirical results

4.1 Point and interval forecasts

We begin our analysis by assessing the accuracy of point forecasts and of forecast confidence intervals delivered by the competing models, relative to those of a random walk.

Point forecasts from model j are compared with those from a random walk through their relative mean squared forecast error:

$$RMSFE_{i,h}^{j,rw} = \frac{\sum_{t=1}^{T_f} (\hat{y}_{i,t+h|t}^{j} - y_{i,t+h})^2}{\sum_{t=1}^{T_f} (\hat{y}_{i,t+h|t}^{rw} - y_{i,t+h})^2}, \tag{6}$$

where T_f is the number of forecasts, while i and h respectively index the variable and the horizon. Table 2 reports the relative mean squared forecast errors of the competing models, for selected horizons and forecast subsamples. Significance is established by means of a Diebold and Mariano test at a 5% and 10% significance levels, modified using the small sample size correction of Harvey et al. (1998).

Over the whole forecast sample, no model systematically delivers smaller point forecast errors than a random walk. A relevant exception are time-varying parameter models, and in particular the univariate one, which deliver significantly smaller forecast errors at a one-month horizon for roughly half of the exchange rates analysed. This advantage is however small and disappears at longer horizons. No significant advantage in terms of point forecast accuracy is instead gained by controlling for macroeconomic predictors.

To analyse the performance of the models during the Great Financial Crisis period, we split the sample into two. The cutoff date is 2007m1, so before the actual burst of the financial crisis. In the pre-crisis sample (1986m2-2006m12), there are some gains from the TVP FFVAR model for Switzerland, Norway and New Zealand, but only at a one month horizon. The same model woks well for Japan also at longer horizons. In the forecast subsample that includes the financial crisis, smaller point forecast errors can instead be obtained either by modelling exchange rates independently but with time-varying parameters (for the Uk, Norway and Sweden), or by modelling exchange rates jointly but allowing for no parameter time variation (for Switzerland, the Euro Area, the Uk and Norway, at horizons greater or equal than 6 months).

Table 2: **Mean squared forecast errors of the competing models relative to a random walk** for different forecast samples and horizons. Two (one) stars denote significantly different RMSFE at a 5% (10%) significance level according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. (1998). The forecast subsamples are defined as follows: PRE-CRISIS SAMPLE (1986:m2 - 2006:m12), CRISIS SAMPLE (2007:m1 - 2013:m6) and FULL SAMPLE (1986:m2 - 2013:m6).

	PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
MODEL ↓ / h →	1	6	12	24	1	6	12	24	1	6	12	24
AUS												
tvp ar	0.99	1.05	1.13	1.30	0.93^{**}	0.96	0.96	0.99	0.96^{**}	1.00	1.06	1.21
tvp ffvar	1.01	1.12	1.25	1.55	1.05	1.13	1.22	1.38	1.03	1.12	1.22	1.44
bvar	1.02	1.08	1.20	1.39	1.00	1.02	1.04	1.09	1.01	1.05	1.13	1.28
tvp ffvar macro	0.99	0.99	1.04	1.53	1.06	1.21	1.53	2.19	1.02	1.12	1.31	1.80
bvar macro	1.01	1.07	1.16	1.33	1.00	1.02	1.04	1.07	1.01	1.04	1.12	1.27
CAD												
tvp ar	1.02	1.11	1.20	1.29	1.00	0.99	0.98	0.93^{*}	1.01	1.04	1.09	1.21
tvp ffvar	1.09	1.53	1.81	2.10	0.97	1.15	1.38	1.95	1.02	1.29	1.56	1.97
bvar	1.01	1.09	1.16	1.27	1.00	1.02	1.05	1.12	1.01	1.05	1.10	1.20
tvp ffvar macro	1.03	1.19	1.35	2.00	1.00	1.27	1.92	4.11	1.01	1.23	1.62	2.48
bvar macro	1.01	1.09	1.13	1.16	1.00	1.02	1.05	1.06	1.01	1.05	1.09	1.15
CFH												
tvp ar	1.02	1.11	1.15	1.22	1.02	1.08	1.15	1.29	1.02	1.10	1.15	1.24
tvp ffvar	0.94^{**}	1.03	1.15	1.53	1.05	1.46	1.70	1.82	0.97	1.14	1.28	1.60
bvar	1.00	1.00	1.01	1.08	1.00	0.98	0.96^{**}	0.87^{**}	1.00	0.99	1.00	1.03
tvp ffvar macro	0.97	1.13	1.46	3.23	1.00	1.22	1.49	1.80	0.98	1.16	1.49	3.00
bvar macro	0.99	0.99	0.99	1.13	1.00	0.99	1.00	0.85^{*}	0.99	0.99	0.99	1.08
EUR												
tvp ar	0.99	1.09	1.11	1.14	0.95^{**}	0.99	1.01	0.99	0.98^{**}	1.06	1.09	1.14
tvp ffvar	1.06	1.14	1.34	1.71	1.12	1.15	1.43	2.26	1.07	1.14	1.36	1.76
bvar	1.00	1.00	1.01	1.05	1.00	0.99^{**}	0.98^{**}	0.95^{**}	1.00	0.99	1.00	1.03
tvp ffvar macro	0.98	1.08	1.26	2.22	1.07	1.40	2.30	6.61	1.00	1.19	1.57	2.79
bvar macro	0.99	0.97	0.95^{*}	1.04	1.02	1.09	1.27	2.35	1.00	1.01	1.03	1.17
GBP												
tvp ar	1.02	1.08	1.17	1.34	0.98^{*}	0.90^{**}	0.82^{**}	0.66^{**}	1.01	1.02	1.06	1.14
tvp ffvar	1.01	1.11	1.21	1.76	0.98	1.15	1.48	1.72	1.00	1.13	1.29	1.74
bvar	1.01	1.06	1.13	1.22	1.00	0.99^{*}	0.99	0.97^{**}	1.01	1.03	1.08	1.14
tvp ffvar macro	1.02	1.24	1.47	2.78	1.01	1.32	1.97	2.88	1.02	1.27	1.65	2.78
bvar macro	1.01	1.07	1.14	1.33	1.01	1.02	1.04	0.98	1.01	1.05	1.11	1.23
JPY												
tvp ar	1.02	1.09	1.18	1.70	1.00	1.00	1.00	1.01	1.01	1.07	1.12	1.45
tvp ffvar	0.99	0.98	0.86^{**}	0.92^{*}	1.03	1.15	1.15	0.97	1.00	1.02	0.94	0.94
bvar	1.00	1.01	1.04	1.14	1.00	1.00	0.99	1.02	1.00	1.01	1.03	1.10
tvp ffvar macro	0.98	1.10	1.14	1.76	1.01	1.09	1.10	1.51	0.99	1.10	1.14	1.74
bvar macro	0.99	0.99	0.98	1.08	1.01	1.06	1.10	1.20	1.00	1.01	1.02	1.12
NOK												
tvp ar	1.01	1.04	1.07	1.12	0.99	0.95	0.91^{**}	0.78^{**}	1.00	1.00	1.02	1.08
tvp ffvar	0.96^{**}	1.13	1.34	1.75	0.93^{**}	1.04	1.20	1.24	0.95^{**}	1.09	1.28	1.62
bvar	1.01	1.04	1.09	1.15	1.00	0.99^{*}	0.98^{**}	0.92^{**}	1.01	1.02	1.05	1.10
tvp ffvar macro	1.00	1.25	1.52	2.85	0.96	1.15	1.66	2.70	0.98	1.21	1.61	2.86
bvar macro	1.00	1.02	1.06	1.16	1.00	1.02	1.07	1.06	1.00	1.03	1.07	1.17
NZD												
tvp ar	1.01	1.03	1.05	1.05	1.01	0.98	0.99	1.18	1.01	1.01	1.04	1.09
tvp ffvar	0.96^{*}	1.08	1.26	1.48	0.94	1.09	1.18	1.09	0.95^{**}	1.09	1.22	1.39
bvar	1.02	1.08	1.17	1.27	1.00	1.00	1.00	0.99	1.01	1.04	1.11	1.21
tvp ffvar macro	0.97	0.91^{*}	0.86^{**}	1.29	1.00	1.33	1.85	2.71	0.99	1.13	1.29	1.67
bvar macro	1.01	1.03	1.10	1.25	1.00	1.03	1.05	0.97	1.01	1.04	1.10	1.23
SEK												
tvp ar	1.01	1.02	1.04	1.08	0.99	0.94^{**}	0.89^{**}	0.78^{**}	1.00	0.99	0.99	1.02
tvp ffvar	1.02	1.19	1.36	1.85	0.96	1.10	1.34	1.55	1.00	1.16	1.35	1.74
bvar	1.01	1.06	1.10	1.16	1.00	1.00	0.99	0.97	1.01	1.03	1.07	1.13
tvp ffvar macro	1.00	1.19	1.39	2.41	0.99	1.25	1.87	3.27	1.00	1.22	1.59	2.71
bvar macro	1.01	1.04	1.07	1.14	1.00	1.01	1.06	0.99	1.00	1.03	1.08	1.16

We proceed to examine how the competing models fare in delivering accurately calibrated forecast confidence intervals. We focus on 68% and 95% intervals, as these are the two most commonly used in empirical studies. We compute coverage rates for each competing model; i.e., the percentage of times in which the actual exchange rate is contained in the forecast confidence interval. As discussed previously, 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.

Table 3 reports actual coverage rates of the competing models, at different horizons over the whole forecast sample.[9] Values in bold denote coverage rates which are not statistically different from their 68% and 95% nominal counterparts, according to the unconditional coverage test described in Christoffersen (1998) and in Clements (2005).[10]

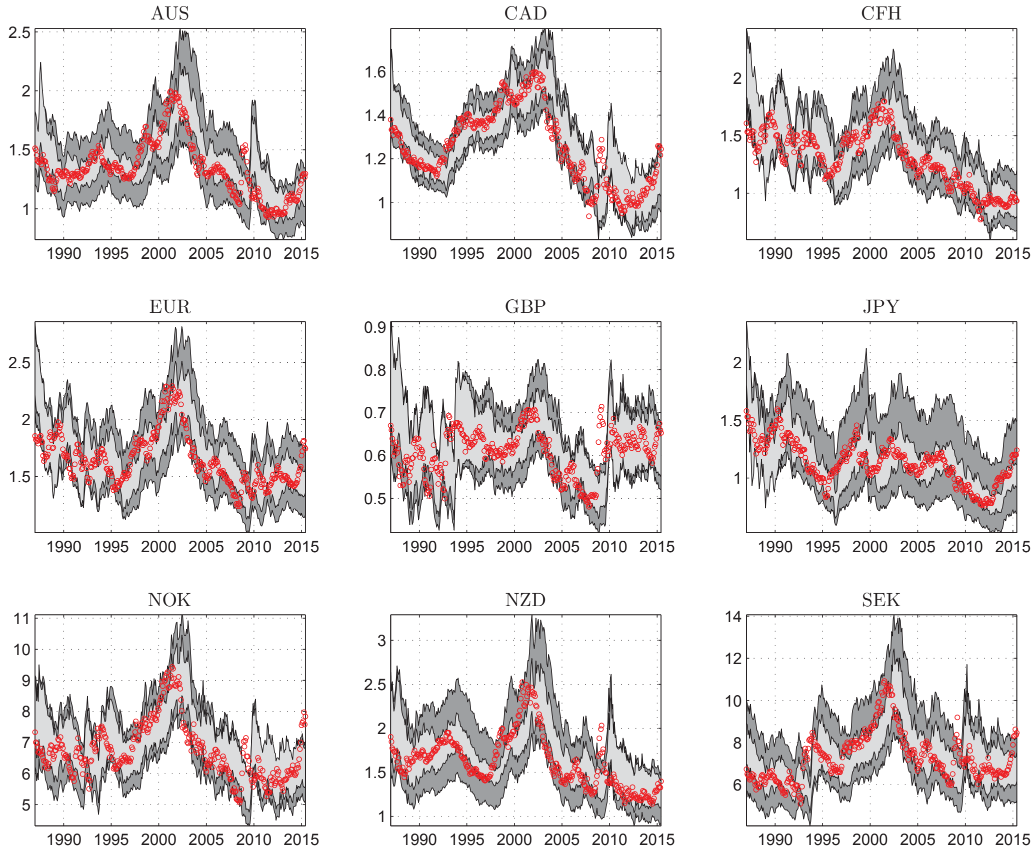
Despite yielding the best point forecasts, a random walk with recursively estimated variance delivers coverage rates which are systematically different from their nominal counterparts. To clarify this result, figure 2 shows the 68% forecast confidence intervals at a one-year horizon of the random walk (dark grey area) and of the TVP AR (light grey area), along with the actual exchange rates (dotted lines). It is evident that with very few exceptions, the random walk overestimates forecast uncertainty and delivers confidence intervals which are excessively large. On the contrary, the TVP AR model does a better job at estimating forecast uncertainty. Table 3 shows that at short horizons, the actual coverage rates are not systematically different from their nominal counterparts. However, the performance worsens as the forecast horizon increases. Controlling for all exchange rates and for macro fundamentals in a time-varying framework marginally improves the coverage rates at medium and long horizons, especially at a 95% coverage level, as the results of the TVP FFVAR and TVP FFVAR MACRO models show.

Overall, all three constant parameter models tend to overestimate forecast uncertainty, as

[9] No systematic differences are detected when splitting the sample in 2007m1, as for the point forecasts.

[10] An independence test as in Christoffersen (1998) systematically rejects the null hypothesis of correct *conditional* coverage for all forecast models, and the results are therefore not reported. This indicates that all models suffer from a problem of clustered outliers: i.e., the probability of a wrong coverage is likely to be affected by past performance.

Figure 2: **Forecast confidence intervals**: 68% forecast confidence intervals delivered by the random walk (dark grey area) and by the TVP AR model (light grey area) at a one-year horizon. Dotted lines denote actual exchange rate levels.



shown by the coverage rates of the BVAR, BVAR MACRO and RW models, which are larger on average than their empirical counterparts.

4.2 Density forecasts

Though point and interval forecasts are clearly interesting, it remains to be assessed which model provides the best characterisation of the forecast distribution of the nine exchange rates. This is relevant in order to attach reliable probabilities to all possible future realisations of the variables analysed. To this purpose, we rank the density forecasts of the competing models based on two scoring rules commonly used in the literature: logarithmic scores and continuous ranked probability scores.

Table 3: **Coverage rates: Unconditional coverage test results**. Main table values denote actual coverage rates, corresponding to nominal coverages of 68% and 95%, for different models and forecast horizons h over the full forecast sample. Values in bold are not significantly different from the nominal counterpart, according to a test of unconditional coverage at a 5% significance level.

		68%									95%								
MODEL	h	AUS	CAD	CFH	EUR	GBP	JPY	NOK	NZD	SEK	AUS	CAD	CFH	EUR	GBP	JPY	NOK	NZD	SEK
RW	1	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	6	0.98	0.95	0.99	1.00	0.96	0.98	0.95	0.98	0.98	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
	12	0.94	0.88	0.94	0.94	0.91	0.97	0.86	0.96	0.94	1.00	0.99	1.00	1.00	0.99	1.00	1.00	1.00	1.00
	24	0.81	**0.65**	0.87	0.83	0.79	0.95	0.76	0.86	0.85	1.00	0.98	0.99	0.99	0.97	0.98	0.99	1.00	1.00
TVP AR	1	**0.70**	**0.69**	**0.69**	**0.73**	**0.71**	**0.70**	**0.71**	**0.73**	**0.71**	**0.95**	**0.93**	**0.95**	**0.95**	**0.96**	**0.97**	**0.95**	**0.94**	**0.95**
	6	0.59	**0.67**	**0.67**	0.54	**0.72**	**0.69**	**0.66**	**0.68**	**0.66**	0.91	0.91	**0.97**	0.93	**0.95**	**0.94**	**0.93**	**0.94**	0.93
	12	0.59	0.60	**0.69**	0.58	**0.69**	**0.69**	**0.64**	0.60	**0.63**	0.87	0.91	0.99	0.92	**0.95**	**0.96**	**0.94**	0.91	0.91
	24	0.54	0.54	**0.68**	**0.65**	**0.67**	**0.64**	0.61	0.52	0.57	0.86	0.88	0.98	0.92	**0.93**	**0.95**	0.98	0.85	**0.93**
TVP FFVAR	1	**0.71**	**0.71**	**0.71**	**0.72**	0.76	**0.69**	**0.71**	0.75	**0.72**	**0.96**	**0.93**	**0.96**	**0.96**	**0.97**	**0.96**	**0.95**	**0.94**	**0.96**
	6	0.62	**0.65**	**0.67**	0.59	0.76	**0.68**	**0.70**	0.75	**0.72**	0.91	0.92	**0.93**	0.92	**0.93**	**0.96**	0.91	**0.95**	**0.93**
	12	0.60	0.62	**0.67**	0.56	**0.69**	0.74	**0.66**	**0.68**	**0.70**	0.89	**0.93**	**0.97**	**0.93**	**0.95**	0.98	0.90	0.92	0.91
	24	0.62	0.59	**0.71**	**0.70**	**0.71**	0.77	**0.65**	**0.69**	**0.67**	0.92	**0.95**	1.00	1.00	**0.93**	0.98	0.97	**0.93**	**0.97**
BVAR	1	0.73	**0.67**	0.80	0.77	0.79	0.76	**0.70**	0.77	**0.73**	**0.97**	0.91	0.97	**0.97**	0.98	0.98	**0.94**	**0.96**	**0.95**
	6	**0.72**	0.75	0.87	0.62	0.85	0.74	**0.72**	0.78	**0.70**	**0.96**	**0.93**	1.00	**0.96**	0.97	**0.97**	**0.94**	**0.96**	**0.96**
	12	0.75	0.81	0.95	**0.67**	0.89	0.77	0.74	0.76	**0.69**	**0.95**	**0.96**	1.00	0.99	0.98	0.99	**0.97**	0.97	**0.96**
	24	0.83	0.86	1.00	0.78	0.92	0.79	0.83	0.77	0.73	1.00	0.99	1.00	1.00	1.00	0.99	1.00	0.99	0.98
TVP FFVAR MACRO	1	0.74	**0.71**	0.73	**0.71**	0.76	0.73	**0.72**	0.74	**0.69**	**0.97**	**0.93**	**0.95**	**0.96**	**0.96**	**0.97**	**0.94**	**0.94**	**0.96**
	6	**0.66**	**0.72**	0.77	**0.66**	0.80	0.73	0.76	0.76	0.77	**0.94**	**0.95**	**0.97**	**0.93**	**0.95**	**0.97**	**0.93**	**0.97**	**0.94**
	12	**0.66**	0.76	0.79	0.62	0.83	0.82	0.75	0.80	0.75	**0.96**	**0.96**	0.99	0.97	**0.96**	0.99	**0.94**	**0.96**	**0.94**
	24	0.81	0.83	0.86	0.82	0.85	0.90	0.82	0.85	0.78	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.98
BVAR MACRO	1	**0.72**	**0.68**	0.79	0.84	0.81	0.76	**0.72**	0.76	0.73	**0.97**	0.91	0.97	0.99	0.98	0.97	**0.94**	**0.96**	**0.95**
	6	**0.66**	0.79	0.86	0.95	0.88	0.75	0.80	0.79	**0.72**	**0.94**	**0.94**	1.00	1.00	0.97	**0.97**	**0.96**	0.97	**0.95**
	12	**0.66**	0.83	0.97	0.99	0.94	0.83	0.84	0.80	**0.72**	**0.94**	**0.96**	1.00	1.00	0.98	0.98	0.99	**0.97**	**0.95**
	24	0.77	0.86	1.00	1.00	0.93	0.82	0.90	0.84	0.77	**0.95**	1.00	1.00	1.00	1.00	0.98	1.00	0.98	**0.97**

Logarithmic scores: Logarithmic scoring rules are based on the evolution over time of log-predictive likelihoods, i.e. the log likelihood of observing the actual realisation of the variable, given a forecast model:

$$\log g_{j,\,h,\,t}(y_{t+h}\,|\,\mathcal{F}_{j,\,t}) \ , \tag{7}$$

where $g_{j,\,h,\,t}(\cdot)$ denotes the predictive likelihood of model j at horizon h (possibly time-varying and thus depending on time t), y is a target variable, and $\mathcal{F}_{t,\,j}$ is the information set of model j which includes the actual value of the variables up to time t and the forecasts of the remaining variables. Of interest is the cumulative difference between two competing prediction models:

$$\mathcal{S}_{j,h} = \sum_{t=1}^{T_f-h} \left[\log g_{1,\,h,\,t}(y_{t+h}\,|\,\mathcal{F}_{1,\,t}) - \log g_{2\,h\,t}(y_{t+h}|\,\mathcal{F}_{g_j,\,t}) \right] . \tag{8}$$

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 first model $g_{1,\,h,\,t}(\cdot)$. Moreover, the statistic in (8) can be interpreted as the summed difference in density forecast errors and can be justified in terms of the Kullback-Leibler distance (KLIC).[11] Hence, when two different predictive densities are compared, the average difference between their logarithms is directly related to their relative KLIC distance. Among a class of alternative models, choosing the one with the highest average log-predictive likelihood entails selecting the model with the minimal distance to the true data generating process.

Figures 3 and 4 plot the statistic $\mathcal{S}$ in equation (8) at a one-month and two-year forecast horizons,[12] where the benchmark is the time-varying forgetting factor univariate model. Since all models perform systematically better than a random walk, this comparison is not shown but available upon request. At short horizons, all models perform better than the univariate time-varying variant, suggesting that controlling for additional information matters for generating accurate density forecasts. Time variation instead appears not to be important, as time-varying parameter models perform similarly to constant-parameter ones.

At longer horizons the time-varying parameter model with only exchange rates delivers higher predictive likelihoods for half the currencies in the sample.

[11] Under some regularity conditions, the average of the sample quantities of the true log predictive density $f_t(\cdot)$, and of the predictive density of model j, $g_{j,t}(\cdot)$, yields a consistent estimator of the KLIC distance. For a more detailed discussion, see Hall and Mitchell (2007).

[12] Results for other horizons are not shown as they are very similar to those presented in the text, but are available upon request.

Figure 3: **Cumulative differences in log-predictive likelihoods,** $h = 1$, relative to the TVP AR model. Increases in the plotted statistics denote observations in favour of the TVP FFVAR (solid line), BVAR (dashed dotted line), TVP FFVAR MACRO (thicker dashed line), and the BVAR MACRO (thicker dotted line).

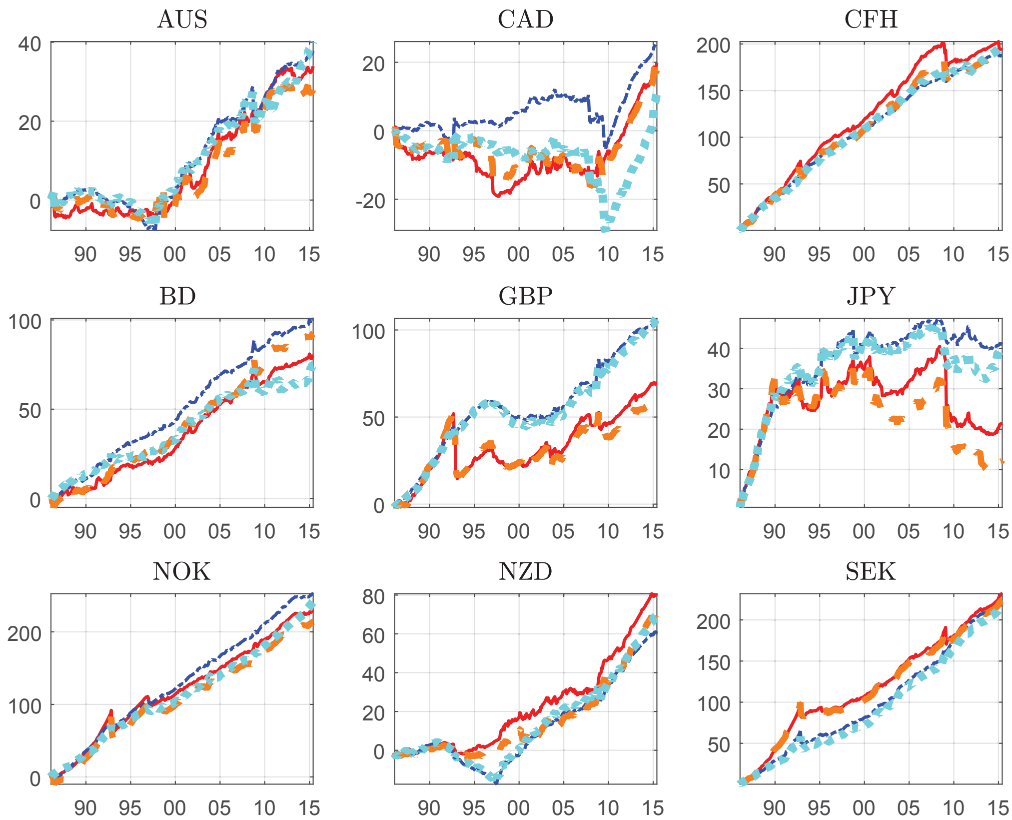
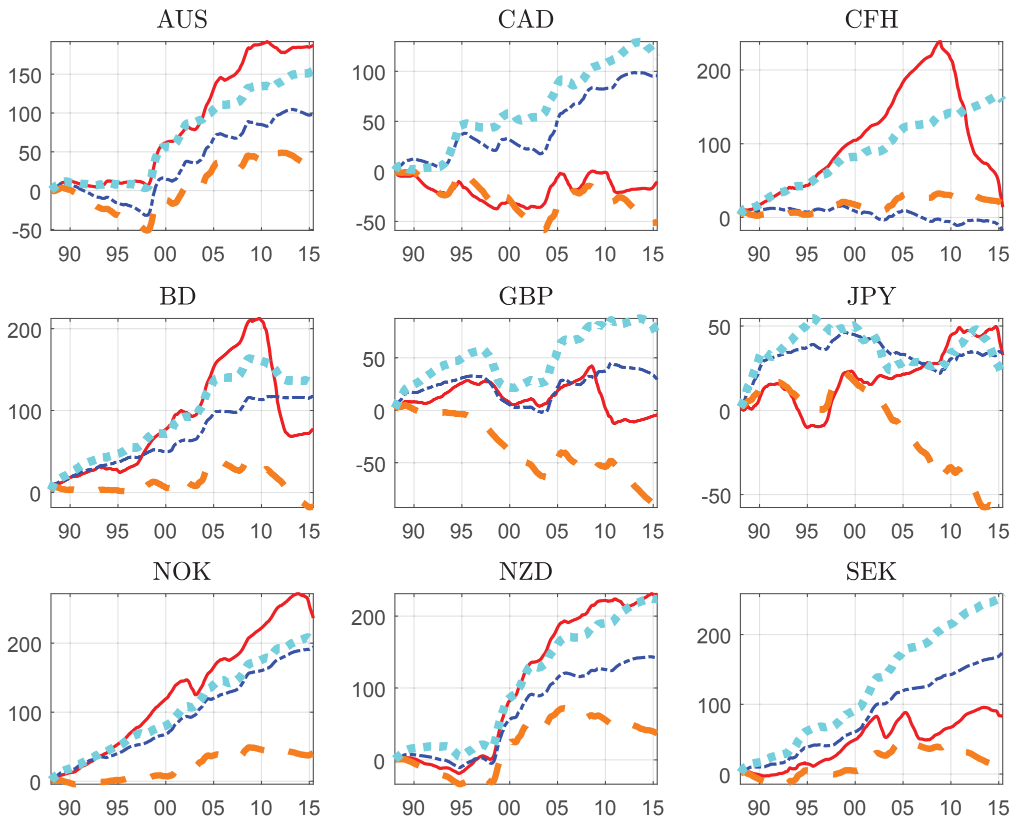


Figure 4: **Cumulative differences in log-predictive likelihoods,** $h = 24$, relative to the TVP AR model. Increases in the plotted statistics denote observations in favour of the TVP FFVAR (solid line),BVAR (dashed dotted line), TVP FFVAR MACRO (thicker dashed line), and the BVAR MACRO (thicker dotted line).



The better performance is however limited to the high-volatility period between 2000 and 2010 and, for some currencies, it disappears after 2010. Controlling for macroeconomic predictors yields better density forecasts if and only if there is no time variation in the data generating process. A constant-parameter Bayesian VAR augmented with macro fundamentals and estimated recursively over the forecast sample generates in fact higher predictive likelihoods than the competing models for one third of the currencies.

Continuous ranked probability scores: Logarithmic scores tend to be sensitive to the distance between the centre of the forecast density and the realised outcome. To overcome this problem, we employ a different scoring rule, the continuous ranked probability score (CRPS), defined in terms of predictive cumulative distribution functions. As surveyed in Gneiting et al. (2007), this scoring rule rewards the "sharpness" of a forecasting rule, i.e. the concentration of the forecast density around its centre, and is less sensitive to distance. Given a realised outcome y_{t+h} and the cumulative predictive density $G_{j,\,h}$ of model j at horizon h, the associated CRPS with negative orientation is expressed as:

$$\mathrm{CRPS}(G_{j,\,h}, y) = E_{G_{j,\,h}} |\hat{y}_{j,t+h} - y_{t+h}| - \frac{1}{2} E_{G_{j,\,h}} |\hat{y}_{j,t+h} - \hat{y}'_{j,t+h}| \; , \tag{9}$$

where $E_{G_{j,\,h}}$ is the expectation for the predictive density of model j at horizon h, while $\hat{y}_{j,t+h}$ and $\hat{y}'_{j,t+h}$ are h-step ahead forecasts for the realisation y_{t+h} obtained from model j. Note that the rule above nests the mean absolute error in the case of point, rather than probabilistic, forecasts.

We compute the CRPS for the different forecast models using a rolling window of five years and, following Ravazzolo and Vahey (2014), construct the relative weight of each model in period t according to the rule:

$$w_{j,\,h,\,t} = \frac{\sum_{\tau=t-s}^{t} \Gamma_{j,\,h,\,\tau}}{\sum_{j=1}^{N} \left[\sum_{\tau=t-s}^{t} \Gamma_{j,\,h,\,\tau} \right]} \; , \tag{10}$$

where $\Gamma_{j,\,h,\,t}$ is the inverse of equation (9) computed for model j at horizon h over the period $[t - s,\, t]$, $s = 60$, and N is the total number of models considered. A lower CRPS value (lower density forecast error) translates into a higher relative weight of model j.

Figures 5 and 6 show the weights, relative to the time-varying autoregressive model, at a one-month and two-year forecast horizons. At short horizons, all models have similar and relatively higher weights than the univariate time-varying model. Towards the end of the sample however, the TVP AR receives progressively more weight, as the volatility of the currencies increases. The cutoff date varies across currencies but is around the year 2005. At longer horizons, the TVP FF VAR model receives on average a higher weight over the whole sample, and in particular in the relatively more volatile period of the early 1990s and between 2000 and 2010. Models augmented with macroeconomic fundamentals receive on average lower weights than their counterparts with only exchange rates.

Both scoring rules have highlighted the difficulties in finding a model that performs uniformly better than the others across forecast periods, horizons and exchange rates. This calls for potentially positive gains from a forecast combination that assigns a weight on each model, proportional to its forecast performance. To explore this issue, we use the CRPS-based weights in (10) to construct point forecasts by optimally combining the competing models. In particular, we compute an optimal forecast combination of all models (DMA MACRO), and of the models containing only exchange rates (DMA). We distinguish between the two in order to isolate the contribution of macroeconomic predictors. Point forecasts relative to a random walk are reported in Table 4 and highlight two main results. First, the two dynamic forecast combinations deliver more accurate point forecasts than a random walk only at a one-month ahead horizon. Second, controlling for the two models augmented with macroeconomic predictors lowers point forecast errors on average. This is particularly evident in the forecast subsample that precedes the financial crisis where controlling for the macro models improves the point forecasts of the Euro, the New Zealand dollar and, to a lesser extent, of the Swiss Franc and the Yen.

Figure 5: **CRPS-based weights,** $h = 1$: Weights relative to the TVP AR model, computed using a 5-year window. Higher weights denote a better performance of the TVP FFVAR (solid line), BVAR (dashed dotted line), TVP FFVAR MACRO (thicker dashed line), and the BVAR MACRO (thicker dotted line).

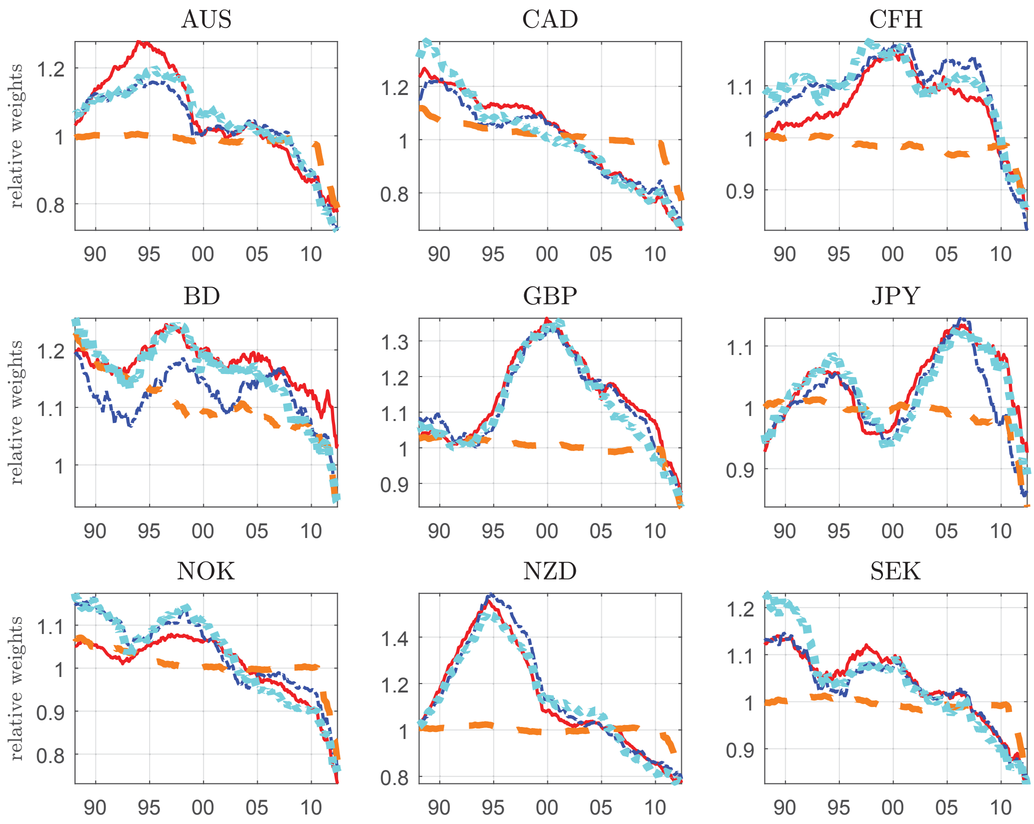


Figure 6: **CRPS-based weights,** $h = 24$: Weights relative to the TVP AR model, computed using a 5-year window. Higher weights denote a better performance of the TVP FFVAR (solid line), BVAR (dashed dotted line), TVP FFVAR MACRO (thicker dashed line), and the BVAR MACRO (thicker dotted line).

Table 4: **Mean squared forecast errors of the forecast combinations relative to a random walk** for different forecast samples and horizons. Forecast combinations pertain to all models (DMA MACRO), or to the models with only exchange rates (DMA). Two (one) stars denote significantly different RMSFE at a 5% (10%) significance level according to a Diebold-Mariano test, modified using the small-sample correction of Harvey et al. (1998). The forecast subsamples are defined as follows: PRE-CRISIS SAMPLE (1986:m2 - 2006:m12), CRISIS SAMPLE (2007:m1 - 2013:m6) and FULL SAMPLE (1986:m2 - 2013:m6).

		PRE-CRISIS SAMPLE				CRISIS SAMPLE				FULL SAMPLE			
	MODEL $\downarrow$ / $h \rightarrow$	1	6	12	24	1	6	12	24	1	6	12	24
AUS	dma	1.00	1.04	1.13	1.32	0.98**	0.99	0.97*	0.89**	0.99*	1.01	1.06	1.17
	dma macro	0.99	1.00	1.05	1.21	0.99	1.02	1.04	0.94**	0.99	1.01	1.06	1.14
CAD	dma	1.03	1.18	1.29	1.44	0.99**	1.02	1.06	1.07	1.00	1.08	1.17	1.32
	dma macro	1.02	1.12	1.18	1.29	0.99	1.04	1.15	1.21	1.00	1.07	1.16	1.26
CFH	dma	0.98**	1.01	1.05	1.14	1.00	1.10	1.16	1.13	0.99*	1.03	1.07	1.15
	dma macro	0.97**	0.96	0.96	1.09	0.99	1.05	1.09	0.88**	0.97**	0.98	0.99	1.08
EUR	dma	1.00	1.03	1.08	1.12	1.01	1.02	1.08	1.12	1.00	1.03	1.08	1.13
	dma macro	0.97**	0.94*	0.93*	1.01	1.02	1.08	1.25	1.51	0.99	0.99	1.02	1.10
GBP	dma	1.00	1.00	0.98	1.06	0.98	0.98	1.02	0.95**	0.99	1.00	1.00	1.04
	dma macro	0.99	1.02	0.98	1.05	0.99	1.04	1.17	1.17	0.99	1.03	1.05	1.11
JPY	dma	1.00	1.00	0.96	1.07	0.99	1.00	0.98	0.90**	1.00	1.00	0.96	1.01
	dma macro	0.98	0.96	0.89**	1.00	0.99	1.00	0.97**	0.98	0.99	0.97	0.92**	0.99
NOK	dma	0.98**	1.03	1.08	1.18	0.97**	0.98**	0.97*	0.79**	0.98**	1.00	1.04	1.11
	dma macro	0.98**	1.01	1.02	1.09	0.97**	1.00	1.08	0.93**	0.97**	1.01	1.06	1.11
NZD	dma	0.98*	1.01	1.09	1.19	0.97*	0.96	0.92**	0.72**	0.98**	0.99	1.03	1.10
	dma macro	0.98**	0.94**	0.95**	1.04	0.98*	1.02	1.07	0.83**	0.98**	0.98	1.02	1.03
SEK	dma	1.00	1.04	1.06	1.13	0.98*	0.99	1.00	0.84**	0.99	1.02	1.04	1.08
	dma macro	0.99	1.01	0.99	1.05	0.98*	1.02	1.11	0.98	0.99**	1.02	1.04	1.07

4.3 Discussion: decomposing parameter time variation

We have seen that accounting for parameter time-variation is better suited in the case of the model with only exchange rates. In addition, its performance relative to the constant-parameter BVAR varies over time. We explore here which source of parameter time variation matters the most: whether it is in the parameters of the slope coefficients, or in the volatility of the innovations. To this purpose, we experiment with different parameterisations of the TVP FFVAR models and switch off the variation either in the slope coefficients or in the volatilities. This is achieved by either setting the forgetting factor λ to 1, or the decay factor κ to 0.99.[13]

Figures 7a and 7b report the joint predictive likelihoods of the nine currencies from the TVP FFVAR model with only exchange rates, relative to those from a model with no time variation in the slope parameters, or with little time variation in the volatilities. Results show that the forecast improvement derived from allowing time variation in the slope parameters is negligible, as there are very little differences in the log-predictive likelihoods of the two models. On the contrary, modelling time-varying volatility increases the log-predictive likelihood, but is less important after the financial crisis.

Allowing for time variation in the slope parameters is actually detrimental for the forecasting performance of the model with exchange rates and fundamentals, as it can be seen by the mostly negative values in Figure 7a. This is in line with the good forecasting results of the BVAR MACRO model. Instead, allowing for a higher degree of variation in the volatilities improves the forecast performance, but only up to one-year ahead (see Figure 8b).

5 Economic evaluation

So far, we have relied on purely statistical criteria to evaluate the competing exchange-rate forecast models. However, an evaluation based on economic criteria might be of interest as well, particularly if the statistical models are to be used in real-world applications. To this purpose, we follow Della Corte et al. (2009) and assume that the different forecast models are used to build dynamic trading strategies with a one-month horizon, which are then evaluated in a mean-variance framework. In particular, we take the perspective of a US based investor, who in each period optimally allocates his wealth by buying a portfolio of home and foreign bonds that redeem in the next period. We outline the different steps in what follows.

[13]The closest κ is to unity, the closest is the covariance matrix of the VAR innovations to its initial value.

Figure 7: TVP FFVAR: Cumulative differences in joint log-predictive likelihoods (increases denote observations in favour of the TVP FFVAR model)

(a) Relative to the constant slope variant (b) Relative to the constant volatility variant

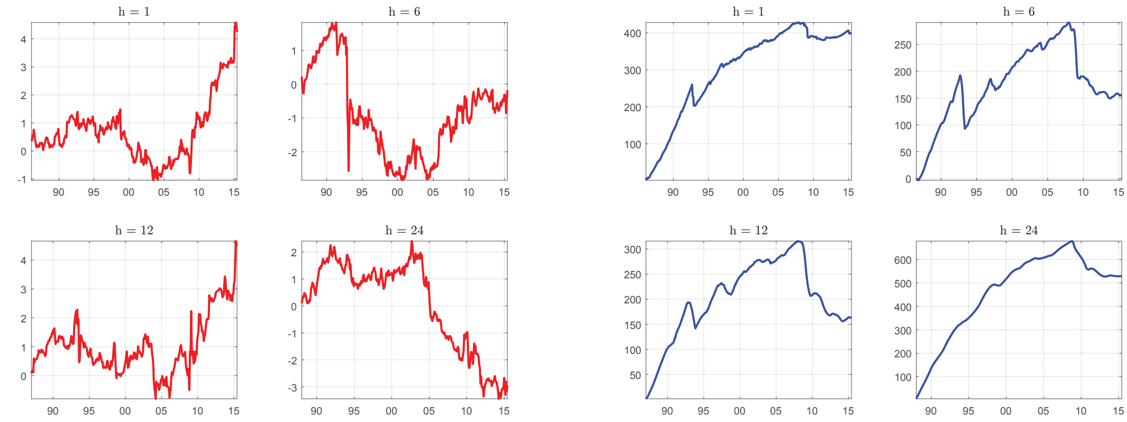
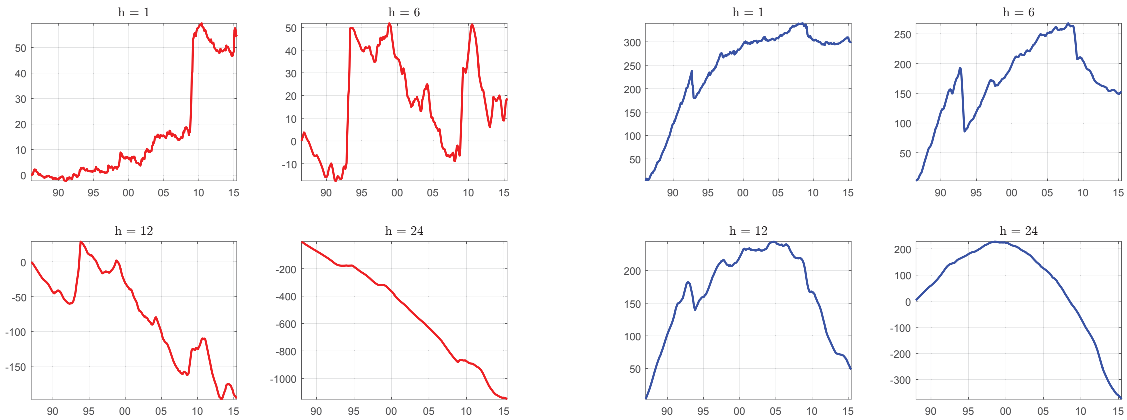


Figure 8: TVP FFVAR MACRO: Cumulative differences in joint log-predictive likelihoods (increases denote observations in favour of the TVP FFVAR MACRO model)

(a) Relative to the constant slope variant (b) Relative to the constant volatility variant



The return on home and foreign bonds is assumed to be safe and equalised across countries, so that the riskiness of foreign bonds originates solely from the future value of the exchange rate. Based on the expectation of the conditional mean and variance of next period's exchange rates, the investor will optimally allocate his wealth in the various bonds by solving the following problem:

$$\max_{w_t} w_t' \mu_{t+1|t} + (1 - w_t' \iota) r_{t+1}^f \quad \text{s.t.} \;\; w_t' \Sigma_{t+1|t} w_t = \bar{\sigma}^2 \tag{11}$$

where w_t is a $M \times 1$ vector of portfolio weights on the risky assets, r_{t+1}^f is the log return on the safe asset and $\bar{\sigma}^2$ is a target conditional volatility of the portfolio returns, which we fix to be 10%. We denote now with Δs_{t+1} the percentage exchange-rate returns, where the exchange rate $s_t = \log e_t$ is the number of currency units needed to purchase 1 dollar (in logs). Then, each forecast model provides the information needed to construct the conditional expected mean and variance of the risky assets returns, respectively denoted as $\mu_{t+1|t}$ and $\Sigma_{t+1|t}$ and defined by:

$$\mu_{t+1|t} = E_t[r_{t+1}] = E_t[r^f_{t+1} - \Delta s_{t+1}] \tag{12}$$

$$\Sigma_{t+1|t} = E_t[(r^f_{t+1} - \Delta s_{t+1} - \mu_{t+1|t})(r^f_{t+1} - \Delta s_{t+1} - \mu_{t+1|t})' \tag{13}$$

The solution to the optimisation problem yields the optimal portfolio composition, i.e. the weights on the risky assets:

$$w_t = \bar{\sigma}\left[(\mu_{t+1|t} - \iota r^f_{t+1})'\Sigma^{-1}_{t+1|t}(\mu_{t+1|t} - \iota r^f_{t+1})\right]^{-\frac{1}{2}}\Sigma^{-1}_{t+1|t}(\mu_{t+1|t} - \iota r^f_{t+1}) \tag{14}$$

We define the trading strategy associated to forecast model j as the optimal choice of portfolio weights, given the forecast model predictions of $\mu_{t+1|t}$ and $\Sigma_{t+1|t}$. The different trading strategies (portfolio choices), are then evaluated within a mean-variance framework. Following Della Corte et al. (2009) and West et al. (1993), we assume that the utility of the investor is increasing in wealth and decreasing in its variance, with constant degree of relative risk aversion δ, which we fix to 6. This enables us to consistently estimate the expected utility generated by an initial level of wealth W_0 through the average realised utility $\bar{U}$:

$$\bar{U} = W_0 \sum_{t=0}^{T-1}\left[R_{j,\,t+1} - \frac{\delta}{2(1+\delta)}R^2_{j,\,t+1}\right], \tag{15}$$

where $R_{j,\,t+1} = 1 + (1 - w'_{j,t}\iota)r^f_{t+1} - w'_{j,t}r_{t+1}$ is the gross realised return on portfolio j . Assuming a constant risk aversion further simplifies the problem by allowing the initial level of wealth W_0 to be normalised to unity.[14]

A strategy j is better than a benchmark if it leads to a higher average realised utility. So, following Della Corte et al. (2009), we calculate the maximum fee Φ that the investor is willing to pay to switch from the benchmark trading strategy b to strategy j as the value that equalises the average realised utilities obtained from the two strategies:

$$\sum_{t=0}^{T-1}\left[(R_{j,\,t+1} - \Phi) - \frac{\delta}{2(1+\delta)}(R_{j,\,t+1} - \Phi)^2\right] = \sum_{t=0}^{T-1}\left[R_{b,\,t+1} - \frac{\delta}{2(1+\delta)}R^2_{b\,t+1}\right], \tag{16}$$

[14] As shown in West et al. (1993), a constant relative risk aversion implies in fact that the expected utility is linearly homogeneous in wealth.

where the benchmark strategy is a random walk, i.e., it entails buying the safe domestic asset.

Additionally, since all trading strategies beside the benchmark are dynamic, we need to take into account transaction costs. The positive realised utility might in fact be undone, if a given strategy requires many portfolio changes and if transaction costs are sufficiently high. Following Della Corte and Tsiakas (2013) and Han (2006) we calculate the break-even transaction cost τ that makes investors indifferent between two strategies. A strategy is preferred to the benchmark only if actual transaction costs are lower than the breakeven cost.

Table 5 reports for each strategy the maximum fee Φ, the break-even transaction cost τ, as well as the average realised portfolio return $\bar{R}_j$, the portfolio standard deviation $\sigma_{R,j}$ and its Sharpe ratio (SR).[15] Controlling for time-varying parameters and macroeconomic fundamentals yields the highest portfolio returns. In addition, all time-varying models yield higher returns than their constant-parameter counterparts, and conditioning on additional information improves the univariate-based strategy. However, the portfolios based on time-varying parameter models are on average more volatile than the benchmark strategy, and this is reflected on the latter having the highest Sharpe ratio. Moreover, since the differences in portfolio returns are small, both the maximum performance fees and breakeven transaction costs are also rather low. Nevertheless, results point towards a positive contribution of macroeconomic fundamentals. First, the strategy with the highest performance fee is the one based on the time-varying model augmented with fundamentals. Second, the strategy with the highest breakeven transaction cost is the one based on the forecast combination that employs all competing models. In particular, breakeven transaction costs increase by 45% once models augmented with macro fundamentals are accounted for in the forecast combination.

6 Conclusions

A big puzzle in the foreign exchange literature is the inability to predict the future behaviour of exchange rates. Moreover, most of the literature evaluates competing forecast models based on their point forecasts relative to a random walk. Little attention has so far been paid to how well the models characterise forecast uncertainty, a feature likely to be of interest for many real-world applications.

In this paper we conduct a comprehensive statistical and economic evaluation of exchange

[15]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.

Table 5: **Trading strategies results** based on the different forecast models. The statistics reported are the average portfolio return $\bar{R}_j$, the portfolio standard deviation $\sigma_{R,\,j}$ and Sharpe ratio ($SR_j = \frac{\bar{R}_j}{\sigma_{R,\,j}}$), the maximum performance fee Φ an investor is willing to pay to switch from the benchmark random walk strategy (RW) to strategy j (in annual percentage points), and the percentage transaction costs τ that would have to be paid in each period to cancel the positive utility derived from strategy j (in basis points, and reported only if positive). The target conditional volatility of the portfolio returns $\bar{\sigma}^2$ is 10%, and the constant degree of relative risk aversion δ is fixed to 6. Bold values denote the best-performing strategy for each statistic.

MODEL	$\bar{R}$	σ_R	SR	Φ	τ
tvp ar	4.96	2.85	1.74	0.17	1.64
tvp ffvar	5.04	2.95	1.71	0.32	5.43
bvar	4.90	2.85	1.72	0.11	–
tvp ffvar macro	**5.07**	3.03	1.68	**0.43**	3.94
bvar macro	4.95	2.89	1.72	0.19	0.68
dma	5.02	2.83	1.78	0.21	7.54
dma macro	5.05	2.87	1.76	0.28	**10.91**
rw	4.92	**2.68**	**1.83**	–	–

rate forecast models, and gauge whether exchange-rate unpredictability can be resolved by modelling time variation in the parameters of the underlying stochastic processes. In particular, we assess whether the point, interval and density forecasts of nine major currencies vis-a-vis the US dollar can be improved by controlling for other currencies and for macroeconomic fundamentals, as well as by allowing for parameter time-variation in the slope and volatility parameters. Time variation is modelled through a parsimonious set up based on forgetting factors recently proposed by Koop and Korobilis (2013), which determines the amount of parameter time variation on the basis of current forecast errors.

We find that modelling parameter time variation significantly improves the estimation of forecast uncertainty. Time-varying parameter models deliver in fact 68% and 95% forecast confidence intervals which are on average accurately calibrated. Conversely, forecast confidence intervals from constant-parameter models are excessively large, thus overestimating forecast uncertainty. Moreover, density forecast results indicate an advantage of time-varying parameter models in high-volatility periods, especially in the one between 2000 and 2010. On the other hand, we do not find evidence that parameter time variation helps to unravel exchange rate predictability by macroeconomic fundamentals. Models augmented with macroeconomic differentials perform in fact better when no parameter time variation is allowed and, at long horizons, yield higher predictive likelihoods for one third of the currencies in the sample. An optimal forecast combination that uses weights based on the relative density forecast performances further reveals that controlling for macroeconomic predictors lowers point forecast errors for half the currencies in the sample, though often not significantly.

Lastly, we have used the competing models to build trading strategies, evaluated within a mean-variance framework. Results show that allowing for parameter time variation and controlling for macroeconomic fundamentals leads to higher portfolios returns, and to higher values for investors, though the differences with respect to a benchmark random walk strategy are not economically large.

References

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

Amisano, G. and Geweke, J. (2013). Prediction using several macroeconomic models. Working Paper Series 1537, European Central Bank.

Bacchetta, P. and Wincoop, E. V. (2009). On the unstable relationship between exchange rates and macroeconomic fundamentals. NBER Working Papers 15008, National Bureau of Economic Research, Inc.

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

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

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

Cheung, Y.-W. and Chinn, M. D. (2001). Currency traders and exchange rate dynamics: a survey of the us market. *Journal of International Money and Finance*, 20(4):439–471.

Cheung, Y.-W., Chinn, M. D., and Pascual, A. G. (2005). Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24(7):1150–1175.

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, 39(4):841–62.

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

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

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

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

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

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

Gneiting, T., Balabdaoui, F., and E. Raftery, A. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268.

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

Han, Y. (2006). Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model. *Review of Financial Studies*, 19(1):237–271.

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

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

Koop, G. and Korobilis, D. (2013). Large time-varying parameter vars. *Journal of Econometrics*, 177:185–198.

Meese, R. and Rogoff, K. (1983). 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.

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

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

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

Ravazzolo, F. and Vahey, S. P. (2014). Forecast densities for economic aggregates from disaggregate ensembles. *Studies in Nonlinear Dynamics & Econometrics*, 18(4):15.

Rogoff, K. S. and Stavrakeva, V. (2008). The continuing puzzle of short horizon exchange rate forecasting. NBER Working Papers 14071, National Bureau of Economic Research, Inc.

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

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, 51(4).

West, K. D., Edison, H. J., and Cho, D. (1993). A utility-based comparison of some models of exchange rate volatility. *Journal of International Economics*, 35(1-2):23–45.

Yongmiao, H., Haitao, L., and Feng, Z. (2007). 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.

Appendix: Database description

Table 6: **Database description:** All series below are monthly data and dowloaded from Datastream. Exchange rates are monthly averages of daily data. The yen is scaled down by 100. The German Mark series, after January 1999, is equal to the Euro scaled by the fixed conversion rate of 1.96. Macro variables are expressed as differentials from the US counterpart. Where needed, seasonality has been removed using a 13-term moving average filter. Industrial production indexes for Australia and Canada are back-casted using the EM algorithm. Finally, we drop the industrial production index for New Zealand on the account of data length.

VARIABLE	MNEMONIC	UNIT	SOURCE
Australian Dollar to US Dollar	AUXRUSD	–	Bank of England
Canadian Dollar to US Dollar	CNUSBOE	–	Bank of England
Swiss Franc to US Dollar	SWUSBOE	–	Bank of England
German Marks to US Dollar	BDXRUSD.	–	Bank of England
British Pound to US Dollar	BDXRUSD.	–	Bank of England
Japanese Yen to US Dollar	JPUSBOE	–	Bank of England
Norwegian Krone to US Dollar	NWUSBOE	–	Bank of England
New Zealand Dollar to US Dollar	NZUSBOE	–	Bank of England
Swedish Krona to US Dollar	SDUSBOE	–	Bank of England
CPI, Australia	AUCCPI..E	Price Index (SA)	Australian Bureau of Statistics
CPI, Canada	CNCONPRCF	Price Index (NSA)	CANSIM - Statistics Canada
CPI, Switzerland	SWCONPRCE	Price Index (SA)	KOF - Swiss Economic Institute
CPI, Germany	BDCONPRCE	Price Index (SA)	Deutsche Bundesbank
Retail Sales, UK	UKRETTOTG	Price Index (SA)	ONS, UK
CPI, Japan	JPCONPRCF	Price Index (NSA)	Statistics Bureau, Japan
CPI, Norway	JPCONPRCF	Price Index (NSA)	Statistics Norway
CPI, New Zealand	NZCCPI..E	Price Index (SA)	Statistics New Zealand
CPI, Sweden	SDCONPRCF	Price Index (NSA)	SCB - Statistics Sweden
CPI, United States	USCONPRCE	Price Index (SA)	Bureau of Labor Statistics
Industrial Prod., Australia	AUCIND..G	Price Index (SA)	Australian Bureau of Statistics
Industrial Prod., Canada	CNI66..CE	Price Index (SA)	IMF
Industrial Prod., Switzerland	SWM66..XR	YoY Change (NSA)	IMF
Industrial Prod., Germany	BDI66..CE	Price Index (SA)	IMF
Industrial Prod., UK	UKI66..CE	Price Index (SA)	IMF
Industrial Prod., Japan	JPI66..CE	Price Index (SA)	IMF
Industrial Prod., Norway	NWI66..CE	Price Index (SA)	IMF
Industrial Prod., Sweden	SDI66..CE	Price Index (SA)	IMF
Industrial Prod., United States	USI66..CE	Price Index (SA)	IMF
Gov. Bond Rate (Short-Term), Australia	AUI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Canada	CNI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Switzerland	SWI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Germany	BDI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, UK	UKI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Norway	NWI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Sweden	SDI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, Japan	JPI61A..	Percentage (NSA)	IMF
Gov. Bond Rate, United States	USI61A..	Percentage (NSA)	IMF