Time-varying uncertainty and exchange rate predictability*

Knut Are Aastveit† Francesco Ravazzolo‡ Herman K van Dijk§

Preliminary version: November 29, 2017

Abstract

We study the predictive ability of macroeconomic fundamentals for monthly exchange rates using a Bayesian predictive combination approach. Our approach combines the entire predictive densities of the individual models, rather than only their point forecasts, and extends earlier combination approaches that have been applied to exchange rate models, by allowing for two key features. First, our method features time-varying combination weights, and explicitly factors into the model combination the inherent uncertainty surrounding the estimation of the combination weights. Second, our method allows for model incompleteness, i.e., that the true model is not necessarily a part of the model space. In an empirical exercise, we study the forecasting performance of our combination approach relative to other combination approaches and common benchmarks for seven major exchange rates vis-à-vis the US dollar over the period 2000-2014. Overall, we find that our combination scheme produces markedly more accurate predictions than the existing alternatives, both in terms of statistical and economic measures of out-of-sample predictability.

JEL-codes: C11, C53, E44, F37, G11

Keywords: Exchange rate forecasting, Forecast combination, Instabilities, Model uncertainty, Bayesian filtering

---

*We thank Graham Elliott, Per Mykland, Hashem Pesaran, Jim Stock, Mike West, as well as seminar and conference participants at Duke University, Macquarie University, the 27th Annual Meeting of the Midwest Econometrics Group, the 7th ESOBE meeting, the 4th IAAE conference, Keio University, the Market Microstructure and High Frequency Data conference at the University of Chicago Stevanovich center, Norges Bank, Reserve Bank of New Zealand for useful comments on a preliminary version of this paper. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. This paper is part of the research activities at the Centre for Applied Macro and Petroleum economics (CAMP) at the BI Norwegian Business School.

†Norges Bank and BI Norwegian Business School, Knut-Are.Aastveit@norges-bank.no

‡Free University of Bozen/Bolzano and BI Norwegian Business School, francescoravazzolo@gmail.com

§Econometric Institute, Erasmus University Rotterdam and Norges Bank, hkvandijk@ese.eur.nl

1 Introduction

In a seminal paper, Meese and Rogoff (1983) identified that exchange rate fluctuations are difficult to predict using standard economic models and that a simple a-theoretical model, such as the random walk, is frequently found to generate better exchange rate forecasts than standard fundamental-based economic models.[1] While a vast amount of empirical studies claim to have found a relation between exchange rates and macroeconomic fundamentals, Rossi (2013), in a thorough survey of the recent literature, points out that the answer is not clear cut and that academics and practitioners are yet to find a definite answer as to whether or not macroeconomic variables have predictive content. In her survey, Rossi points out that exchange rate predictability is affected by decisions regarding the choice of predictors, forecast horizons, forecasting model as well as the methods for forecast evaluation. Specifically, the predictive power seems to be specific to some countries in certain periods, signalling the presence of instability in the models' forecasting performance. Several papers, such as Bacchetta and Wincoop (2004, 2013), Sarno and Valente (2009) and Bacchetta et al. (2010), have discussed the issue of instability, but as pointed out by Rossi (2013), models that take account of these instabilities by either allowing for time-variation in the coefficients or by combining forecasts from various models, do not greatly succeed in outperforming the random walk benchmark in out-of-sample forecasting.

In this paper, we introduce a novel Bayesian model combination technique that accounts for time-varying uncertainty of several model and data features in order to provide more accurate and complete density forecasts for exchange rates. We stress that our approach combines the entire predictive densities of the individual models, rather than only their point forecasts. Our combination method extends earlier approaches that have been applied to exchange rate models, by allowing for two key features. First, our method features time-varying combination weights, and explicitly factors into the model combination the inherent uncertainty surrounding the estimation of the combination weights. Second, our method allows for model incompleteness, i.e., that the true model is not necessarily a part of the model space. We conjecture that the lack of of predictive power, earlier found in the literature, is related to these two features.

We examine the out-of-sample performance from our forecasting combination approach in

[1] This finding is often referred to as the Meese and Rogoff puzzle.

both statistical and economic terms. We use monthly data for the observation period 1986M1-2014:M3 and evaluate the forecasting performance for the period 2000:M4-2014M3. We study the forecasting performance for 7 currency pairs, with the US dollar (USD) as the numeraire, including the Australian dollar, Canadian dollar, Norwegian krone, the Euro, the Japanese yen, the Swiss franc, and UK pound sterling.

We find that our combination approach systematically outperforms all benchmarks we compare it to, both in terms of statistical (point and density forecasts) and economic terms. At the one month horizon, the magnitude of reduction in terms of mean squared prediction error (MSPE) and continuous ranked probability score (CRPS) relative to the driftless Random Walk benchmark always exceeds 7% for MSPE and 18% for CRPS. We find the largest gains when we apply our combination approach to single equation models that accounts for time-varying parameters and stochastic volatilities. For several of the countries we then find improvements that are substantially larger, exceeding 15% for MSPE and 35% for CRPS. While the various individual models show considerable instabilities in predictive performance over time, the performance from our combination approach is far more robust, yielding a steady improvement over the various benchmarks over different time periods.

While accounting for time-varying weights and weight uncertainty plays a role in improving the density forecasting performance, the main bulk of the gains, both in terms of point and density forecasting performance, stems from allowing for model incompleteness in the combination scheme. We find that there is a relation between range volatility and the standard deviation of the combination residuals. The latter can be viewed as a measure of model incompleteness. This suggests that the lack of predictability from standard fundamental-based models could be related to these models not being able to appropriately account for volatility dynamics. Finally, in terms of economic performance, based on a mean-variance dynamic asset allocation exercise, we find that our combination approach yields gross returns that are substantially higher than, and in most cases more than double of, the ones from the other benchmarks.

Our paper is related to the research on the role of instabilities in obstructing model forecasting performance and the literature on forecast combinations. Among papers focusing on pooling exchange rate forecasts, we note contributions by Wright (2008), Sarno and Valente (2009), Della Corte et al. (2009), Li et al. (2015), Foroni et al. (2015) and Beckmann and

Schüssler (2016). While these papers focus on combining point forecasts, our approach is motivated by recent advances in macroeconomic forecasting as we combine the entire predictive densities of the individual models, rather than only their point forecasts, see e.g., Hall and Mitchell (2007), Jore et al. (2010), Geweke and Amisano (2011), Aastveit et al. (2014), Billio et al. (2013), Casarin et al. (2015), Pettenuzzo and Ravazzolo (2016), Aastveit et al. (2016) and McAlinn and West (2017).[2]

Among articles studying the importance of instabilities in an exchange rate setting, our paper is related to Bacchetta et al. (2010), Rossi and Sekhposyan (2011), Byrne et al. (2016, 2017) and Kouwenberg et al. (2017). Rossi and Sekhposyan (2011) decompose measures of out-of-sample forecasting performance into components of relative predictive ability. Their results point to a lack of predictive content and time-variation in forecasting performance as the main obstacles to exchange rate models' forecasting ability. On the contrary, Byrne et al. (2017) find that time-variation in parameters of the models is the main cause of time-variation in forecasting performance. Using a framework that allows to pin down several sources of instability that might affect the out-of-sample forecasting performance of exchange rate models, they show that models which embed a high-degree of coefficient variability yield forecast improvements at horizons beyond 1-month. A different approach is taken by Kouwenberg et al. (2017), who design a dynamic model selection rule that captures the current set of fundamentals that best predicts the exchange rate. They allow model weights to vary through time to capture the dynamic relation between exchange rates and fundamentals. In their approach, variables with insignificant forecasting power are eliminated by backward regression, so that only a few, the most relevant, fundamentals remain in the forecasting equation. While we share similarities with all of these papers, we differ in the sense that we allocate the improved forecasting performance to using a framework where we combine the entire predictive densities of the individual models using time-varying combination weights, where we explicitly factor into the model combination the inherent uncertainty surrounding the estimation of the combination weights as well as allowing for model incompleteness.

The structure of the paper is as follows. Section 2 introduces our combination approach. Section 3 describes the essential features of our empirical exercise. In Section 4 we report our

[2]The theoretical foundation for combining predictive densities builds on earlier work in statistics by West and Crosse (1992) and West (1992).

main results. Section 5 concludes.

2 Econometric Methodology

2.1 Predictive Regression

In our analysis we are primarily interested in forecasting the h-month-ahead change in the exchange rate. Following a vast amount of papers, see e.g. Mark (1995), Engel et al. (2007), Rossi (2013) and Byrne et al. (2016, 2017), we focus on single equation models where we model the exchange rate as a function of its deviation from its fundamental's implied value. As explained in Mark (1995), this accords with the notion that exchange rates frequently deviate from the level implied by fundamentals, particularly in the short-run. We let $e_{t+h} - e_t \equiv \Delta e_{t+h}$ denote the h-step ahead change in the log of the exchange rate, and f_t a set of exchange rate fundamentals. For simplicity, we then define $s_{t+h} \equiv \Delta e_{t+h}$. While most earlier studies on exchange rate predictability focus on linear models, see Rossi (2013) for a survey, some recent papers, see e.g. Byrne et al. (2016, 2017), have documented that the instability in models' forecasting performance can be improved upon by using models that accounts for time-varying parameters and stochastic volatility. We therefore consider predictions from both linear single equation models as well as from single equation models with time-varying parameters and stochastic volatility.

2.1.1 Constant coefficients

We consider predictive regressions with constant coefficients of the following form:

$$s_{t+h} = X_t\beta + \upsilon_{t+h}, \quad \upsilon_{t+h} \sim N(0, \sigma^2), \tag{1}$$

where $X_t = [1, z_t]$ and

$$z_t = f_t - e_t \tag{2}$$

As equation (2) suggests, f_t signals the exchange rate's fundamental value, hence z_t is the deviation from the fundamental's implied level. When the spot exchange rate is higher than its fundamental's implied level, the spot exchange rate is expected to decrease. We assume that $\upsilon_{t+h} \sim N(0, \sigma^2)$. The benefit of this is that an exact analytical expression is available. Bayesian inference is applied with weak informative conjugate priors to restrict regression coefficients to zero. We use a normal-inverse-γ prior with mean of β equal to zero and variances equal to 100. For the variance V we use an inverse-γ with degrees of freedom set equal to the number of regressors including the intercept.

2.1.2 Time-varying parameters and stochastic volatility

Following Byrne et al. (2017), we also consider predictive regressions with time-varying parameters and stochastic volatility of the following form:

$$s_{t+h} = \Delta e_{t+h} = X_t \theta_t + \upsilon_{t+h}, \quad \upsilon_{t+h} \sim N(0, V_t), \tag{3}$$

$$\theta_t = \theta_{t-1} + \varpi_t, \quad \varpi_t \sim N(0, W_t), \tag{4}$$

with X_t and z_t as defined above.

The predictive regression given by the system of equations (3) and (4) allows the coefficients linked to the disequilibrium term (z_t) and to the constant, to change over time. Equation (4) suggests that we assume a random walk process for the parameter θ_t, following Rossi (2006). We further consider that the disturbance terms, υ_{t+h} and ϖ_t, are uncorrelated and normally distributed with mean zero and time-varying matrices V_t and W_t, respectively.

The degree of time-variation in the regression coefficients is determined by the matrix W_t. In periods of low variance the estimation error shrinks towards zero as more data become available. In contrast in periods of high variance the estimation error increases, affecting the prediction. In light of this, a common practice is to impose some structure on W_t; see, for example, Dangl and Halling (2012), Koop and Korobilis (2012), and West and Harrison (1997). We follow West and Harrison (1997) ch. 4 and let W_t be proportional to the estimation variance of the coefficients at time t

$$W_t = \frac{1-\delta}{\delta} S_t C_t^*, \quad 0 < \delta \leq 1 \tag{5}$$

where S_t is the estimate of the variance of the error term in the observation equation, C_t^* is the estimated conditional covariance matrix of θ_t in equation (3), and δ is a discount factor that controls the degree of time-variation in coefficients (see also Dangl and Halling, 2012). In equation (5), setting $\delta = 1$ implies that $W_t = 0$, and therefore the coefficients are effectively constant over time. By contrast, specifying $0 < \delta < 1$ is consistent with time-varying coefficients, with the underlying variability determined by the magnitude of increase in the variance by a ratio of $1/\delta$.

2.2 Bayesian Model Averaging

While allowing for time-varying coefficients addresses one potential source of instability in predictive ability, the literature on exchange rate predictability also indicates that the relevant set of predictors appears to change over time, see Bacchetta and Wincoop (2004), Sarno and Valente (2009) and Rossi (2013). In this context, model combination methods in a Bayesian setting offers a valuable alternative as it allows to incorporate parameter and model uncertainty into the estimation and inference steps. Therefore forecasts based on model combination should be more robust to model misspecifications than forecasts from a single model. More specifically, the Bayesian approach assigns posterior probabilities to a the set of competing exchange rate prediction models. It then uses the probabilities as weights on the individual models to obtain a composite-weighted model.

For example, assume that a forecaster wants to predict the h-step ahead change in the exchange rate at time t and has N competing models (M_1, M_2,...,M_N) available. After eliciting prior distributions on the parameters of each model, she can derive posterior estimates on all such parameters, and ultimately obtain N distinct predictive distributions, one for each model entertained. We denote with $\left\{p(s_{t+h}|M_i, \mathcal{D}^t)\right\}_{i=1}^{N}$ the N predictive densities for s_{t+h}, where $\mathcal{D}^t$ stands for the information set available at time t, i.e. $\mathcal{D}^t = \{s_{\tau+h}, x_\tau\}_{\tau=1}^{t-h} \cup x_t$, where $\tau = 1, ..., t-h$.

When using Bayesian Model Averaging (BMA, henceforth) the individual predictive densities are combined into a composite-weighted predictive distribution $p(s_{t+h}|\mathcal{D}^t)$, given by

$$p(s_{t+h}|\mathcal{D}^t) = \sum_{i=1}^{N} P\left(M_i|\,\mathcal{D}^t\right) p(s_{t+h}|M_i, \mathcal{D}^t) \tag{6}$$

where $P\left(M_i|\,\mathcal{D}^t\right)$ is the posterior probability of model i, derived by Bayes' rule,

$$P\left(M_i|\,\mathcal{D}^t\right) = \frac{P\left(\mathcal{D}^t|\,M_i\right)P\left(M_i\right)}{\sum_{j=1}^{N} P\left(\mathcal{D}^t|\,M_j\right)P\left(M_j\right)}, \qquad i = 1, ..., N \tag{7}$$

and where $P\left(M_i\right)$ is the prior probability of model M_i, with $P\left(\mathcal{D}^t\middle|\,M_i\right)$ denoting the corresponding marginal likelihood.

Although several papers have found that BMA is useful for improving exchange rate predictability, it suffers from particularly two important drawbacks. First, and perhaps the most important one is that BMA assumes that the true model is included in the model set. Indeed, under such an assumption, it can be shown that the combination weights in equation (7) converge (in the limit) to select the true model. However, as noted by Diebold (1991), all models could be false, and as a result the model set could be misspecified. Geweke (2010) labels this problem *model incompleteness*. Second, BMA does not account for the uncertainty of the weights attached to each model. In a a forecasting environment characterized with large model uncertainty and large instability in the various models' performances, the weight uncertainty can be very large.

2.3 Bayesian density combination with time-varying weights and model misspecifaction

We propose to make use of the general density combination approach developed by Billio et al. (2013) and Casarin et al. (2015) to study exchange rate predicability. This is a Bayesian combination approach which accounts for several sources of uncertainty such as time-varying weights and weight uncertainty as well as model incompleteness. The approach has recently been used to successfully improve GDP nowcasts and stock return forecasts, see Aastveit et al. (2016) and Pettenuzzo and Ravazzolo (2016), respectively. We now turn to explaining the model combination scheme in more detail.

We assume that at a generic point in time t, the forecaster has available N different models to predict exchange rates at time $t+h$, each model producing a predictive distribution $p\left(s_{t+h}|\,M_i, \mathcal{D}^t\right)$, $i = 1, ..., N$. For example, the forecaster may be considering N alternative predictors for exchange rates, leading to N univariate models, each one in the form of equation (1) (or equations (3) and (4)) and including as right-hand-side one of the N available predictors.

To ease the notation, we aggregate the N predictive distributions $\left\{p\left(\left. s_{t+h}\right| M_i, \mathcal{D}^t\right)\right\}_{i=1}^N$ into the pdf $p\left(\left.\widetilde{\mathbf{s}}_{t+h}\right| \mathcal{D}^t\right)$. Next, the composite predictive distribution $p(s_{t+h}|\mathcal{D}^t)$ is given by

$$p\left(\left. s_{t+h}\right| \mathcal{D}^t\right) = \int\int p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathcal{D}^t)p\left(\left.\widetilde{\mathbf{s}}_{t+h}\right| \mathcal{D}^t\right) d\widetilde{\mathbf{s}}_{t+h} d\mathbf{w}_{t+h} \tag{8}$$

where $p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)$ denotes the combination scheme based on the N predictive densities $\widetilde{\mathbf{s}}_{t+h}$ and the combination weights $\mathbf{w}_{t+h} \equiv (w_{1,t+h}, \ldots, w_{N,t+h})'$, and $p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathcal{D}^t)$ denotes the posterior distribution of the combination weights $\mathbf{w}_{t+h}$. Equation (8) generalizes equation (6), taking into account the limitations discussed in the previous section. First, by specifying a stochastic process for the model combination scheme, $p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)$, our approach explicitly allows for either model misspecification or model incompleteness to play a role. Second, by introducing a proper distribution for the model combination weights $\mathbf{w}_{t+h}$, $p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathcal{D}^t)$, we gain two important advantages. On the one hand, our method can allow for time-varying combination weights. On the other hand, we have flexibility in how to model the dependence of the combination weights on the individual models' performance, and are no longer confined to have the weights depend on some measure of the individual models' statistical fit. We note, inter alia, that in addition to addressing the limitations discussed above, the combination scheme in equation (8) allows to factor into the composite predictive distribution the uncertainty over the model combination weights, a feature that should prove useful in the context of exchange rate predictions, where there is significant uncertainty over the identity of the best model(s) for predicting exchange rates. We label the density combination in equation (8) as DeCo. We now turn to describing in more details how the individual terms in equation (8) are obtained.

2.3.1 Combination scheme

We start by explaining the first term on the right hand side of equation (8), $p\left(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t\right)$, denoting the combination scheme adopted in our model combination. We note that since both the N original densities $\left\{p\left(\left. s_{t+h}\right| M_i, \mathcal{D}^t\right)\right\}_{i=1}^N$ and the combination weights $\mathbf{w}_{t+h}$ are in the form of densities, the combination scheme for $p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)$ is based on a convolution mechanism. We follow Billio et al. (2013), and apply a Gaussian combination scheme,

$$p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \sigma_\kappa^{-2}) \propto \exp\left\{-\frac{1}{2}\left(s_{t+h} - \widetilde{\mathbf{s}}'_{t+h}\mathbf{w}_{t+h}\right)' \sigma_\kappa^{-2}\left(s_{t+h} - \widetilde{\mathbf{s}}'_{t+h}\mathbf{w}_{t+1}\right)\right\} \quad (9)$$

The combination relationship is assumed to be linear and explicitly allows for model misspecification, possibly because all models in the combination may be false (incomplete model set or open model space). The combination residuals are estimated and their distribution follows a Gaussian process with mean zero and standard deviation σ_κ, providing a probabilistic measure of the incompleteness of the model set. In other words, the model that is specified in equation (9) can be written as:

$$s_{t+h} = \widetilde{\mathbf{s}}'_{t+h}\mathbf{w}_{t+h} + \kappa_{t+h} \quad (10)$$

with $\kappa_{t+h} \sim \mathcal{N}(0, \sigma_\kappa^2)$.

2.3.2 Combination weights

We now turn to describing how we specify the conditional density for the combination weights, $p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}_{t+h}}, \mathcal{D}^t)$. First, in order to have the weights $\mathbf{w}_{t+h}$ belong to the simplex $\Delta_{[0,1]^N}$, we introduce a vector of latent processes $\mathbf{z}_{t+h} = (z_{1,t+h}, \dots, z_{N,t+h})'$, where N is the total number of models considered in the combination scheme, and[3]

$$w_{i,t+h} = \frac{\exp\{z_{i,t+h}\}}{\sum_{l=1}^N \exp\{z_{l,t+h}\}}, \quad i = 1, \dots, N \quad (11)$$

Next, in order to obtain the combination weights we need to make additional assumptions on how the vector of latent processes $\mathbf{z}_{t+h}$ evolves over time and how it maps into the combination

[3]Under this convexity constraint, the weights can be interpreted as discrete probabilities over the set of models entering the combination.

weights $\mathbf{w}_{t+h}$. One possibility is to specify a Gaussian random walk process for $\mathbf{z}_{t+h}$,[4]

$$\begin{aligned} \mathbf{z}_{t+h} &\sim p(\mathbf{z}_{t+h}|\mathbf{z}_t, \Lambda) \\ &\propto |\Lambda|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\mathbf{z}_{t+h} - \mathbf{z}_t\right)' \Lambda^{-1} \left(\mathbf{z}_{t+h} - \mathbf{z}_t\right)\right\} \end{aligned} \tag{12}$$

with Λ an $(N \times N)$ diagonal matrix.

Effectively, equations (12) and (11) imply time-varying combination weights, where time $t+h$ combination weights depend in a non-linear fashion on time t combination weights. Alternatively, we could allow the combination weights to depend on a learning function based on the past performance of the N individual prediction models entering the combination. While Aastveit et al. (2016) and Pettenuzzo and Ravazzolo (2016) found that allowing the combination weights to depend on past performance is useful for nowcasting GDP and forecasting stock returns, we could not find similar results for exchange rate predictability. Thus, we consider time-varying combination weights that has a 1st order Markovian structure, but that do not depend on past predictive performance.

2.3.3 Individual models

Finally we explain how we specify the last term on the right-hand side of (8), $p\left(\widetilde{\mathbf{s}}_{t+h}\middle|\, \mathcal{D}^t\right)$, which we remind is short-hand for the set of individual predictive distributions $\left\{p\left(s_{t+h}\middle|\, M_i, \mathcal{D}^t\right)\right\}_{i=1}^N$ entering the model combination. As discussed in section 2.1, we consider two type of model classes: linear single equation models and single equation models with time-varying parameters and stochastic volatility. For both model classes we projects the h-step ahead change in the log exchange rate s_{t+h} on a lagged predictor, x_t, where x_t is assumed to be a scalar.[5]

For the two model classes, once we obtain draws from the posterior distributions of β and σ^2 in the case for linear single equation models or θ, V_t and W_t in the case of single equation models with time-varying parameters and stochastic volatility, we use them to form a predictive density for s_{t+h}. Then by repeating the process for the N individual models entering the combination we obtain the set of N individual predictive distributions $\left\{p\left(s_{t+h}\middle|\, M_i, \mathcal{D}^t\right)\right\}_{i=1}^N$.

[4]We assume that the variance-covariance matrix Λ of the process $\mathbf{z}_{t+h}$ governing the combination weights is diagonal. We leave for further research the possibility of allowing for cross-correlation between model weights.

[5]In our setting we consider only one predictor at the time, thus x_t is a scalar. It would be possible to include multiple predictors, but we follow the bulk of the literature on exchange rate predictability and focus on a single predictor.

2.3.4 Algorithm

We conclude this section by briefly describing how we estimate the posterior distributions $p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)$ and $p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathcal{D}^t)$. Equations (8), (9), (11), and (12), as well as the individual model predictive densities $p\left(\widetilde{\mathbf{s}}_{t+h}|\, \mathcal{D}^t\right)$ are first grouped into a non-linear state space model.[6] Because of the non-linearity, standard Gaussian methods such as the Kalman filter cannot be applied. We instead apply a Sequential Monte Carlo method, using a particle filter to approximate the transition equation governing the dynamics of $\mathbf{z}_{t+h}$ in the state space model, yielding posterior distributions for both $p(s_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathbf{w}_{t+h}, \mathcal{D}^t)$ and $p(\mathbf{w}_{t+h}|\widetilde{\mathbf{s}}_{t+h}, \mathcal{D}^t)$. For additional details, see Billio et al. (2013) and Casarin et al. (2015).

3 Forecasting Environment

3.1 Choice of Regressors

The range of macroeconomic predictors that we consider is guided by standard fundamental-based empirical exchange rate models; see, among others, Engel et al. (2007), Engel and West (2005), Molodtsova and Papell (2009) and Rossi (2013). Engel and West (2005) show that the majority of these models fit within the asset-pricing framework, such that the exchange rate can be expressed as a present-value of a linear combination of fundamentals and random noise. When combined with rational expectations and a random walk process for the fundamentals, the spot exchange rate becomes a function of current observable fundamentals and unexpected noise. We consider five different observable fundamentals as predictors:

- The Monetary Model (MM):

$$\mathbf{z}_{t,MM} \equiv (m_t - m_t^*) - (y_t - y_t^*) - s_t, \tag{13}$$

  where m_t is the log of money supply, y_t is the log of income, and asterisks denote foreign country variables. Note that we have assumed an income elasticity of one in the monetary model ($\mathbf{z}_{t,MM}$), following Mark (1995) and Engel and West (2005).

[6] The non-linearity is due to the logistic transformation mapping the latent process $\mathbf{z}_{t+h}$ into the model combination weights $\mathbf{w}_{t+h}$.

- Purchasing Power Parity (PPP) condition:

$$\mathbf{z}_{t,PPP} \equiv p_t - p_t^* - s_t, \tag{14}$$

where p_t is the log of price level. PPP is a long-run condition stating that the national price levels should be equal when expressed in a common currency.

- Uncovered Interest Rate Parity (UIP) condition:

$$\mathbf{z}_{t,UIP} \equiv i_t - i_t^*, \tag{15}$$

with i_t denoting the short-term nominal interest rate. UIP holds under risk neutrality and rational expectations, and implies that the forward rate is an unbiased estimator of the future spot rate. Moreover, UIP implies that the expected exchange rate return is equal to the interest differentials.

- A symmetric and an asymmetric Taylor rule (TRsy and TRasy, respectively):

$$\mathbf{z}_{t,TRsy} \equiv 1.5(\pi_t - \pi_t^*) + 0.5(\overline{y}_t - \overline{y}_t^*), \tag{16}$$

$$\mathbf{z}_{t,TRasy} \equiv 1.5(\pi_t - \pi_t^*) + 0.1(\overline{y}_t - \overline{y}_t^*) + 0.1(s_t + p_t^* - p), \tag{17}$$

where π_t is the domestic inflation rate, π_t^* is the foreign inflation rate, $\overline{y}_t$ the domestic output gap and $\overline{y}_t^*$ the foreign output gap.[7] We use conventional parameter values for the inflation difference, output difference and the real exchange rate, equal to the ones in Engel et al. (2015) (equation (16)) and Li et al. (2015) (equation (17)).

- Random Walk Benchmark Since Meese and Rogoff (1983), the random walk (RW) model has commonly been used as the benchmark in assessing exchange rate predictability. Thus, we use the RW model as the benchmark which we compare all models. The RW captures

[7] We follow the procedure in Foroni et al. (2015). Monthly industrial production (IP) is used as a proxy for output. The output gap is obtained by applying the Hodrick and Prescott (1997) filter recursively to the output series using the conventional smoothing parameter for monthly data - 14400. To correct for the uncertainty about these estimates at the recursive sample end-points, we follow the method in Watson (2007). We estimate bivariate VAR(ℓ) regressions on the first difference of inflation and the change in the log IP, with the lag length in the VAR determined by AIC. These regressions are then used to forecast and backcast three years worth of monthly data on IP, and the filter is applied to the resulting extended series.

the prevailing view in the FX literature that exchange rates are not predicable when condition on economic fundamentals, especially at short horizons.

3.2 Data and Forecasting Mechanics

We use monthly data from 1986M9 to 2014M3 for the following seven: Australia (AUD), Canada (CAD), Norway (NOK), Euro-area (EUR) Japan (JPN), Switzerland (CHE) and the United Kingdom (UK). The exchange rate is defined as the end-of-month value of the national currency per U.S. dollar. We employ the central bank's policy rate when this is available for the entire sample period and, alternatively, the discount rate or the money market rate. The inflation rate is calculated as the change in the log of monthly consumer price index (CPI). We use monthly industrial production (IP) as the proxy for output. Finally, we measure money supply by the aggregate M1.[8] The data are mostly obtained from the IMF's International Financial Statistics (IFS) and OECD Main Economic Indicators (MEI) and in a few cases supplemented with national central bank sources.

We use a direct, rather than an iterative, method to forecast the h-month-ahead change in the exchange rate for $h = [1, 3]$. The forecasting exercise is based on a recursive approach using data available up to the time the forecast is made. For example, a 3-months ahead forecast of the change in exchange rate for 2010M1 is made using data available up to 2009M10. Our forecasting window begins in 2000M4+h for all regressions.

3.3 Statistical evaluation of exchange rate predictability

We employ the mean squared forecast error (MSFE) as a statistical measure of out-of-sample point forecast accuracy. We use the driftless random walk (RW) model as the benchmark and compute the ratio of the MSFE of the fundamental-based models to the RW.

$$\text{Relative RMSFE} = \frac{\sqrt{\frac{1}{P}\sum_{\bar{p}=1}^{P} fe_{i,\bar{p}}^2}}{\sqrt{\frac{1}{P}\sum_{\bar{p}=1}^{P} fe_{RW,\bar{p}}^2}}, \tag{18}$$

where P is the number of out-of-sample forecasts, fe_i^2 and fe_{RW}^2 are the squared forecast errors of our model i and the RW, respectively. Values of the relative MSFE below one are

[8] In cases where the M1 aggregate is unavailable, we use a broader aggregate.

consistent with a more accurate point forecast of model i against the RW. To evaluate whether the differences in the MSFE between our models and the RW are significant we use the Clark and West (2006, 2007) test, hereafter CW-test. To examine the forecasting performance of our models over time in terms of point forecast, we compute the relative RMSFE recursively over the out-of-sample period. To examine whether the forecast performance is driven by specific periods, we also compute the cumulative sum of squared prediction error difference (CSSED) over time.

Our use of Bayesian methods allow us to fully exploit the information in the predictive density, rather than focusing exclusively on point forecast. We consider two different metrics of predictive performance. First, following Amisano and Giacomini (2007), Geweke and Amisano (2010), and Hall and Mitchell (2007) we compute the mean log-score differentials (MLSD):

$$MLSD = P^{-1} \sum_{\overline{p}=1}^{P} (LS_{i,\overline{p}} - LS_{RW,\overline{p}}), \tag{19}$$

where $LS_{i,\overline{p}}$ and $LS_{RW,\overline{p}}$ are the log-scores of our model i and the RW, respectively. Positive values of $MLSD$ are consistent with more accurate density forecasts of model i relative to the RW. We also calculate the cumulative log-score differentials (CLSD) of our regressions relative to those of the RW over the out-of-sample period. Positive values of the CLSD indicate that our fundamentals-based regressions produce more accurate density forecasts than the RW benchmark.

Lastly, we follow Gneiting and Raftery (2007) Gneiting and Ranjan (2011) and Groen et al. (2013), and consider the average continuously ranked probability score differential (CRPSD),

$$CRPSD_m = \frac{\sum_{\tau=\underline{t}}^{\overline{t}} (CRPS_{PM,\tau} - CRPS_{m,\tau})}{\sum_{\tau=\underline{t}}^{\overline{t}} CRPS_{PM,\tau}} \tag{20}$$

where $CRPS_{m,\tau}$ ($CRPS_{PM,\tau}$) measures the average distance between the empirical cumulative distribution function (CDF) of r_τ (which is simply a step function in r_τ), and the empirical CDF that is associated with model m's (PM's) predictive density. Gneiting and Raftery (2007) explain how the CRPSD measure circumvents some of the problems of the logarithmic score, most notably the fact that the latter does not reward values from the predictive density that are close but not equal to the realization. We also calculate the cumulative CRPS differentials

(CCRPSD) of our regressions relative to those of the RW over the out-of-sample period. Negative values of the CCRPSD indicate that our fundamentals-based regressions produce more accurate density forecasts than the RW benchmark.

4 Empirical Results

4.1 Out-of-Sample Statistical Evaluation

Tables 1 and 2 present results on the accuracy of both point and density forecasts for a variety of single predictor models and model combinations, including an equal-weighted scheme, BMA and DeCo. In Table 1 we report results using model specifications with constant coefficients, while results for TVPSV are reported in Table 2. For MSPE and CRPS a value below 1 indicate that the alternative model or model combination perform better than the RW benchmark model. For LS, a positive value indicate superior forcecasts to the RW benchmark model. The key findings can be summarized as follows. First, for all countries DeCo provides more accurate point and density forecasts than the individual models, the alternative model combinations and the RW benchmark.

Second, none of the other models or approaches systematically outperforms the RW benchmark for all countries. Consistent with the earlier literature the various fundamental-based forecasting models and model combinations yields results that are very similar to the RW benchmark.

Third, the magnitude of reduction in MSPE and CRPS relative to the RW benchmark always exceeds 7% for MSPE and 18% for CRPS. For some countries the improvements are substantially larger exceeding 15% for MSPE and 30% for CRPS.

Fourth, in line with results in Byrne et al. (2017), we find that TVPSV models provides more accurate forecasts than models with constant coefficients. Accordingly, DeCo applied to models with TVPSV provide more accurate forecasts than DeCo applied to models with constant coefficients.

Table 1. Relative MSPE, Relative CPRS and Log-score Differentials for various models and model combinations, constant coefficients models and $h = 1$

Model	AUS	CAN	NOR	EUR	JPN	CHE	GBP
				MSPE			
DeCo	**0.923**	**0.912**	**0.910**	**0.884**	**0.821**	**0.871**	**0.834**
BMA	1.002	0.996	0.993	0.987	0.997	0.989	0.987
EW	1.006	1.000	1.003	0.998	1.002	1.001	0.991
MM	1.034	1.021	1.040	1.041	1.016	1.053	1.013
PPP	1.007	0.994	0.985	0.981	1.014	0.996	1.015
TR1	1.006	0.996	1.004	0.994	1.016	1.003	0.997
TR2	0.993	0.993	0.992	0.990	1.011	0.996	0.997
UIP	1.004	1.007	1.011	1.008	0.994	0.999	1.006
				CRPS			
DeCo	**0.819**	**0.800**	**0.778**	**0.727**	**0.647**	**0.710**	**0.664**
BMA	1.052	1.053	1.028	1.024	0.990	1.012	0.992
EW	1.118	1.113	1.083	1.063	1.029	1.054	1.024
MM	1.024	1.005	1.020	1.021	1.004	1.021	1.010
PPP	1.004	0.996	0.994	0.992	1.005	1.003	1.019
TR1	0.996	0.997	0.998	1.001	1.002	1.000	1.004
TR2	0.991	0.994	0.992	0.996	1.002	1.000	1.001
UIP	1.000	1.001	1.003	1.003	0.991	0.999	1.005
				LS			
DeCo	**3.414**	**3.414**	**3.414**	**4.761**	**5.504**	**5.013**	**5.334**
BMA	-1.088	-0.547	-0.946	-0.863	-0.066	-0.326	-0.632
EW	-7.701	-7.701	-5.336	-3.114	-0.809	-2.094	-1.777
MM	0.079	0.020	-0.023	-0.036	-0.011	0.001	-0.007
PPP	0.141	0.101	-0.039	0.011	-0.016	0.010	-0.013
TR1	0.147	0.254	-0.011	0.008	-0.007	0.015	-0.021
TR2	-0.068	0.175	-0.001	0.018	-0.010	-0.003	-0.020
UIP	-0.001	0.126	-0.092	-0.017	0.001	0.020	-0.017

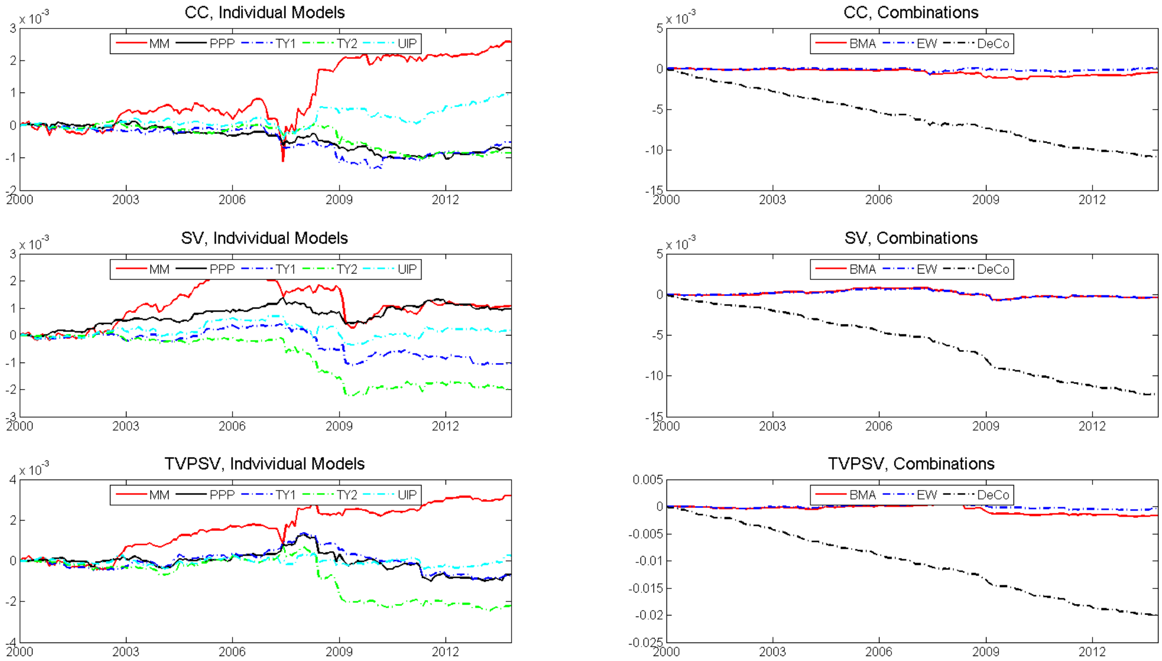
Note: The table reports forecasting performance for the various single predictor models and model combinations. The table presents results for models with constant coefficients. For point forecasting we report Mean Squared Forecast Error (MSFE) of the fundamental-based forecasting models or model combinations relative to the MSFE of the driftless Random Walk (RW). Values less than 1 (one) indicate that a lower MSFE than the RW, hence, providing better forecasts than the RW. For density forecasting we first report Continues Ranked Probability Scores (CRPS) of the fundamental-based forecasting models or model combinations relative to the CRPS of the RW. Values less than 1 indicate a lower CRPS than the RW, hence, providing better forecasts than the RW. As a second measure for density forecasting performance we present the average log-score differentials between the same models relative to the RW. Positive values indicate that a model or model combination improves upon the RW in terms of density forecasts.

Table 2. Relative MSPE, Relative CPRS and Log-score Differentials for various models and model combinations, TVP-SV models and $h = 1$

Model	AUS	CAN	NOR	EUR	JPN	CHE	GBP
				MSPE			
DeCo	**0.938**	**0.845**	**0.863**	**0.827**	**0.787**	**0.824**	**0.793**
BMA	0.998	1.000	0.997	0.994	0.999	1.001	0.996
EW	0.998	1.000	0.997	0.994	0.999	1.001	0.996
MM	1.019	1.032	1.051	1.039	1.034	1.046	1.018
PPP	1.011	1.001	0.981	0.992	1.032	0.999	1.014
TR1	1.001	1.000	1.020	0.980	1.017	0.994	1.010
TR2	0.993	0.988	1.003	0.988	1.010	0.999	1.014
UIP	1.007	1.008	1.020	1.002	1.010	1.010	0.993
				CRPS			
DeCo	**0.818**	**0.649**	**0.674**	**0.648**	**0.591**	**0.642**	**0.601**
BMA	1.004	1.019	1.008	1.001	0.984	1.003	0.974
EW	1.012	1.022	1.037	1.016	0.998	1.014	0.987
MM	1.020	1.048	1.058	1.068	1.087	1.072	1.088
PPP	1.017	1.039	1.028	1.049	1.078	1.050	1.084
TR1	1.008	1.036	1.040	1.043	1.073	1.041	1.080
TR2	1.005	1.030	1.038	1.043	1.069	1.049	1.082
UIP	1.013	1.038	1.045	1.052	1.066	1.055	1.075
				LS			
DeCo	1.400	2.642	2.302	2.154	2.352	2.159	2.134
BMA	0.245	0.504	-0.092	-0.320	0.010	-0.181	-0.064
EW	0.002	-0.992	-0.869	-0.684	-0.112	-0.494	-0.359
MM	0.244	0.564	-0.073	-0.119	-0.174	-0.109	-0.169
PPP	0.252	0.553	-0.042	-0.100	-0.168	-0.095	-0.159
TR1	0.247	0.574	-0.052	-0.103	-0.162	-0.089	-0.159
TR2	0.256	0.586	-0.057	-0.104	-0.164	-0.093	-0.160
UIP	0.255	0.556	-0.063	-0.109	-0.162	-0.097	-0.153

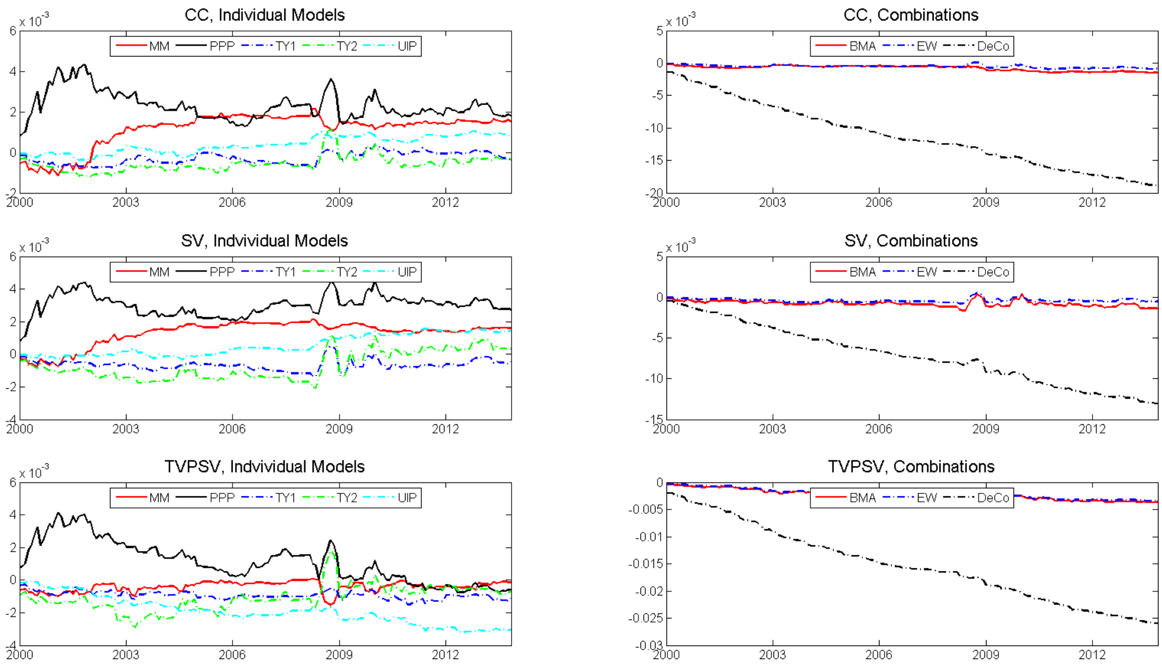
Note: The table reports forecasting performance for the various single predictor models and model combinations. The table presents results for models with time-varying parameters and stochastic volatility. For point forecasting we report Mean Squared Forecast Error (MSFE) of the fundamental-based forecasting models or model combinations relative to the MSFE of the driftless Random Walk (RW). Values less than 1 (one) indicate that a lower MSFE than the RW, hence, providing better forecasts than the RW. For density forecasting we first report Continues Ranked Probability Scores (CRPS) of the fundamental-based forecasting models or model combinations relative to the CRPS of the RW. Values less than 1 indicate a lower CRPS than the RW, hence, providing better forecasts than the RW. As a second measure for density forecasting performance we present the average log-score differentials between the same models relative to the RW. Positive values indicate that a model or model combination improves upon the RW in terms of density forecasts.

Figure 1. Cumulative MSPE differentials, Canada



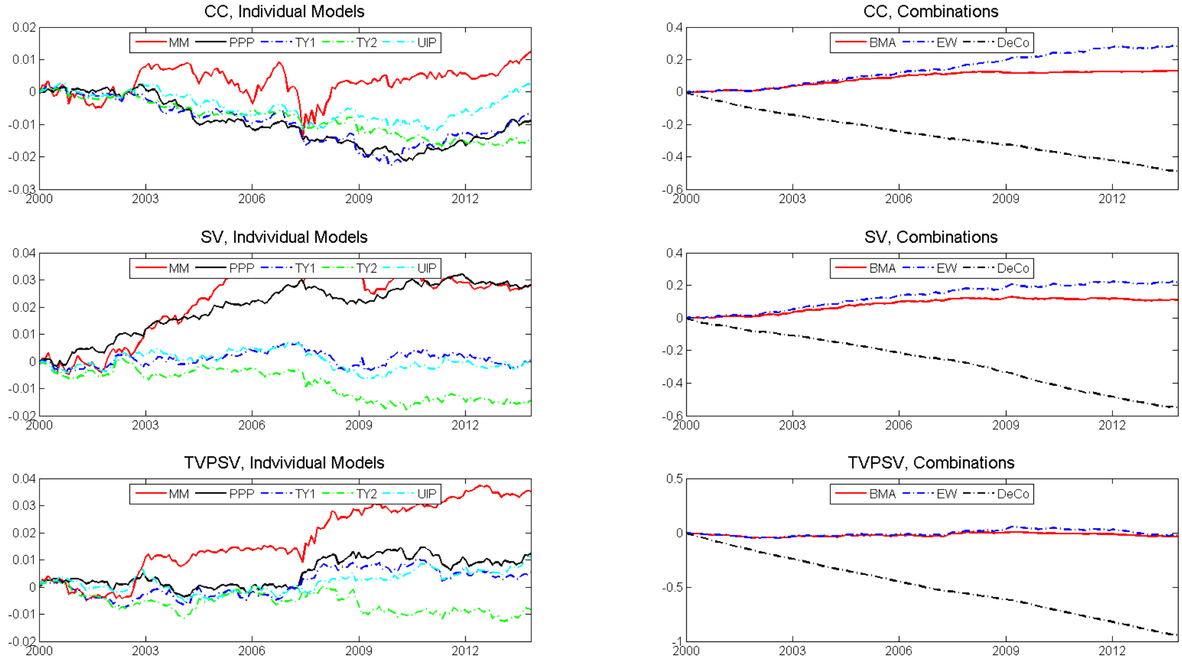
The figure plots the Cumulative MSPE differentials from the various predictive regressions relative to those of the RW over the out-of-sample period for Canada.

Figure 2. Cumulative MSPE differentials, United Kingdom



The figure plots the Cumulative MSPE differentials from the various predictive regressions relative to those of the RW over the out-of-sample period for the United Kingdom.

Figure 3. Cumulative CRPS differentials, Canada



The figure plots the Cumulative CRPS differentials from the various predictive regressions relative to those of the RW over the out-of-sample period for Canada.

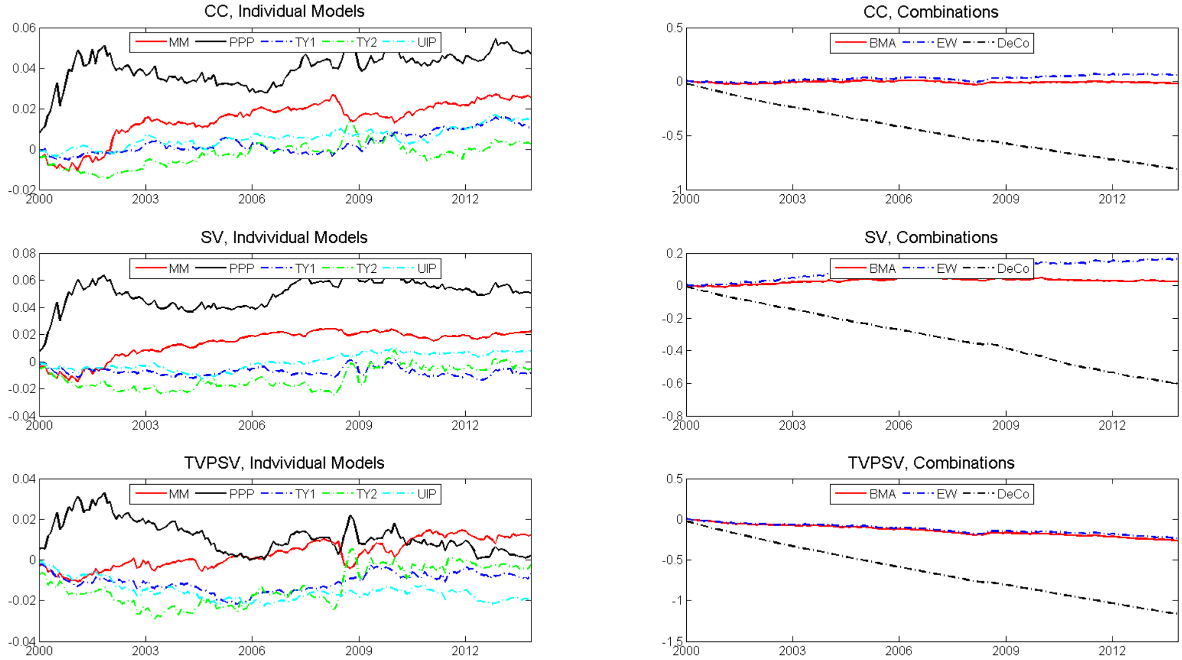
Rossi (2013) shows that there is considerable instabilities in predictive performance over time. To address whether our improved out-of-sample forecast performance is limited to a certain time period or driven by some outliers, we report the cumulative MSPE, CRPS and LS differentials of our regressions relative to those of the RW for Canada and the UK over the out-of-sample period in figures 1 to 4.[9] While the various individual models show considerable instabilities in predictive performance over time, the performance from our combination approach, DeCo, is far more robust, yielding a steady improvement over the various benchmarks over different time periods.

4.1.1 The importance of time-varying weight uncertainty and model incompleteness

Our novel Bayesian combination approach accounts for several sources of uncertainty. Particularly, there are two important features we include that has not been accounted for in the

[9]We report results for the additional countries in the appendix. There we also formally test for forecasting ability over time in presence of instabilities by implementing the Giacomini and Rossi (2010) one-sided fluctuation test.

Figure 4. Cumulative CRPS differentials, United Kingdom



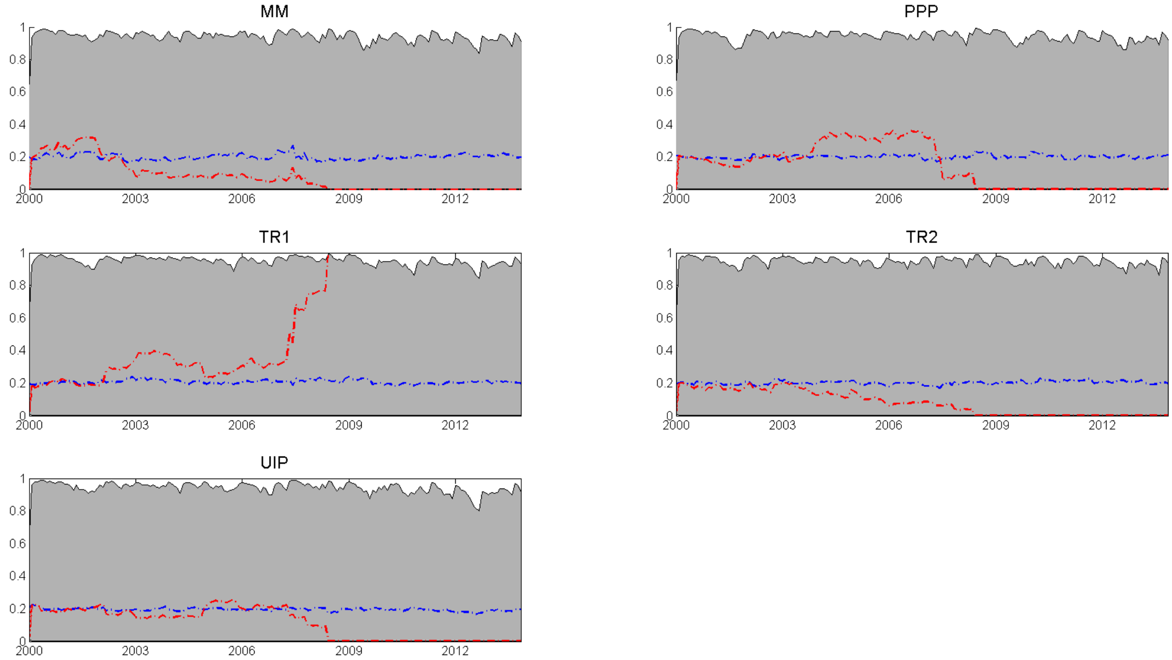
The figure plots the Cumulative CRPS differentials from the various predictive regressions relative to those of the RW over the out-of-sample period for the United Kingdom.

earlier literature on exchange rate predicability. The first is that we explicitly allow for either model misspecification or model incompleteness to play a role. Second, by introducing a proper distribution for the model combination weights we gain two important advantages. On the one hand, our method can allow for time-varying combination weights. On the other hand, we have flexibility in how to model the dependence of the combination weights on the individual models' performance, and are no longer confined to have the weights depend on some measure of the individual models' statistical fit. In what follows, we will shed light on which of these mechanisms are the most important for driver for our results.

Figures 5 and 6 show the weights associated with the five fundamental-based models for the Canada and the UK countries.[10] The median weights (blue dotted line) are fairly stable over time and close to being equal for each of the models. However, notice the large uncertainty around the weights. An important feature of our combination approach is that the large weight uncertainty is accounted for in our combined forecasts. The red dotted line in each figure shows the corresponding weights obtained by the BMA approach. Comparing the DeCo weights with

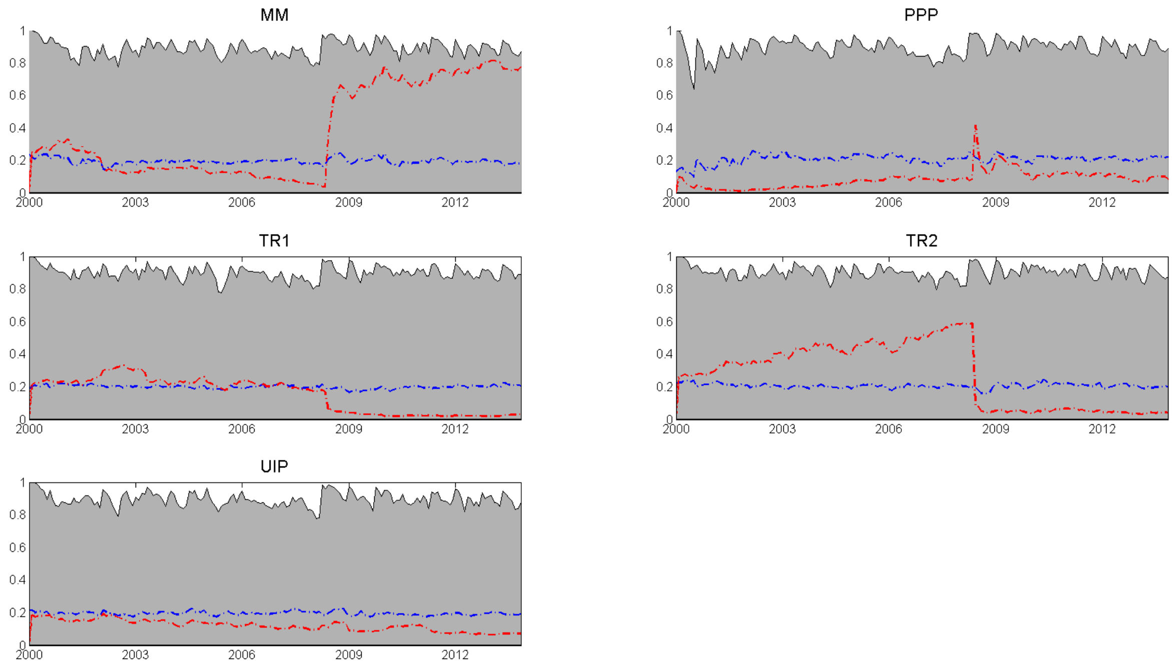
[10]We report additional results for the other countries in the appendix

Figure 5. Time-varying weights, Canada



The figures plot the 90% credibility intervals of the model posterior weights and their medians (blue dotted lines). The red dotted line shows the weights attached to each model using BMA.

Figure 6. Time-varying weights, United Kingdom



The figures plot the 90% credibility intervals of the model posterior weights and their medians (blue dotted lines). The red dotted line shows the weights attached to each model using BMA.

the BMA weights, we see two interesting differences. First, the medians of the DeCo weights and BMA weights differ substantially, with much larger movements over time from the BMA weights. Second, BMA selects much more extreme weights, attaching almost all the weights to one single model, consistent with findings in Amisano and Geweke (2013). The main difference between DeCo and BMA is that our weighting scheme allows for model incompleteness (the BMA weights based on predictive likelihood will also take into account past predictive performance scores).

Table 3 report results from our preferred combination DeCo, as well as for DeCo without incompleteness and Deco without model incompleteness and without time-varying weights. The results suggest that both model incompleteness and time-varying weights are important drivers for our results, accounting for roughly half of the improvements each.

Table 3. The importance of incompleteness and weight uncertainty

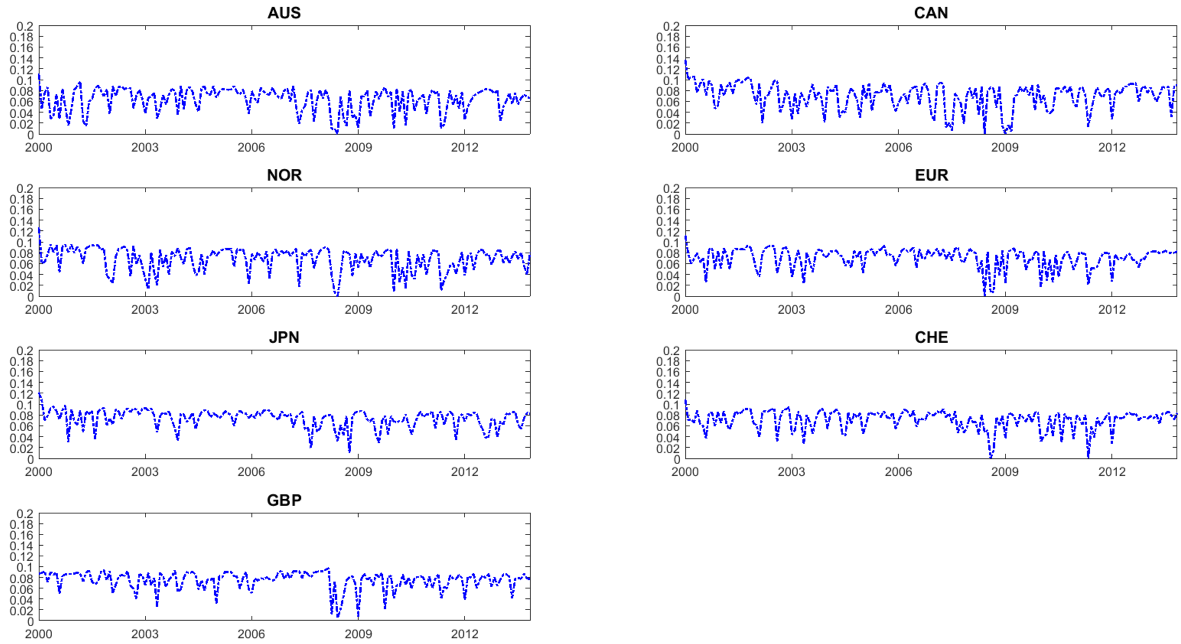
Model	AUS	CAN	NOR	EUR	JPN	CHE	GBP
				MSPE			
DeCo	**0.923**	**0.912**	**0.910**	**0.884**	**0.821**	**0.871**	**0.834**
No Inc	0.970	0.974	0.962	0.953	0.948	0.957	0.926
No Inc and TVW	1.002	0.999	0.999	0.987	1.002	0.998	0.987
				CRPS			
DeCo	**0.819**	**0.800**	**0.778**	**0.727**	**0.647**	**0.710**	**0.664**
No Inc	0.903	0.902	0.875	0.847	0.800	0.841	0.800
No Inc and TVW	1.013	1.018	1.006	0.988	0.979	0.990	0.982

Note: The table reports forecasting performance for our preferred combination DeCo, as well as DeCo without model incompleteness and DeCo without model incompleteness and without time-varying weights. For point forecasting we report Mean Squared Forecast Error (MSFE) of the fundamental-based forecasting models or model combinations relative to the MSFE of the driftless Random Walk (RW). Values less than 1 (one) indicate that a lower MSFE than the RW, hence, providing better forecasts than the RW. For density forecasting we first report Continues Ranked Probability Scores (CRPS) of the fundamental-based forecasting models or model combinations relative to the CRPS of the RW. Values less than 1 indicate a lower CRPS than the RW, hence, providing better forecasts than the RW.

Figure 7 shows the standard deviations of the combination residuals for the incomplete model sets, see equation (10), over time. The figure shows the standard deviations of the combination residuals is somewhat noisy and fluctuates a lot over time. For all countries there are several downward spikes, where some of them are common, such as during the Great Recession, when market were very volatile, the standard deviations of the combination residuals spikes down for all countries.

We study the relation between range volatility and model incompleteness. We construct a monthly measure of range volatility using daily exchange rate observations and measure model

Figure 7. Standard deviation of the combination residuals



The figure plots the standard deviation of the combination residuals for incomplete model sets from equation 10 for the seven countries.

incompleteness as the standard deviation of the combination residual. For each draw we regress contemporaneous range volatility on our measure of model incompleteness (computed across particles). In table 4 we report the median and 95% credible interval for regression coefficients. The results suggests that incompleteness is related to range volatility and for 6 over 7 exchange rates zero is outside the 95% credible interval and the relation is negative. Note, however, that the range volatility is not observed at the time when the forecast is made. Thus, one may interpret that the model incompleteness measure from our combination approach is a proxy for contemporaneous range volatility or risk.

4.2 Economic evaluation of exchange rate predictability

We follow Della Corte et al. (2009) and assume that different forecast models are used to build dynamic trading strategies with one month horizon. Consider a US based investor that in each period optimally allocates her wealth by buying a portfolio of home and foreign bonds that redeem in the next period. The return on home and foreign bonds is assumed to be safe and equalized across countries, so that the riskiness of foreign bonds originates solely from the future

Table 4. Relation between range volatility and model incompleteness

Country	Median	95% CI
AUS	-0.285	[-0.426, -0.092]
CAN	-0.196	[-0.301 , -0.086]
NOR	-0.217	[-0.338 , -0.060]
EUR	-0.201	[-0.320 , -0.036]
JPN	-0.116	[-0.234 , 0.010]
CHE	-0.173	[-0.301 , -0.028]
GBP	-0.126	[-0.241 , 0.019]

Note: The table reports the median regression coefficient and 95% credible intervals from regressing contemporaneous range volatility on model incompleteness.

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:

$$max_{w_t} \quad r_{p,t+1|t} = w_t' r_{t+1|t} + (1 - w_t'\iota) r_f, \qquad s.t. \quad w_t' \Sigma_{t+1|t} w_t = \bar{\sigma}^2 \tag{21}$$

where w_t is $m \times 1$ vector of portfolio weights on the risky asset, r_f is the log return on the safe asset and $\bar{\sigma}^2$ is the target conditional volatility of the portfolio returns, fixed equal to 10%.

Each forecast model provides the information needed to construct the conditional expected mean and variance of the risky assets returns, respectively denoted as $r_{t+1|t}$ and $\Sigma_{t+1|t}$ and defined by:

$$r_{t+1|t} = E_t[r_{t+1}] \tag{22}$$

$$\Sigma_{t+1|t} = E_t[(r_{t+1} - r_{t+1|t})(r_{t+1} - r_{t+1|t})'] \tag{23}$$

The solution to this optimization problem delivers the risky asset weights:

$$w_t = \frac{\bar{\sigma}}{\sqrt{C_t}} \Sigma_{t+1|t}^{-1} (r_{t+1|t} - \iota r_f) \tag{24}$$

where $C_t = (r_{t+1|t} - \iota r_f)' \Sigma_{t+1|t}^{-1} (r_{t+1|t} - \iota r_f)$.

In Table 5 we report gross returns of the trading strategy starting with an initial wealth of 1. Note that transaction costs and risk are not accounted for. The table shows two interesting results. First, our combination approach yields gross returns that are substantially higher than, and in most cases more than double of, the ones from the other models or combinations. Second, DeCo based on time-varying models yield higher returns than DeCo based on constant-parameter counterparts

Table 5. Trading strategies results based on the different forecast models.

Model	CC	TVP-SV
DeCo	**1.755**	**1.846**
BMA	1.366	1.365
EW	1.367	1.367
RW	1.341	1.266
MM	1.226	1.245
PPP	1.340	1.308
TR1	1.269	1.265
TR2	1.305	1.295
UIP	1.295	1.319

Note: The table reports the gross return from trading strategies based on different forecasting models. Initial wealth is set equal to 1 and transaction costs and risk are not accounted for.

Fra. We should add transaction costs.

5 Conclusion

In this paper, we introduce a novel Bayesian model combination technique that accounts for time-varying uncertainty of several model and data features in order to provide more accurate and complete density forecasts for exchange rates. Our approach combines the entire predictive densities of the individual models, rather than only their point forecasts, and extends earlier combination approaches that have been applied to exchange rate models, by allowing for two key features. First, our method features time-varying combination weights, and explicitly factors into the model combination the inherent uncertainty surrounding the estimation of the combination weights. Second, our method allows for model incompleteness, i.e., that the true model is not necessarily a part of the model space.

In an empirical exercise, we study the forecasting performance of our combination approach relative to other combination approaches and common benchmarks for seven major exchange rates vis-à-vis the US dollar over the period 2000-2014. We find that our combination approach systematically outperforms all benchmarks we compare it to, both in terms of statistical (point and density forecasts) and economic terms. At the one month horizon, the magnitude of reduction in terms of mean squared prediction error (MSPE) and continuous ranked probability score (CRPS) relative to the driftless Random Walk benchmark always exceeds 7% for MSPE and 18% for CRPS. We find the largest gains when we apply our combination approach to single equation models that accounts for time-varying parameters and stochastic volatilities. For several of the countries we then find improvements that are substantially larger, exceeding 15% for MSPE and 35% for CRPS. While the various individual models show considerable instabilities in predictive performance over time, the performance from our combination approach is far more robust, yielding a steady improvement over the various benchmarks over different time periods.

While accounting for time-varying weights and weight uncertainty plays a role in improving the density forecasting performance, the main bulk of the gains, both in terms of point and density forecasting performance, stems from allowing for model incompleteness in the combination scheme. Finally, we find that there is a relation between range volatility and the standard deviation of the combination residuals. The latter can be viewed as a measure of of model incompleteness. This suggests that the lack of predictability from standard fundamental-based models could be related to these models not appropriately being able to account for volatility.

References

Aastveit, K., K. Gerdrup, A. Jore, and L. Thorsrud (2014). Nowcasting GDP in real time: A density combination approach. *Journal of Business & Economic Statistics 32*(1), 48–68.

Aastveit, K. A., F. Ravazzolo, and H. K. van Dijk (2016). Combined Density Nowcasting in an uncertain economic environment. *Journal of Business & Economic Statistics Forthcoming.*

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

Amisano, G. and R. Giacomini (2007). Comparing density foreasts via weighted likelihood ratio tests. *25*(2), 177–190.

Bacchetta, P., E. van Wincoop, and T. Beutler (2010). *Can Parameter Instability Explain the Meese-Rogoff Puzzle?*, pp. 125–173. NBER International Seminar on Macroeconomics 2009. University of Chicago Press.

Bacchetta, P. and E. V. Wincoop (2004). A scapegoat model of exchange-rate fluctuations. *American Economic Review 94*(2), 114–118.

Bacchetta, P. and E. V. Wincoop (2013). On the unstable relationship between exchange rates and macroeconomic fundamentals. *Journal of International Economics 91*(1), 18–26.

Beckmann, J. and R. Schüssler (2016). Forecasting exchange rates under parameter and model uncertainty. *Journal of International Money and Finance 60*(C), 267–288.

Billio, M., R. Casarin, F. Ravazzolo, and H. K. van Dijk (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics 177*(2), 213–232.

Byrne, J. P., D. Korobilis, and P. J. Ribeiro (2016). Exchange rate predictability in a changing world. *Journal of International Money and Finance 62*(C), 1–24.

Byrne, J. P., D. Korobilis, and P. J. Ribeiro (2017). On the Sources of Uncertainty in Exchange Rate Predictability. *International Economic Review* (Forthcoming).

Casarin, R., S. Grassi, F. Ravazzolo, and H. K. van Dijk (2015). Parallel sequential Monte Carlo for efficient density combination: The DeCo MATLAB toolbox. *Journal of Statistical Software 68(3).*

Clark, T. E. and K. D. West (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics 135*(1), 155–186.

Clark, T. E. and K. D. West (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics 138*(1), 291–311.

Dangl, T. and M. Halling (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics 106*(1), 157 – 181.

Della Corte, P., L. Sarno, and I. Tsiakas (2009). An Economic Evaluation of Empirical Exchange Rate Models. *Review of Financial Studies 22*(9), 3491–3530.

Diebold, F. (1991). A note on bayesian forecast combination procedures. In A. Westlund and P. Hackl (Eds.), *Economic Structural Change: Analysis and Forecasting*, pp. 225–232. New York: Springer-Verlag.

Engel, C., N. C. Mark, and K. D. West (2015). Factor Model Forecasts of Exchange Rates. *Econometric Reviews 34*(1-2), 32–55.

Engel, C., N. C. Mark, K. D. West, K. Rogoff, and B. Rossi (2007). Exchange rate models are not as bad as you think. *NBER Macroeconomics Annual 22*, 381–473.

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

Foroni, C., F. Ravazzolo, and P. J. Ribeiro (2015, October). Forecasting commodity currencies: the role of fundamentals with short-lived predictive content. Working Paper 2015/14, Norges Bank.

Geweke, J. (2010). *Complete and Incomplete Econometric Models*. Princeton: Priceton University Press.

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

Geweke, J. and G. Amisano (2011). Optimal prediction pools. *Journal of Econometrics 164*(1), 130 – 141.

Giacomini, R. and B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics 25*(4), 595–620.

Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102*, 359–378.

Gneiting, T. and R. Ranjan (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *Journal of Business & Economic Statistics 29*(3), 411–422.

Groen, J. J. J., R. Paap, and F. Ravazzolo (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics 31*(1), 29–44.

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

Hodrick, R. J. and E. C. Prescott (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit and Banking 29*(1), 1–16.

Jore, A. S., J. Mitchell, and S. P. Vahey (2010). Combining forecast densities from vars with uncertain instabilities. *Journal of Applied Econometrics 25*(4), 621–634.

Koop, G. and D. Korobilis (2012). Forecasting Inflation Using Dynamic Model Averaging. *International Economic Review 53*(3), 867–886.

Kouwenberg, R., A. Markiewicz, R. Verhoeks, and R. C. J. Zwinkels (2017). Model Uncertainty and Exchange Rate Forecasting. *Journal of Financial and Quantitative Analysis 52*(01), 341–363.

Li, J., I. Tsiakas, and W. Wang (2015). Predicting exchange rates out of sample: Can economic fundamentals beat the random walk? *Journal of Financial Econometrics 13*(2), 293–341.

Mark, N. C. (1995). Exchange Rates and Fundamentals: Evidence on Long-Horizon Predictability. *American Economic Review 85*(1), 201–218.

McAlinn, K. and M. West (2017). Dynamic Bayesian predictive synthesis in time series forecasting. *Journal of Econometrics Forthcoming.*

Meese, R. A. and K. Rogoff (1983). Empirical exchange rate models of the seventies : Do they fit out of sample? *Journal of International Economics 14*(1-2), 3–24.

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

Pettenuzzo, D. and F. Ravazzolo (2016). Optimal portfolio choice under decision-based model combinations. *Journal of Applied Econometrics 31*(7), 1312.

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), 1063–1119.

Rossi, B. and T. Sekhposyan (2011). Understanding models forecasting performance. *Journal of Econometrics 164*(1), 158–172.

Sarno, L. and G. Valente (2009). Exchange rates and fundamentals: footloose or evolving relationship? *Journal of the European Economic Association 7*(4), 786–830.

Watson, M. W. (2007). How accurate are real-time estimates of output trends and gaps? *Economic Quarterly 93*(2), 143–161.

West, M. (1992). Modelling agent forecast distributions. *Journal of the Royal Statistical Society (Ser. B) 54*, 553–567.

West, M. and J. Crosse (1992). Modelling of probabilistic agent opinion. *Journal of the Royal Statistical Society (Ser. B) 54*, 285–299.

West, M. and J. Harrison (1997). *Bayesian Forecasting and Dynamic Models.* New York: Springer.

Wright, J. H. (2008). Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics 146*(2), 329–341.