On the Sources of Uncertainty in Exchange Rate Predictability

Joseph P. Byrne[*,†], Dimitris Korobilis[‡,§], Pinho J. Ribeiro[‡]

September 13, 2016

Abstract

In a unified framework, we examine four sources of uncertainty in exchange rate forecasting models: (i) random variations in the data, (ii) estimation uncertainty, (iii) uncertainty about the degree of time-variation in coefficients, and (iv) uncertainty regarding the choice of the predictor. We find that models which embed a high-degree of coefficient variability yield forecast improvements at horizons beyond 1-month. At the 1-month horizon, and apart from the standard variance implied by unpredictable fluctuations in the data, the second and third sources of uncertainty listed above are key obstructions to predictive ability. The uncertainty regarding the choice of the predictors is negligible.

JEL Classification: C53, C55, E44, F37, G11.

Keywords: Exchange rate forecasting, Instabilities, Time-varying parameter models, Bayesian model selection, Forecast combination

[*]We benefited from discussions with Roberto Casarin, Pasquale Della Corte, Steven Englander, John Hydeskov, Gary Koop, Ronald MacDonald, Michael Melvin, Francesco Ravazzolo, Dagfinn Rime, Barbara Rossi, Christian Schumacher, Allan Timmermann, Shaun Vahey, Herman K. van Dijk, participants at the Norges Bank Research Spring Meeting (Oslo, 2015), the 25$^{\text{th}}$ EC^2 Conference, the 4$^{\text{th}}$ Workshop on "Financial Determinants of Exchange Rates", and the 2014 Glasgow SIRE Econometric Workshop. We also thank Pasquale Della Corte, Lucio Sarno, and Giulia Sestieri for making the data associated with their paper (Della Corte et al., 2012) publicly available.

[†]Heriot-Watt University, School of Management and Languages, Edinburgh, EH14 4AS, UK.

[‡]University of Glasgow, Adam Smith Business School, Glasgow, G12 8QQ, UK.

[§]Corresponding author, E-mail: dimitris.korobilis@glasgow.ac.uk; Tel: +44 (0)141 330 2950.

1 Introduction

Thirty years on since Meese and Rogoff (1983) identified that exchange rate fluctuations are difficult to predict using standard economic models, academics and practitioners are yet to find a definite answer as to whether or not macroeconomic variables have predictive content. In a thorough survey of the recent literature, Rossi (2013) points out that the answer is not clear-cut. Decisions regarding the choice of the predictor, forecast horizon, forecasting model, and methods for forecast evaluation, all exert influence in exchange rate predictability. Ultimately, the predictive power appears to be specific to some countries in certain periods, signalling the presence of instability in the models' forecasting performance (Rogoff and Stavrakeva, 2008; Rossi, 2013). The issue of instability was also pointed out by Meese and Rogoff (1983) and is echoed in other recent papers including, Bacchetta and van Wincoop (2004, 2013), Bacchetta *et al.* (2010), Sarno and Valente (2009), among others. However, as Rossi (2013) notes, models that take into account these instabilities, by allowing for time-variation in the coefficients for instance, do not greatly succeed in outperforming a random walk benchmark in an out-of-sample forecasting exercise.

In this paper, we employ a framework that allows us to pin down several sources of instability that might affect the out-of-sample forecasting performance of exchange rate models. The starting point of our analysis is the exact conjecture by Meese and Rogoff (1983), that time-variation in parameters may play a significant role in explaining the predictive power of these models. However, unlike previous attempts to explain this conjecture, we do not assume ex-ante that coefficients in the forecasting regressions change in the same fashion over time (e.g., Rossi, 2006). Instead, we allow for a range of possible degrees of time-variation in coefficients, encompassing moderate to sudden changes, and even no-change in coefficients. We then use a likelihood-based approach to identify what degree of time-variation in coefficients is consistent with the data. In this framework we can infer, for example, whether allowing for sudden changes in coefficients leads to a better forecasting performance, relative to situations where coefficients change gradually or remain constant over time.

In light of the hypotheses advanced in recent papers, not only are coefficients in an exchange rate model likely to change over time, but the relevant set of exchange rate determinants may also differ at each point in time. See, for example, the scapegoat theory of exchange rates of Bacchetta and van Wincoop (2004, 2013), as well as the empirical evidence in Berge (2013), Fratzscher *et al.* (2015), Markiewicz (2012), and Sarno and Valente (2009). Hence in our setting, in addition to allowing for varying degrees of coefficient adaptivity over time, we also entertain the possibility that a different fundamental may be relevant at each point in time. In this unified framework, we can examine whether models with a certain configuration, characterized by a specific degree of time-variation in coefficients and choice of predictor, can forecast well.

Our key contribution in this paper goes much further than merely establishing whether our models outperform a random walk benchmark. As the evidence on time-varying forecasting performance suggests, the possibility that a model with a specific configuration can forecast well in a certain period and country, and not in another setting, introduces uncertainty regarding the ex-ante choice of the model. In this context, our unified approach provides the ideal framework to analyze the sources of exchange rate models' prediction uncertainty. Within it, we can distinguish between (i) model uncertainty due to random or unpredictable fluctuations in the data, (ii) model uncertainty due to errors when estimating the coefficients, (iii) model uncertainty originating from time-variation in coefficients, and (iv) model uncertainty due to a time-varying set of exogenous predictors. To the best of our knowledge, this is the first exchange rate prediction paper that seeks to accomplish these goals. We investigate, for example, how relevant is the issue of time-variation in coefficients relative to the choice of fundamentals when forecasting out-of-sample.

We use dynamic linear models of the sort considered in Dangl and Halling (2012) when examining stock returns, which we also extend to allow for a time-changing volatility. These models not only allow for a time-varying relationship between exchange rates and fundamentals, but also facilitate assigning posterior probability weights to specifications that differ in the selected fundamental and in the degree of time-variation in coefficients, in light of the relevant evidence. We can then find the specification supported by the data at each point in time, based on these weights. The methodology is also flexible enough in that it enables us to decompose the prediction variance of the exchange rate into its constituent components, highlighting the origins of the prediction uncertainty.

Our predictive regressions employ information sets from classic and the more recently proposed fundamentals. Engel and West (2005) show that exchange rate models based on these fundamentals are consistent with the present-value asset pricing framework. Rossi (2013) surveys the out-of-sample predictive content of several fundamentals-based exchange rate models, including the Taylor rule-based model and the monetary model. The evidence of predictability is more apparent with Taylor rule fundamentals at short horizons and monetary fundamentals at long horizons. Nevertheless, consistent with the idea that the relevant set of fundamentals may change over time, in Rossi's (2013) findings predictability appears momentarily for some currencies and certain forecasting horizons. For example, monetary fundamentals were better predictors of monthly changes in the Japanese yen/U.S. dollar exchange rate around 2006-2008 but not before or after this period. At longer horizons, predictability based on the same fundamentals and for the same currency is never found.

In terms of the empirical design, our dataset consists of monthly data spanning 1977M1 - 2013M5 on eight OECD currency exchange rates relative to the U.S. dollar. We use a direct method to forecast recursively, the period-ahead change in the exchange rate at 1-, 3-, and 12-months horizons. The forecasts from our fundamentals-based models are compared to those of the toughest benchmark – the driftless random walk (RW) (Rossi, 2013). To evaluate the statistical significance of the differences in the forecasts we use the Diebold and Mariano (1995) and West (1996) tests. In order to take account of concerns about data-mining in light of our search over multiple predictors, we employ critical values computed using a data-mining robust bootstrap technique proposed in Inoue and Kilian (2005), which we adjust to allow for cross-country correlation. An additional measure of relative forecast accuracy is based on predictive likelihoods (Geweke and Amisano, 2010).[1]

[1]We also examine the ability of our models to generate economic value in a stylized asset portfolio management setting, following Della Corte *et al.* (2012) and Li *et al.* (2015). We compute indicators such as Sharpe ratios, maximum performance fees, excess premium returns, and break-even transaction costs that render an investor indifferent between using our models and the RW.

Apart from the research on the role of instabilities in obstructing model forecasting performance, our paper is related to the literature on forecast combinations. Among articles studying the importance of instabilities in an exchange rate setting, we compare our paper to Rossi and Sekhposyan (2011), Bacchetta *et al.* (2010), and Giannone (2010). 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. However, while they mention that time-variation in parameters of the models might cause time-variation in forecasting performance, they do not explicitly examine the influence of the former upon the latter. Thus, our study complements theirs, as time-variation in parameters is an integral part of our scrutiny.[2]

Among papers focusing in pooling exchange rate forecasts, we note contributions by Wright (2008), Sarno and Valente (2009), Beckmann and Schuessler (2015), and Li *et al.* (2015). The main difference with our contribution is their emphasis on finding whether combined forecasts from several models with a certain configuration are superior to those from univariate models and to the random walk benchmark. Instead, we extend the analysis in the above papers to examine the origins of model prediction uncertainty. An additional difference is our use of a data-mining robust bootstrap procedure when evaluating our models' forecasting performance.[3]

To preview our results, we find that models which allow for the relevant set of predictors to change over time and with varying degrees of coefficients adaptivity forecast well. For the majority of the currencies we examine, these models generate a lower root mean squared forecast error than the benchmark at all, but 1-month forecasting horizon. At horizons greater than 1-month, regressions with a high degree of time-variation in coefficients dom-

[2]Bacchetta *et al.* (2010) use a theoretical reduced-form model of exchange rate calibrated to match the moments of the data to examine whether parameter instability could rationalize the Meese-Rogoff puzzle. They conclude that it is not time-variation in parameters, but small sample estimation error that explains the puzzle. However, Giannone (2010) disputes these findings and points out that both, time-variation in parameters and estimation uncertainty, are important in accounting for the puzzle. As we noted above, we extend the analysis to consider other sources of instabilities, quantify their relative importance, and our approach is entirely data-based.

[3]Sarno and Valente (2009) use a Reality Check procedure to account for data-mining.

inate regressions with constant and moderately time-varying coefficients. Importantly, we find that apart from the common uncertainty related to random fluctuations in the data, the uncertainty in coefficient estimation and the uncertainty regarding the correct level of time-variation in coefficients are key obstructions to exchange rate prediction. When the models successfully embed these two sources of uncertainty, they yield a satisfactory out-of-sample forecasting performance and economic value. The uncertainty about the choice of the predictors appears to be small. Our findings are therefore consistent with the simulation-based results of Giannone (2010) and they provide supportive evidence for Rossi and Sekhposyan's (2011) conjectures on the causes of time-variation in the models' predictive ability.

The paper is organized as follows. In the next section we lay out our econometric methodology. Section 3 describes the forecasting environment, including the competing models, the menu of predictors, data, forecasting mechanics, and the criteria for forecast evaluation. Results are reported in Section 4, followed by robustness checks in Section 5. Section 6 concludes.

2 Econometric Methodology

2.1 Predictive Regression

In line with the majority of the papers in exchange rate forecasting, we model the exchange rate as a function of its deviation from its fundamental's implied value.[4] As advanced by Mark (1995), this fits with the notion that in the short-run, exchange rates frequently deviate from their long-run fundamental's implied level. More precisely, let $e_{t+h} - e_t \equiv \Delta e_{t+h}$ be the h-step-ahead change in the log of the exchange rate, and f_t a set of exchange rate fundamentals. Then, we consider predictive regressions of the following state space form:

$$\Delta e_{t+h} = X_t \theta_t + v_{t+h}, \quad v_{t+h} \sim N(0, V_t), \text{ (observation equation);} \tag{1}$$

[4] See, for example, Byrne *et al.* (2016), Cheung *et al.* (2005), Engel *et al.* (2007), Mark (1995), Molodtsova and Papell (2009), Rossi and Inoue (2012), and Rossi (2013).

$$\theta_t = \theta_{t-1} + \varpi_t, \quad \varpi_t \sim N(0, W_t), \text{ (transition equation);} \tag{2}$$

with $X_t = [\mathbf{1}, z_t]$, where

$$z_t = f_t - e_t. \tag{3}$$

As equation (3) indicates, z_t measures the disequilibrium between the exchange rate's spot value and the level of the fundamentals. When the spot exchange rate is higher than its fundamental's implied level, then the spot rate is expected to decrease, as long as the coefficient attached to z_t in equation (1) is less than one. In the next section we discuss the set of fundamentals contained in f_t. Here we note that the predictive regression given by the system of equations (1) and (2) allows the coefficients linked to the disequilibrium term (z_t) and to the constant, to change over time. Equation (2) suggests that we assume a random walk process for the parameter θ_t, following Rossi (2006) and Wolff (1987). We further consider that the disturbance terms, v_{t+h} and ϖ_t, are uncorrelated and normally distributed with mean zero and time-varying matrices V_t and W_t, respectively.

The variance of the error term in the transition equation, W_t, is crucial in determining the degree of time-variation in the regression's coefficient. Setting this matrix to zero implies that the coefficients are constant over time, and therefore equation (1) nests a constant-coefficient predictive regression. In contrast, if the variance increases, the shocks to the coefficients also increase. While this renders more flexibility to the model, the increased variability of the coefficients translates into high prediction variance, which increases the prediction error. 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, Ch. 4). We define this structure together with the description of the estimation methodology below.

We use Bayesian methods to estimate the parameters of our dynamic linear model, following Dangl and Halling (2012) and West and Harrison (1997, Ch. 3&4). The estimation is based on a full conjugate Bayesian analysis, implying that when prior information on the unknown parameters is combined with the likelihood function, we obtain a posterior with the same distribution as the prior, hence no simulation algorithms are required. Specifically, let the prior for the coefficients vector θ_t be normally distributed, and the prior for the observational variance V_t be derived from an inverse-gamma distribution. In a conjugate analysis, the posteriors are jointly normally/inverse-gamma distributed. In the supplementary appendix online we provide details on the updating scheme of the system of equations at some arbitrary time $t+1$, given the information available at time t (D_t). This information set contains the exchange rate variations, the predictors, and the prior parameters at time-zero. i.e., $D_t = [\Delta e_t, \Delta e_{t-1}, ..., \Delta e_{1,}, X_t, X_{t-1} ..., X_1, \text{Priors}_{t=0}]$. For the prior parameters at $t = 0$, we follow Fernandez *et al.* (2001) and use a benchmark conjugate g-prior specification:

$$\theta_0 | D_0, V_0 \sim N\left(0, S_0 \left[gX'X\right]^{-1}\right), \tag{4}$$

$$V_0 | D_0 \sim IG\left[\frac{T_0}{2}, \frac{S_0}{2}\right], \tag{5}$$

where $S_0 = (T-1)^{-1} \Delta e' (I - X(X'X)^{-1}X')\Delta e$, and $T_0 = 1$.

The prior for the coefficient vector in equation (4) is a diffuse prior centered around the null hypothesis of no predictability, with g as the scaling factor that conveys the confidence assigned to this hypothesis. The coefficients' variance-covariance matrix is a multiple of the OLS estimate of the variance in coefficients, S_0. The fact that this matrix is multiplied by a large scalar translates into an uninformative prior, implying that the estimation procedure adapts quickly to the empirical pattern. This is consistent with our objective of examining which instabilities are supported by the data. In our empirical exercise, we adopt a common procedure and set our g-prior based on estimates from the entire sample (see, e.g., Dangl and Halling, 2012 and Wright, 2008). Following the recommendations in Fernandez *et al.* (2001), we set $g = 1/T$ for the main results and examine cases of relatively more informative prior in the robustness checks.

The other crucial element in the methodology we employ is the predictive density. This is obtained by integrating the conditional density of Δe_{t+h} over the space spanned by θ_t and V_t. West and Harrison (1997, Ch. 4) show that it is a Student $t-$distribution with n_t

degrees-of-freedom, mean $\widehat{\Delta e}_{t+h}$, variance Q_{t+h}, evaluated at Δe_{t+h}:

$$f(\Delta e_{t+h}|D_t) = \mathbf{t}_{n_t}(\Delta e_{t+h}; \widehat{\Delta e}_{t+h}, Q_{t+h}). \tag{6}$$

Using this predictive distribution we can recursively forecast Δe_{t+h}; for details, see the supplementary appendix online.

As we pointed out, the degree of time-variation in the regressions' coefficients is determined by the matrix W_t in equation (2). Given that the coefficients are exposed to random shocks that follow a normal distribution with mean zero and variance W_t, when the variance is low, 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. To capture this direct relationship between the coefficients' estimation error and the variance, we follow West and Harrison (1997, Ch. 4) and let W_t be proportional to the estimation variance of the coefficients at time t:

$$W_t = \frac{1-\delta}{\delta} S_t C_t^*, \quad 0 < \delta \leq 1; \tag{7}$$

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 (1), and δ is a discount factor that controls the degree of time-variation in coefficients (see also Dangl and Halling, 2012).

Effectively, setting $\delta = 1$ implies that $W_t = 0$, and therefore the coefficients are assumed 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 $1/\delta$. For instance, with $\delta = 0.98$ the variance increases by 62% within 24 months. Further reducing δ to 0.90, translates into 52% increase in the variance in just five months, suggesting remarkably abrupt changes in coefficients. The value of δ can alternatively be interpreted as implying that observations m periods in the past carry a weight of δ^m in the estimation. Based on this interpretation, setting δ to 0.90 means that observations two years in the past have just 8% as much weight as last period's observation. Further lowering it to 0.80 for example, would imply a weight close to zero for those observations.

In our empirical work we consider $\delta = [0.90, 0.92, 0.94, 0.96, 0.98, 1.00]$ as the possible support points for time-variation in coefficients, which appears to cover all the sensible values for δ given our monthly frequency data. Beckmann and Schüssler (2015) in an exchange rate study, and Dangl and Halling (2012) in a stock returns context consider a range of $[0.96, 0.97, 0.98, 0.99, 1.00]$. West and Harrison (1997, Ch. 2) examine cases of $[0.70, 0.80, 0.90, 1.00]$ for the U.S. dollar/U.K. pound sterling exchange rate and find evidence of large support from the data for 0.90. They also point out that a lower limit consisting of 0.70 or a smaller value is unreasonable for monthly data, as it implies extremely adaptive and implausible models.[5] From the range of time-variation in coefficients that we consider, we then examine empirically which support point is consistent with the data in a Bayesian model averaging approach that we discuss in the next subsection.

2.2 Dynamic Model Averaging and Selection

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 (Bacchetta and van Wincoop, 2004; Rossi, 2013; and Sarno and Valente, 2009). To address this latter source of instability, we allow for the possibility that from a set of k potential predictors, one applies at each time period. If we let d be the number of possible discrete support points for time-variation in coefficients as defined by each δ, then our range of possible models is $d.k$.

Selecting one specific model, characterized by a specific choice of predictor and degree of time-variation in coefficients, and using it to forecast at time t requires a method. Bayesian model selection is a methodical approach that tests the validity of all $d.k$ models against the observed data. The approach involves assigning prior probabilities to each candidate predictor, as well as prior probability to each possible support point for time-variation in

[5] Koop and Korobilis (2012) consider alternative values of $\delta = 0.95$ or $\delta = 0.99$ when forecasting inflation with quarterly data. Using similar discounting factor concepts, Longerstaey and Spencer (1996, Ch. 5) estimate decay factors based on measures of forecast performance from 460 financial series and find 0.94 to be an optimal value for daily data and 0.97 for monthly data.

parameters. Then based on the realized likelihood of the model's prediction, the posterior probability of each of the $d.k$ models is updated according to Bayes rule. In the supplementary appendix we provide details on the exact formulae. Note that we set diffuse conditional prior probability for each model M_i; and equally, an uninformative prior for the range of support points for the degree of time-variation in coefficients. In notation, the prior probabilities are $P(M_i|\delta_j, D_0) = 1/k$ and $P(\delta_j|D_0) = 1/d$, respectively. Hence, at the beginning of the forecast window, each predictor and model setting has the same chance of becoming probable.

The overall model's predictive density is the posterior probability weighted average predictive density of all $d.k$ models. In this sense, we perform Bayesian Model Averaging (BMA) in a setting with varying degrees of time-variation in coefficients. The flexibility of the approach implies, for instance, that we can implement Bayesian Model Selection (BMS), thus selecting the single model with the highest probability at each point and using it to forecast. We can further let $\delta = 1$, such that all the models exhibit constant coefficients and then average over models with this characteristic (BMA excluding time-varying coefficients). We can alternatively keep $\delta = 1$, but select the best model at each time-period (BMS excluding time-varying coefficients).[6] Furthermore, the approach permits us to track all sources of uncertainty with respect to the prediction in a variance decomposition framework. We elaborate on this framework in what follows.

2.3 Variance Decomposition and Sources of Instability

We use the law of total variance to decompose the variance of the random variable (Δe_t) into its constituent parts. Following Dangl and Halling (2012), we begin with the decomposition with respect to different values of δ:

$$Var(\Delta e_{t+h}) = E_{\delta_j}(Var(\Delta e_t|\delta_j)) + Var_{\delta_j}(E(\Delta e_t|\delta_j)), \tag{8}$$

[6] In fact, as we detail in subsection 3.1, we can analyze several other cases depending on model specification and choice of degree of time-variation, including cases of BMA over time-varying coefficients with single predictors.

where E_{δ_j} and Var_{δ_j} indicate the expected value and the variance with regards to δ_j. Because the expected value of the variance of δ_j is conditional on specific choice of model M_i, it can be also decomposed as follows:

$$Var(\Delta e_t|\delta_j) = E_{M_i}(Var(\Delta e_t|M_i, \delta_j)) + Var_{M_i}(E(\Delta e_t|M_i, \delta_j)). \tag{9}$$

Dangl and Halling (2012) show that using equation (9) to substitute for the expected value of the variance of δ_j in equation (8), and employing the corresponding expressions of these variances - see the supplementary appendix online - yields the following decomposition:

$$\begin{aligned} Var(\Delta e_{t+h}) \quad = \quad & \sum_j \left[\sum_i (S_t|M_i, \delta_j, D_t) P(M_i|\delta_j, D_t) \right] P(\delta_j|D_t) \\ & + \sum_j \left[\sum_i (X_t' R_t X_t|M_i, \delta_j, D_t) P(M_i|\delta_j, D_t) \right] P(\delta_j|D_t) \\ & + \sum_j \left[\sum_i (\widehat{\Delta e}^j_{t+h,i} - \widehat{\Delta e}^j_{t+h})^2 P(M_i|\delta_j, D_t) \right] P(\delta_j|D_t) \\ & + \sum_j (\widehat{\Delta e}^j_{t+h} - \widehat{\Delta e}_{t+h})^2 P(\delta_j|D_t). \end{aligned} \tag{10}$$

The four individual terms in equation (10) highlight the sources of uncertainty in the prediction. The first term is the expected variance of the disturbance term in the observation equation, with $(S_t|M_i, \delta_j, D_t)$ measuring the time t estimate of the variance, V_t, given the choice of the predictor and degree of time-variation in coefficients. This provides a measure of random fluctuations in the data relative to the predicted trend component. The second term captures the expected variance from errors in the estimation of the coefficients. It can be referred to as estimation uncertainty. The third term characterizes model uncertainty with respect to the choice of the predictor. The last term also characterizes model uncertainty, but with respect to time-variability of the coefficients. Hence both, the third and fourth term, capture model uncertainty. Implementing this four-fold variance decomposition represents an innovation in the exchange rate literature. We now turn to our forecasting environment.

3 Forecasting Environment

3.1 Competing Models

We first consider alternative model specifications derived from our Bayesian approach, namely:

- BMA including Time-varying Coefficients (BMA incl. Tvar-coeffs): Based on Bayesian model averaging over individual models and with varying degrees of coefficient adaptivity.

- BMA excluding Time-varying Coefficients (BMA excl. Tvar-coeffs): This is a restricted version of the above, as it is based on BMA over individual models that exclude time-variation in coefficients. It corresponds to conventional BMA.

- BMS including Time-varying Coefficients (BMS incl. Tvar-coeffs): This is determined by the individual models that receive the highest posterior probability, among all individual models and with varying degrees of coefficient variation.

- BMS excluding Time-varying Coefficients (BMS excl. Tvar-coeffs): This specification is nested in the BMS including Tvar-coeffs model. It includes the individual models that receive the highest posterior probability, among all individual models excluding time-variation in coefficients.

- Single Predictor including Time-varying Coefficients (Single Predictor and BMA incl. Tvar-coeffs): These models consider only a single predictor at a time, but average over the range of all degrees of time-variation in coefficients.

- Single Predictor excluding Time-varying Coefficients (Single Predictor excl. Tvar-coeffs): This is a restricted version of the Single Predictor including Tvar-coeffs model. It includes only one predictor at a time in a setting excluding time-variation in coefficients.

In addition to these specifications, we look at forecast combination methods based on frequentist approaches. In this case the forecast combination of Δe_{t+h} made at time t, is a weighted average of the k individual models' forecast based on OLS estimates of simple linear regressions, excluding time-varying coefficients. That is,

$$\widehat{\Delta e}_{t+h}^{c} = \sum_{i=1}^{k} \omega_{i,t} \widehat{\Delta e}_{t+h}^{i}, \tag{11}$$

where $\{\omega_{i,t}\}_{i=1}^{k}$ are the ex-ante combining weights formed at time t. Following Stock and Watson (2004) and Rapach *et al.* (2010) we consider the following combination methods:

- Mean Combination: The combined forecasts are obtained by using the following constant weighting scheme: $\omega_{i,t} = 1/k$, for $i = 1, ..., k$ in equation (11).
- Median Combination: The median combination forecasts is the median of $\{\widehat{\Delta e}_{t+h}^{i}\}_{i=1}^{k}$.
- Trimmed Mean Combination: The combined forecasts are obtained by setting $\omega_{i,t} = 0$ for the smallest and largest individual forecasts, and $\omega_{i,t} = 1/(k-2)$ for the remaining forecasts in equation (11). As in the median combination and the DMSPE combination method below, the weights change over time.
- DMSPE Combination: In this method, the weights are related to the historical forecasting performance of the individual models in a holdout-out-of-sample period. The discount mean squared prediction error (DMSPE) method uses the following weights: $\omega_{i,t} = \Phi_{i,t}^{-1} / \sum_{i=1}^{k} \Phi_{i,t}^{-1}$, where $\Phi_{i,t} = \sum_{so}^{t-1} \vartheta^{t-1-so} (\Delta e_{so+h} - \widehat{\Delta e}_{so+h}^{i})^2$ and so is the end of the in-sample portion. The parameter ϑ denotes the discount factor applied to the mean squared prediction error. Based on results in Rapach *et al.* (2010) and Stock and Watson (2004), we set its value to 0.9.[7] This is consistent with attaching greater weight to the individual models that performed better in the holdout-out-of-sample period. We set this holdout-out-of-sample period to five years, implying that for this combination method, the forecast evaluation period starts five years later relative to that of the other models.

[7] In Rapach *et al.* (2010) and Stock and Watson (2004) the best forecasting performance is achieved with a discount factor of 0.9, in a set that includes 0.95 and 1.0.

By examining combined forecasts from models which exclude time-variation in coefficients, we can further check the sources of differences in performance and better understand the importance of time-variation in coefficients.

3.2 Menu of Predictors

The range of predictors we consider are derived from the conventional and the more recently proposed fundamentals-based empirical exchange rate models. 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.

Our first class of predictors is constituted by five observable fundamentals, as follows:

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

$$\begin{aligned} f_{t,TRsy} &= 1.5(\pi_t - \pi_t^*) + 0.5(\overline{y}_t - \overline{y}_t^*) + e_t, & (12) \\ f_{t,TRasy} &= 1.5(\pi_t - \pi_t^*) + 0.1(\overline{y}_t - \overline{y}_t^*) + 0.1(e_t + p_t^* - p_t) + e_t, & (13) \end{aligned}$$

  where π_t is the inflation rate, $\overline{y}_t$ the output gap in the home country, p_t is the log of the domestic price level, and e_t the log exchange rate. Asterisks denote identical variables for the foreign country.

- Monetary fundamentals:

$$f_{t,MM} = (m_t - m_t^*) - \rho(y_t - y_t^*), \quad (14)$$

  where m_t is the log of money supply, and y_t is the log of income. Following Mark (1995), we assume an income elasticity parameter of $\rho = 1$. Note that we do not include other variants of the monetary model in line with the vast literature that ensued following

Mark's (1995) findings that the performance of this model is robust to changes in the income elasticity parameter from 0 to 1; see also Sarno and Valente (2009) for congruent findings.

- Fundamentals implied by Purchasing Power Parity (PPP) and Uncovered Interest Rate Parity (UIRP):

$$f_{t,PPP} = (p_t - p_t^*), \tag{15}$$

$$f_{t,UIRP} = (i_t - i_t^*) + e_t. \tag{16}$$

The class of fundamentals described in equations (12-16) is typical in exchange rate studies, including in Engel *et al.* (2007, 2015), Molodtsova and Papell (2009), Rossi (2013), Li *et al.* (2015), among others.[8]

The second class of fundamentals is derived from non-observable information embedded in exchange rates as propounded by Engel *et al.* (2015). Accordingly, we extract three factors from the exchange rates in our sample and define our last predictors as:

$$f_{t,Factor} = \widehat{l}_{1n}\widehat{F}_{1t} + \widehat{l}_{2n}\widehat{F}_{2t} + \widehat{l}_{3n}\widehat{F}_{3t}, \tag{17}$$

where $\widehat{F}_{1t}$, $\widehat{F}_{2t}$, and $\widehat{F}_{3t}$ are factors estimated by principal components analysis, and $\widehat{l}_{1n}, \widehat{l}_{2n}, \widehat{l}_{3n}$ are the corresponding loadings for the currency n, $n = 1, ..., 8$. The choice of the number of factors is guided by the evidence in Engel *et al.* (2015), who find three factors to be advantageous in terms of forecasting performance.

[8] We are aware of the cyclical external imbalances model of Gourinchas and Rey (2007) and the Chen and Tsang (2013) yield factors-based exchange rate model. We exclude them from our main analysis mainly because data to construct these fundamentals for the countries we examine are not readily available over the span of our sample. We consider these fundamentals later, in a robustness analysis, focusing on a smaller sample and fewer exchange rates.

3.3 Data and Forecasting Mechanics

This paper uses monthly data from 1977M1 to 2013M5 for eight countries: Australia, Canada, Euro-area/Germany, Japan, Norway, Sweden, Switzerland, and the United Kingdom (U.K.). Greenaway-McGrevy *et al.* (2015) note that the currencies of these countries constitute the top eight, after the U.S. dollar, in terms of average (1998-2013) foreign exchange market turnover, allowing us to set the stylized dynamic asset allocation exercise for a sensible basket of currencies. The home country is taken as the United States. The data are obtained from the IMF's International Financial Statistics (IFS), supplemented by national central banks sources. The exchange rate is defined as the end-of-month value of the U.S. dollar price of a unit of national currency. We measure the money supply by the aggregate M1.[9]

Computation of information sets from Taylor rules requires data on the short-run central bank nominal interest rate, the inflation rate, and the output gap. We employ the central bank's policy rate when 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. Following a common practice in the literature, the output gap is obtained by applying the Hodrick and Prescott (1997) filter recursively to the output series, with the standard smoothing parameter for monthly data frequency (i.e., 14400). We equally correct for the uncertainty about these estimates at our recursive sample end-points by following Watson's (2007) method. This entails estimating bivariate VAR(ℓ) models that include the first difference of inflation and the change in the log IP, with ℓ determined by Akaike Information criterion. These VARs are used to forecast and backcast three years of monthly data-points of IP, and the HP filter is applied to the resulting extended series.[10] The data on money supply, IP, and CPI were seasonally adjusted by taking the mean over twelve months

[9] In cases where the M1 aggregate is unavailable, we use a broader aggregate - see the supplementary appendix online for exact details on these aggregates and data sources. Note too that data limitations prevent us from using real-time data for the countries we consider.

[10] We have also experimented with estimating an AR(ℓ) model for $\Delta \ln(IP_t)$ instead of a VAR(ℓ) model. The resulting output gap series were similar to those based on the VAR forecasts, suggesting small differences in the forecast precision between the two models.

following Engel *et al.* (2015).[11]

We use a direct, rather than an iterative, method to forecast the h-month-ahead change in the exchange rate for $h = [1, 3, 12]$. Wright (2008) notes that both methods lead to qualitatively similar conclusions. 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 1995M1 is made using data available up to 1994M10. Our forecasting window begins in 1987M12+h for all regressions, except for the DMSPE combination method, which requires a holdout out-of-sample period.[12]

3.4 Criteria for Forecast Evaluation

We first apply statistical measures of forecasting performance. We compute the ratio of the Root Mean Squared Forecast Error (RMSFE) of our models relative to the RMSFE of the driftless random walk (RW). Models that perform better than the RW benchmark have a value of this ratio, also known as the Theil's U, less than one. According to Rossi (2013), the RW is the most appropriate benchmark. In the supplementary appendix online we use economic metrics for forecast evaluation.

To assess the statistical significance of the differences in the forecasts, many papers employ the Diebold and Mariano (1995) and West (1996) tests (hereafter DMW), and/or the Clark and West (2006, 2007) test (hereafter CW). The DMW tests whether two competing forecasts are identical under general conditions (Diebold, 2015). The CW tests whether the benchmark model is equivalent to the competing model in population. However, Clark and West (2006) show that when comparing nested models, the DMW test is undersized, hence, the RMSFE differential should be adjusted by a term that accounts for the bias introduced by the larger model. On the other hand, Rogoff and Stavrakeva (2008) make the case for using the bootstrapped DMW test, rather than the CW test, arguing that the latter does not always test for a minimum mean squared forecast error. Additionally, they recall that the asymptotics of the CW test are well-defined when forecasting in a rolling, instead of a

[11] CPI and IP data for Australia are only available at a quarterly frequency. We obtain monthly data via a quadratic-match-average interpolation method from quartely data, as in Molodtsova and Papell (2009).

[12] Note that the exchange rate factors are also constructed without using future information.

recursive framework.

In light of the preceding discussion, we construct bootstrapped critical values for the one-sided DMW test in the spirit of Rogoff and Stavrakeva (2008). Given our search over multiple predictors, however, we modify their procedure to account for data-mining, following Inoue and Kilian (2005) and Rapach and Wohar (2006). We also allow for cross correlation across countries; see, for example, Della Corte and Tsiakas, (2012). In any event, to examine if inference based on the standard critical values alters our conclusions, we further employ the asymptotic DMW. The exact bootstrap steps are described in a supplementary appendix online.[13]

4 Empirical Results

We begin by examining the forecasting performance of our competing methods. We then proceed and study in detail the characteristics of the BMA including time-varying coefficients. In this respect, we analyze the sources of prediction uncertainty, the degree of time-variation in coefficients consistent with the data, and which macroeconomic fundamentals are highly informative about exchange rates movements.[14]

4.1 Out-of-Sample Statistical Evaluation

Table 1 shows in three panels the forecasting results of the predictive regressions that allow predictors and coefficients to change over time, and all the restricted versions that take into account multiple predictors. The key findings can be summarized as follows. First, regressions which allow for time-changing sets of predictors and varying degrees of coefficients adaptivity yield a smaller Root Mean Squared Forecast Error (RMSFE) than the RW at the

[13]The Diebold and Mariano (1995) and West (1996) test is computed as: $DMW = \overline{f}\sqrt{P}/[$sample variance of $\widehat{f}_{t+h} - \overline{f}]^{1/2}$; where P is the number of out-of-sample forecasts, $\widehat{f}_{t+h} = \widehat{fe}^{2}_{1,t+h} - \widehat{fe}^{2}_{2,t+h}$, with $\widehat{fe}_{1,t+h}$ denoting the h-step-ahead forecast error of the RW, $\widehat{fe}_{2,t+h}$ the corresponding forecast error of the alternative model, and $\overline{f}$ is the mean of $\widehat{f}_{t+h}$. Note that in computing the tests, including in our bootstrap, we use Newey and West (1987) HAC standard errors with a lag truncation parameter of $int\{Sample^{0.25}\}$, following Rossi (2013).

[14]To preserve space, we report the results for economic evaluation in a supplementary appendix online.

Table 1: Statistical Evaluation of Bayesian and Forecast Combinations Methods

Panel A: Models that allow predictors and coefficients to change over time

	h=1	h=3	h=12	h=1	h=3	h=12
	BMA incl. TVar-Coeffs			BMS incl. TVar-Coeffs		
AUD	1.029	**0.893****	**0.626***^{a}**	1.007	**0.886****	**0.620**^{a}**
CAD	1.031	**0.966**	**0.823***	1.022	**0.876**^{a}**	**0.761****
GBP	1.040	**0.977**	**0.756****	1.001	**0.956**	**0.676**^{a}**
JPY	1.040	**0.961***	**0.825*****	1.019	**0.944**^{b}**	**0.802**^{a}**
NOK	1.032	1.008	**0.876****	1.018	**0.962**	**0.843****
SEK	1.018	**0.962**	**0.944**	1.005	**0.955**	**0.758****
CHF	1.045	**0.953**	**0.754***^{a}**	1.033	**0.925**^{b}**	**0.726**^{b}**
EUR	1.014	**0.935**^{a}**	**0.817***^{a}**	1.008	**0.906***^{c}**	**0.789**^{b}**

Panel B: Models that allow predictors to change over time, excl. TVar-Coeffs

	BMA excl. TVar-Coeffs			BMS excl. TVar-Coeffs		
AUD	1.015	1.029	1.005	1.011	1.022	**0.985**
CAD	1.009	1.014	1.097	1.000	**0.995**	1.094
GBP	1.018	**0.984**	**0.869***	**0.982^{a}**	**0.977**	**0.871***
JPY	1.023	1.030	1.058	1.005	1.013	1.043
NOK	1.020	1.021	**0.970**	1.009	1.001	**0.957****
SEK	1.024	1.017	**0.937****	1.013	1.012	**0.926****
CHF	1.029	1.035	**0.935**	1.011	1.017	**0.919***
EUR	1.005	1.010	**0.870*****	1.006	1.000	**0.859****

Panel C: Combined forecasts, excl. TVar-Coeffs

	OLS-Mean Combination			OLS-Median Combination		
AUD	1.003	1.008	1.030	1.004	1.007	1.025
CAD	1.002	1.007	1.027	1.003	1.008	1.032
GBP	1.000	1.000	**0.986**	1.000	1.001	**0.986**
JPY	1.000	**0.999**	**0.989**	**0.999**	**0.998**	1.016
NOK	1.003	1.011	1.029	1.004	1.012	1.032
SEK	1.002	1.010	1.009	1.000	1.002	**0.993**
CHF	1.000	**0.997**	**0.960***	**0.998**	**0.992**	**0.937****
EUR	**0.999**	**0.998**	**0.972***	**0.997**	**0.993**	**0.961***
	OLS-Trimmed Mean			OLS-DMSPE Combination		
AUD	1.003	1.007	1.026	1.001	1.002	**0.991**
CAD	1.002	1.007	1.029	1.001	1.003	**0.992**
GBP	**0.999**	**0.998**	**0.979**	**0.999**	**0.994^{a}**	**0.932****
JPY	1.000	**0.999**	**0.988**	1.001	**0.996**	**0.889**^{b}**
NOK	1.003	1.011	1.026	1.003	1.007	1.001
SEK	1.001	1.004	**0.992**	1.002	1.005	**0.965****
CHF	**0.999**	**0.995**	**0.961***	**0.999**	**0.989**	**0.877**^{b}**
EUR	**0.998**	**0.996**	**0.971**	**0.999**	**0.993**	**0.908**^{b}**

Notes: Relative RMSFE of the Bayesian Model Averaging (BMA) and Selection (BMS) including or excluding Tvar-coeffs and simple forecast combinations methods. For all methods, the driftless Random Walk (RW) constitutes the benchmark model. Hence, values below one indicate that the method under scrutiny generates a lower RMSFE than RW. The Table also reports the one-sided DMW test-statistic, with p-values based on a data-mining robust semi-parametric bootstrap and standard critical values. The superscripts a for 10%, b for 5%, and c for 1% indicate the level of significance at which the null hypothesis of equal RMSFE is rejected using the bootstrapped critical values, favouring the alternative that the fundamental-based method yields a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) indicate the rejection of the same null, using standard critical values. Currency codes denote: AUD - Australian dollar; CAD - Canadian dollar; GBP - Pound sterling; JPY - Japanese yen; NOK - Norwegian krone; SEK - Swedish krona; CHF - Swiss franc; EUR - euro. The forecast evaluation period begins in 1987M12+h in all, but the OLS-DMSPE Combination case (1992M12+h).

3- and 12-months horizons, but not at the 1-month horizon. In Panel A, BMA including Tvar-coeffs, for example, generates more accurate forecasts than the RW for seven out of the eight currencies considered at the 3-months horizon, and all the currencies at the 12-months horizon. Qualitatively similar findings apply to BMS including Tvar-coeffs, which produces smaller RMSFE than the RW for all currencies at the two horizons.

Second, the magnitude of the reduction in the RMSFE, and the cases of statistical significant differences in the RMSFE relative to the RW increase with the forecasting horizon. For instance, at the 3-months horizon and under the asymptotic DMW, the magnitude of significant improvements upon the RW is of at least 5.6% and a maximum of 12.4% in BMS including Tvar-coeffs. At the 12-months horizon, the gains raise to a minimum of 15.7% and a maximum of 38.0%. And the differences in the RMSFE are statistically significant for over 42% of the currencies at the 3-months horizons and 83% at the 12-months when using the asymptotic DMW test. Instead, if we consider the critical values implied by our data-mining robust bootstrap, the proportion of significant cases is of at least 50% for BMS including Tvar-coeffs at the two horizons. This proportion drops to less than 30% for BMA including Tvar-coeffs.

Third, none of the approaches which exclude time-variation in parameters achieve the magnitudes of improvements described above at horizons beyond 1-month. At best, in Panel B and C, methods such as BMA and BMS excluding Tvar-coeffs predominantly outperform the RW at the longest horizon we consider. And in these cases, the improvement is mostly for fewer currencies when compared with the Bayesian methods with Tvar-coeffs, and the reduction in the RMSFE never exceeds 1.6% or 14.1% at the 3- or 12-months horizon, respectively. Moreover, the differences in RMSFE are rarely significant at the 3-months horizons, regardless of whether bootstrapped or asymptotic critical values are employed. At the longest horizon, BMS excluding Tvar-coeffs and the DMSPE combination significantly outperform the RW for at least half of the currencies when using asymptotic critical values. But only the results from DMSPE combination method remain robust to using bootstrapped critical values for three currencies. Finally, we equally note that at the 1-month horizon, these constant-coefficients methods largely fail to outperform the RW benchmark.[15]

Table 2 presents the results for Bayesian methods based on single predictors including Tvar-coeffs and excluding Tvar-coeffs. At the 1-month horizon, none of the predictive regressions improves upon the RW. In contrast, at the 3-and 12-months horizons there is at least one or more predictors that yield a smaller RMSFE than the RW for each currency in regressions with Tvar-coeffs. This is the case for BMA and BMS including Tvar-coeffs in panels A and B. Examining the cases of regressions excluding Tvar-coeffs in Panel C and for similar horizons, there are several currencies for which none of them forecasts better than the RW. The exchange rates of five currencies, namely the AUD, CAD, JPY, SEK, and CHF at the 3-months horizon, and the CAD and JPY at the 12-months horizon are a case in point. And among the cases where these regressions beat the RW, the number of statistically significant improvements tends to be smaller relative to regressions with Tvar-coeffs. This applies irrespective of whether standard or bootstrapped critical values are used.

Table 2: Statistical Evaluation of Single-Predictor Models Including or Excluding TVar-Coeffs

Panel A: Single Predictor and Bayesian Model Averaging (BMA) incl.TVar-Coeffs

	AUD	CAD	GBP	JPY	NOK	SEK	CHF	EUR
				h=1				
TRsy	1.032	1.030	1.031	1.041	1.025	1.025	1.038	1.014
TRasy	1.019	1.063	1.030	1.029	1.023	1.022	1.041	1.016
PPP	1.039	1.030	1.043	1.038	1.028	1.034	1.041	1.026
UIRP	1.021	1.031	1.028	1.035	1.027	1.009	1.041	1.038
MM	1.014	1.031	1.036	1.040	1.033	1.040	1.050	1.026
F1	1.033	1.016	1.033	1.037	1.025	1.030	1.025	1.029
F2	1.026	1.037	1.074	1.041	1.050	1.068	1.036	1.028
F3	1.040	1.041	1.024	1.043	1.029	1.035	1.033	1.031
				h=3				
TRsy	1.044	**0.959**	**0.973**	**0.957**^{**b}	**0.997**	1.001	**0.965**	**0.924**^{**b}
TRasy	1.001	**0.984**	**0.988**	1.007	**0.989**	1.007	1.044	**0.963**
PPP	1.068	**0.964**	**0.994**	**0.954**^{**b}	**0.981**	1.019	**0.939**^{**b}	**0.949**^{**a}
UIRP	**0.935**	**0.972**	**0.959**	**0.983**	**0.963**	**0.942**	**0.994**	**0.977**
MM	**0.909**	**0.940***	**0.949***	1.014	**0.990**	1.069	1.003	**0.962**
F1	1.020	**0.961**	**0.948**	**0.968**	1.011	**0.963**	**0.995**	1.024

[15] Our results on the performance of BMA excluding TVar-Coeffs are coherent with those from Wright (2008). He finds that in a setting of tighter priors and shrinkage towards the null of no predictability, BMA excluding TVar-coeffs improves upon the RW, although the improvement in terms of reduction in the RMSFE is small. However, with loose priors and less shrinkage, it fails to improve upon the RW.

Table 2 Continued

	AUD	CAD	GBP	JPY	NOK	SEK	CHF	EUR
F2	**0.972**	1.075	1.048	**0.991**	1.131	1.119	1.005	1.011
F3	**0.966**	1.046	**0.992**	**0.990**	1.018	**0.991**	**0.975**	**0.974**
					h=12			
TRsy	1.008	**0.957**	1.260	1.295	**0.992**	**0.802****	**0.735*****^{c}	**0.816*****^{c}
TRasy	1.052	1.714	1.007	3.210	1.352	1.209	1.051	1.370
PPP	1.033	**0.950**	1.208	1.203	**0.868*****^{a}	**0.789*****^{b}	**0.766*****^{b}	**0.820*****^{c}
UIRP	**0.897***	**0.968**	1.108	**0.819*****^{b}	**0.850*****^{b}	1.277	1.399	1.710
MM	**0.811*****	**0.723*****^{a}	**0.706*****^{a}	1.112	1.476	**0.873***	**0.911****^{a}	**0.832*****^{a}
F1	**0.624*****	1.147	**0.704*****^{b}	1.383	**0.961**	**0.804*****^{b}	1.015	1.079
F2	1.682	1.645	**0.872***	1.081	1.027	1.004	**0.854****^{a}	**0.917**
F3	1.000	**0.896**	1.643	1.184	1.012	**0.841****	**0.856***	**0.884***

Panel B: Single Predictor and Bayesian Model Selection (BMS) incl.TVar-Coeffs

					h = 1			
	AUD	CAD	GBP	JPY	NOK	SEK	CHF	EUR
TRsy	1.022	1.033	1.035	1.041	1.024	1.023	1.042	1.015
TRasy	**0.982**	1.052	1.020	1.010	1.002	1.011	1.039	1.014
PPP	1.034	1.026	1.044	1.039	1.026	1.033	1.042	1.026
UIRP	1.011	1.029	1.028	1.030	1.027	1.012	1.041	1.035
MM	1.010	1.028	1.036	1.039	1.032	1.034	1.051	1.027
F1	1.026	1.015	1.039	1.036	1.021	1.029	1.024	1.023
F2	1.022	1.035	1.064	1.042	1.035	1.067	1.035	1.025
F3	1.021	1.038	1.023	1.043	1.026	1.037	1.034	1.029
					h=3			
TRsy	1.014	**0.957**	**0.959**	**0.955****^{b}	**0.987**	**0.972**	**0.963**	**0.921****^{b}
TRasy	**0.996**	**0.981**	**0.980**	1.004	**0.987**	1.008	1.037	**0.959***
PPP	1.044	**0.964**	**0.980**	**0.952****^{b}	**0.968**	**0.998**	**0.936****^{b}	**0.947****^{a}
UIRP	**0.934**	**0.964**	**0.951**	**0.981**	**0.960**	**0.939**	**0.992**	**0.975**
MM	**0.907**	**0.939***	**0.946***	**0.975**	**0.976**	1.036	**0.989**	**0.960**
F1	1.003	**0.956**	**0.934**	**0.966**	1.009	**0.961**	**0.994**	1.009
F2	**0.955**	1.018	1.014	**0.986**	1.064	1.065	**0.992**	**0.993**
F3	**0.955**	1.045	**0.992**	**0.987**	**0.995**	**0.976**	**0.967**	**0.974**
					h=12			
TRsy	**0.982**	**0.920**	**0.936**	1.284	**0.984**	**0.782*****^{a}	**0.716*****^{c}	**0.805*****^{c}
TRasy	1.021	1.709	**0.997**	3.202	1.349	1.202	1.032	1.362
PPP	**0.987**	**0.915**	**0.978**	1.186	**0.858*****^{a}	**0.768*****^{b}	**0.764*****^{b}	**0.803*****^{c}
UIRP	**0.890**	**0.967**	1.094	**0.811*****^{b}	**0.843*****^{b}	**0.925**	1.340	1.624
MM	**0.803*****	**0.715*****^{a}	**0.702*****^{a}	1.070	1.464	**0.854****	**0.902****^{a}	**0.805*****^{a}
F1	**0.623*****	1.143	**0.699*****^{b}	1.371	**0.944**	**0.796*****^{b}	**0.993**	1.052
F2	1.673	1.639	**0.867***	1.069	1.019	1.001	**0.844****^{a}	**0.903**
F3	**0.981**	**0.844****	1.622	1.158	1.007	**0.817*****^{a}	**0.848***	**0.873***

Table 2 Continued

	AUD	CAD	GBP	JPY	NOK	SEK	CHF	EUR
Panel C: Single Predictor excl. TVar-Coeffs								
	AUD	CAD	GBP	JPY	NOK	SEK	CHF	EUR
					h=1			
TRsy	1.024	1.009	1.023	1.028	1.002	1.019	1.029	1.020
TRasy	1.004	1.010	1.011	1.009	1.006	1.005	1.021	1.007
PPP	1.026	1.009	1.037	1.030	1.010	1.019	1.031	1.024
UIRP	1.011	1.015	1.036	1.020	1.011	1.017	1.021	1.021
MM	1.030	1.019	1.043	1.029	1.038	1.022	1.047	1.022
F1	1.026	1.005	1.052	1.024	1.011	1.014	1.015	1.016
F2	1.026	1.008	1.059	1.034	1.037	1.039	1.014	1.016
F3	1.014	1.008	1.029	1.027	1.013	1.013	1.015	1.018
					h=3			
TRsy	1.035	1.006	**0.981**	1.041	**0.986**	1.012	1.044	1.003
TRasy	1.020	1.031	**0.987**	1.071	1.001	1.001	1.059	**0.990**
PPP	1.056	1.009	1.007	1.029	1.018	1.013	1.036	1.013
UIRP	1.017	1.049	1.056	1.025	1.011	1.027	1.023	1.026
MM	1.043	1.023	1.023	1.059	1.175	1.025	1.090	1.017
F1	1.029	1.004	1.013	1.025	1.037	1.017	1.025	1.021
F2	1.043	1.035	1.046	1.075	1.113	1.052	1.013	1.019
F3	1.028	1.017	1.052	1.029	1.022	1.047	1.019	1.016
					h=12			
TRsy	1.189	1.128	**0.857***	1.494	1.120	**0.897****	**0.950**	**0.885***^{c}**
TRasy	**0.945***	1.709	1.010	3.194	1.554	1.229	1.110	1.355
PPP	1.196	1.166	**0.850****	1.440	**0.956****	**0.949**	**0.964**	**0.887***^{b}**
UIRP	1.040	1.083	1.136	1.019	1.032	**0.999**	3.121	2.440
MM	1.276	1.922	**0.984**	1.315	2.005	**0.975**	**0.952***	**0.996**
F1	1.050	1.041	**0.879****	1.633	**0.988**	**0.996**	1.036	1.046
F2	1.780	1.884	1.040	1.547	1.036	1.016	1.051	1.123
F3	1.057	1.096	2.068	1.502	1.231	1.266	1.251	1.104

Notes: RMSFE of the single-predictor models including or excluding time-varying coefficients, relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the model under under scrutiny generates a lower RMSFE than the RW. The description of the predictors is as follows: TRsy and TRasy correspond to the symmetric and asymmetric Taylor rules, respectively; MM- fundamentals from the Monetary Model; PPP - Purchasing Power Parity; UIRP- Uncovered Interest Rate Parity; and F1, F2, F3 the first, second, and third factor respectively. The superscripts a for 10%, b for 5%, and c for 1% denote the level of significance of the one-sided DMW test based on a semi-parametric bootstrap, for the null hypothesis of equality in the RMSFE. Asterisks (*10%, **5%, ***1%) indicate the rejection of the same null, using standard critical values. The forecast evaluation period begins in 1987M12+h.

All in all, therefore, our findings suggest that at the 1-month horizon the Meese and Rogoff's (1983) exchange rate predictability puzzle prevails, even with an approach as flexible as the Bayesian method we consider. The approach, however, is beneficial in terms of improving the out-of-sample forecasting performance at the 3-and 12-months horizons.[16] And the fact that it still delivers relatively better forecasts than models with constant coefficients also when using each fundamental in isolation, hints at the view that allowing for a flexible time-variation in parameters is comparatively more critical than allowing for time-changing fundamentals. We scrutinize this view and other key characteristics of our flexible Bayesian method in the next subsection.

4.2 Characterization of BMA Including Time-Varying Coefficients

The results from our flexible Bayesian method emanate from a complex combination of the performance of multiple individual models. Understanding the method's underlying characteristics is therefore useful in explaining the sources of difference in forecasting ability relative to other competing models and across forecasting horizons. This constitutes a key contribution of this paper.

4.2.1 Sources of Prediction Uncertainty

We begin by analyzing the sources of prediction uncertainty through a variance decomposition process. As noted in our methodological section, we decompose the total variance into observational variance, variance due to errors in the estimation of the coefficients, variance due to model uncertainty with respect to the choice of the predictor, and variance associated with model uncertainty with regards to the choice of the degree of time-variation in coefficients. We plot the recursive decomposition at the 1- and 12-months horizons for a representative selection of four currencies - CAD, GBP, JPY, and EUR. The plots for the 3-months horizon exhibit patterns of intermediate cases, hence, they are omitted here to conserve space but shown in the supplementary appendix online. The main characteristic of the plots at the 3-months horizon is that they resemble the pattern of the ones at the 1-month horizon when BMA with Tvar-Coeffs fails to improve upon the RW, such as for the NOK, or when the magnitude of improvement is somewhat small - GBP and CAD. When

[16] The evidence we report in the supplementary appendix online also shows that the approach yield concrete economic gains at the 3-and 12-months horizons.

the improvement is significant, the plots are similar to those at the 12-months horizon.

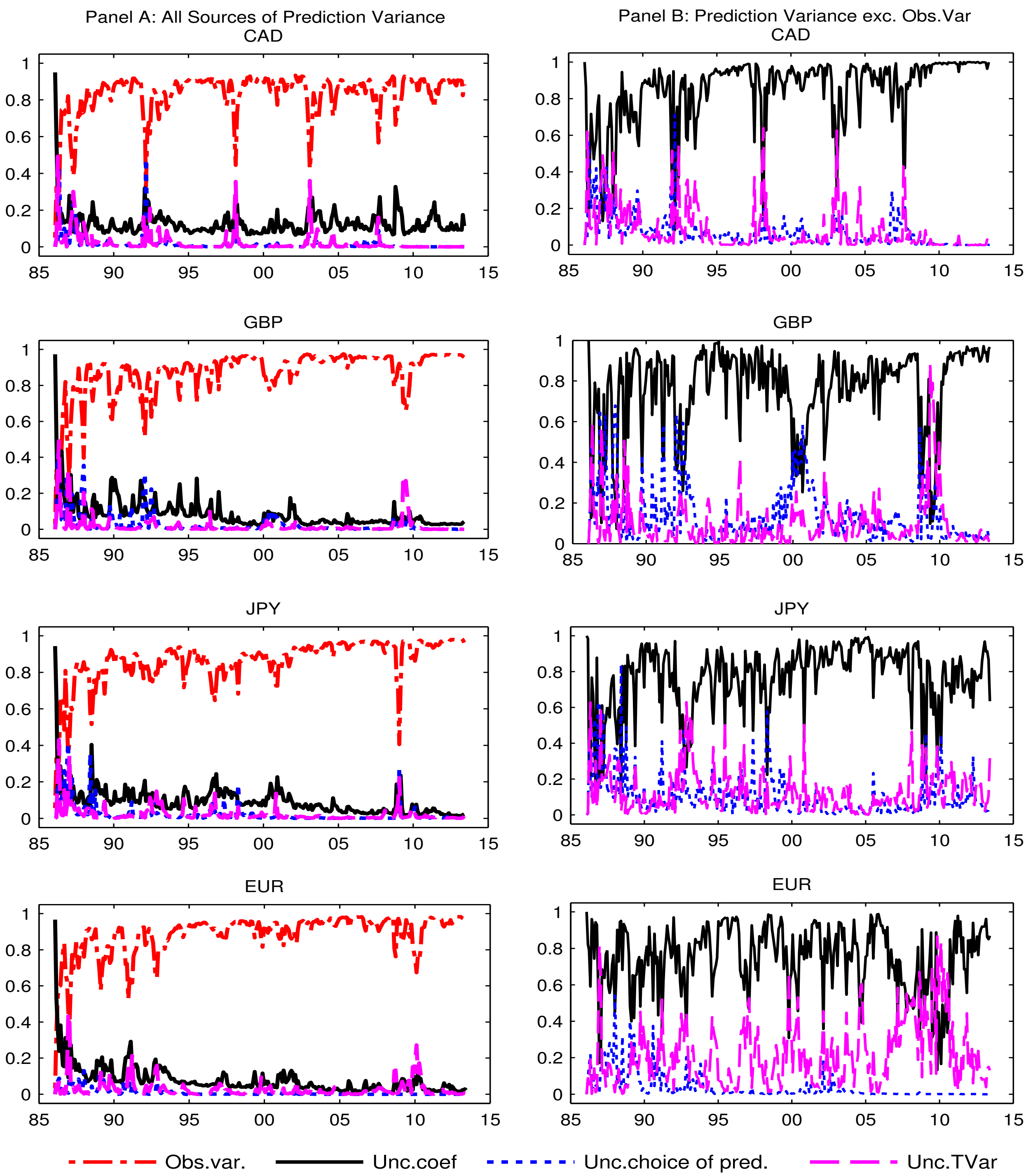
Figures 1 and 2 depict the variance decomposition for the two horizons and countries we focus on. In both figures, Panel A illustrates the relative weight of each of the components of prediction variance in the total variance. For all countries and horizons, the predominant source of uncertainty is observational variance. Dangl and Halling (2012) point out that this is normal for asset prices, as they frequently fluctuate randomly over their expected values. These fluctuations are expected to be noticeable for the horizons that we are considering and to dominate the predicted trend component. In Panel B of the same figures we exclude the observational variance allowing us to focus upon the relative weights of the remaining three sources of prediction uncertainty. The variance from errors in the estimation of the coefficients is now the dominant source of prediction uncertainty at both horizons. Thus, estimation uncertainty is the second most important source of prediction uncertainty and, therefore, one of the main factors hindering model forecasting performance.

Between the 1-month and the 12-months horizons, however, there are differences with respect to the extra two sources of prediction variance in Panel B. At the 1-month horizon, the uncertainty about which predictor is more informative about changes in exchange rates and, most notably, the uncertainty regarding the correct degree of time-variation in the regressions coefficients are detectable throughout the forecasting sample. Taking the case of the JPY as an example, in various periods, the uncertainty regarding the choice of the degree of time-variation in the coefficients represents over one tenth of the total variance excluding observational variance, peaking in certain periods, such as the 2008 financial crisis. In contrast, at the 12-months horizon the two sources of uncertainty are clustered at the beginning of the out-of-sample period, which corresponds to the initial data-points in the expanding window of the forecasting procedure.[17] As more evidence is accumulated, they remain low or episodic for the most part of the forecasting window.

We note, though, that while at the 1-month horizon the uncertainty about the choice of

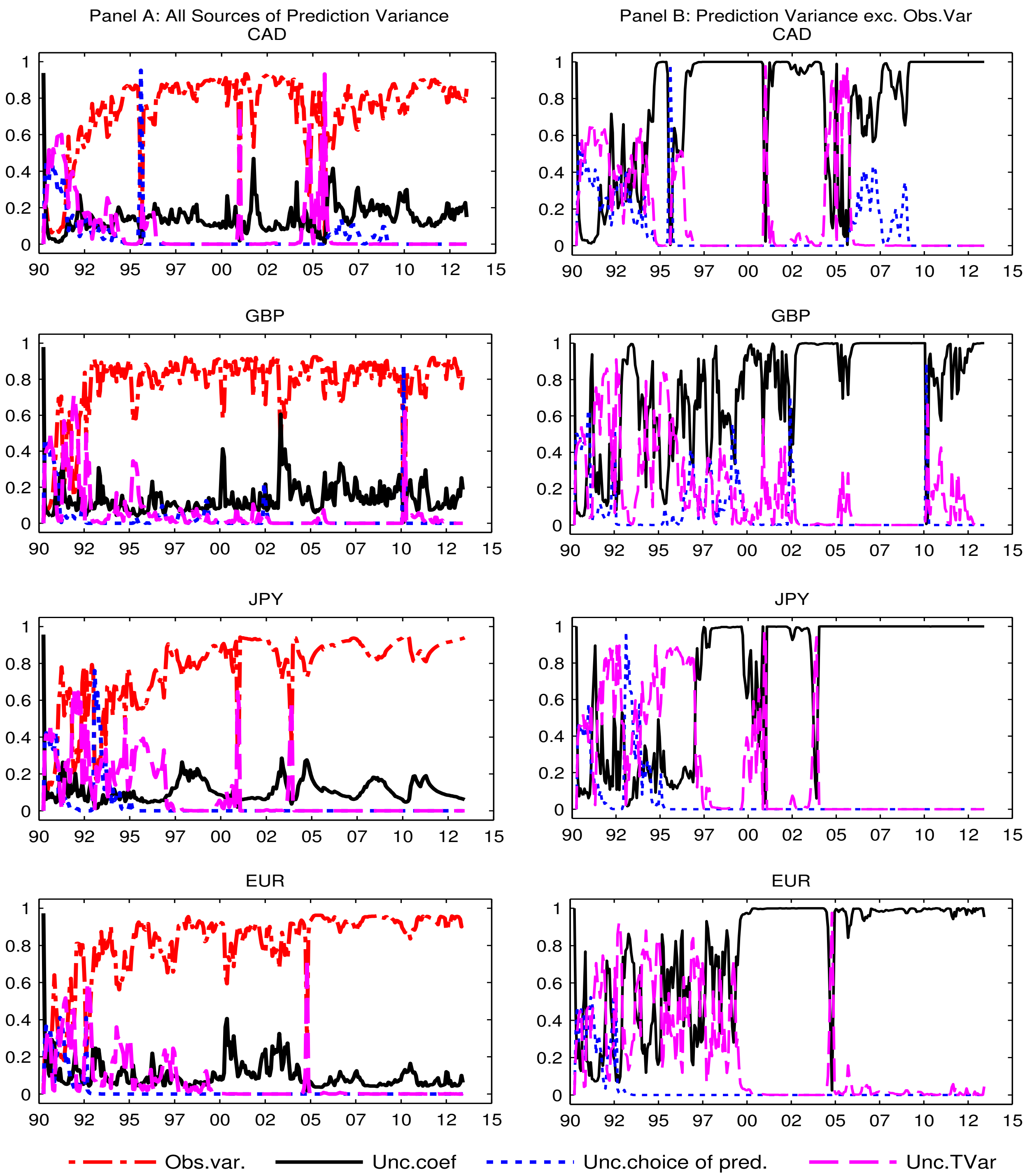
[17] Geweke and Amisano (2010) point out that it is customary for the results in a Bayesian analysis to be sensitive at the beginning of the out-of-sample period, as this reflects sensitiveness to the prior density. As they emphasize, nonetheless, after a number of observations have been accumulated the results become invariant to substantial changes in the prior density distribution.

Figure 1: Sources of Prediction Variance at 1-Month Forecasting Horizon



Notes: Decomposition of the prediction variance into its constituent parts at the 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure 2: Sources of Prediction Variance at 12-Months Forecasting Horizon



Notes: As in Figure 1, except that here the forecasting horizon is 12-months.

the relevant predictor is apparently non-negligible in the cases of the GBP and the JPY, this appears to be less critical than estimation uncertainty and the uncertainty about the level of time-variation in coefficients in terms of influencing forecasting performance. As we will illustrate when examining the importance of individual predictors, the reason is that for the most part of the forecasting window, effectively up to three predictors are selected and used in the regression. And the switches between these predictors are largely infrequent.

We interpret these findings as suggesting that although estimation uncertainty is a key obstructing factor at both horizons, at the 1-month horizon our flexible models fail to improve upon the RW because of the additional uncertainty regarding the precise level of time-variation in coefficients necessary to capture instabilities present in the data. Put differently, there is no certainty about the exact degree of time-variation in coefficients to embed in the model, in order to offset the loss in forecasting performance emanating from estimation uncertainty. On the contrary, as the forecasting horizon increases our models successfully embed the level of time-variation in coefficients present in the data. They therefore counterbalance the loss in the precision in the coefficients' estimation with increased variability of the coefficients, leading to more accurate forecasts than RW. This means that beyond the typical uncertainty associated with random fluctuations in the data, both, estimation uncertainty and coefficient instability obstruct exchange rate models forecasting performance. And our BMA and BMS including Tvar-coeffs adapt to the pattern in the data at longer forecasting horizons.

We relate these findings to Bacchetta *et al.* (2010) and Giannone (2010). Bacchetta *et al.* (2010) calibrate a theoretical reduced-form model of the exchange rate on actual data to examine whether parameter instability rationalizes the Meese and Rogoff (1983) result of exchange rate unpredictability. They find that estimation uncertainty is the main factor that hinders exchange rate models forecasting performance and not time-variation in coefficients. Giannone (2010) disputes these findings and shows that both, estimation uncertainty and parameter instability, might explain the Meese-Rogoff puzzle.

Giannone (2010) also examines the trade-offs between estimation error and parameter uncertainty. He observes three stages when forecasting in an expanding window of data.

The first stage is characterized by a high forecast error with the first few observations of the forecasting period. In the second stage, the forecast accuracy increases as the estimation window is expanded beyond these few initial observations, signalling reduction in the coefficients' estimation error. In the third stage, however, further increasing the estimation window deteriorates the forecasting performance, as gains from reduced estimation error are compensated by losses due to the presence of structural instabilities. Accordingly, the recursive ratio of the relative RMSFE over the forecast window loosely resembles a $\cup$ shape.

The figures of our recursive relative RMSFE for the BMA excluding Tvar-coeffs and BMA including Tvar-coeffs are relegated to the supplementary appendix. The noteworthy fact from the figures is that they elucidate Giannone's (2010) observations for the first and second stages, but not the third stage. In our case, further increasing the window does not lead to a complete deterioration of the forecasting performance relative to the RW. For the most part of the forecasting period and horizons greater than 1-month the relative RMSFE is below one, favouring our flexible models. When read in conjunction with the main sources of instabilities we detect, this reinforces our conclusions that BMA and BMS including TVar-Coeffs successfully capture the degree of time-variation in parameters necessary to offset the loss in forecast accuracy due to estimation uncertainty.

To add up to these results, in the supplementary appendix online we focus in another measure of forecast accuracy that underlies our Bayesian approach, namely the predictive likelihoods, see Geweke and Amisano (2010). The two main insights from the measure are as follows. First, the measure confirms that models with time-varying coefficients are empirically plausible, especially as the forecasting horizon increases. Second, observations around the 2008 financial crisis contribute highly to the evidence in favor of the time-varying coefficients models. Overall, our findings remain invariant to this measure of forecast accuracy and they further rule out the $\cup$ pattern reported in Giannone (2010).

4.2.2 Fluctuation Test

In addition to the recursive relative RMSFE and the cumulative log predictive likelihoods, we formally test for forecasting ability over time in presence of instabilities by implementing the

Giacomini and Rossi's (2010) one-sided Fluctuation test (Ft-test). Under the null hypothesis, the Ft-test gauges whether the local relative forecasting performance (based on DMW test) of the Fundamentals-based model and the RW is equal at each point in time. The alternative is that the Fundamentals-based model forecasts better than the RW. Hence, when the Ft-test statistic is above its critical value at the 10% level of significance, the Fundamental-based model forecasts significantly better than the RW at that point in time; otherwise, if the Ft-test is below its critical value, the evidence is consistent with absence of forecasting ability of the model. To compute the test, we follow the recommendations in Giacomini and Rossi (2010) and set the size of moving local window to a third of the in-sample observations.[18]

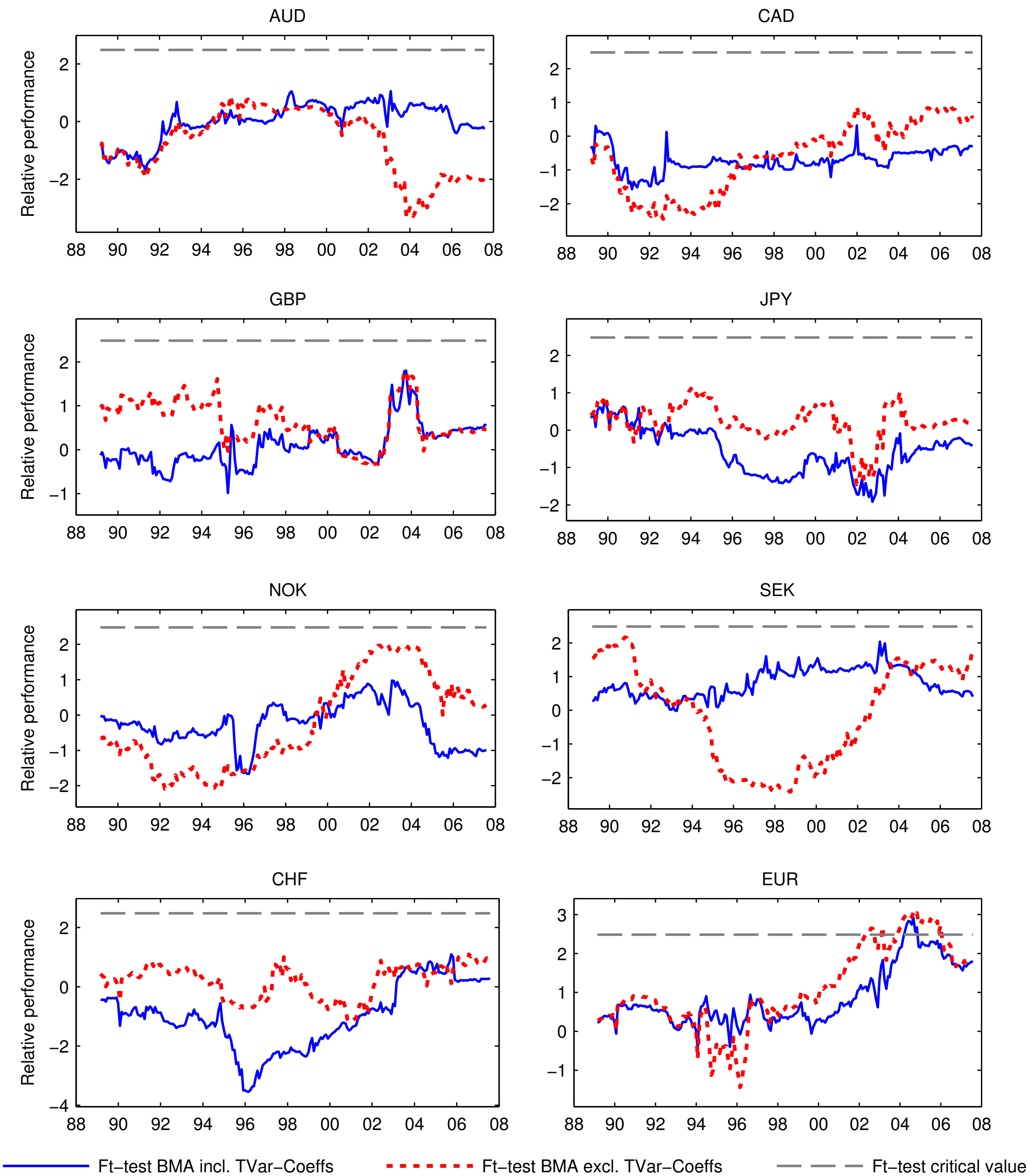
The Ft-test for the 1-month and 12-months forecasting horizons are reported in Figures 3 and 4. Clearly, at the 1-month horizon, our BMA including or excluding Tvar-coeffs never display significant forecasting ability, as the Ft-test is always below its critical value. On the other hand, at 12-months the Ft-test for the BMA including Tvar-coeffs is consistently above the critical value for all currencies; while for the BMA excluding Tvar-coeffs forecasting ability shows up in some periods, but not others. Overall, this is a further piece of evidence regarding the importance of systematically taking into account time-variation in parameters of the exchange rate models.

4.2.3 Analysis of the Degree of Time-Variation in Coefficients

Our analysis of the sources of prediction uncertainty indicates that the uncertainty regarding the degree of time-variation in parameters is not trivial at 1-month horizon, while at longer forecasting horizons it becomes low as more evidence is gathered. But the precise amount of time-variation was not mentioned. Figure 5 provides this information focusing on the same representative set of currencies - CAD, GBP, JPY, and EUR. It depicts the total posterior probability of each of the support points for time-variation in coefficients, δ. At the 1-month horizon in Panel A, at least three support points for time-variation in coefficients are fairly

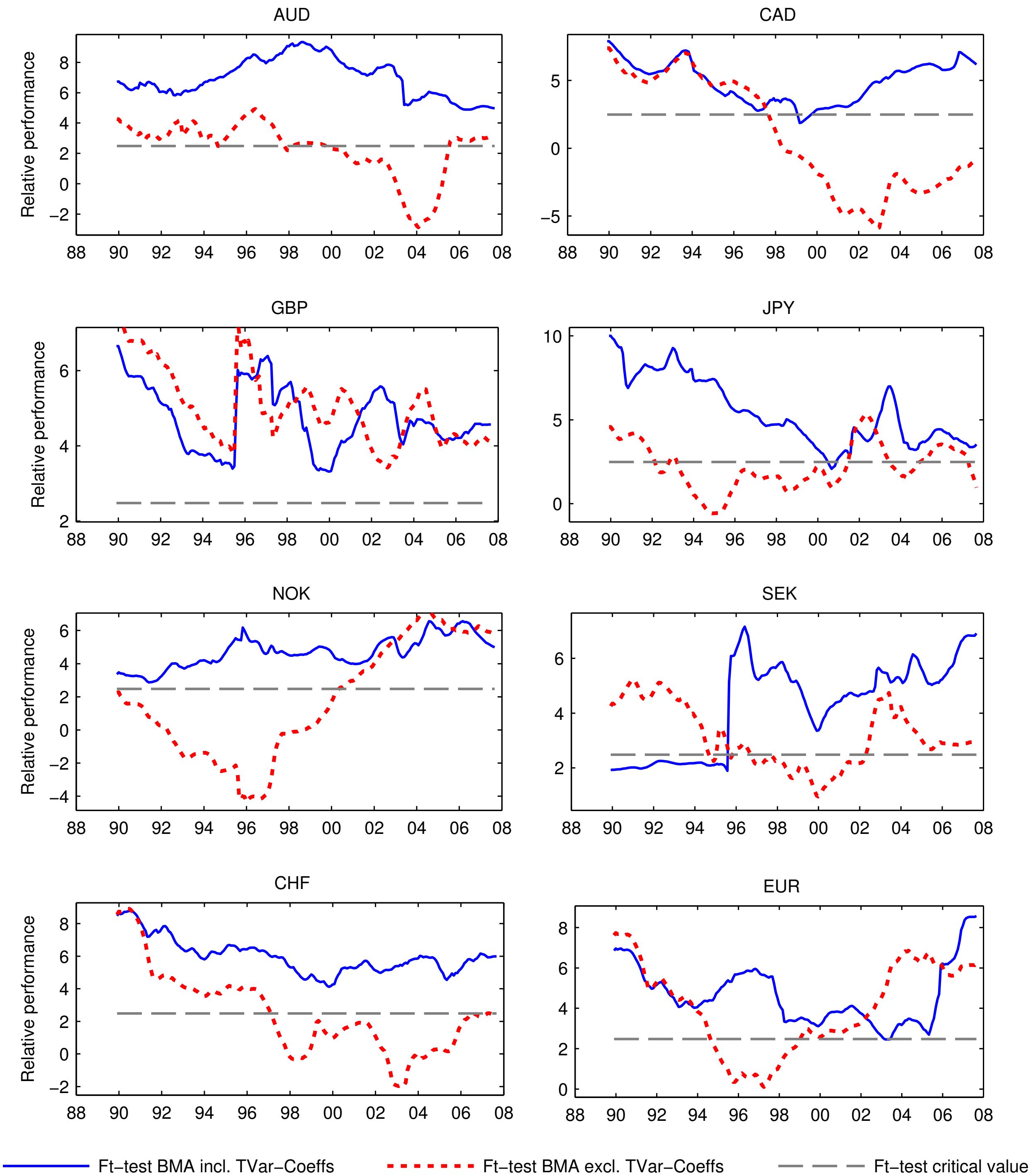
[18] According to Giacomini and Rossi's (2010) Monte Carlo evidence, the Fluctuation test has good properties when implemented using a local moving window size that is a small, such as a third of the in-sample estimation window; see also Rossi (2013). For the critical values of the Fluctuation test, see Table I in Giacomini and Rossi (2010).

Figure 3: Fluctuation Test at 1-Month Forecasting Horizon



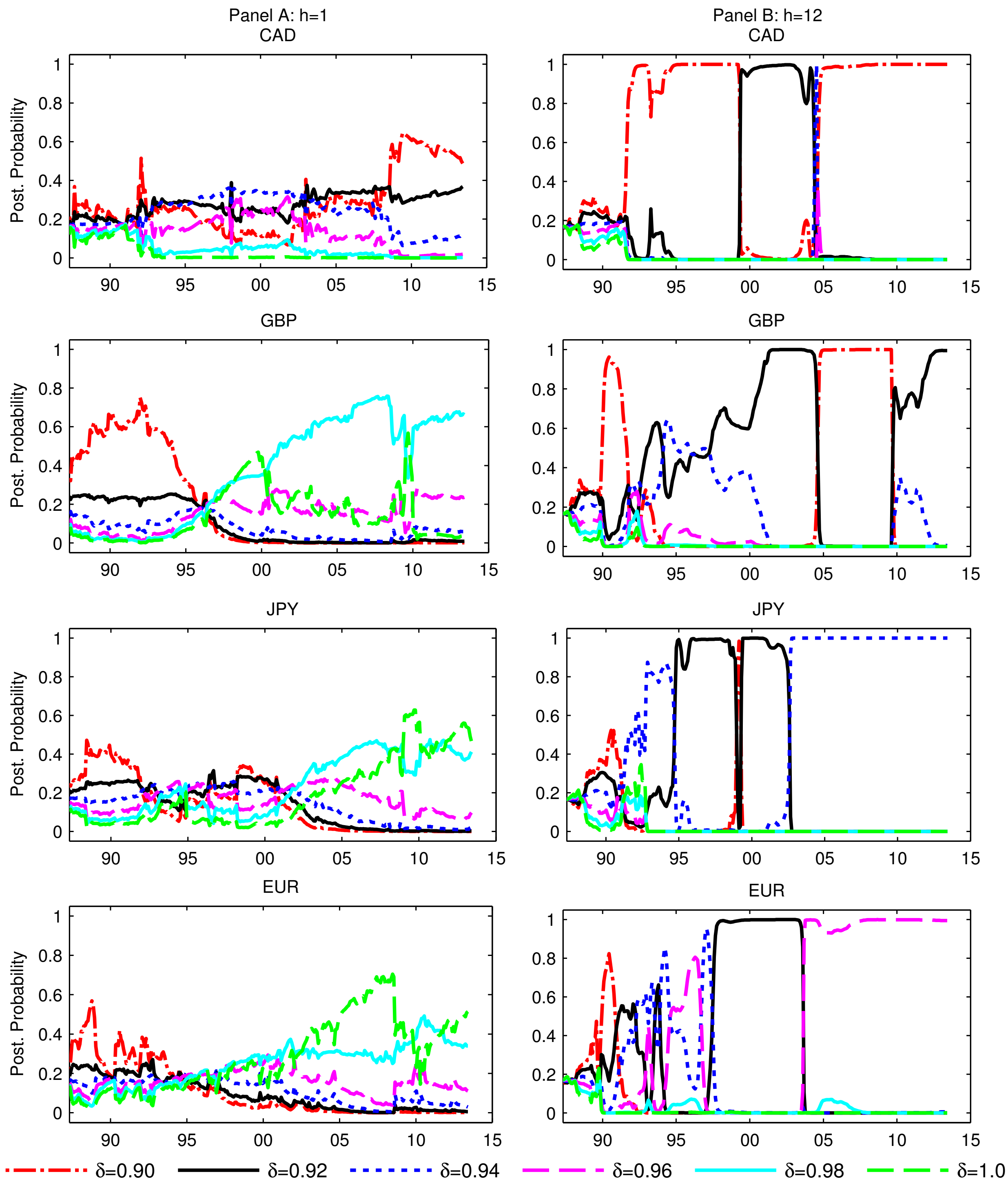
Notes: The Figure shows the Giacomini and Rossi's (2010) one-sided Fluctuation test (Ft-test) based on DMW-test for the BMA including Tvar-coeffs and the BMA excluding Tvar-coeffs. The benchmark model is the driftless Random Walk (RW). It also displays the one-sided Ft-test critical value at 10% level of significance. When the Ft-test statistic is above its critical value, we reject the null of equal local relative forecasting performance between our methods and the RW, and conclude that the method under scrutiny forecasts significantly better than the RW at that point in time. When the Ft-test is below its critical value, the evidence is consistent with absence of forecasting ability of the method under consideration. The forecasting horizon is $h = 1$.

Figure 4: Fluctuation Test at 12-Months Forecasting Horizon



Notes: As in Figure 3, except that here the forecasting horizon is 12-months.

Figure 5: Posterior Probabilities of Degrees of Time-variation in Coefficients (δ)



Notes: Posterior probabilities of values of δ (support points for time-variation in coefficients) for a representative selection of countries. Panel A is for the 1-month forecasting horizon and Panel B for 12-months. These are the weights employed to produce the average forecasts in the BMA including Tvar-coeffs.

likely, as reflected in the magnitude of their weights over the out-of-sample window. In the case of the JPY and the EUR, for example, models with constant and moderate degree of time-variation in coefficients ($\delta = [1.00, 0.98]$) are, on average, as likely as the ones with high degree of time-variation ($\delta = 0.96$), both with probability varying around 20%. For the CAD, the weights are approximately distributed evenly between models with high degree of time-variation in coefficients ($\delta = [0.90, 0.92, 0.94, 0.96]$). And in all cases, there are frequent changes in the most likely support point to embed in the model over the forecasting window. These recurrent changes are reflected in the relatively high uncertainty about the correct degree of time-variation in coefficients at the 1-month forecasting horizon.

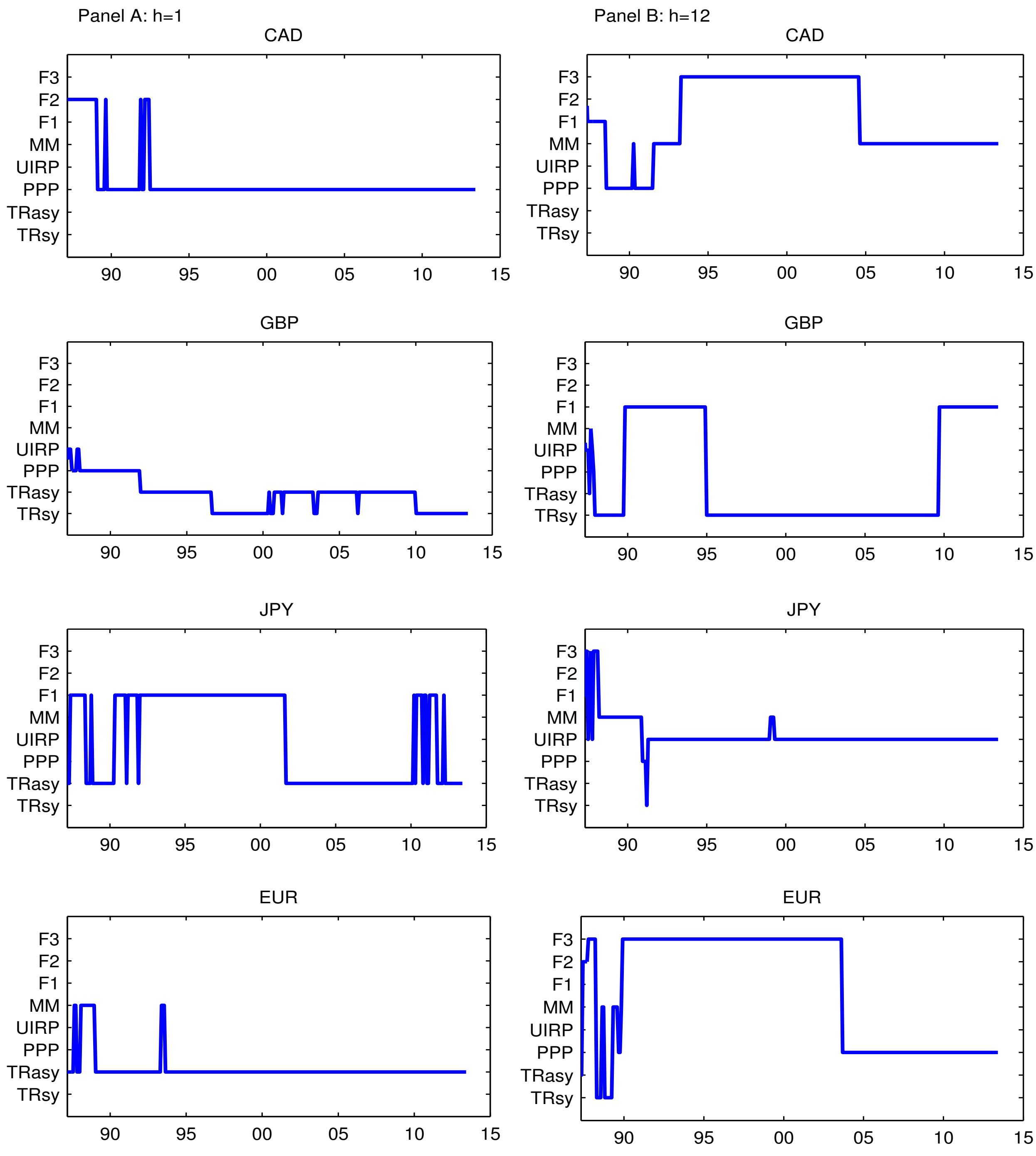
By contrast, at the 12-months horizon in Panel B, up to two very dynamic models - with δ ranging from 0.90 to 0.96 - attract the largest support from the data. And the switches in the most likely support points over the forecasting window are occasional, once the influence of the initial data-points is discounted. As an example, for the JPY, models with $\delta = 0.92$ were empirically plausible up to 2003, while from this period onwards, $\delta = 0.94$ is the most likely degree of time-variation supported by the data. The preponderance of a certain value of δ for a prolonged period and with occasional shifts within the forecasting window, is mirrored in the low uncertainty with respect to the degree of time-variation in coefficients at this horizon.

Interestingly, Giannone (2010) finds that to match the pattern of exchange rate unpredictability present in the data, a significant amount of time-variation in coefficients were necessary in his simulations. Beckmann and Schuessler (2015) show in a Monte Carlo Simulation that a time-varying parameter model like ours, i.e., which allows for gradual to high degree of time-variation in coefficients, is well suited to recover the patterns in the data. As well, in an application to equity returns, Dangl and Halling (2012) find that models with moderate ($\delta = 0.98$) to high ($\delta = 0.96$) degree of time-variation are empirically appropriate.

4.2.4 Analysis of the Importance of Individual Predictors

A final characteristic that we explore in our flexible models is the importance of individual predictors. Figure 6 shows which predictors accumulate the highest probability at each

Figure 6: Predictors with the Largest Posterior Probability at Each Period



Notes: Predictors with the highest probability at each point in time for a representative selection of countries. Panel A is for 1-month forecasting horizon and Panel B for 12-months. The forecasts from the BMS including Tvar-coeffs are based on these predictors.

point in time. For each horizon-currency, there are at most three predictors that are highly informative about movements in the currency's exchange rate. But for the same currency rate, the relevant predictors often change with the forecasting horizon. Focusing on the EUR as an example, while fundamentals from the asymmetric Taylor rule (TRasy) are dominant over the entire forecasting period at 1-month horizon, the third exchange rate factor and PPP fundamentals are more relevant in explaining variations in the USD/EUR rate at the 12-months horizon.

Furthermore, the shifts between the currency-specific fundamentals at each horizon are largely infrequent after ignoring the effect of the initial observations. The case of the EUR currency discussed above also exemplifies this point. After the initial swings among several predictors at the beginning of the forecasting window, the third exchange rate factor attracts the largest support from the data up to 2004M12. From this point, PPP fundamentals become the most relevant determinants of variations in the USD/EUR rate at the 12-months horizon. All in all, the tight number of predictors that are empirically plausible for each currency-horizon, together with the few or clustered swings across these predictors, translates into negligible uncertainty with respect to the choice of the predictor.

5 Robustness Checks and Extensions

We verify the robustness of the empirical findings in the previous section in multiple dimensions. These include: (1) performing our previous analyses in a setting with additional predictors; (2) allowing for an autoregressive process in the coefficient's law of motion; (3) changing the priors; (4) forecasting in a rolling window approach; (5) examining the forecasting performance at the 6- and 24-months forecasting horizons; (6) extending the analyses to other exchange rates; (7) verifying the sensitiveness of our results to the method used in the seasonal adjustment; and (8) changing the base currency. In all cases, we focus in the BMA including Tvar-coeffs and the BMA excluding Tvar-coeffs, and relying exclusively on statistical metrics for forecast evaluation. Where applicable, statistical significance is based on the DMW test with asymptotic critical values. Results for the four initial checks are reported below, and the remaining robustness are tabulated in the supplementary appendix online. In essence, our results hold up to a large extent.

5.1 Inclusion of Additional Set of Predictors

In our baseline analysis the choice of fundamentals is motivated by their preponderance in the exchange rate literature and data availability. Here we repeat the analysis including four additional fundamentals, but focusing on a relatively smaller sample and fewer currencies - CAD, GBP, JPY, and EUR- for which we are able to obtain data.

The first three extra fundamentals are constructed from the Nelson-Siegel (1987) relative factors, as recently put forward by Chen and Tsang (2013):

$$f_{t,LR} = L_t^{NS} + e_t, \qquad f_{t,SR} = S_t^{NS} + e_t, \qquad f_{t,CR} = C_t^{NS} + e_t,$$

where L_t^{NS} is the relative level factor, S_t^{NS} is the relative slope factor, C_t^{NS} is the relative curvature factor, and e_t denotes the nominal exchange rate. Also in line with Chen and Tsang (2013), we obtain the relative factors period by period from OLS estimates of the following equation:

$$i_t^m - i_t^{m*} = L_t^{NS} + S_t^{NS}\left(\frac{1-e^{-\lambda m}}{\lambda m}\right) + C_t^{NS}\left(\frac{1-e^{-\lambda m}}{\lambda m} - e^{-\lambda m}\right) + \varepsilon_t^m \tag{18}$$

with i_t^m denoting the continuously compounded zero-coupon nominal domestic yield on an m-month bond, i_t^{m*} is its foreign counterpart, λ is a parameter set to 0.0609 as typical in the literature, and ε_t^m is an estimation error.

Our use of each relative factor as a regressor, rather than all the factors jointly, is guided by Chen and Tsang (2013) and Berge (2013), who uncover evidence of distinct predictive ability of each factor. The relative factors are constructed from zero-coupon bonds yields for various maturities, typically of 3, 6, 12, 24, 36, 60, 72, 84, 96, 108, and 120 months. The data are obtained from national central banks, except for Japan where the source is the ministry of finance.[19]

The last additional fundamental is derived from the Gourinchas and Rey's (2007) external balance model:

$$f_{t,nxa} = nxa_t^{(n)} + e_t \tag{19}$$

where $nxa_t^{(n)}$ is the bilateral measure of cyclical external imbalances between the home and foreign country n. We first construct $nxa_t^{(n)}$ using quarterly data from Della Corte *et al.* (2012) and also employing their instrumental variable approach.[20] We then derive monthly observations via a quadratic-match-average interpolation method. Adding these regressors to the eight previous ones, our model space now contains 12 regressors. Because data on bond yields are available from 1990M1, and the data in Della Corte *et al.* (2012) extend up to 2007Q3, our effective sample runs from 1990M1 to 2007M9. We use the first 120 observations for in-sample estimation, leaving the remaining for OOS analysis.

Forecast results for the four currencies are shown in Panel I of Table 3, and variance decomposition results based on the representative case of the CAD are displayed in Panel I of Figure 7. Our previous conclusions remain unaltered. Models with time-varying coefficients deliver larger forecast gains than the RW and the constant-coefficients models at $h = 3$ and $h = 12$. As well, the uncertainty in the coefficient's estimation and the uncertainty about the precise degree of coefficients variability to allow in the model at $h = 1$ remain the key obstructing factors to forecast improvements.

[19]The level factor L_t^{NS} has a constant loading in the yield curve and picks up aspects that shift the relative yield curve (e.g., inflation expectations). The slope factor S_t^{NS} captures the short end of the yield curve and is connected with dynamics in monetary policy. The curvature factor C_t^{NS} has its largest effect in the middle of the yield curve, and a zero loading at maturity $m = 0$ as well as at the extreme maturities (Berge, 2013)

[20]The approach in Della Corte *et al.* (2012) is as follows. First, the global (nxa_t), rather than the bilateral, measure of cyclical external imbalances is regressed on a collection of instruments. These instruments include the nxa_t of the foreign country and the bilateral detrended net exports between the home and foreign countries, constructed as a linear combination of the stationary components of log bilateral exports and imports to wealth ratios. The fitted value from this regression is then the proxy for $nxa_t^{(n)}$. See Della Corte *et al.* (2012) for extra details.

Table 3: Statistical Evaluation of Forecasting Performance - Various Robustness Checks

Panel I: Inclusion of Additional Set of Predictors						
	h=1	h=3	h=12	h=1	h=3	h=12
	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
CAD	1.021	**0.954**	**0.610*** **	1.013	**0.974**	**0.772*** **
GBP	1.032	**0.923***	**0.659*** **	1.033	1.004	**0.841** **
JPY	1.066	**0.924***	**0.884**	1.063	**0.974**	1.009
EUR	1.040	**0.966**	**0.715*** **	1.044	1.000	**0.789*** **
Panel II: Change in Coefficient's Law of Motion to Autoregressive Coefficients						
	BMA incl. TVar-Coeffs			BMA incl. TVar-Coeffs		
	G=0.95			G=0.90		
AUD	1.009	**0.899**	1.171	1.005	**0.875***	**0.961**
CAD	1.009	**0.924*** **	**0.855** **	1.009	**0.917*** **	**0.830*** **
GBP	1.012	**0.924*** **	**0.797** **	1.012	**0.916** **	**0.806*** **
JPY	1.011	**0.954** **	**0.728*** **	1.008	**0.948*** **	**0.739*** **
NOK	1.007	**0.943***	**0.920** **	1.006	**0.933** **	**0.861*** **
SEK	1.011	**0.986**	**0.840*** **	1.007	**0.953**	**0.809*** **
CHF	1.016	**0.918*** **	**0.920** **	1.013	**0.925*** **	**0.879*** **
EUR	1.009	**0.922*** **	**0.877*** **	1.006	**0.920*** **	**0.853*** **
	G=0.50			G=0.20		
AUD	1.001	**0.920*** **	**0.921*** **	1.001	**0.971**	**0.976**
CAD	1.004	**0.944*** **	**0.884*** **	1.002	**0.980**	**0.963**
GBP	1.002	**0.938*** **	**0.905*** **	1.000	**0.984**	**0.979**
JPY	**0.999**	**0.958*** **	**0.877*** **	**0.999**	**0.989**	**0.961***
NOK	**0.999**	**0.938*** **	**0.897*** **	**0.999**	**0.977**	**0.977**
SEK	**0.994**	**0.925*** **	**0.899*** **	1.000	**0.979**	**0.972**
CHF	1.001	**0.933*** **	**0.917*** **	**0.999**	**0.978***	**0.981***
EUR	1.000	**0.937*** **	**0.902*** **	**0.999**	**0.981***	**0.982** **
Panel III: Change in Priors						
	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
AUD	1.007	**0.968**	**0.819** **	0.999	**0.994**	**0.952*** **
CAD	1.014	**0.953**	**0.942** **	1.002	**0.998**	**0.992**
GBP	1.007	**0.944***	**0.837** **	1.002	**0.998**	**0.956*** **
JPY	1.010	**0.982**	**0.966**	1.005	1.002	**0.948*** **
NOK	1.007	**0.998**	**0.845*** **	1.000	1.003	**0.981***
SEK	1.016	**0.988**	**0.912**	1.007	**0.997**	**0.998**
CHF	1.008	**0.949***	**0.955**	1.002	**0.996**	1.002
EUR	1.009	**0.958***	**0.825*** **	1.005	**0.997**	**0.966*** **
Panel IV: Forecasting in a Rolling Window Approach						
	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
AUD	1.017	**0.978**	**0.980**	1.008	1.006	**0.931** **

Table 3 Continued

	h=1	h=3	h=12	h=1	h=3	h=12
CAD	1.024	**0.943***	**0.975**	1.009	**0.995**	1.024
GBP	1.023	**0.964***	**0.890***	1.020	**0.966***	**0.873****
JPY	1.022	**0.986**	**0.944*****	1.017	1.009	**0.951**
NOK	1.015	1.003	**0.857*****	1.010	**0.998**	**0.972**
SEK	1.019	**0.971**	**0.863*****	1.018	1.002	1.001
CHF	1.027	**0.947***	1.066	1.020	1.008	1.053
EUR	1.019	**0.945****	**0.874*****	1.018	**0.983**	**0.892*****

Notes: The table presents forecasting performance results for various robustness checks. The entries constitute the RMSFE of the specified fundamentals-based method relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the method under scrutiny generates a lower RMSFE than RW. The Table also reports one-sided DMW test-statistic for the null of equal forecasting performance, based on asymptotic critical values. Asterisks (*10%, **5%, ***1%) indicate the level of significance at which the null hypothesis is rejected, favouring the alternative that the fundamental-based method yields a lower RMSFE than the RW. Currency codes denote: AUD-Australian dollar; CAD - Canadian dollar; GBP - Pound sterling; JPY - Japanese yen; NOK - Norwegian krone; SEK - Swedish krona; CHF - Swiss franc; EUR - euro. The forecast evaluation period begins in 1987M12+h in all, but the OLS-DMSPE Combination case (1992M12+h).

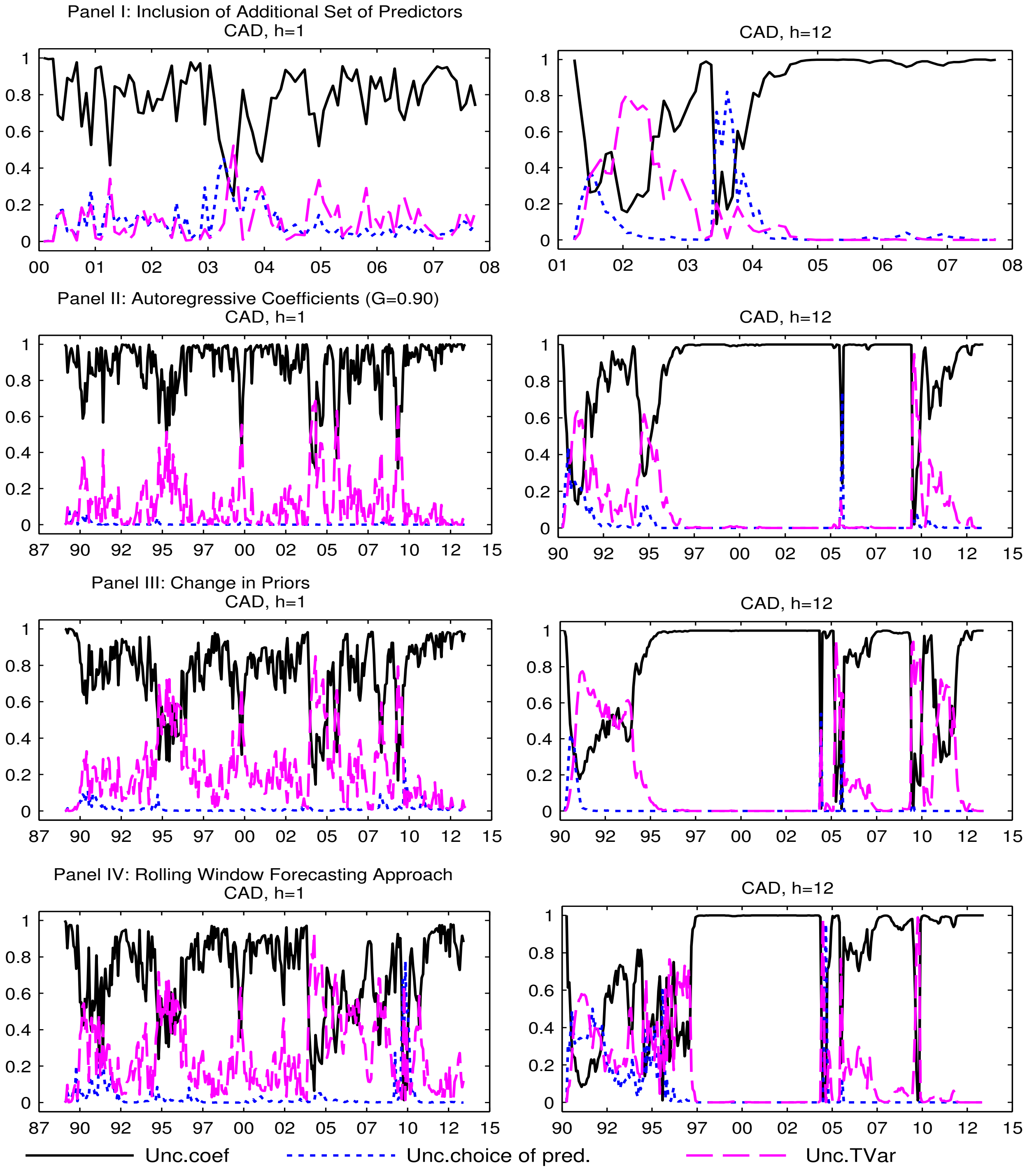
5.2 Autoregressive Coefficients

The coefficients' law of motion in our state-space system follows a random-walk process. To verify whether the results are sensitive to this law of motion we experiment with an autoregressive process. Following Dangl and Halling (2012), we introduce autoregression in a simple manner by re-writing our transition equation as:

$$\theta_t = G\mathbf{I}\theta_{t-1} + \varpi_t, \qquad 0 < G < 1; \tag{20}$$

with G denoting a scalar, and $\mathbf{I}$ an identity matrix. We fix values of G to several alternatives, including $G = 0.95, 0.90, 0.50, 0.20$. As shown in Panel II of Table 3 and Figure 7, our results remain coherent with our previous findings. With the first three values of G, the BMA including Tvar-coeffs still yield more precise forecasts than the RW. As the value of G approximate to zero, our forecasts tend to become indistinguishable from those of the RW as the parameters estimates are shrunk towards zero. The variance decomposition based on a model with $G = 0.90$ as an example, displays the pattern documented in our main findings.

Figure 7: Sources of Prediction Variance Excl. Observational Variance - Various Robustness Checks



Notes: Decomposition of the prediction variance excluding observational variance for a representative selection of currency - the CAD. (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

5.3 Change in Priors

Our benchmark priors are sensible and relatively diffuse, as they are set to allow the data to dominate the prior information. Here we focus on a setting that assigns equal weights to the prior and the data in the posterior covariance matrix by fixing $g = 1$ and $T_0 = 1$ in equation (4), see Koop (2003, Ch. 11). Panel III of Table 3 shows the point forecast results from models estimated with this prior elicitation setting, and Panel III of Figure 7 the corresponding prediction variance decomposition for the CAD. The baseline line conclusions on the forecasting performance of the BMA including Tvar-coeffs and on the sources of uncertainty hold up to a great extent.

5.4 Rolling Window Forecasting Approach

To examine whether the forecasting scheme drives the findings we obtain, we experiment with a rolling window estimation approach. In this case, our forecasting models are now conditioned on sets of information (D_t) constructed in a 10-year moving window, from which we generate the same number of forecasts as in the recursive forecasting method. We see in Panel IV of Table 3 that models with time-varying coefficients are still preferable at $h = 3$ and $h = 12$ relative to the RW and to models with constant coefficients. As before, estimation uncertainty and uncertainty about the degree of time-variation in coefficients, remain the main factors hindering forecasting ability.

6 Conclusion

The exchange rate literature indicates that the out-of-sample predictive power of the empirical exchange rate models is erratic. Models that forecast well for certain currencies and periods, often fail when applied to other exchange rates and samples (Rogoff and Stavrakeva, 2008; Rossi, 2013). While this signals the presence of instabilities, attempts to account for them, for example by considering regressions with time-varying coefficients, have not yet produced overwhelming results (Rossi, 2013). In this paper we employ a systematic approach to properly account for time-variation in the coefficients of exchange rate forecasting regressions. The approach also incorporates the idea that the relevant set of regressors may change at each point in time; as articulated, for example, by Bacchetta and van Wincoop (2004, 2013), Berge (2013), and Sarno and Valente (2009). Inspired by recent advances in Bayesian methods, we further employ our framework to investigate all sources of uncertainty in the predictive models, through a variance decomposition procedure.

What we find is that exchange rate models generate more accurate forecasts than the driftless random walk benchmark for most currencies at all the forecasting horizons we consider, except at the 1-month horizon. The key to improving upon the benchmark is forecasting with predictive regressions that capture both, the possibly changing set of explanatory variables, and the appropriate time-varying weights associated with these variables. At horizons beyond 1-month, i.e., $h = 3$ and $h = 12$, our regressions successfully embed these characteristics. Models which allow for switching sets of regressors and sudden, rather than smooth, changes in the time-varying weights of these regressors are empirically plausible. By contrast, at 1-month forecasting horizon our predictive regressions fail to successfully capture the suitable time-varying weights associated with the regressors, yielding poor performance.

We then proceed and track the sources of uncertainty in the regressions, in an effort to pin down the origins of the weak performance. Here we find that beyond the typical uncertainty associated with unpredictable fluctuations in the data, the uncertainty in the estimation of the models' coefficients and the uncertainty about the level of time-variation in coefficients to incorporate in the model, are the main factors hindering the models' predictive ability. When the uncertainty emanating from these two sources is low or is successfully embedded in the model, the out-of-sample forecasting performance of the models is satisfactory. In further characterization of our models, we find that the set of variables that are more informative about exchange rate movements generally differ between forecasting horizons and between countries. But within a specific country-horizon often few variables matter. As a consequence, the uncertainty regarding the choice of the predictors appears negligible.

Overall, we view our results as providing direct evidence towards the prevalent conjectures or simulation based findings that time-variation in parameters of the models might cause time-variation in the models' forecasting performance (Giannone, 2010; Meese and Rogoff, 1983; Rossi, 2013; Rossi and Sekhposyan, 2011).

References

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

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

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

Beckmann, J. and Schussler, R. (2015). Forecasting exchange rates under parameter and model uncertainty. *Journal of International Money and Finance*, 60:267–288.

Berge, T. J. (2013). Forecasting disconnected exchange rates. *Journal of Applied Econometrics*, 29(5):713–735.

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

Chen, Y.-c. and Tsang, K. P. (2013). What does the yield curve tell us about exchange rate predictability? *Review of Economics and Statistics*, 95(1):185–205.

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.

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

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

Della Corte, P., Sarno, L., and Sestieri, G. (2012). The predictive information content of external imbalances for exchange rate returns: How much is it worth? *Review of Economics and Statistics*, 94(1):100–115.

Della Corte, P. and Tsiakas, I. (2012). *Statistical and Economic Methods for Evaluating Exchange Rate Predictability*, chapter 8, pages 221–264. John Wiley & Sons, Inc.

Diebold, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. *Journal of Business & Economic Statistics*, 33(1):1–9.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

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

Engel, C., Mark, N. C., and West, K. D. (2015). Factor model forecasts of exchange rates. *Econometric Reviews*, 34(1-2):32–55.

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

Fernandez, C., Ley, E., and Steel, M. F. (2001). Benchmark priors for Bayesian model averaging. *Journal of Econometrics*, 100(2):381–427.

Fratzscher, M., Rime, D., Sarno, L., and Zinna, G. (2015). The scapegoat theory of exchange rates: The first tests. *Journal of Monetary Economics*, 70:1–21.

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

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

Giannone, D. (2010). *Comment on "Can Parameter Instability Explain the Meese-Rogoff Puzzle?"*, pages 180–190. NBER International Seminar on Macroeconomics 2009. University of Chicago Press.

Gourinchas, P.-O. and Rey, H. (2007). International financial adjustment. *Journal of Political Economy*, 115(4):665–703.

Greenaway-McGrevy, R., Mark, N., Sul, D., and Wu, J.-L. (2015). Identifying exchange rate common factors. Working Papers 212012, Hong Kong Institute for Monetary Research.

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

Inoue, A. and Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23(4):371–402.

Koop, G. (2003). *Bayesian Econometrics*. John Wiley & Sons, Ltd., London.

Koop, G. and Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, 53(3):867–886.

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

Longerstaey, J. and Spencer, M. (1996). *RiskMetrics Technical Document*. J.P. Morgan, fourth edition.

Mark, N. C. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *The American Economic Review*, 85(1):201–218.

Markiewicz, A. (2012). Model uncertainty and exchange rate volatility. *International Economic Review*, 53(3):815–844.

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

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.

Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, 60(4):473–89.

Newey, W. K. and West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, 23(2):821–862.

Rapach, D. E. and Wohar, M. E. (2006). In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance*, 13(2):231 – 247.

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

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

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

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

Stock, J. H. and Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, 23(6):405–430.

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

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica*, 64(5):1067–1084.

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

Wolff, C. C. P. (1987). Time-varying parameters and the out-of-sample forecasting performance of structural exchange rate models. *Journal of Business & Economic Statistics*, 5(1):87–97.

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

On the Sources of Uncertainty in Exchange Rate Predictability

Supplementary Appendix

Joseph P. Byrne*, Dimitris Korobilis†,‡, Pinho J. Ribeiro‡

September 13, 2016

This document contains supplementary material to the above titled paper and is organized in nine parts. In part A we examine the ability of our models to generate economic value in a stylized asset portfolio management setting. We describe the criteria for such evaluation and the corresponding results. Part B presents figures on variance decomposition and other results associated with the BMA including time-varying coefficients at the 3-months forecasting horizon. Part C contains figures on the sources of prediction uncertainty for four currencies whose figures were omitted from the main text to conserve space (AUD, NOK, SEK, and CHF). Part D shows the recursive relative Root Mean Squared Forecast Error (RMSFE) for the Bayesian model averaging excluding and including time-varying coefficients. In part E we use predictive likelihoods to measure relative forecasting performance of the Bayesian model selection (BMS) excluding time-varying coefficients relative to BMS including time-varying coefficients. Part F reports results for additional robustness checks. We present the details regarding the estimation of the dynamic linear model that we consider in part G. Part H describes the data sources and the last part elaborates on the procedure to construct bootstrapped critical values for the DMW test.

*Heriot-Watt University, School of Management and Languages, Edinburgh, EH14 4AS, UK.
†University of Glasgow, Adam Smith Business School, Glasgow, G12 8QQ, UK.
‡Corresponding author, E-mail: dimitris.korobilis@glasgow.ac.uk; Tel: +44 (0)141 330 2950.

A Economic Evaluation of Forecasting Performance

A.1 Criteria for Economic Evaluation

A limitation of the statistical measures of forecasting performance is their inability to convey the economic gains associated with better forecasting performance. To address this limitation we follow Della Corte *et al.* (2012) and Li *et al.* (2015) and use economic evaluation criteria. Inspired by these studies, we consider a stylized dynamic asset allocation strategy in which a US investor rebalances her portfolio by allocating her assets between a US bond and eight foreign bonds ($B8$). She can rebalance her portfolio either monthly, in three-months period or in twelve-months period, which correspond to the horizons for which she generates exchange rates forecasts ($h = 1, 3, 12$). At every period, the $B8$ bonds carry a risk-free return in local currency but a risky return r_{t+h} in US dollars terms. The yields on the bonds consist of the interest rates at the respective adjustment periods. When pursuing an investment strategy in the $B8$ bonds, she expects a return of $r_{t+h|t} = i_t^* + \Delta e_{t+h|t}$, where $r_{t+h|t} = E_t[r_{t+h}]$ is the conditional expectation of r_{t+h}; i_t^* denotes the nominal interest rate in the corresponding countries; and $\Delta e_{t+h|t} = E_t[\Delta e_{t+h}]$ is the conditional expectation of Δe_{t+h}. Given that the interest rate is known at time t, the return that the investor projects from time t to $t+h$ is only exposed to the exchange rate risk.

In her dynamic asset allocation process the investor wishes to minimize the foreign exchange risk exposure by finding the optimal portfolio weights. At each time period and rebalancing horizon, she uses our models' forecasts to rebalance her portfolio by calculating new optimal weights on each bond, taking into account the portfolio's mean return and variance. Essentially, her problem is to find the optimal portfolio weights subject to a target volatility of the portfolio returns. Della Corte *et al.* (2012) show that the solution to this problem implies the following weights to the risky bonds:

$$w_t^g = \frac{\sigma_p^*}{\sqrt{C_t^g}} \Sigma_{t+h|t}^{-1} (u_{t+h|t}^g - \iota r_f); \tag{A.1}$$

where $C_t^g = (u_{t+h|t}^g - \iota r_f)' \Sigma_{t+h|t}^{-1} (u_{t+h|t}^g - \iota r_f)$; σ_p^* is the target volatility; $\Sigma_{t+h|t} = [(r_{t+h} - u_{t+h|t}^g)(r_{t+h} - u_{t+h|t}^g)']$ is the conditional covariance matrix; r_{t+h} is the $(B8 \times 1)$ vector of risky asset returns; and $u_{t+h|t}^g = E_t[r_{t+h}]$ defines the conditional expectation of r_{t+h}. The weight on the risk-free asset is $(1 - w_t'\iota)$.

The gross return on the portfolio is given by:

$$R_{p,t+h} = 1 + w_t^{g\prime} r_{t+h} + (1 - w_t^{g\prime}\iota) r_f = R_f + w_t^{g\prime}(R_t^g - \iota R_f); \tag{A.2}$$

where R_t^g is the $(B8 \times 1)$ vector of gross returns on risky bonds and R_f are the gross returns on the risk-free bond. In line with Della Corte *et al.* (2012) and Li *et al.* (2015), the investor replaces the unconditional covariance matrix with the conditional one when solving her portfolio allocation problem: $\Sigma_{t+h|t} = \overline{\Sigma}$. This ensures that the optimal weights will vary only to the extent of the differences in our models' forecasting ability.

We can employ this framework to examine our models' ability to produce tangible economic benefits, compared to a strategy that uses forecasts from the driftless random walk (RW). We compute the following indicators of economic value:

- Sharpe ratio (SR): Defined as the ratio of the average realized excess portfolio returns relative to the portfolio returns standard deviation. We compute Sharpe ratios associated with our models' forecasts and the forecasts from the RW. Higher Sharpe ratios are preferred to lower ones. To assess if differences in Sharpe ratios are statistically significant, we apply the bootstrap method propounded by Ledoit and Wolf (2008).

- Maximum performance fee (pf): Proposed by Fleming *et al.* (2001), this is the fee that a risk-averse investor with quadratic utility is willing to pay to use our models rather than the RW. The measure is founded on the principle that forecasts from a specific model are superior to those from an alternative one, if investment decisions that rely on the specific model yield greater average realized utility than the alternative. The starting point is the assumption that a portfolio based on the RW for example, generates the same average utility as compared to a portfolio based on an alternative model that is subject to expenses pf, at h-month(s) horizon. Because the investor would be neutral between the two strategies, pf is interpreted as the maximum performance fee that she is ready to pay to swap from the RW to the alternative model. It is often expressed as a fraction of the wealth invested and obtained by solving:

$$\sum_{t=0}^{T-1}\left\{(R^*_{p,t+h}-pf)-\frac{RRA}{2(1+RRA)}(R^*_{p,t+h}-pf)^2\right\}=\sum_{t=0}^{T-1}\left\{R_{p,t+h}-\frac{RRA}{2(1+RRA)}R^2_{p,t+h}\right\}; \tag{A.3}$$

  where $R^*_{p,t+h}$ is the gross return from using our models, $R_{p,t+h}$ is the gross return from the benchmark RW, and RRA is the investor's constant degree of relative risk aversion. Larger values of pf suggest that the investor wishes to pay more to swap from the RW to our models (we report pf in annualized basis points - bps).

- Excess premium (pr): Based on the manipulation-proof performance measure of Goetzmann *et al.* (2007), this indicator captures the portfolio's premium return after adjusting for risk. To compute the premium, first obtain the manipulation-proof performance measure:

$$M(R_p) = \frac{1}{1 - RRA} \ln \left\{ \frac{1}{T} \sum_{t=0}^{T-1} \left(\frac{R_{p,t+h}}{R_f} \right)^{1-RRA} \right\}; \tag{A.4}$$

where $M(R_p)$ is the risk adjusted portfolio's premium return from the RW, while a similar measure for our models is $M(R_p^*)$. The excess premium return from our models relative to the RW is therefore:

$$pr = M(R_p^*) - M(R_p). \tag{A.5}$$

Higher values of pr are indicative of greater premium returns of our models relative to the RW after accounting for risk. We equally present pr in annualized *bps*.

- Break-even transaction costs: These are the proportional transaction costs that eliminate any positive performance fee obtained by conditioning on our models. When the investor reaches this point, she becomes indifferent between using the RW and our models. To compute the cost, we follow Han (2006) and Della Corte *et al.* (2012), and assume that transaction costs constitute a fixed fraction (T_r) of the value traded in each bond. Therefore, the costs are: $T_r \left| w_t - w_{t-h}(R_t^g / R_p) \right|$. In cases where the investor's transaction costs are below the break-even transaction cost level, T^{be}, she will continue to prefer using our models; alternatively she would keep the RW. The value of T^{be} is reported in monthly *bps* for $h = 1$, quarterly *bps* for $h = 3$, and annual *bps* for $h = 12$.

A.2 Economic Evaluation Results

As described above, the economic evaluation of our models builds upon the maximum expected return strategy, which is often used in active currency management (Della Corte *et al.*, 2012 and Li *et al.*, 2015). We recall that we focus on four measures, namely, the Sharpe ratio (SR), the performance fee (pf), the excess premium return (pr), and the break-even transaction cost (T^{be}). In line with results in Li *et al.* (2015) our investor targets an annualized volatility of $\sigma_p^* = 10\%$ and her degree of relative risk aversion is $RRA = 6$.[1]

Table A.1 presents results from portfolio allocation schemes based on forecasts from each of the methods we examine and the RW. In general, results are consistent with the findings from the statistical evaluation. At the 3- and 12-months horizons, strategies conditioning on forecasts from Bayesian model averaging or selection including Tvar-coeffs yield economic gains above those accruing from the RW, regardless of the specific economic indicator. The

[1]We also experiment with different values of σ_p^* and RRA, and find that the conclusions are insensitive to changes in these parameters.

Table A.1: Economic Evaluation of Bayesian and Forecast Combinations Methods

Exchange rate forecast based on:	$\overline{Rt}(\%)$	Vol (%)	SR	$pf(bps)$	$pr(bps)$	$T^{be}(bps)$
	monthly rebalancing period (i.e., h=1)					
Random Walk	14.61	13.02	0.82			
BMA incl. TVar-Coeffs	6.61	11.57	0.23	-686	-585	-
BMS incl. TVar-Coeffs	10.14	12.11	0.51	-374	-284	-
BMA excl. TVar-Coeffs	7.67	10.15	0.37	-481	-387	-
BMS excl. TVar-Coeffs	10.44	10.05	0.64	-198	-105	-
OLS-Mean Combination	13.93	12.09	0.83	7	54	2
OLS-Median Combination	14.10	11.79	0.86	46	94	7
OLS-Trimmed Mean	13.94	11.94	0.84	19	60	5
OLS-DMSPE Combination	11.42	13.33	0.62	-55	-105	-
	3-months rebalacing period (i.e., h=3)					
Random Walk	13.67	14.88	0.65			
BMA incl. TVar-Coeffs	20.36	12.29	1.34	928	1411	83
BMS incl. TVar-Coeffs	22.26	12.64	1.45*	1086	1576	95
BMA excl. TVar-Coeffs	7.32	11.63	0.29	-321	108	-
BMS excl. TVar-Coeffs	9.22	11.02	0.48	-82	369	-
OLS-Mean Combination	13.35	13.91	0.68	71	309	34
OLS-Median Combination	12.08	13.50	0.60	-15	285	-
OLS-Trimmed Mean	12.64	13.82	0.63	10	243	5
OLS-DMSPE Combination	12.60	15.37	0.61	-2	105	-
	12-months rebalancing period (i.e., h=12)					
Random Walk	10.57	16.94	0.40			
BMA incl. TVar-Coeffs	22.05	18.24	1.00**	538	1896	274
BMS incl. TVar-Coeffs	22.57	18.34	1.02**	591	577	270
BMA excl. TVar-Coeffs	12.11	21.06	0.39	-455	-2464	-
BMS excl. TVar-Coeffs	12.59	20.74	0.42	-407	-1908	-
OLS-Mean Combination	9.41	17.03	0.33	-143	-565	-
OLS-Median Combination	7.29	16.09	0.21	-131	-1051	-
OLS-Trimmed Mean	8.29	16.03	0.28	-19	-505	-
OLS-DMSPE Combination	13.73	17.67	0.60**	386	470	306

Notes: Economic gains generated by Bayesian Model Averaging (BMA) and Bayesian Model Selection (BMS) including or excluding time-varying coefficients and simple forecast combinations methods. Using forecasts from these methods an investor constructs a strategy of maximum expected return, conditional on portfolio volatility target of Vol = 10%. Every h-month (s) period, she dynamically adjusts her portfolio investing in US bonds and seven foreign bonds. The Table shows the gains obtained by the investor by conditioning on the forecasts from each method and the driftless Randow Walk (RW) at every rebalancing period. The gains are gauged based on: (i) the annualized mean return- $\overline{Rt}$ (ii) Vol - the annualized volatility, (iii) SR - annualized Sharpe ratio, (iv) pf, the maximum performance fee a risk-averse investor with quadratic utility would be willing to pay to use the corresponding method instead of the RW - in annualized bps (v) pr - the excess premium return (in annualized bps) and (vi) T^{be} - the break-even proportional transation costs that offset any positive performance fee obtained by using the method under consideration - in h-month(s) period bps. Asterisk denotes statistically significant differences in the SR in favor of the method in question relative to the driftless RW at the *10%, **5%, or ***1% level of significance based on the bootstrap procedure of Ledoit and Wolf (2008). The evaluation period is 1987M12+h to 2013M5.

Sharpe ratio implied by BMS including Tvar-coeffs is 1.45 at $h = 3$, significantly higher than that of the RW at 0.65. As the rebalancing horizon extends to 12-months period, the gains remain significant, at a Sharpe ratio of 1.02 against 0.40 of the RW benchmark. Results for BMA including Tvar-coeffs are qualitatively similar. In contrast, nearly all the constant-coefficients forecasting models are dominated by strategies based on the RW at these horizons. The exception is the DMSPE combination method which significantly improves upon the RW at $h = 12$, but the improvement is still inferior to the gains accruing from the Bayesian methods with Tvar-coeffs. Note too that like in the statistical evaluation criteria, at the 1-month rebalancing period, the RW generally delivers the best outcomes. At this horizon, although simple forecast combination methods based on the mean, median, and trimmed mean produce a higher Sharpe ratio relative to the RW, the improvement is statistically insignificant.

Other indicators of economic value also convey the beneficial effects of conditioning on the more flexible models at longer rebalancing periods. At $h = 12$, for example, the performance fee is $pf = 591bps$ when using BMS including Tvar-coeffs, which implies that a risk-averse investor is ready to pay a fee of 5.91% annually to use this forecasting approach rather than the RW. He also obtains an excess premium return of 5.77% annually, whilst the break-even transaction costs are $T^{be} = 270bps$. If the investor's proportional transaction costs are higher than this magnitude of T^{be}, he will continue using the RW. However, as Della Corte *et al.* (2012) point out, this is unlikely as transaction costs in the currency markets are low.

B Variance Decomposition and Characterization of the BMA including Tvar-Coeffs at 3-Months Horizon

This section presents figures on variance decomposition and other characteristics of our BMA including time-varying coefficients at the 3-months horizons. These are:

- Figure B.1: Sources of Prediction Variance at 3-Months Forecasting Horizon;
- Figure B.2: Posterior Probabilities of Degrees of Time-variation in Coefficients, $h = 3$;
- Figure B.3: Predictors with the Largest Posterior Probability at Each Period, $h = 3$;
- Figure B.4: Fluctuation Test at 3-Months Forecasting Horizon.

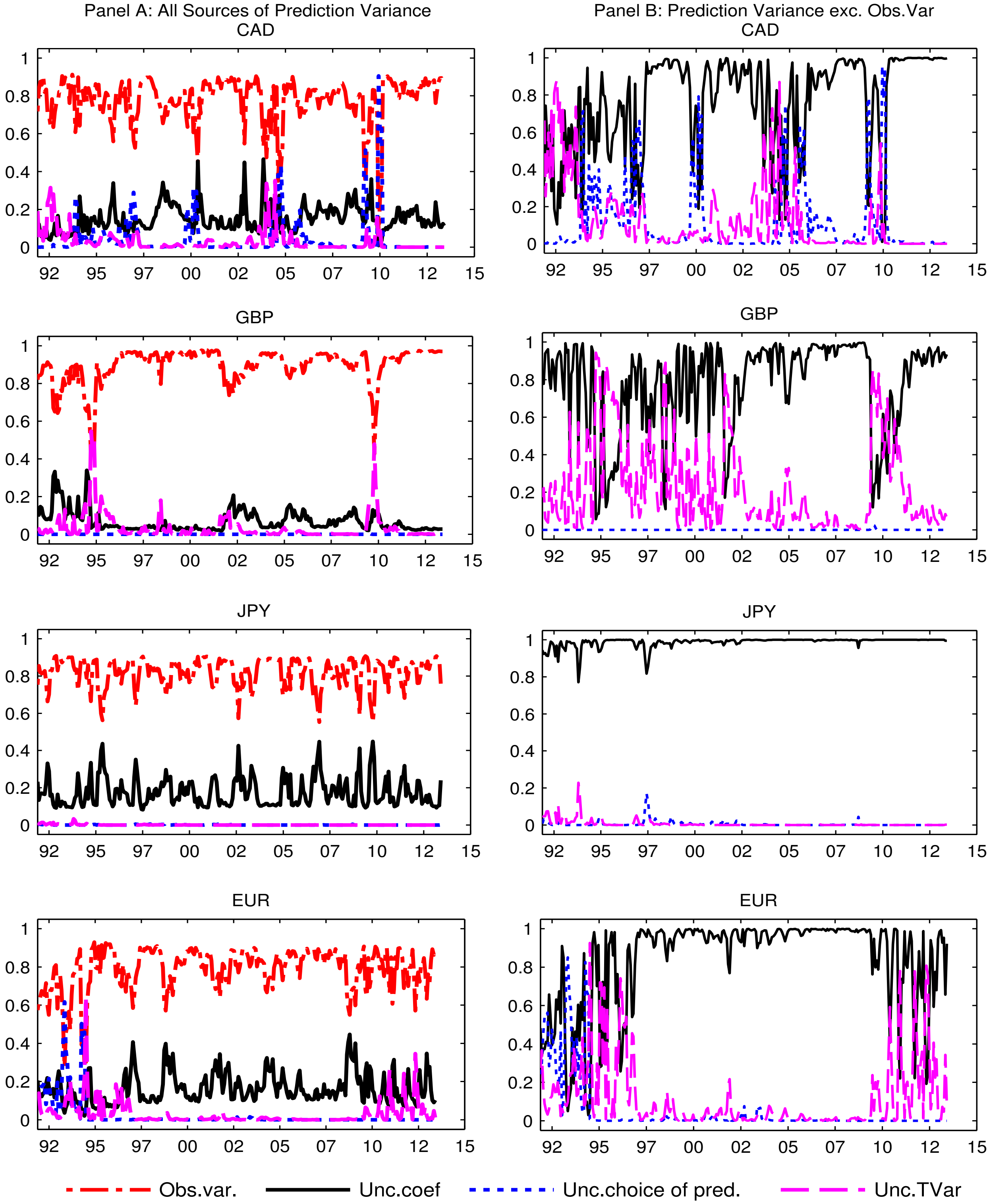
The key aspect of these figures is that they represent an intermediate case of the patterns shown for the figures corresponding to the 1-month and 12-months horizons. For instance, the plots for the NOK resemble the pattern discussed in the main text for 1-month horizon, since our flexible Bayesian method yielded poor performance for this currency. For currencies for which the Bayesian methods including Tvar-coeffs significantly outperformed the RW, like the JPY and the EUR, the patterns of the figures are fairly similar to the ones at the 12-months horizon. This characterization is therefore consistent with the finding that when estimation uncertainty and the uncertainty about the level of time-variation in coefficients is low or is appropriately captured, the model's out-of-sample forecasting performance is satisfactory.

Looking at the weights associated with each degree of time-variation in coefficients in Figure B.2, models with high degree ($\delta = 0.90$) commonly attract the largest weight. And for the majority of the currencies, the shifts in the most likely support points across the forecasting sample are rare if we disregard the initial data-points. The exception is also the NOK and to some extent the GBP, for which the shifts are noticeable over the entire forecasting sample. As a consequence of the frequent shifts in the degree of time-variation supported by the data for these two currencies, the uncertainty with respect to the degree of time-variation in coefficients is relatively high, impairing the model's forecasting ability.

From the figure on the predictors with the largest support from the data at the 3-months horizon - Figure B.3, we can draw the same conclusion reported in the main text. That is, within a specific country-horizon often few variables matter, and the uncertainty about the choice of the predictors is mostly small.

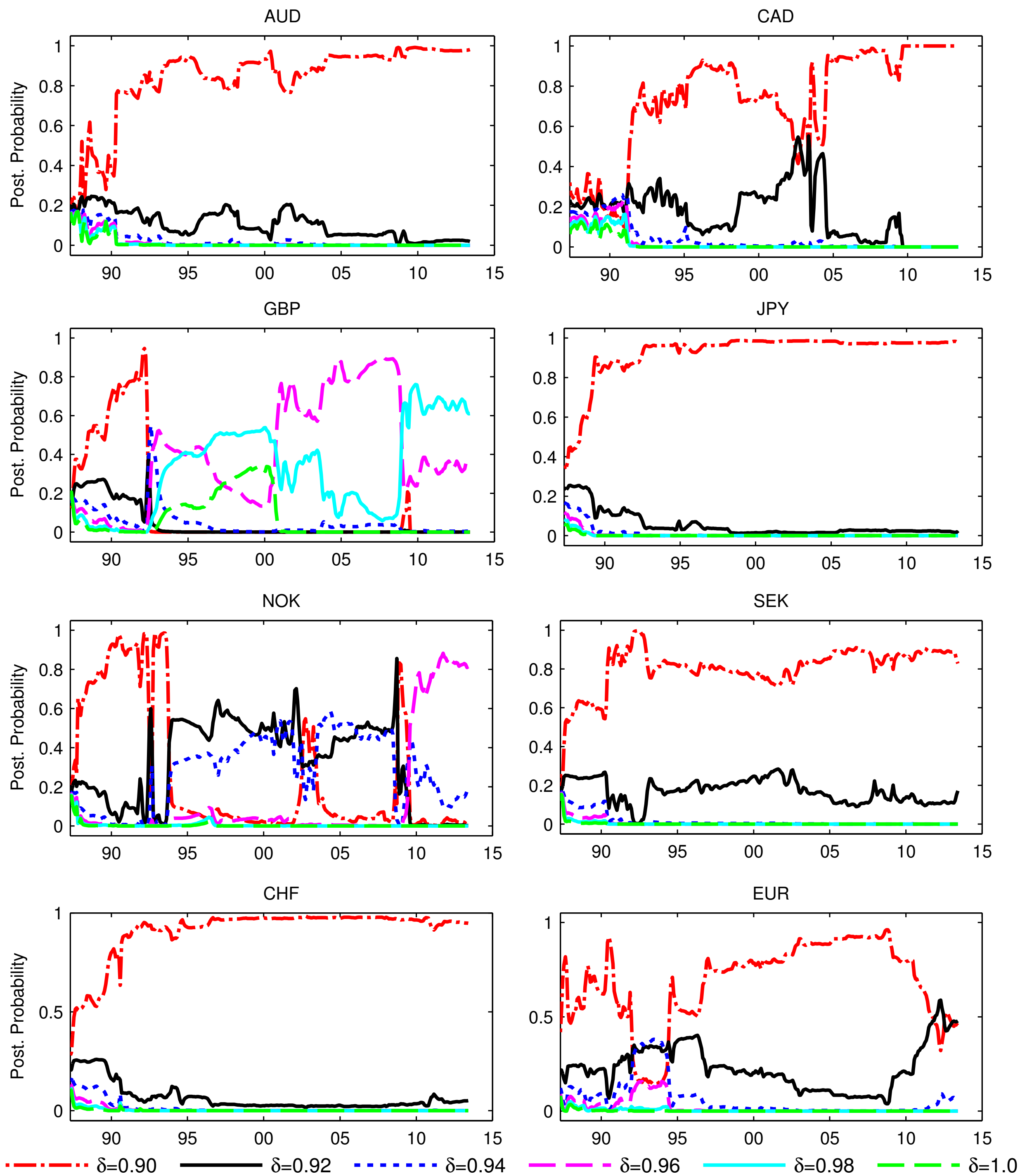
In Figure B.4, the Fluctuation test (Ft-test) shows that our BMA including Tvar-coeffs is preferable to BMA excluding Tvar-coeffs at the 3-months horizon. While models with constant-coefficients rarely display significant forecasting ability, the ones with time-varying coefficients exhibit forecasting power for a large proportion of the forecasting window for several currencies. Examples of these latter cases are vivid for the AUD, SEK, EUR, and CHF. The major exception is the Ft-test for the GBP, for which predictability is barely found regardless of whether models with constant or time-varying coefficients are used.

Figure B.1: Sources of Prediction Variance at 3-Month Forecasting Horizon



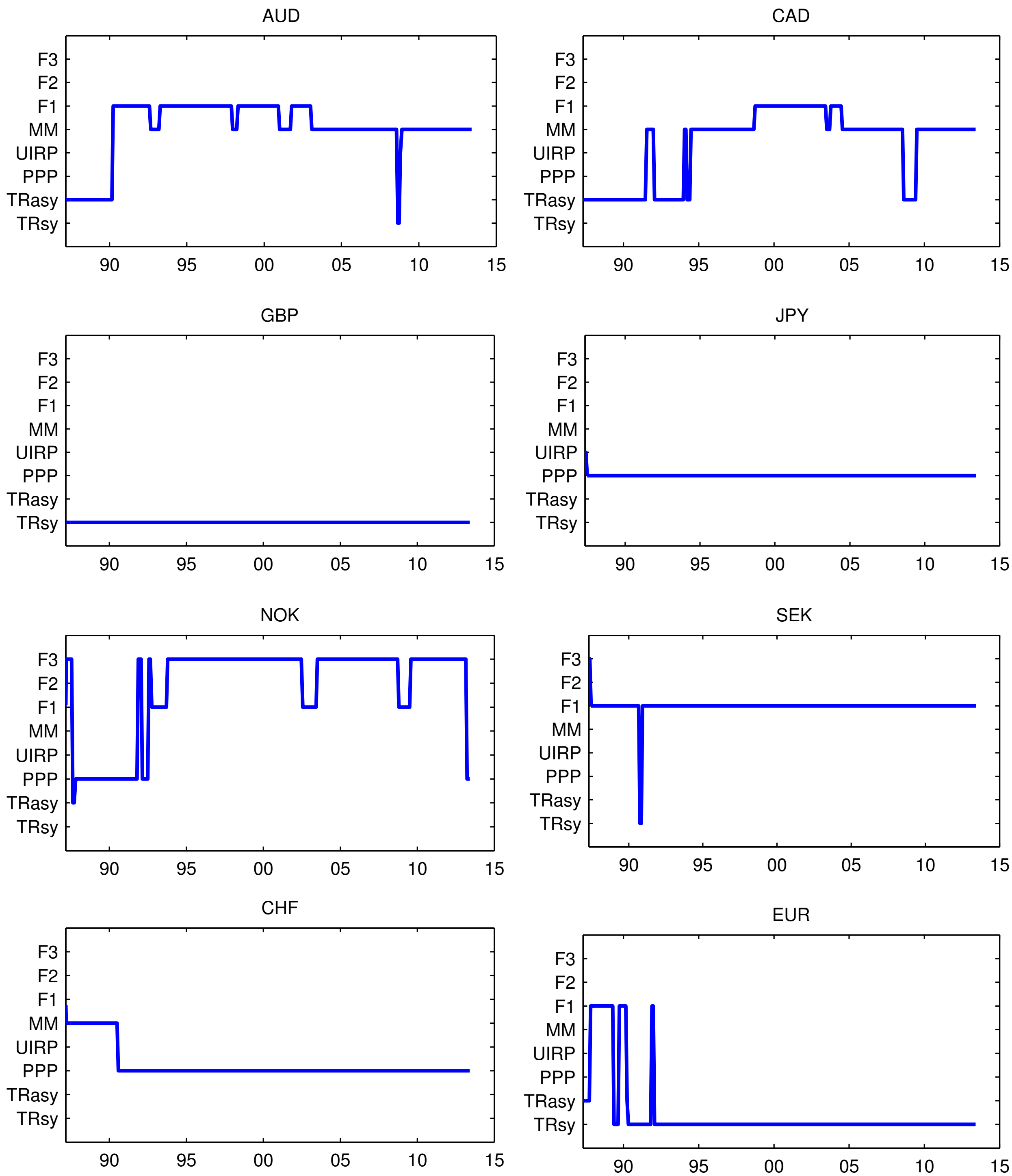
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure B.2: Posterior Probabilities of Degrees of Time-variation in Coefficients, h=3



Notes: Posterior probabilities of values of δ (support points for time-variation in coefficients) at 3-months forecasting horizon. These are the weights employed to produce the average forecasts in the BMA including TVar-Coeffs at this horizon.

Figure B.3: Predictors with the Largest Posterior Probability at Each Period, h=3



Notes: Predictors with the highest probability at each point in time at 3-months forecasting horizon. The forecasts from the BMS including TVar-Coeffs at this horizon are based on these predictors.

Figure B.4: Fluctuation Test at 3-Months Forecasting Horizon

Notes: The Figure shows the Giacomini and Rossi's (2010) one-sided Fluctuation test (Ft-test) based on DMW-test for the BMA including TVar-Coeffs and the BMA excluding TVar-Coeffs. The benchmark model is the driftless Random Walk (RW). It also displays the one-sided Ft-test critical value at 10% level of significance. When the Ft-test statistic is above its critical value, we reject the null of equal local relative forecasting performance between our BMA methods and the RW, and conclude that the method under scrutiny forecasts significantly better than the RW at that point in time. When the Ft-test is below its critical value, the RW produces more accurate forecasts than the competing method. The forecasting horizon is $h = 3$.

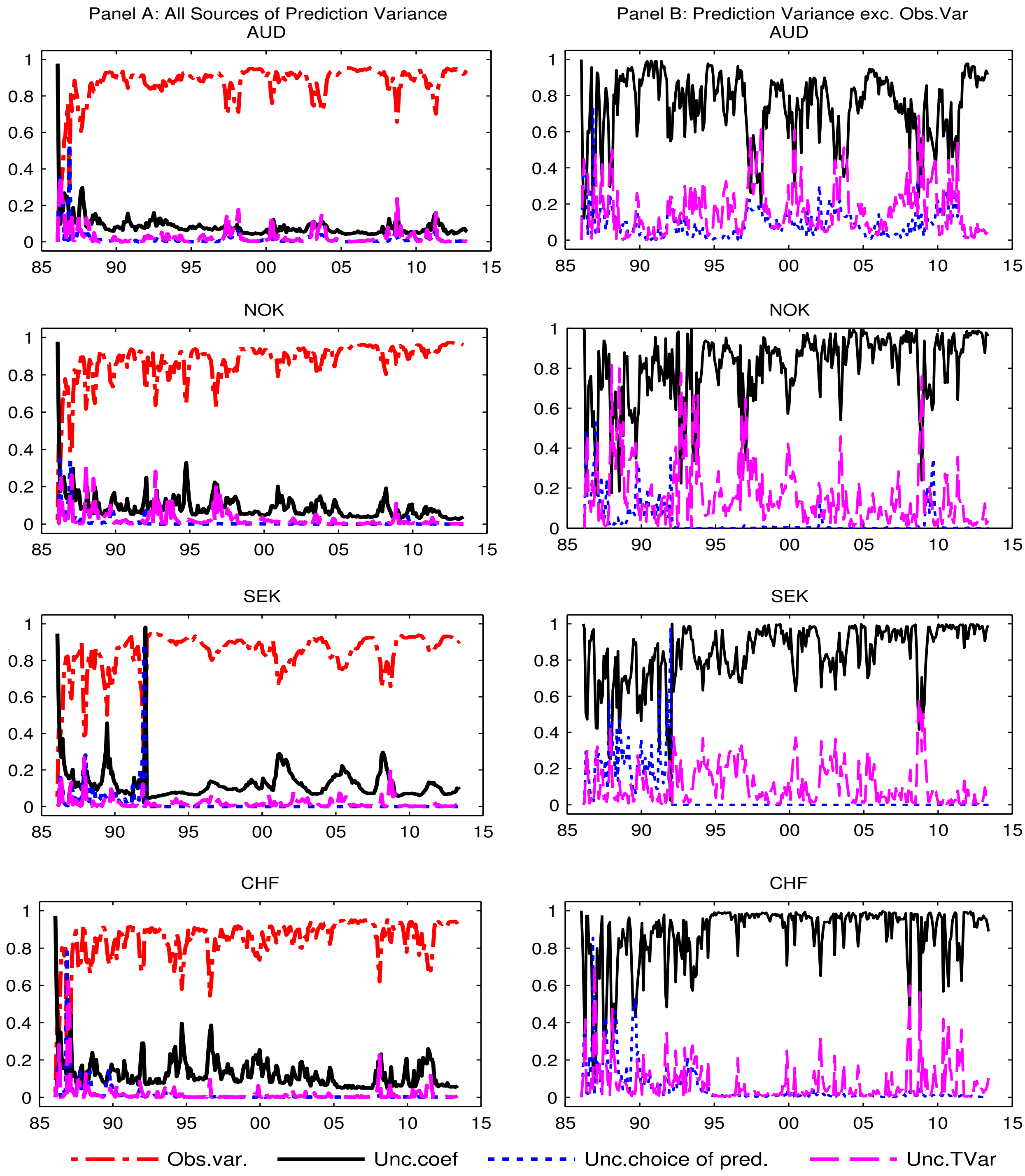
C Sources of Prediction Uncertainty for Additional Currencies

In the main text, we present figures on the sources of prediction uncertainty for the CAD, GBP, JPY, and the EUR. This section contains analogous figures for four extra currencies, namely the AUD, NOK, SEK, and CHF. These are:

- Figure C.1: Sources of Prediction Variance at 1-Month Forecasting Horizon;
- Figure C.2: Sources of Prediction Variance at 3-Months Forecasting Horizon;
- Figure C.3: Sources of Prediction Variance at 12-Months Forecasting Horizon.

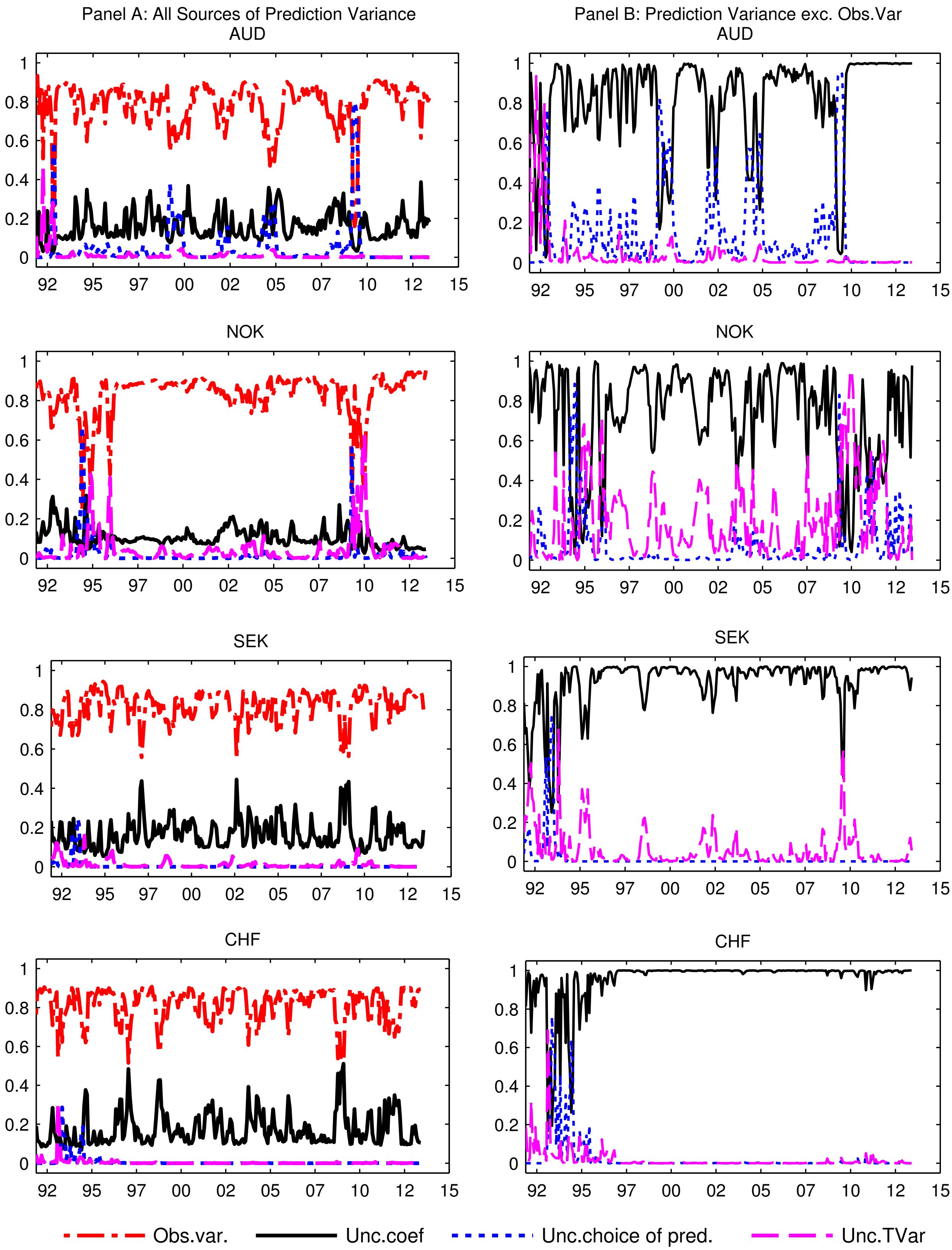
Like the results for the representative selection of currencies discussed in the main text, at the shortest-horizon, estimation uncertainty and uncertainty with respect to the choice of degree of time-variation in coefficients are noticeable throughout the forecasting period. As the forecasting horizon increases, see Figure C.3, estimation uncertainty remains the most prominent source of prediction uncertainty for most currencies, which is successfully embedded in our flexible BMA approach. The main case apart at the 12-months horizon is the SEK, for which the uncertainty with respect to the choice of degree of time-variation in coefficients prevails over the forecasting period. However, note that for this currency the forecast improvement upon the RW is statistically insignificant using either bootstrapped or the asymptotic critical values - see the first table in the main text. Note too that the plots at the 3-months horizon constitute intermediate cases, with currencies like the NOK displaying patterns similar to the 1-month horizon, and the CHF showing patterns akin to the 12-months horizon.

Figure C.1: Sources of Prediction Variance at 1-Month Forecasting Horizon



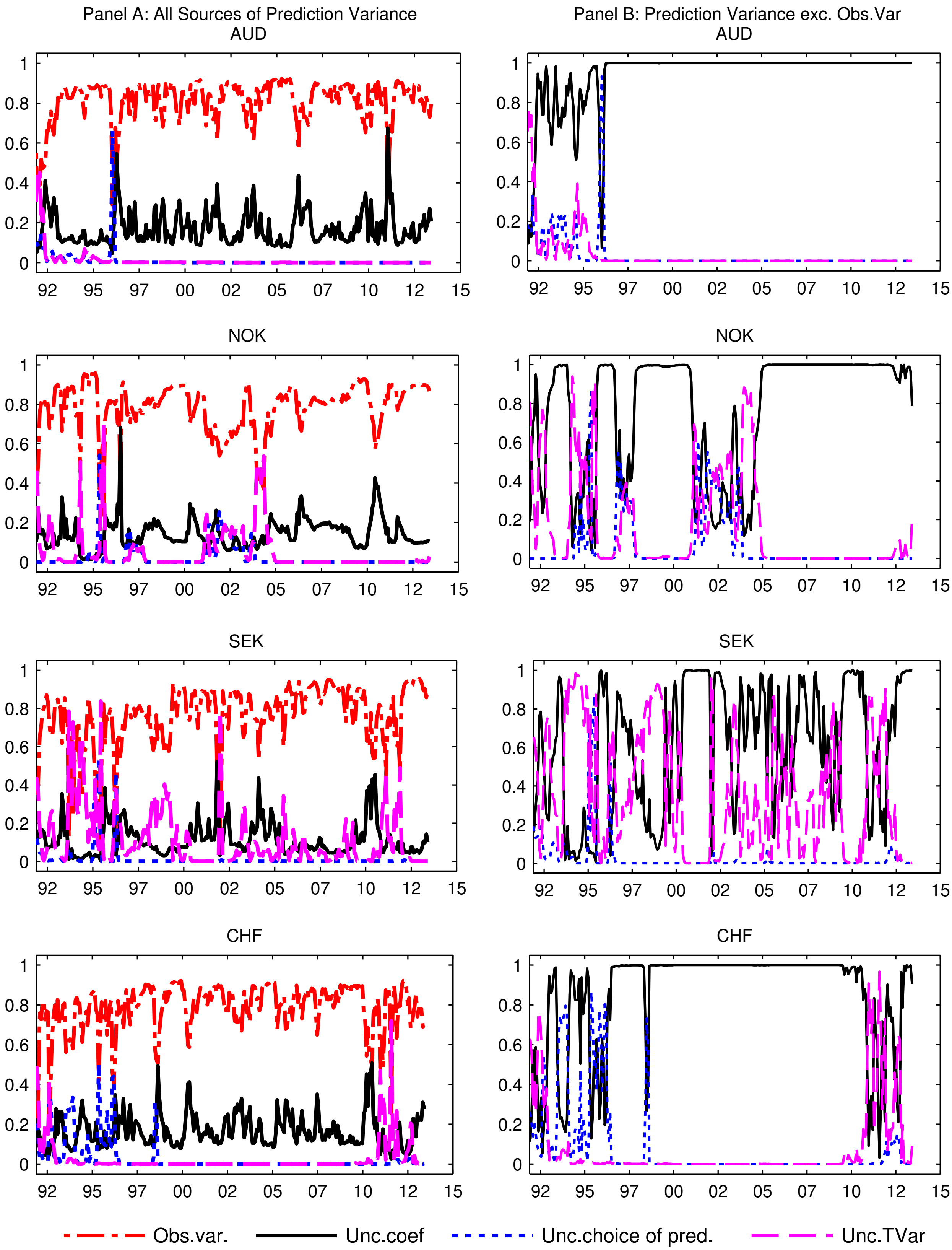
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure C.2: Sources of Prediction Variance at 3-Months Forecasting Horizon



Notes: As in Figure C.1, except that here the forecasting horizon is 3-months.

Figure C.3: Sources of Prediction Variance at 12-Months Forecasting Horizon

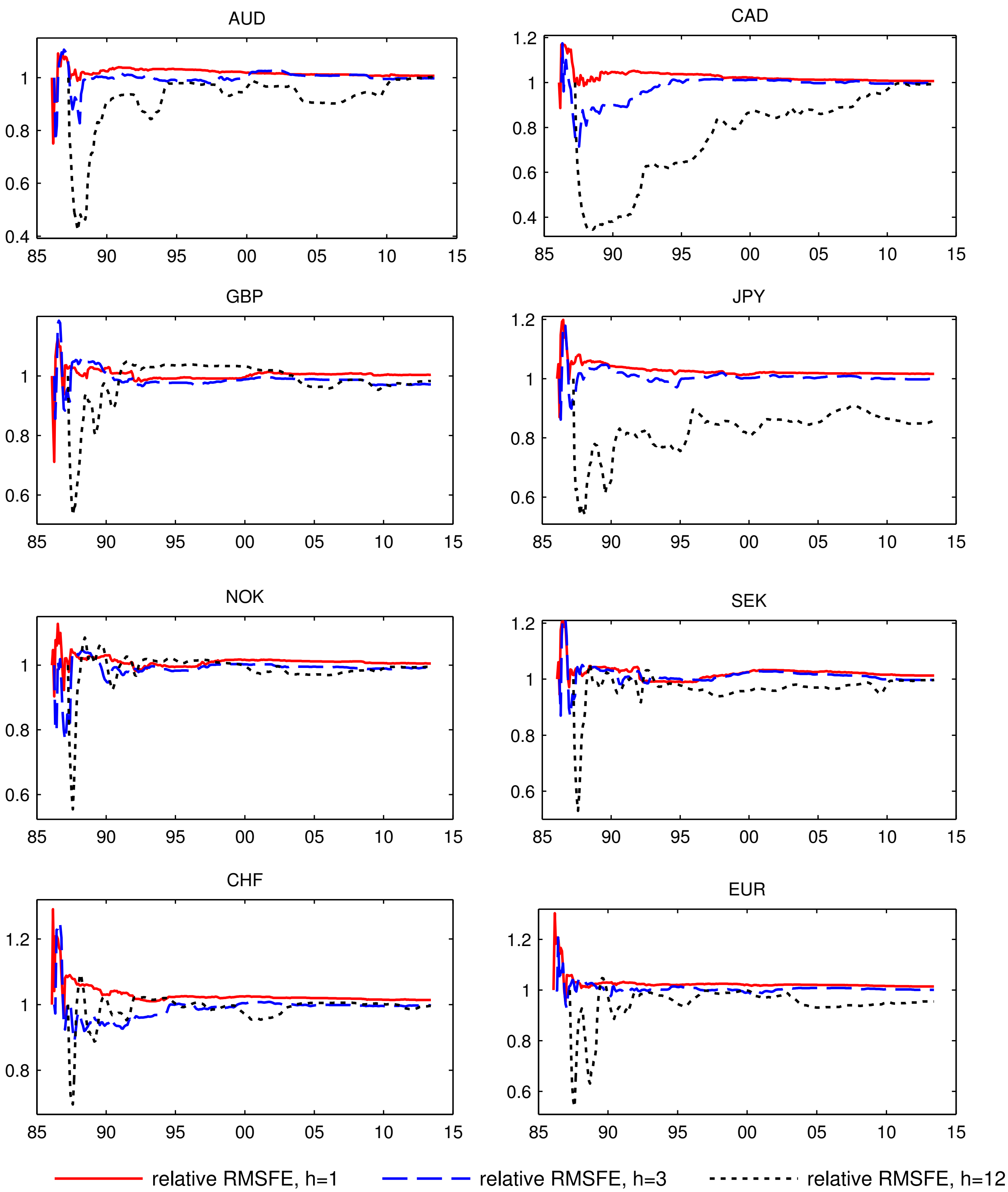


Notes: As in Figure C.1, except that here the forecasting horizon is 12-months.

D Recursive Root Mean Squared Forecast Error (RMSFE)

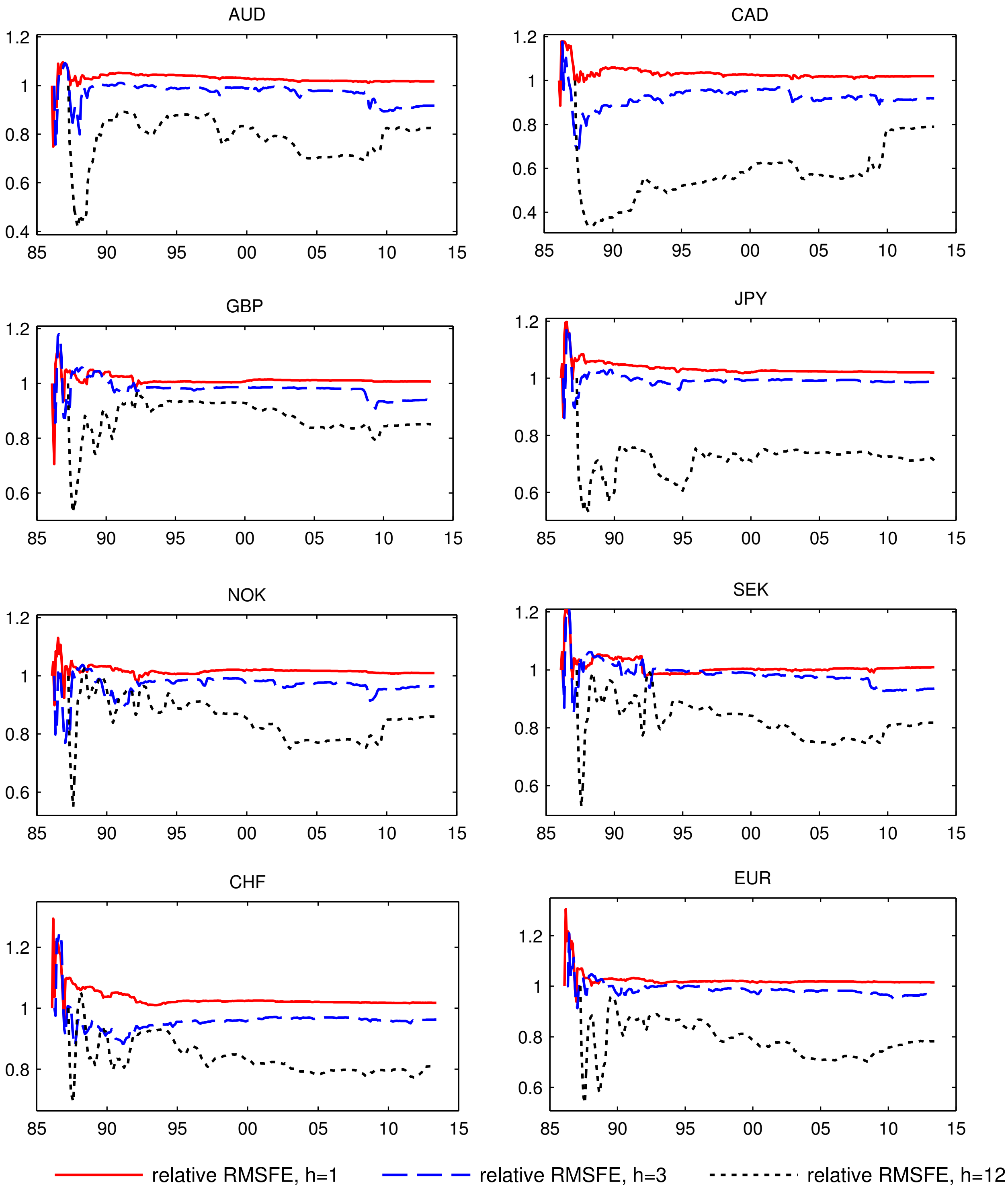
This section contains plots of the recursive relative RMSFE for the BMA excluding Tvar-coeffs and BMA including Tvar-coeffs - Figures D.1 and D.2, respectively. Apart from examining forecasting performance over time, these figures also allow us to shed light on the $\cup$ pattern documented in Giannone (2010) - see also the main text. As the figures illustrate, the pattern documented by Giannone (2010) is clear for the BMA excluding Tvar-coeffs. Taking the vivid example of the AUD and CAD at the 12-months horizon, the relative forecast accuracy improves steeply as the forecasting window is extended by few extra observations - stage two. But just before 1990M12, the performance starts to deteriorate, and by the end of the forecasting window the gains relative to the RW become small or disappear. In contrast, for the BMA including Tvar-coeffs we can confirm Giannone's (2010) observations for the first and second stages, but not the third stage. As shown in Figure D.2, further increasing the window does not lead to a complete deterioration of the forecasting performance relative to the RW. For the most part of the forecasting period and horizons greater than 1-month the relative RMSFE is below one, favouring our flexible models.

Figure D.1: Recursive RMSFE of BMA excl. TVar-Coeffs relative to the RMSFE of the RW



Notes: Recursive RMSFE of Bayesian Model Averaging (BMA) excluding Tvar-coeffs relative to the RMSFE of the Random Walk (RW). A value of one corresponds to equal forecasting performance; values below one are in support of the BMA excluding Tvar-coeffs and values above one are in favour of the RW.

Figure D.2: Recursive RMSFE of BMA incl. TVar-Coeffs relative to the RMSFE of the RW



Notes: Recursive RMSFE of Bayesian Model Averaging (BMA) including Tvar-coeffs relative to the RMSFE of the Random Walk (RW). A value of one corresponds to equal forecasting performance; values below one are in support of models with time-varying coefficients and figures above one are in favour of the RW.

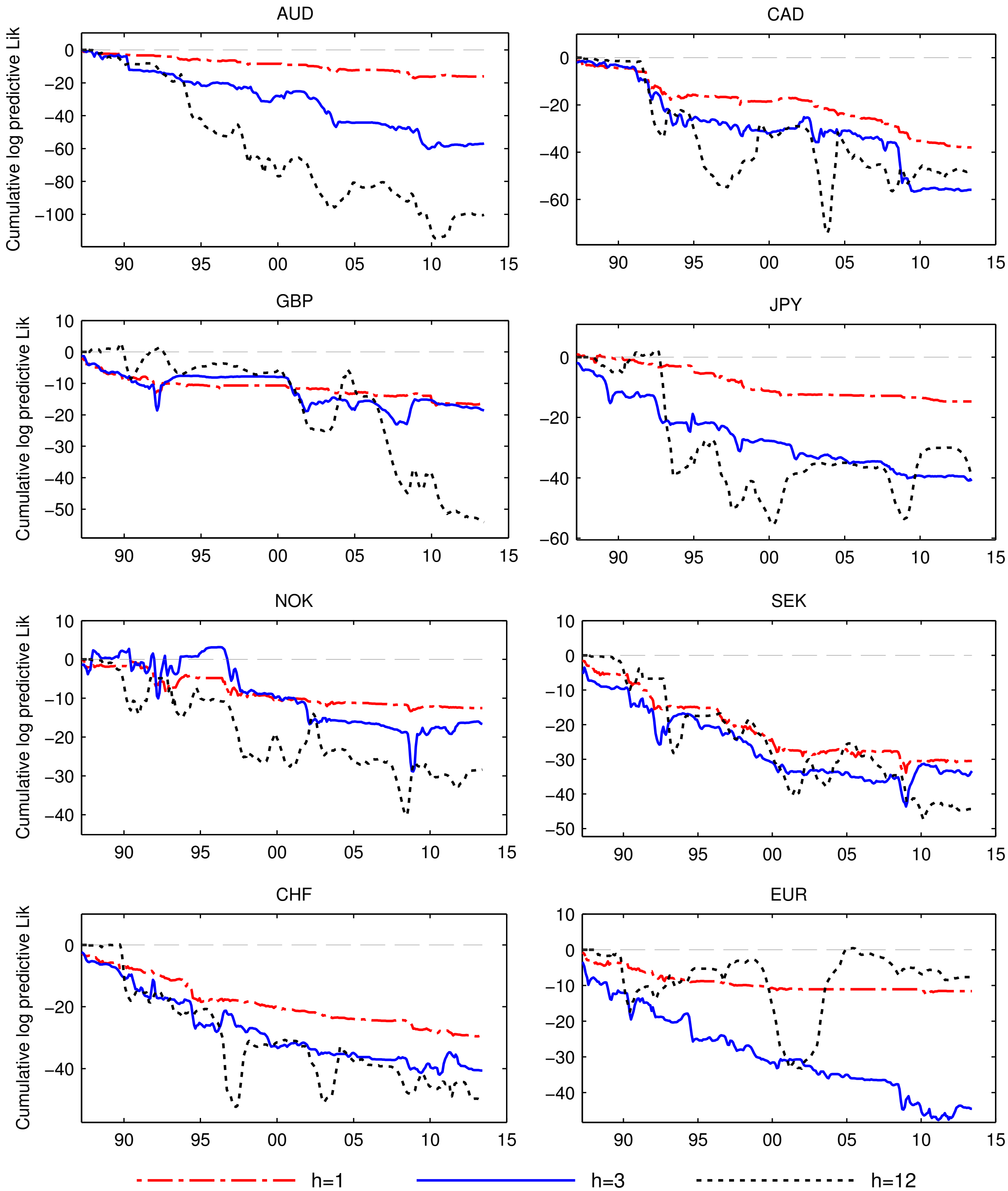
E Cumulative Predictive Likelihoods

To add up to the results on the forecasting performance over time we focus in another measure of forecast accuracy that underlies our Bayesian approach, namely the predictive likelihoods, see Geweke and Amisano (2010). Figure E.1 depicts the cumulative log predictive likelihoods for models with constant-coefficients relative to the ones with time-varying coefficients. A value of zero corresponds to equal marginal support for both models; negative values are in support of models with time-varying coefficients; and positive values are in favour of models with constant-coefficients.

Two main results are apparent in Figure E.1. First, it confirms that models with time-varying coefficients are empirically plausible. The cumulative log predictive likelihoods become negative after a number of out-of-sample data-points have been accumulated.[2] These cumulative predictive likelihoods show a downward trend, consistent with additive evidence favouring models with time-varying coefficients. Second, observations around the 2008 financial crisis, where significant shifts in economic conditions occurred, contribute highly to the evidence in favor of the time-varying coefficients models. This is especially true at the 3- and 12-months horizons. Overall, our findings remain invariant to this measure of forecast accuracy and they further rule out the $\cup$ pattern reported in Giannone (2010).

[2]Geweke and Amisano (2010) point out that it is customary for the results to be sensitive at the beginning of the out-of-sample period, as this reflects sensitiveness to the prior density. As they emphasize, nonetheless, after a number of observations have been accumulated the results become invariant to substantial changes in the prior density distribution.

Figure E.1: Cumulative Log Predictive Likelihoods: BMS excl. TVar-Coeffs/BMS incl. TVar-Coeffs



Notes: Cumulative log predictive likelihoods of the Bayesian Model Selection (BMS) excl. Tvar-coeffs relative to the BMS incl. Tvar-coeffs. A value of zero corresponds to equal marginal support for both models; negative values are in support of models with time-varying coefficients; and positive values are in favour of models with constant coefficients.

F Additional Robustness Checks

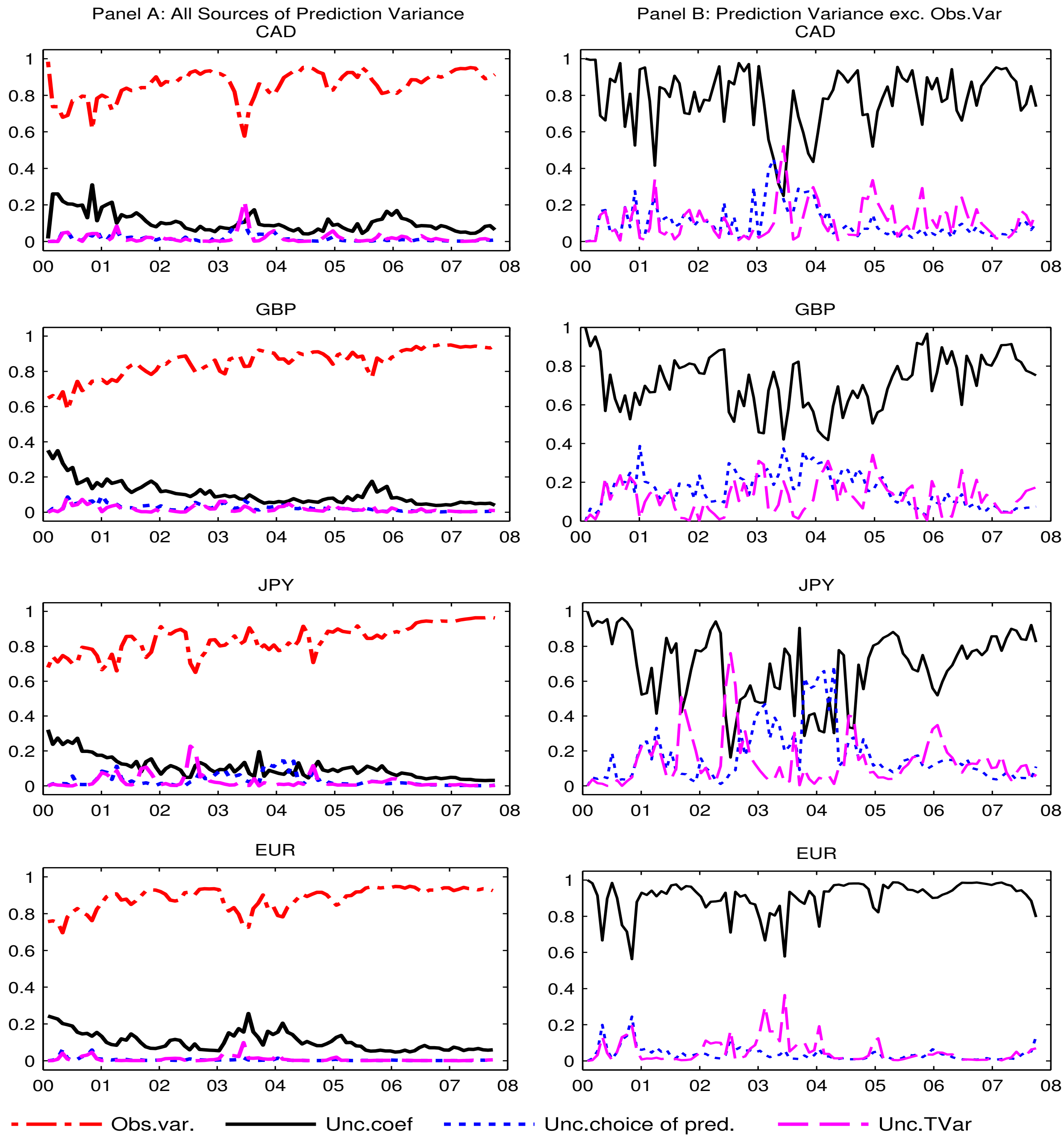
This section presents the figures pertaining to the variance decomposition for the model with an extended set of predictors described in the paper. It further discusses four additional robustness checks: (i) extension to 6- and 24-months forecasting horizons; (ii) extension to other exchange rates; (iii) robustness to the seasonal adjustment method; and (iv) change in the base currency. Unless stated otherwise, in all cases we use the prior setting discussed in the main text in these checks. We also apply the asymptotic one-sided DMW test for statistical inference.

F.1 Variance Decomposition with an Extended Set of Predictors

Figures F.1.1, F.1.2, and F.1.3 present the variance decomposition results for all the four currencies considered when examining the forecasting performance of our methods in a setting with an extended set of predictors. In general, and despite the relatively shorter sample, our findings remain qualitatively similar to the results from the main text. Excluding observational variance, the uncertainty in the estimation of the coefficients and the uncertainty about the true degree of time-variation in coefficients are present over the entire forecasting path at the 1-month and 3-months horizons. On the other hand, at the 12-months horizon in Figure F.1.3, estimation uncertainty is the most dominant source of prediction uncertainty for three of the four currencies, after isolating the influence of initial observations.[3]

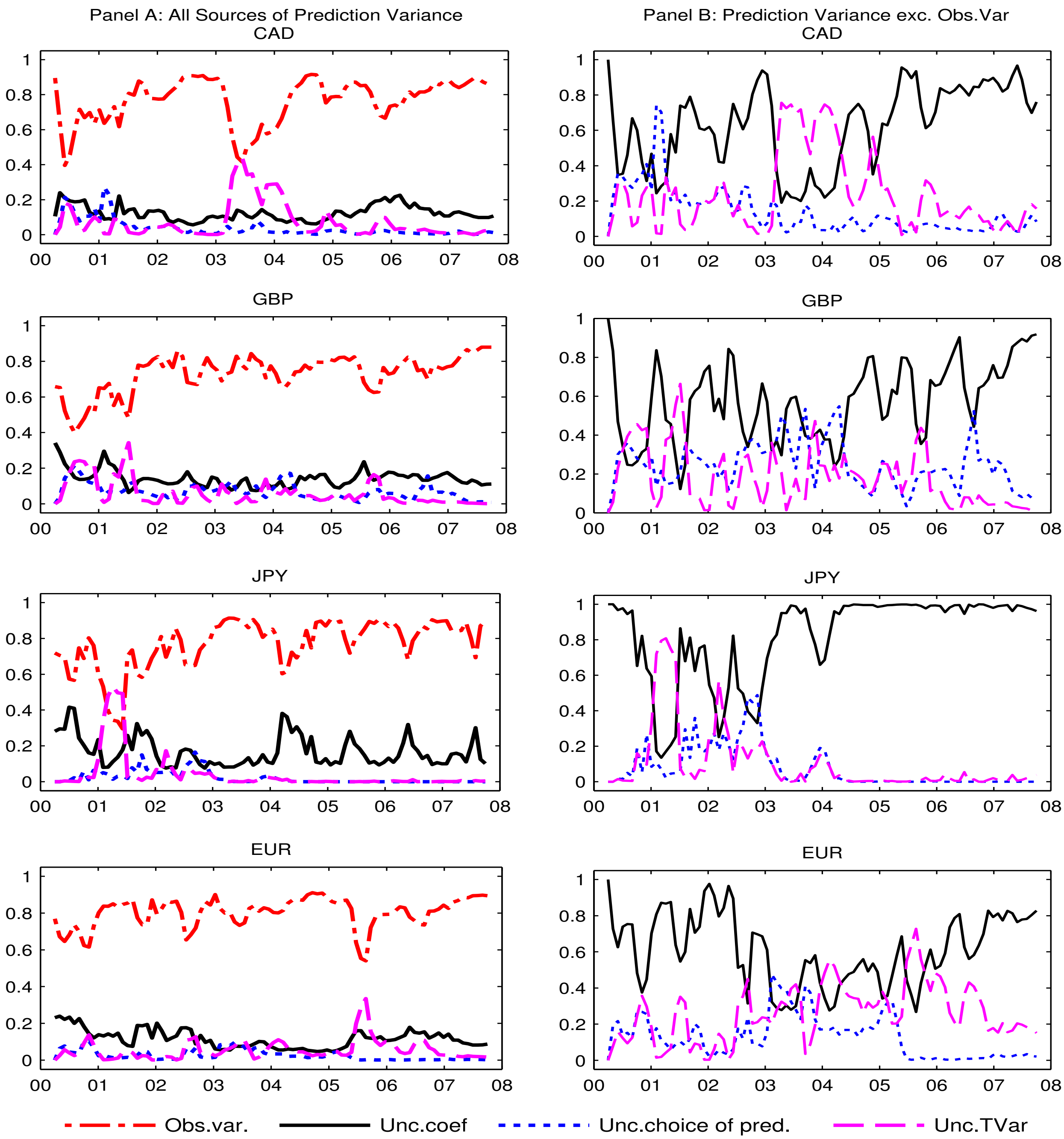
[3]We caution that one should bear in mind that the data on net foreign asset, which is one of the regressor included in our set of predictors here, are likely to be noisy. The data were transformed to monthly observations by interpolation from quartely data, where the quarterly data themselves were also interpolated from annual data by Della Corte et al. (2012).

Figure F.1.1: Sources of Prediction Variance at 1-Month Forecasting Horizon



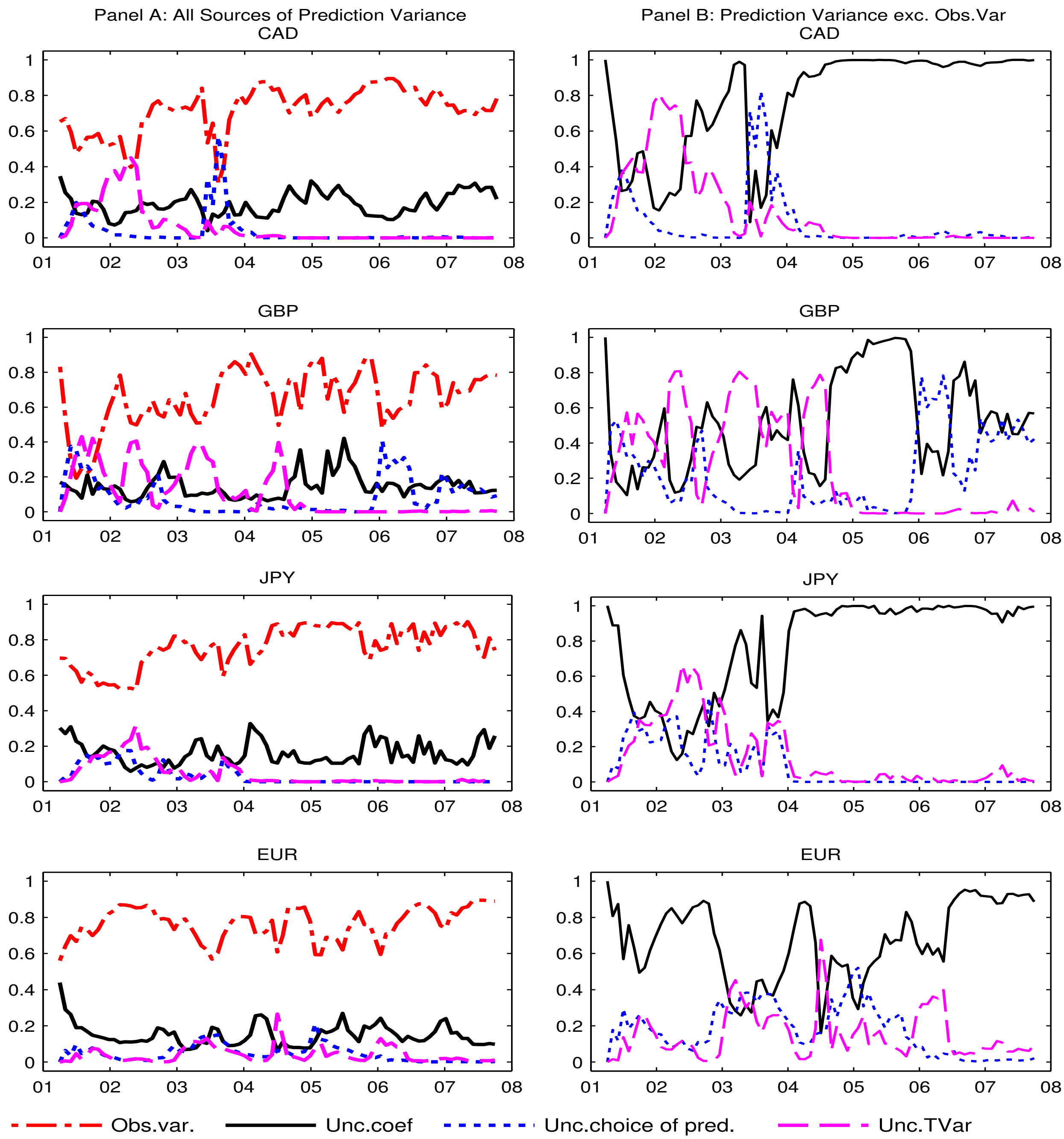
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure F.1.2: Sources of Prediction Variance at 3-Months Forecasting Horizon



Notes: As in Figure F.1.1, except that here the forecasting horizon is 3-months.

Figure F.1.3: Sources of Prediction Variance at 12-Months Forecasting Horizon



Notes: As in Figure F.1.1, except that here the forecasting horizon is 12-months.

F.2 Extension to 6- and 24-Months Forecasting Horizons

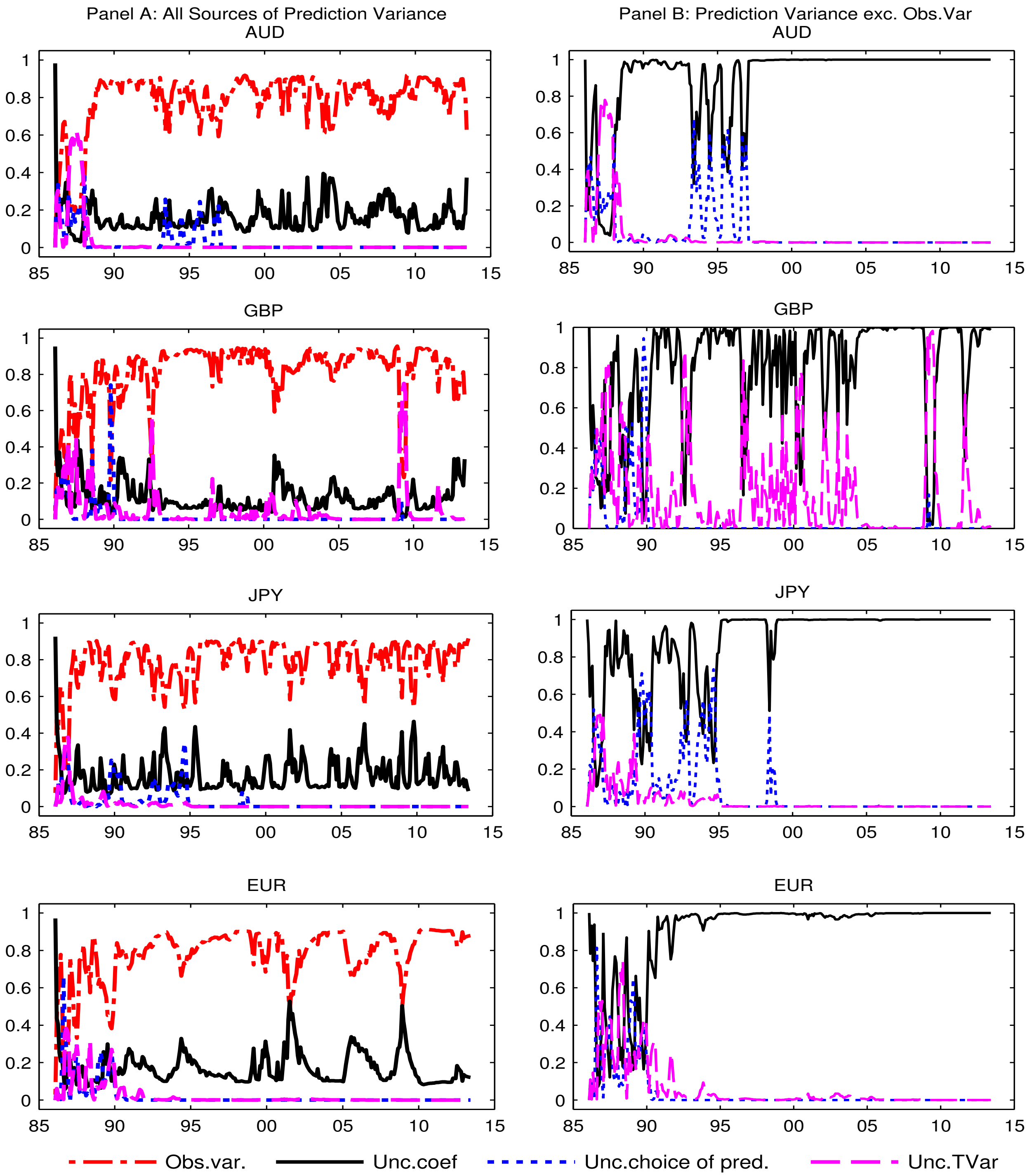
In Table F.2.1 and figures F.2.1-F.2.2 we extend our key analyses to the 6- and 12-months forecasting horizon. Results confirm that BMA including TVar-Coeffs remains preferable to BMA with constant coefficients at these horizons. As well, the figures on variance decomposition mostly reveal the pattern we documented: the uncertainty regarding the true degree of time-variation in the coefficients is episodic or clustered at the initial data-points of the out-of-sample period. By embedding in the model the most likely degree of time-variation present in the data, BMA including Tvar-coeffs offsets the loss in forecasting performance stemming from estimation uncertainty, leading to significant improvements upon the RW.

Table F.2.1: Forecasting Performance at 6- and 24-Months Horizon

	BMA incl. TVar-Coeffs		BMA excl. TVar-Coeffs	
	h=6	h=24	h=6	h=24
AUD	**0.744*** **	**0.842***	1.032	1.134
CAD	**0.860***	**0.568*** **	**0.967***	1.096
GBP	**0.914**	**0.513*** **	**0.908***	**0.832*** **
JPY	**0.862*** **	**0.756*** **	1.014	**0.796*** **
NOK	**0.885**	**0.662*** **	**0.977**	**0.929** **
SEK	**0.961**	**0.930**	**0.989**	1.205
CHF	**0.935***	**0.727*** **	1.005	1.113
EUR	**0.840*** **	**0.693*** **	1.003	1.020

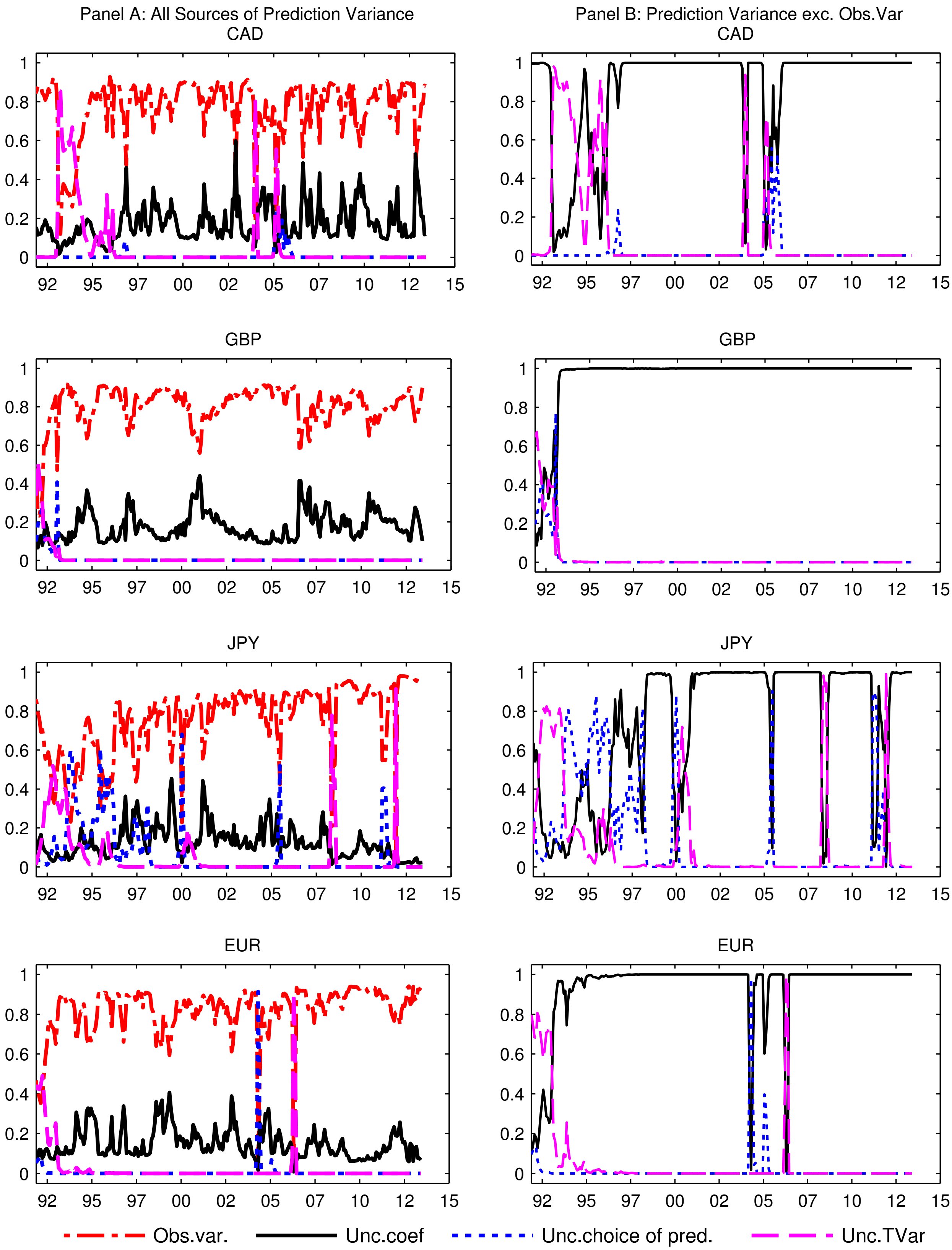
Notes: RMSFE of the BMA including or excluding time-varying coefficients, relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the method under scrutiny generates a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) denote the level of significance of the one-sided DMW test based on standard critical values, for the null hypothesis of equality in the RMSFE.

Figure F.2.1: Sources of Prediction Variance at 6-Months Forecasting Horizon



Notes: Decomposition of the prediction variance into its constituent parts at 6-months forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure F.2.2: Sources of Prediction Variance at 24-Months Forecasting Horizon



Notes: As in Figure F.2.1, except that here the forecasting horizon is 24-months.

F.3 Extension to other Exchange Rates

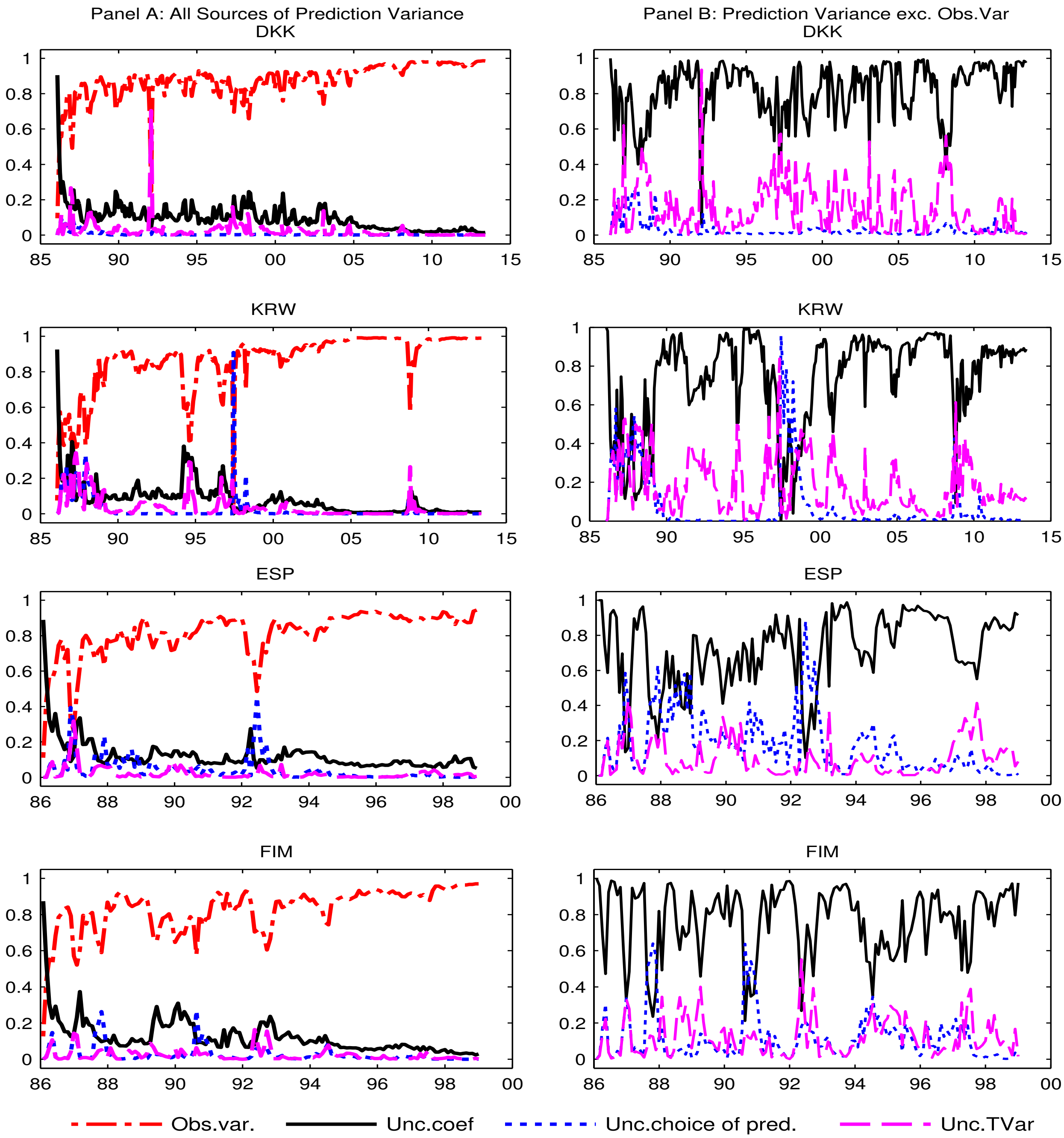
In Table F.3.1 and figures F.2.1-F.2.2 we examine whether our results apply more generally to other exchange rates. Specifically we consider U.S. dollar exchange rates for the following currencies: Danish krone (DKK), Korean won (KRW), Austrian schilling (ATS), Belgium franc (BEF), French franc (FRF), Spanish peseta (ESP), Italian lira (ITL), and the Finnish markka (FIM). The forecasting period is 987M12+h to 2013M5 for the DKK and KRW, while for the countries that adhered to the Euro in 1999, the forecasting period ends in 1998M12. The results suggest that our findings also hold for these currencies.

Table F.3.1: Extension to Other Currencies

	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
	h=1	h=3	h=12	h=1	h=3	h=12
DKK	1.046	**0.958**	**0.903**	1.016	**0.993**	**0.970**
KRW	1.073	**0.948**	**0.583*****	1.001	1.005	**0.991**
ATS	1.032	**0.989**	**0.765****	1.025	**0.988**	**0.912**
BEF	1.035	**0.950**	**0.803****	1.034	**0.996**	**0.915****
FRF	1.062	**0.920**	**0.906**	1.045	**0.973**	**0.955**
ESP	1.028	**0.926**	**0.644*****	1.024	**0.994**	**0.867****
ITL	1.058	**0.996**	**0.789***	1.030	**0.988**	**0.895**
FIM	1.042	**0.990**	**0.530*****	1.016	1.020	**0.752****

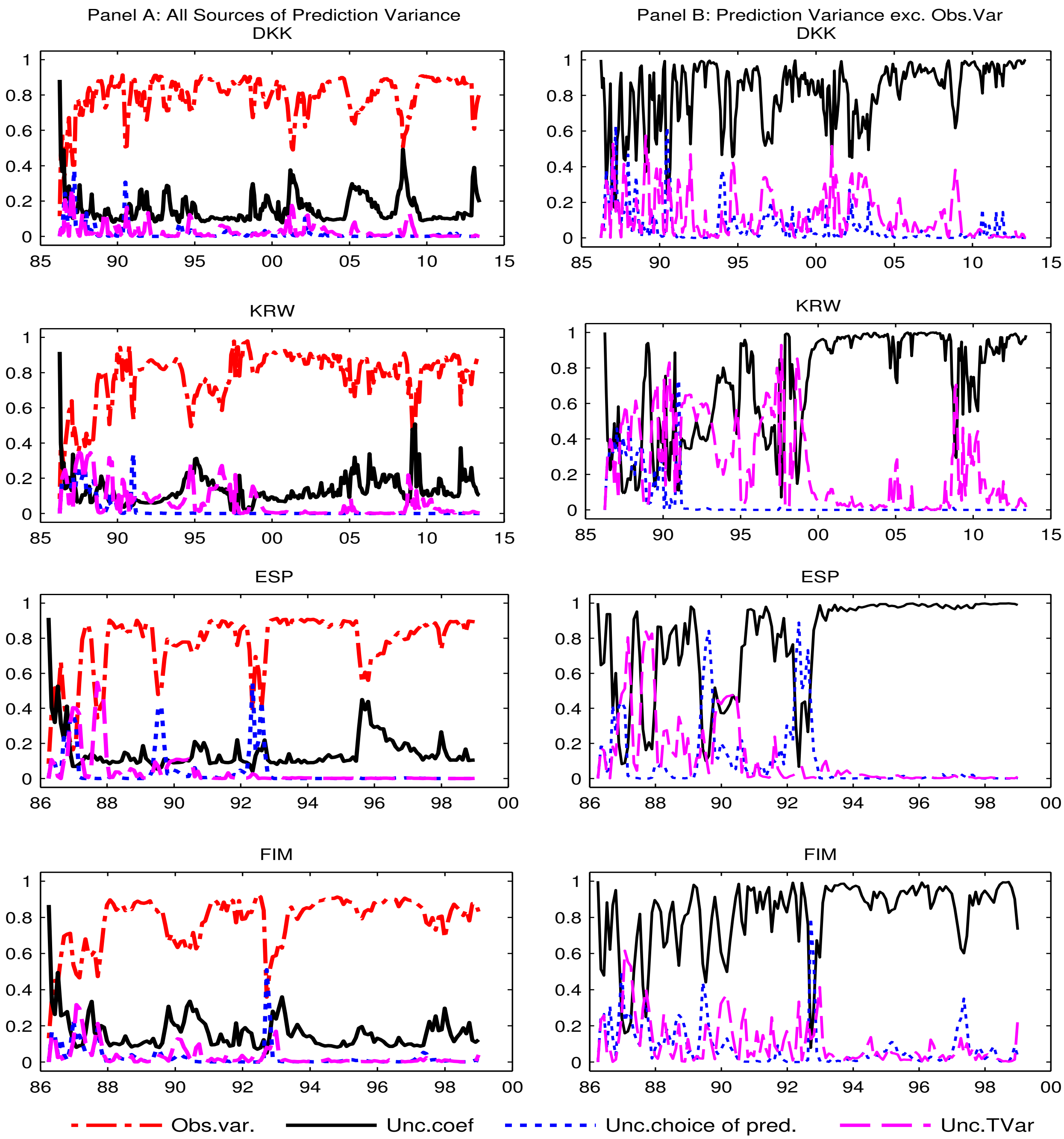
Notes: RMSFE of the BMA including or excluding time-varying coefficients, relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the method under scrutiny generates a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) denote the level of significance of the one-sided DMW test based on standard critical values, for the null hypothesis of equality in the RMSFE. Currency codes denote: DKK-Danish krone; KRW-Korean won; ATS-Austrian schilling; BEF-Belgium franc; FRF-French franc; ESP-Spanish peseta; ITL-Italian lira; and FIM-Finnish markka. The forecasting period is 1987M12+h - 2013M5 for the DKK and KRW, while for the other currencies it ends in 1998M12.

Figure F.3.1: Sources of Prediction Variance at 1-Month Forecasting Horizon



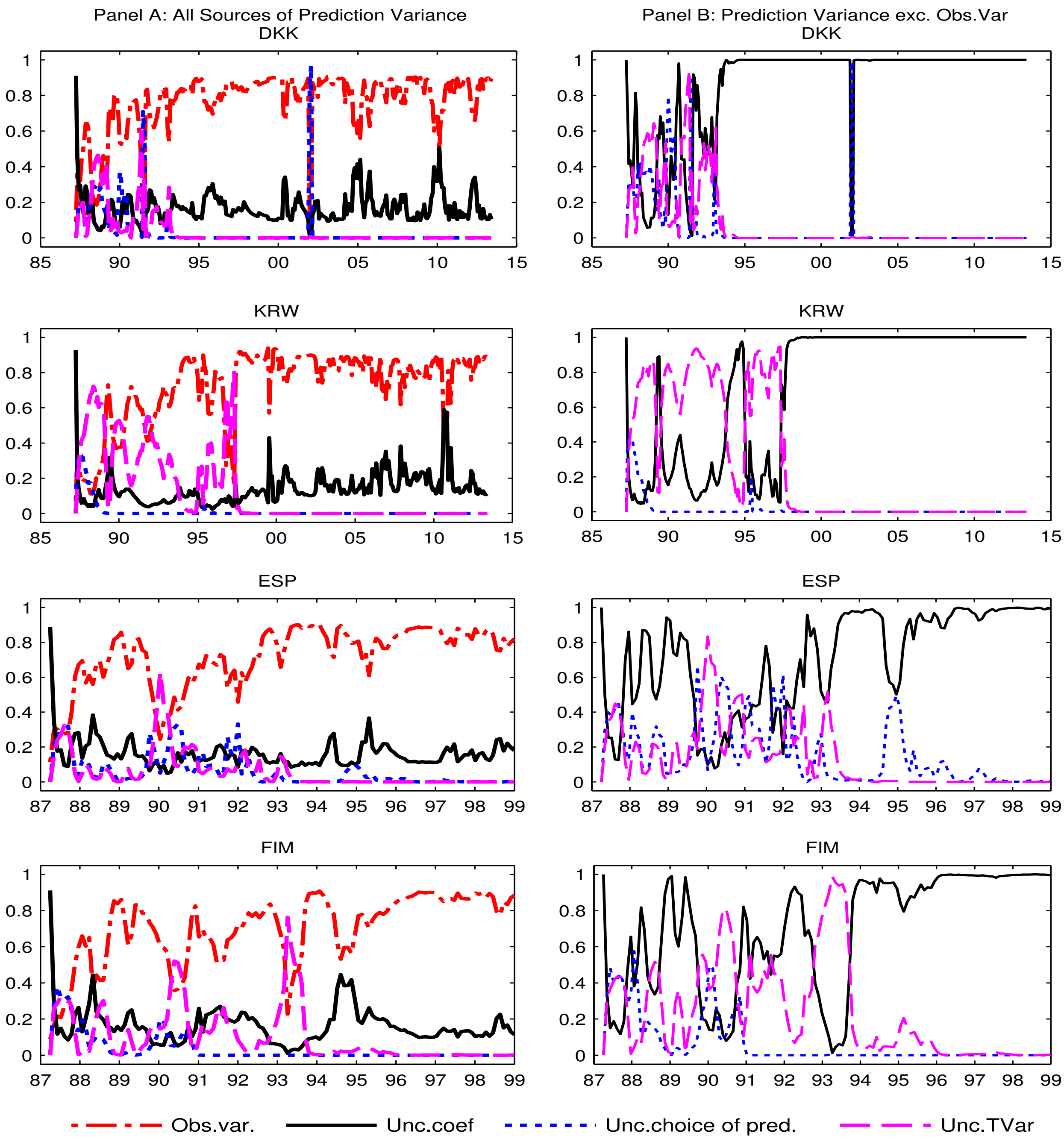
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure F.3.2: Sources of Prediction Variance at 3-Months Forecasting Horizon



Notes: As in Figure F.3.1, except that here the forecasting horizon is 3-months.

Figure F.3.3: Sources of Prediction Variance at 12-Months Forecasting Horizon



Notes: As in Figure F.3.1, except that here the forecasting horizon is 12-months.

F.4 Robustness to the Seasonal Adjustment Method

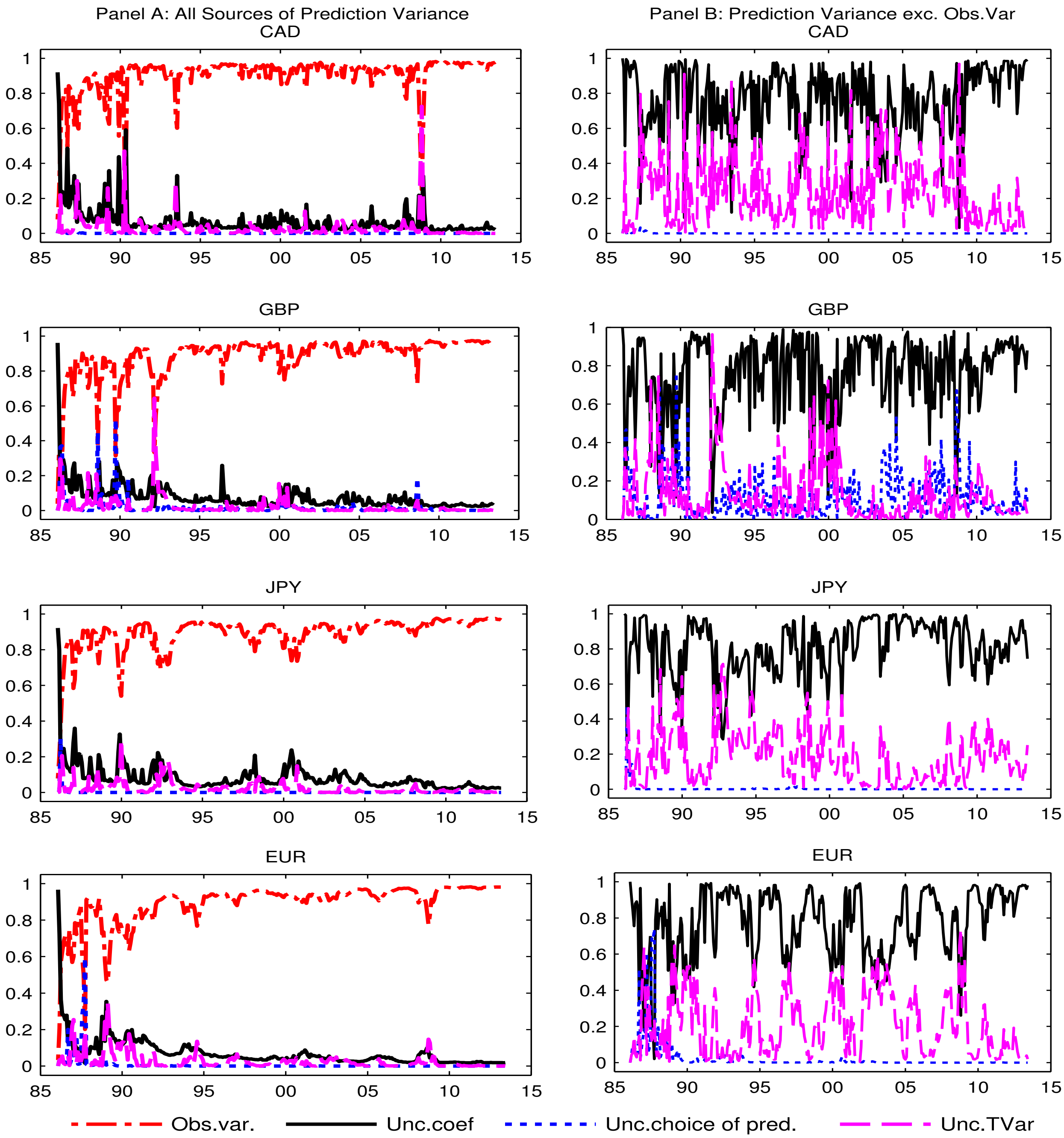
In all our exercises so far we have used data that were seasonally adjusted by taking a moving average of twelve months. To examine the sensitiveness of our results to the method used to remove seasonality we apply the X12 decomposition algorithm of the U.S. Census Bureau. Forecasting results based on these data are reported in Table F.4.1, and variance decomposition patterns for the same representative selection of currencies examined in the main text are in Figures F.4.1-F.4.3. Overall, our results remain largely unaffected.

Table F.4.1: Robustness to Seasonal Adjustment Method (use of X12).

	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
	h=1	h=3	h=12	h=1	h=3	h=12
AUD	1.017	**0.932**	**0.624*****	1.024	1.009	1.024
CAD	1.030	**0.993**	**0.988**	1.010	1.004	1.086
GBP	1.069	**0.967**	**0.691*****	1.049	1.009	**0.891***
JPY	1.034	**0.956****	**0.819*****	1.027	1.018	1.043
NOK	1.020	**0.992**	**0.912****	1.009	1.007	1.010
SEK	1.029	**0.943**	1.159	1.012	**0.999**	**0.917*****
CHF	1.035	**0.955***	**0.794*****	1.034	1.020	**0.990**
EUR	1.033	**0.952**	**0.799*****	1.021	1.021	**0.897****

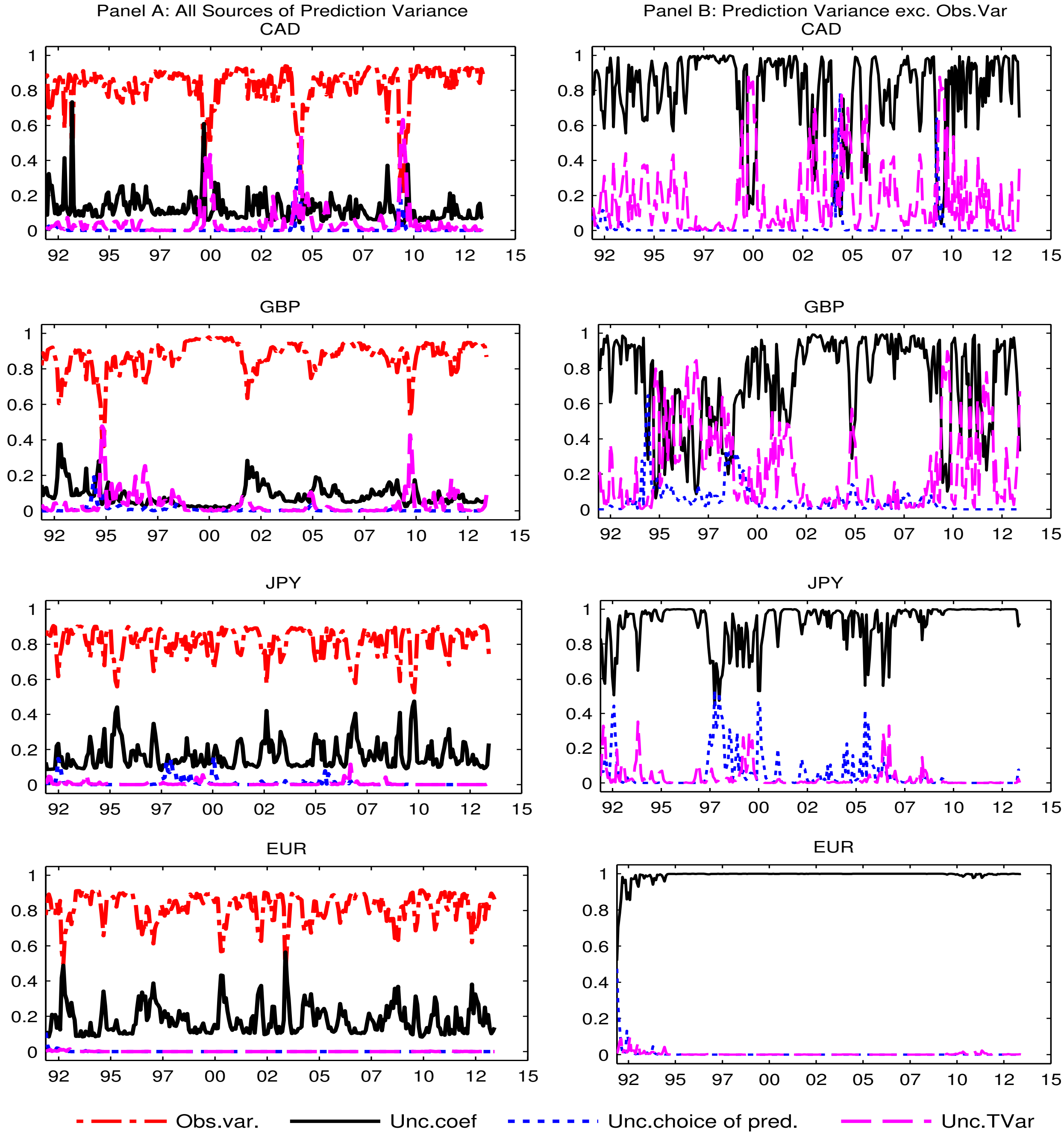
Notes: RMSFE of the BMA including or excluding time-varying coefficients, relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the method under scrutiny generates a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) denote the level of significance of the one-sided DMW test based on standard critical values, for the null hypothesis of equality in the RMSFE. Currency codes denote: AUD - Australian dollar; CAD - Canadian dollar; GBP - Pound sterling; JPY - Japanese yen; NOK - Norwegian krone; SEK - Swedish krona; CHF - Swiss franc; EUR - euro. The forecast evaluation period begins in 1987M12+h in all, but the OLS-DMSPE Combination case (1992M12+h).

Figure F.4.1: Sources of Prediction Variance at 1-Month Forecasting Horizon



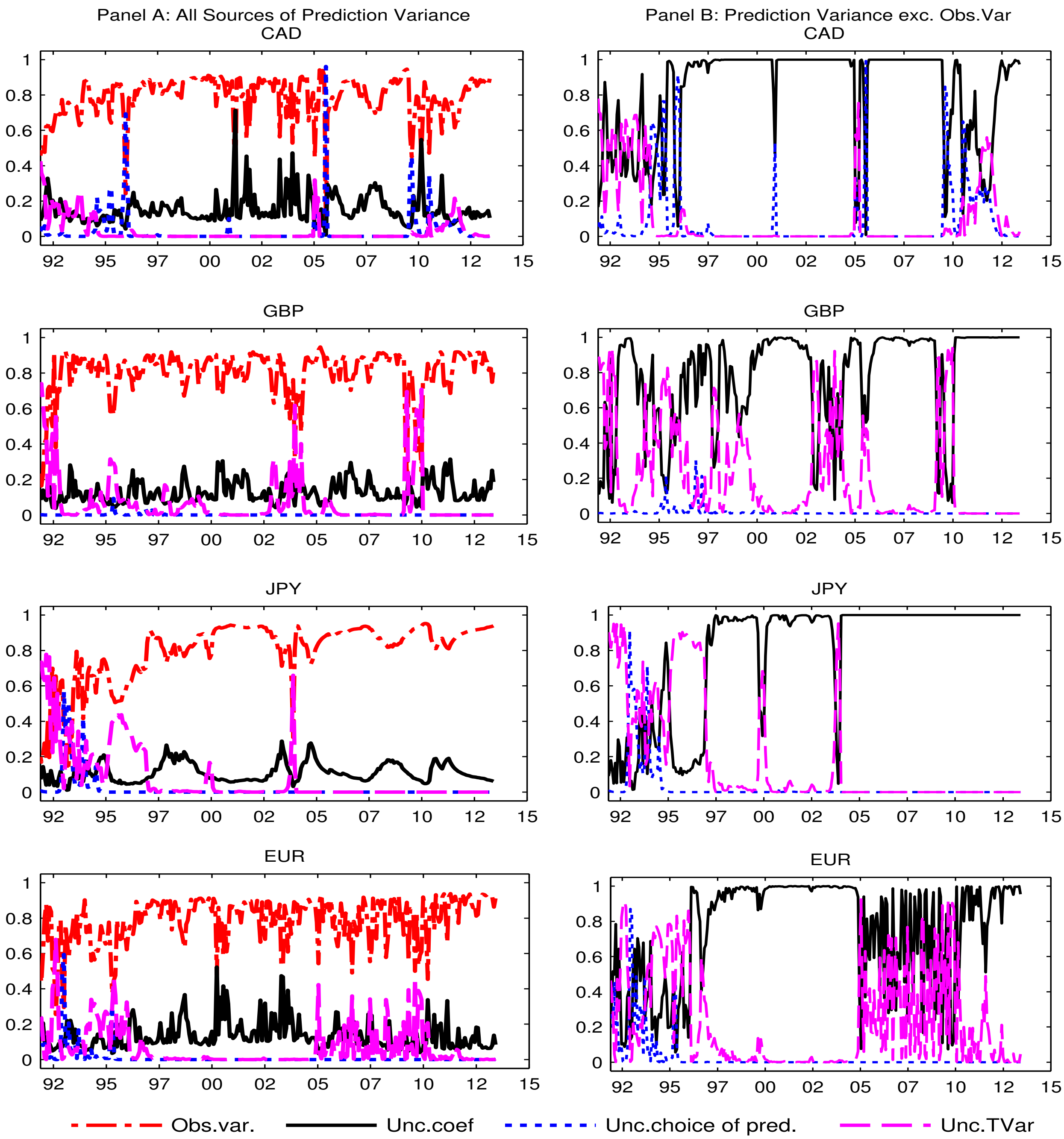
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure F.4.2: Sources of Prediction Variance at 3-Months Forecasting Horizon



Notes: As in Figure F.4.1, except that here the forecasting horizon is 3-months.

Figure F.4.3: Sources of Prediction Variance at 12-Months Forecasting Horizon



Notes: As in Figure F.4.1, except that here the forecasting horizon is 12-months.

F.5 Robustness to Change in Base Currency

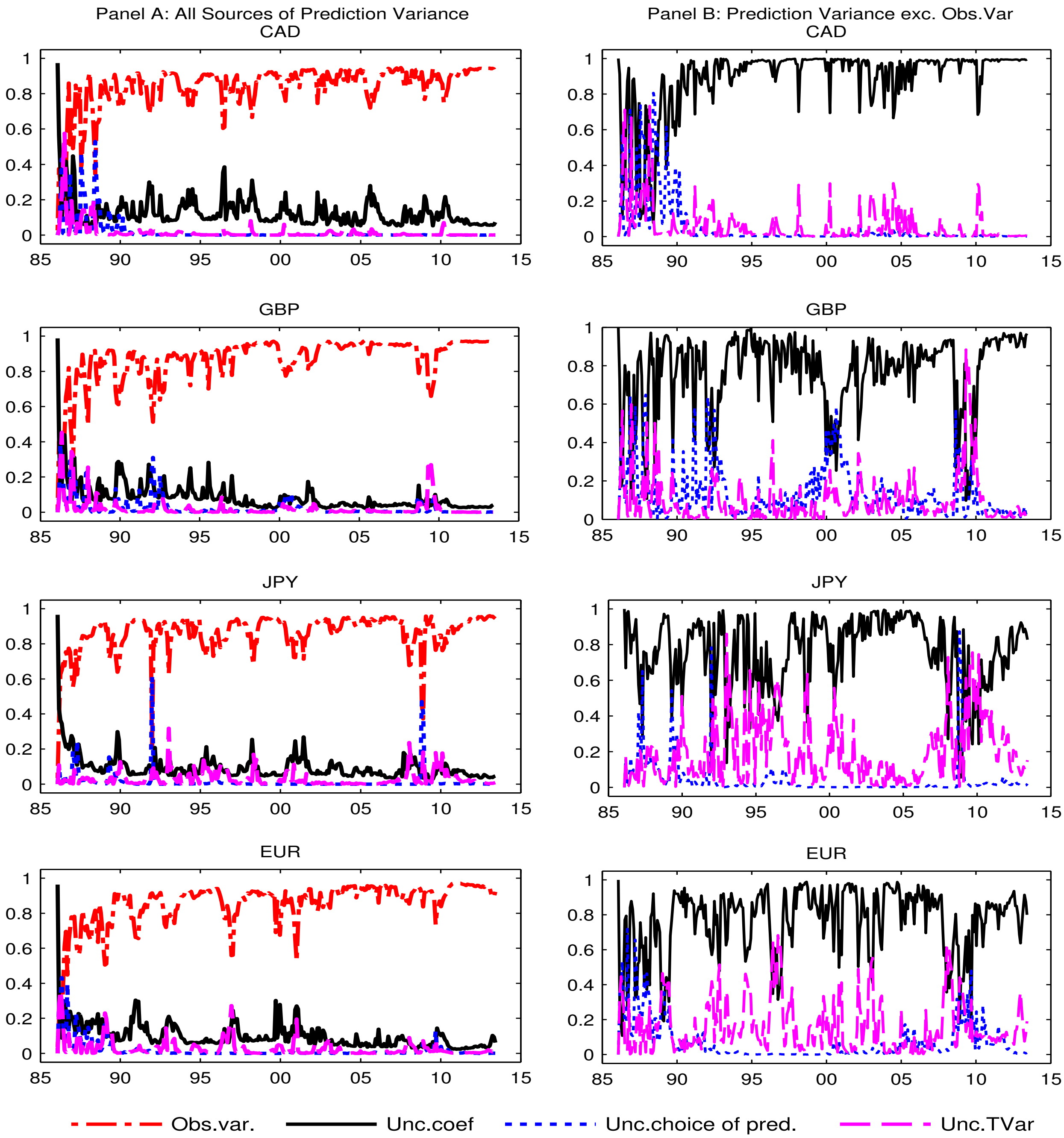
We also changed the base numeraire (home country) to the GBP (U.K.) following Chen *et al.* (2010). Accordingly, we redefined our predictors taking the U.K. as the home country. Forecasting results in Table F.5.1 remain mostly congruent to our early findings. The BMA incl. TVar-Coeffs approach is still advantageous when compared to the RW and the BMA excl. TVar-Coeffs especially at the 12-months horizon. The corresponding analysis of prediction variance for four representative currencies in Figures F.5.1-F.5.3, also reveals coherent findings: estimation uncertainty and uncertainty with regards to the exact degree of time-variation constitute the main obstacles to models forecasting performance.

Table F.5.1: Robustness to Change in Base Currency to the GBP

	BMA incl. TVar-Coeffs			BMA excl. TVar-Coeffs		
	h=1	h=3	h=12	h=1	h=3	h=12
AUD	1.038	**0.953***	**0.885****	1.021	**0.991**	**0.930****
CAD	1.038	1.044	**0.824****	1.039	1.055	**0.979**
GBP	1.039	**0.973**	**0.781****	1.017	**0.984**	**0.864***
JPY	1.038	1.009	**0.715*****	1.015	**0.984**	1.001
NOK	1.044	1.017	**0.960**	1.014	1.001	**0.994**
SEK	1.012	1.020	**0.913**	1.001	1.006	**0.969**
CHF	1.053	**0.919****	**0.556*****	1.023	**0.963***	**0.840****
EUR	1.025	**0.965**	**0.780***	1.011	1.018	**0.961**

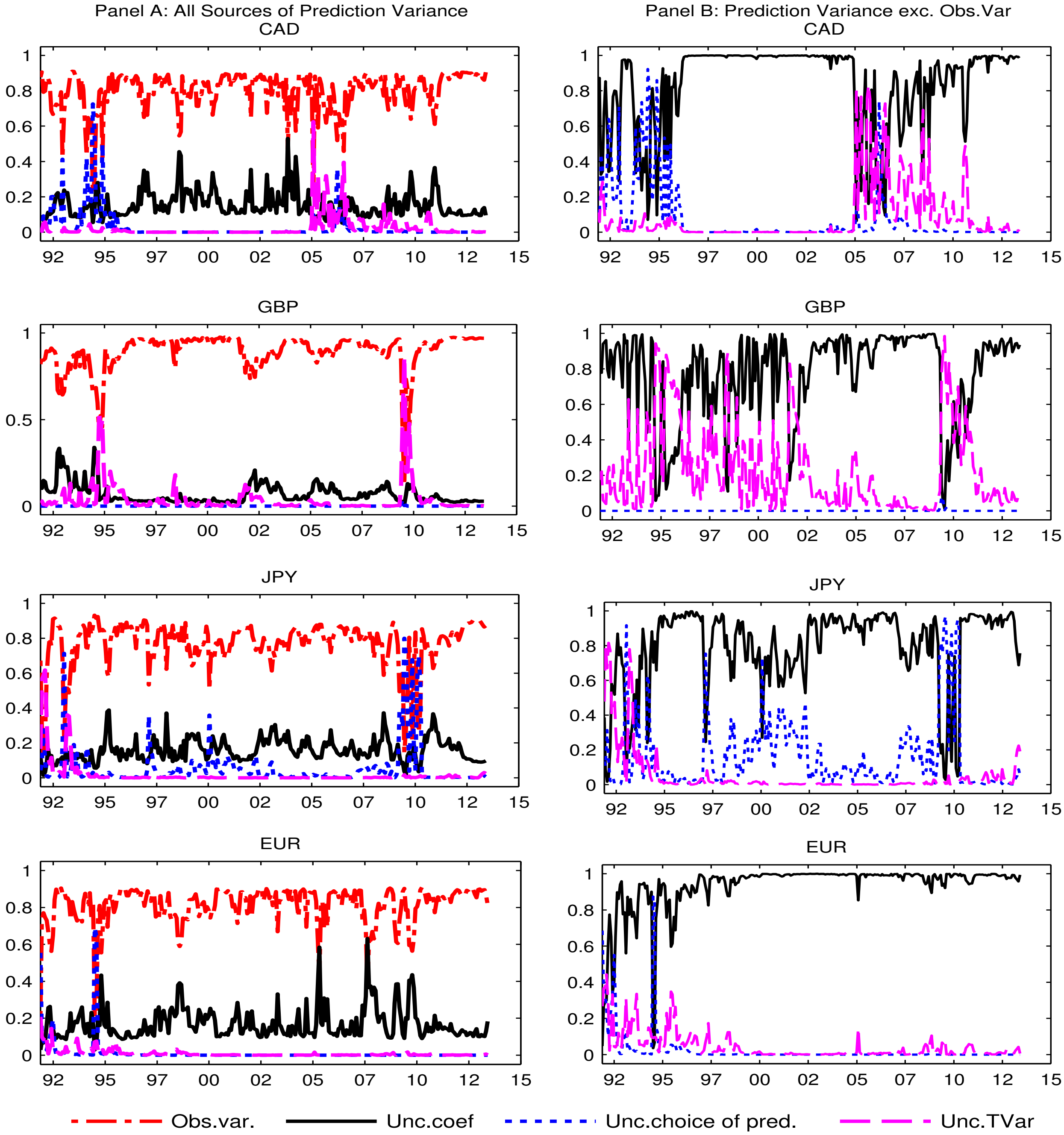
Notes: RMSFE of the BMA including or excluding time-varying coefficients, relative to the RMSFE of the driftless Random Walk (RW). Values below one indicate that the method under scrutiny generates a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) denote the level of significance of the one-sided DMW test based on standard critical values, for the null hypothesis of equality in the RMSFE. Currency codes denote: AUD - Australian dollar; CAD - Canadian dollar; GBP - Pound sterling; JPY - Japanese yen; NOK - Norwegian krone; SEK - Swedish krona; CHF - Swiss franc; EUR - euro. The forecast evaluation period begins in 1987M12+h in all, but the OLS-DMSPE Combination case (1992M12+h).

Figure F.5.1: Sources of Prediction Variance at 1-Month Forecasting Horizon



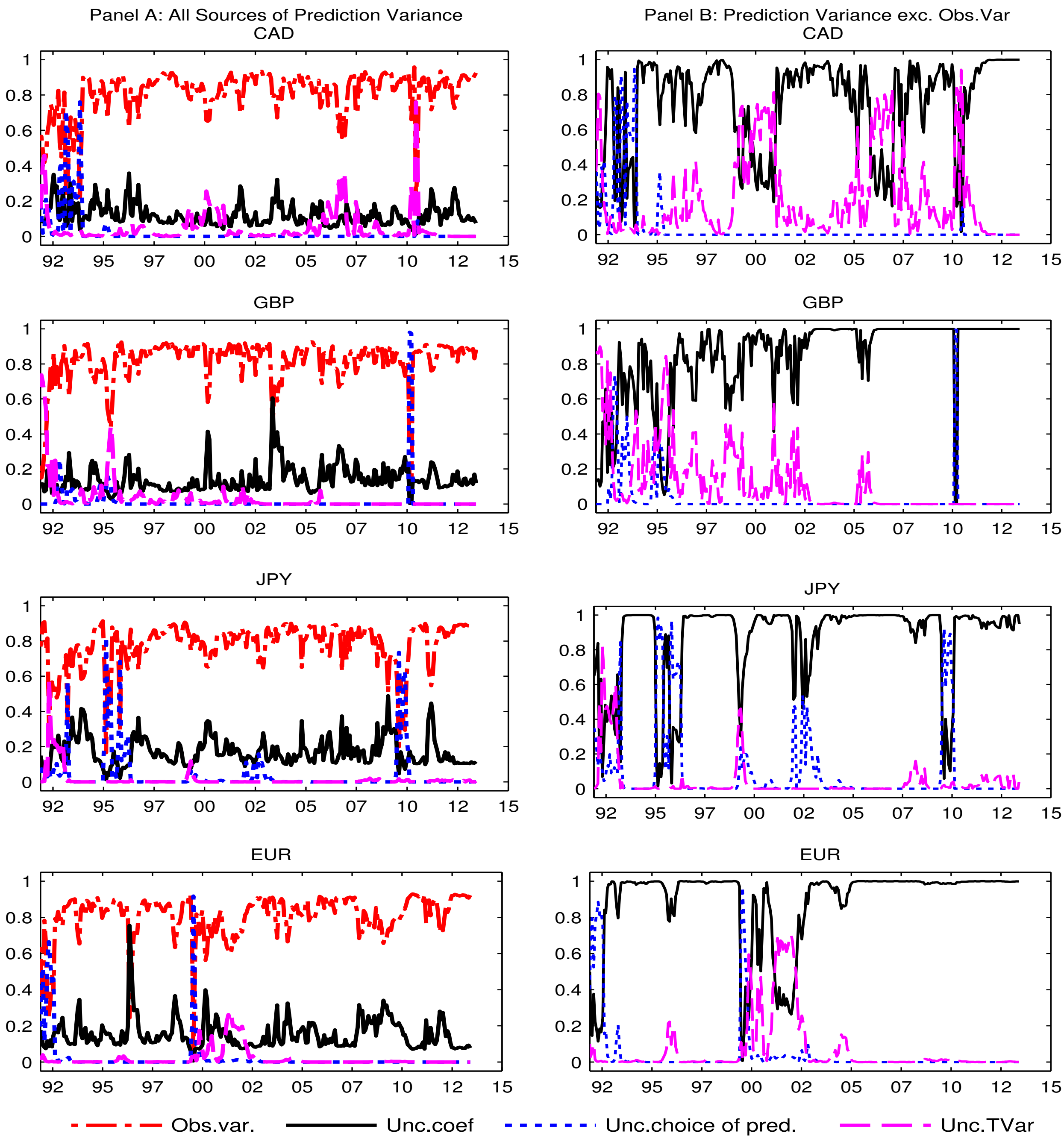
Notes: Decomposition of the prediction variance into its constituent parts at 1-month forecasting horizon. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance due to model uncertainty with respect to the choice of the predictor (Unc.choice of pred); and (iv) variance due to model uncertainty with respect to the choice of the degree of time-variation in coefficients (Unc.TVar). The Panel shows relative proportions of these variances. Panel B excludes the variance due to random fluctuations in the data (Obs.var.) and shows the relative weights of the remaining sources of prediction variance, and hence also sum to one.

Figure F.5.2: Sources of Prediction Variance at 3-Months Forecasting Horizon



Notes: As in Figure F.5.1, except that here the forecasting horizon is 3-months.

Figure F.5.3: Sources of Prediction Variance at 12-Months Forecasting Horizon



Notes: As in Figure F.5.1, except that here the forecasting horizon is 12-months.

G Forecasting and Averaging in a Bayesian Framework

This Appendix provides details on the methods used to forecast with the dynamic linear model defined in the methodological section of the main text, as well as the averaging approach. We draw mainly from expositions in Dangl and Halling (2012) and West and Harrison (1997, Ch. 3&4).

G.1 Bayesian Estimation of the Parameters of the Predictive Regression

For convenience we begin by transcribing the predictive regression from the methodological section in our main text:

$$\Delta e_{t+h} = X_t\theta_t + v_{t+h}, \quad v_{t+h} \sim N(0, V_t), \text{ (observation equation)}; \tag{G.1.1}$$

$$\theta_t = \theta_{t-1} + \varpi_t, \quad \varpi_t \sim N(0, W_t), \text{ (transition equation)}; \tag{G.1.2}$$

The essential components of the Bayesian approach we employ are the priors for V_t and θ_t, along with a method to estimate W_t; the joint or conditional posterior distribution of V_t and θ_t; and in the context of our predictive regression, the predictive density. Finally, we also require an updating scheme for the priors after observing the data.

The approach involves a full conjugate Bayesian analysis. The starting point is the natural conjugate $g-$prior specification set at $t = 0$:

$$V_0|D_0 \sim IG\left[\frac{1}{2}, \frac{1}{2}S_0\right], \tag{G.1.3}$$

$$\theta_0|D_0, V_0 \sim N[0, S_0(gX'X)^{-1}], \tag{G.1.4}$$

where

$$S_0 = \frac{1}{N-1}\Delta e'(I - X(X'X)^{-1}X')\Delta e, \tag{G.1.5}$$

and D_0 indicates the conditioning information at $t = 0$. In general, at any subsequent period, $D_t = [\Delta e_t, \Delta e_{t-1},..., \Delta e_{1,}, X_t, X_{t-1}..., X_1, \text{Priors}_{t=0}]$. That is, D_t contains the exchange rate variations, the predictors, and the prior parameters. At this t period, we can form a posterior belief about the unobservable coefficient $\theta_t|D_t$, and the variance of the observation equation error term (observational variance $V_t|D_t$). The use of a natural conjugate prior implies that the posterior distributions are from the same family as the priors. Specifically, the posteriors are also jointly normally-inverse gamma distributed:

$$V_t|D_t \sim IG\left[\frac{n_t}{2}, \frac{n_t S_t}{2}\right], \tag{G.1.6}$$

$$\theta_t|D_t, V_t \sim N(m_t, V_t C_t^*), \tag{G.1.7}$$

where S_t is the estimate of the observational variance, with n_t as the corresponding number of degrees-of-freedom; m_t is the estimate of coefficient vector (θ_t) conditional on D_t and V_t; and C_t^* corresponds to the conditional variance matrix of θ_t, normalized by the observational variance. Integrating out the distribution given by (G.1.7) with respect to V_t, yields a multivariate $t-$distribution for the coefficients' posterior:

$$\theta_t|D_t \sim T_{n_t}(m_t, S_t C_t^*). \tag{G.1.8}$$

When updating the coefficients vector conditional on information available up to the previous period, the preceding point estimates do not directly become the prior because according to the transition process, the coefficients are exposed to normally distributed random shocks, which widens the variance but maintains the mean. Hence:

$$\theta_t|D_{t-h} \sim T_{n_t}(m_t, S_t C_t^* + W_t). \tag{G.1.9}$$

The predictive density of the $h-$step-ahead change in the exchange rate is obtained by integrating the conditional density of Δe_{t+h} over the space spanned by θ_t and V_t. This yields:

$$f(\Delta e_{t+h}|D_t) = \mathbf{t}_{n_t}(\widehat{\Delta e}_{t+h}, Q_{t+h}), \tag{G.1.10}$$

where $\mathbf{t}(\widehat{\Delta e}_{t+h}, Q_{t+h})$ denotes a $t-$distribution density with n_t degrees-of-freedom, mean $\widehat{\Delta e}_{t+h}$, variance Q_{t+h}, evaluated at Δe_{t+h}. The mean of the predictive distribution is computed as:

$$\widehat{\Delta e}_{t+h} = X_t' m_t, \tag{G.1.11}$$

and the total unconditional variance of the same distribution is given by:

$$Q_{t+h} = X_t' R_t X_t + S_t, \tag{G.1.12}$$

with

$$R_t = S_t C_t^* + W_t, \tag{G.1.13}$$

where R_t is the unconditional variance of the coefficient vector θ_t at time t. The first term in equation (G.1.12) captures the variance arising from uncertainty in the estimation of the coefficient vector θ_t. The last term S_t denotes the estimate of the variance of the disturbance term of the observation equation.

The key components to update the priors on θ_t and V_t are described in equations (G.1.14)-(G.1.18) below. The first element, is the prediction error,

$$\varepsilon_{t+h} = \Delta e_{t+h} - \widehat{\Delta e}_{t+h}, \text{ (prediction error)}, \tag{G.1.14}$$

which is useful in the estimate of the observational variance:[4]

$$S_{t+h} = \kappa S_t + (1-\kappa)\varepsilon^2_{t+h}, \text{ (estimator of observational variance)}. \tag{G.1.15}$$

where κ $(0 < \kappa < 1)$ is a (decay) factor that governs the responsiveness of the estimator to the most recent data. Setting $\kappa = 1$ implies that all the observations receive the same weight in the estimate and, in fact, the estimate of the observational variance remains constant. Smaller values of κ induce more variability of the estimate and hence in the coefficients. We set $\kappa = 0.99$ following the study of Koop and Korobilis (2012).

An additional element that induces changes in the coefficients is the adaptive vector:

$$A_{t+h} = \frac{R_t X_t}{Q_{t+h}}, \text{ (adaptive vector)}. \tag{G.1.16}$$

It characterizes the degree to which the posterior of the coefficient vector θ_t changes to new observation. The numerator of equation (G.1.16) conveys the information content of the current observation, and the denominator measures the precision of the estimated coefficients. With the above elements, we can update the coefficients' point estimate m_t and the covariance matrix C^*_t:

$$m_{t+h} = m_t + A_{t+h}\varepsilon_{t+h}, \text{ (expected coeff. vector estimator)}, \tag{G.1.17}$$

$$C^*_{t+h} = \frac{1}{S_t}\left(R_t - A_{t+h}A'_{t+h}Q_{t+h}\right), \text{ (variance of the coeff. vector estimator)}. \tag{G.1.18}$$

The exposition so far does not include a method to estimate W_t. As described in the main text, to capture the relationship between the coefficients' estimation error and the variance, we let W_t be proportional to the estimation variance $S_t C^*_t$ of the coefficients $\theta_t | D_t$. Formally,

$$W_t = \frac{1-\delta}{\delta} S_t C^*_t, \quad \delta \in \{\delta_1, \delta_2, ..., \delta_d\}, \; 0 < \delta_j \leq 1. \tag{G.1.19}$$

Therefore, the variance of the predicted coefficient vector expressed in equation (G.1.13)

[4] This estimator is known as the Exponentially Weighted Moving Average (EWMA), frequently used to model stochastic volatility in financial aplications.

simplifies to:

$$R_t = S_t C_t^* + \frac{1-\delta}{\delta} S_t C_t^* = \frac{1}{\delta} S_t C_t^*. \tag{G.1.20}$$

This completes the requisites for forecasting with one model. The approach we pursue, however, allows for k candidate predictors and d possible support points for time-variation in coefficients, and therefore, $k.d$ models. We deal with these possibilities in a Bayesian model selection and averaging approach that we outline next.

G.2 Bayesian Dynamic Averaging Over Models and Forgetting Factors

Let M_i constitute a specific selection of a predictor from a set of k candidates, and δ_j a specific choice of degree of time-variation in coefficients from the space $\{\delta_1, \delta_2, ..., \delta_d\}$. The mean of the predictive distribution computed above (see equation (G.1.11)) is influenced by these specific choices. Hence, the point estimate of Δe_{t+h} also becomes conditional on M_i and δ_j:

$$\widehat{\Delta e}_{t+h,i}^{j} = E(\Delta e_{t+h}|M_i, \delta_j, D_t) = X_t' m_t|M_i, \delta_j, D_t. \tag{G.2.1}$$

The starting point in examining which model setting turns out to be important empirically, is to assign prior weights to each individual predictor M_i and each support point δ_j. We begin with a prior that allows each predictor and each support point to have the same chance of becoming probable. That is, for each M_i and δ_j we set uninformative priors:

$$\begin{aligned} P(M_i|\delta_j, D_0) &= 1/k, & \text{(G.2.2)} \\ P(\delta_j|D_0) &= 1/d. & \text{(G.2.3)} \end{aligned}$$

At time t, the posterior probabilities are updated using Bayes's rule. We first update the posterior probability of a certain model, given a value of δ_j:

$$P(M_i|\delta_j, D_t) = \frac{f(\Delta e_t|M_i, \delta_j, D_{t-h}) P(M_i|\delta_j, D_{t-h})}{f(\Delta e_t|\delta_j, D_{t-h})}, \tag{G.2.4}$$

where

$$f(\Delta e_t|\delta_j, D_{t-h}) = \sum_M f(\Delta e_t|M_i, \delta_j, D_{t-h}) P(M_i|\delta_j, D_{t-h}). \tag{G.2.5}$$

The key ingredient is the conditional density:

$$f(\Delta e_t|M_i, \delta_j, D_{t-h}) \sim \frac{1}{\sqrt{Q^j_{t,i}}}\mathbf{t}_{n_{t-1}}\left(\frac{\Delta e_t - \Delta e^j_{t,i}}{\sqrt{Q^j_{t,i}}}\right), \tag{G.2.6}$$

where $\mathbf{t}_{n_{t-1}}$ is the density of a Student$-t-$distribution and $\Delta e^j_{t,i}$ and $Q^j_{t,i}$ are the corresponding point estimates and variance of the predictive distribution of model M_i, given $\delta = \delta_j$ (refer to equation (G.1.10)). The prediction of the average model for each of the specific value of $\delta = \delta_j$ is given by:

$$\widehat{\Delta e}^j_{t+h} = \sum_{i=1}^{k} P(M_i|\delta_j, D_t)\widehat{\Delta e}^j_{t+h,i}. \tag{G.2.7}$$

Essentially, for each specific δ, it is the sum of the forecasts of each of the k models weighted by their posterior probability. If there were only one support point for time-variation in coefficients, such that $d = 1$, then equation (G.2.7) would complete the averaging approach. However, because we are considering several possibilities for δ, we also perform Bayesian averaging over these values.

Starting with the prior probability in equation (G.2.3), the posterior probability of a specific δ is:

$$P(\delta_j|D_t) = \frac{f(\Delta e_t|\delta_j, D_{t-h})P(\delta_j|D_{t-h})}{\sum_\delta f(\Delta e_t|\delta, D_{t-h})P(\delta|D_{t-h})}. \tag{G.2.8}$$

We note that using this probability, we can infer the degree of time-variation in coefficients supported by the data.

We can now find the total posterior probability of a model determined by a specific selection of predictor M_i and degree of coefficient variation δ_j,

$$P(M_i, \delta_j|D_t) = P(M_i|\delta_j, D_t)P(\delta_j|D_t), \tag{G.2.9}$$

and the unconditional average prediction of the average model,

$$\widehat{\Delta e}_{t+h} = \sum_{j=1}^{d} P(\delta_j|D_t)\widehat{\Delta e}^j_{t+h}. \tag{G.2.10}$$

Thus, $\widehat{\Delta e}_{t+h}$ is obtained by averaging over the average models' prediction, over degrees of time-variation in coefficients.

H Data Appendix

This Appendix describes the data used in our analyses. The sample period is 1979M1:2013M5, for nine countries - Table H.0.1. For each country in the first column, the Table indicates the source of information for each variable in the subsequent columns.

Table H.0.1: Data Sources

Country	Nominal exchange rate (USD per National currency)	Industrial prod. index, NSA, 2005=100	Money supply, NSA, National currency (10ˆ9)	Short-term nominal interest rate (%)	Consumer price index NSA, 2005=100
Australia	IFS,193..AE-ZF	IFS, 19366..CZF[a]	M1, OECD, MEI	IFS, 19360B..ZF	OECD, MEI[a]
Canada	IFS, 156..AE.ZF	IFS, 15666..CZF	M1, OECD, MEI	IFS, 15660B..ZF	IFS, 15664...ZF
Germany/Eur	IFS, 134..AE.ZF	IFS, 13466..CZF	M1; Bundesbank	IFS, 13460B..ZF	Bundesbank
Japan	IFS, 158..AE.ZF	IFS, 15866..CZF	M1, OECD, MEI	IFS, 15860B..ZF	IFS, 15864...ZF
Norway	IFS, 142..AE.ZF	IFS, 14266..CZF	M2, OECD, MEI	IFS, 14260...ZF	IFS, 14264...ZF
Sweden	IFS, 144..AE.ZF	OECD MEI	M3, OECD, MEI	IFS, 14460B..ZF	IFS, 14464...ZF
Switzerland	IFS, 146..AE.ZF	IFS, 14666..BZF	M1, OECD, MEI	IFS, 14660...ZF	IFS, 14664...ZF
UK	IFS, 112..AE.ZF	IFS, 11266..CZF	M4, Bank of England	IFS, 11260...ZF	IFS, 11264B..ZF
US		IFS, 11166..CZF	M1, FED	IFS, 11160B..ZF	IFS, 11164...ZF

Notes: The exchange rate is defined as the end-of-month value of the U.S. dollar (USD) price of a unit of national currency. IFS denotes International Financial Statistics as published by the IMF. OECD, MEI denotes the OECD's Main Economic Indicators database. NSA stands for non-seasonally adjusted and and the superscript ([a]) denotes monthly data obtained via quadratic-match-average interpolation method from quartely data.

I The Bootstrap

The bootstrap is primarily based on Kilian (1999) and Rogoff and Stavrakeva (2008); but tailored to account for possible data-mining as proposed by Inoue and Kilian (2005). We use a semi-parametric bootstrap with the data generating process (DGP) for the fundamentals specified in an error correction form. Following Rogoff and Stavrakeva (2008) and Byrne *et al.* (2016), we also assume cointegration between the exchange rate and fundamentals. For each country we postulate the following DGP under the null of no predictability:

$$\Delta e_{it} = v_{it}^{e}, \tag{I.0.1}$$

$$\Delta z_{it} = c_{i0} + t + \Upsilon_i z_{it-1} + \sum_{\ell=1}^{\ell e} B_{i\ell}^{e} \Delta e_{it-\ell} + \sum_{\ell=1}^{\ell z} B_{i\ell}^{z} \Delta z_{it-\ell} + v_{it}^{z}, \tag{I.0.2}$$

where $\Delta e_{it} = e_{it} - e_{it-1}$; $\Delta z_{it} = z_{it} - z_{it-1}$; c_{i0} is a constant, t is a trend, $i = 1, 2, ..., N$, denotes the country subscript, and v_{it}^{e} and v_{it}^{z} are *i.i.d* error terms. We first estimate equations (I.0.1) and (I.0.2) via OLS, with lag orders ℓe and ℓz selected using Akaike's Information Criterion (AIC). The AIC also allows us to determine the inclusion or exclusion of the constant, the trend or both.[5] Subsequently, to preserve the contemporaneous correlation and cross-sectional dependence from the original data, we re-sample with replacement the residuals matrix $([v_{1t}^{e}, v_{2t}^{e}, ..., v_{Nt}^{e}], [v_{1t}^{z}, v_{2t}^{z}, ..., v_{Nt}^{z}])$, in tandem, and in non-overlapping blocks consisting of 12 consecutive observations each. The choice of the block size is guided by Rogoff and Stavrakeva (2008), who also find that results are invariant to a row by row sampling. We have experimented with a block of size one and a block size equal to the length of forecasting horizon h, and results remain qualitatively similar.

In the next step, we use the re-sampled residuals to recursively generate pseudo-samples of e_{it} and z_{it}. The first 100 observations are discarded to avoid potential bias due to using the sample averages as initial values for the recursions. We then employ each of the predictive model (Single Predictor including or excluding TVar-Coeffs, Mean Combination, Median Combination, Trimmed Mean Combination, and DMSPE Combination) to forecast using the pseudo-samples, and calculate the DMW test statistic. We repeat this process 1000 times, providing us with an empirical distribution of the statistic. The p-value is the proportion of the bootstrap statistics that are above the test-statistic calculated using observed data.

The bootstrap procedure just described assumes that each predictor is analyzed in isolation, but our flexible Bayesian approach allows for a search over several potential predictors. To take into account concerns about data-mining we suitably adjust the procedure above following Inoue and Kilian (2005).[6] The adjustment involves assuming that under the null

[5] In equation (I.0.2) the sum of the coefficients of the lags of Δz_t is restricted to one to avoid exploding simulated pseudo data (Rogoff and Stavrakeva, 2008).

[6] See also Rapach and Wohar (2006) for an application to stock returns.

of no predictability the DGP comprises:

$$
\begin{aligned}
\Delta e_{it} &= v_{it}^{e}, && \text{(I.0.3)}\\
\Delta z_{1,it} &= c_{1,i0} + t + \Upsilon_{1i} z_{1,it-1} + \sum_{\ell=1}^{\ell e} B_{1,i\ell}^{e} \Delta e_{it-\ell} + \sum_{\ell=1}^{\ell z} B_{1,\ell}^{z} \Delta z_{1,it-\ell} + v_{1,it}^{z} && \text{(I.0.4)}\\
&\vdots \\
\Delta z_{k,it} &= c_{k,i0} + t + \Upsilon_{ki} z_{k,it-1} + \sum_{\ell=1}^{\ell e} B_{k,i\ell}^{e} \Delta e_{it-\ell} + \sum_{\ell=1}^{\ell z} B_{k,i\ell}^{z} \Delta z_{k,it-\ell} + v_{k,it}^{z}
\end{aligned}
$$

where $v_{it}^{e}, v_{1,it}^{z}$,..., $v_{k,it}^{z}$, are fitting errors. As the system of equations suggests, we are now considering $1, 2, ..., k$ candidate predictors for every i country. Each of the equation is also estimated via OLS. We again collect all the residuals in the same matrix in tandem, and follow the same steps as in the bootstrap procedure above, except that for each bootstrap we store the maximal DMW statistic over the k predictors. This provide us with an empirical distribution of the maximal statistic for each country and from which, after ordering, the 900^{th}, 950^{th}, and 990^{th} values constitute the 10%, 5% and 1% critical values, respectively. The forecasting performance of the BMA (BMS) including or excluding TVar-Coeffs are evaluated using critical values constructed in this manner.[7]

[7]Note that a similar procedure can also be applied to construct bootstrapped critical values and p-values for the Clark and West (2006, 2007) test statistic, and the Theil's U-statistic.

References

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

Chen, Y.-C., Rogoff, K. S., and Rossi, B. (2010). Can exchange rates forecast commodity prices? *The Quarterly Journal of Economics*, 125(3):1145–1194.

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

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

Della Corte, P., Sarno, L., and Sestieri, G. (2012). The predictive information content of external imbalances for exchange rate returns: How much is it worth? *Review of Economics and Statistics*, 94(1):100–115.

Fleming, J., Kirby, C., and Ostdiek, B. (2001). The economic value of volatility timing. *The Journal of Finance*, 56(1):329–352.

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

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

Giannone, D. (2010). *Comment on "Can Parameter Instability Explain the Meese-Rogoff Puzzle?"*, pages 180–190. NBER International Seminar on Macroeconomics 2009. University of Chicago Press.

Goetzmann, W., Ingersoll, J., Spiegel, M., and Welch, I. (2007). Portfolio performance manipulation and manipulation-proof performance measures. *Review of Financial Studies*, 20(5):1503–1546.

Han, Y. (2006). Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies*, 19(1):237–271.

Inoue, A. and Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23(4):371–402.

Kilian, L. (1999). Exchange rates and monetary fundamentals: What do we learn from long-horizon regressions? *Journal of Applied Econometrics*, 14(5):491–510.

Koop, G. and Korobilis, D. (2012). Forecasting inflation using dynamic model averaging. *International Economic Review*, 53(3):867–886.

Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.

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

Rapach, D. E. and Wohar, M. E. (2006). In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance*, 13(2):231 – 247.

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

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