On the Sources of Uncertainty in Exchange Rate Predictability*

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

September 26, 2014

Abstract

We analyse the role of time-variation in coefficients and other sources of uncertainty in exchange rate forecasting regressions. Our techniques incorporate the notion that the relevant set of predictors and their corresponding weights, change over time. We find that predictive models which allow for sudden, rather than smooth, changes in coefficients significantly beat the random walk benchmark in out-of-sample forecasting exercise. Using an innovative variance decomposition scheme, we identify uncertainty in coefficients' estimation and uncertainty about the precise degree of coefficients' variability, as the main factors hindering models' forecasting performance. The uncertainty regarding the choice of the predictor is small.

Keywords: Instabilities; Exchange Rate Forecasting; Time-Varying Parameter Models; Bayesian Model Averaging; Forecast Combination; Financial Condition Indexes; Bootstrap
JEL Classification: C53, C55, E44, F37.

*We benefited from discussions with Roberto Casarin, Gary Koop, Francesco Ravazzolo, Christian Schumacher, Shaun Vahey, and participants at the 2014 Glasgow SIRE Econometric Workshop on Modelling and Forecasting in Central Banks.

†Heriot-Watt University, School of Management and Languages, Edinburgh, EH14 4AS, UK.
‡University of Glasgow, Adam Smith Business School, Glasgow, G12 8QQ, UK.
§University of Glasgow, Adam Smith Business School, Glasgow, G12 8QQ, UK.
¶Corresponding author, E-mail address: p.ribeiro.1@research.gla.ac.uk, Tel: +44 (0)141 330 8506.

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 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, for instance by allowing for time-variation in the coefficients, 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 prior 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 study, for example, whether allowing for sudden changes in coefficients leads to a better forecasting performance, relative to situations where coefficients change gradually over time.

In light of the hypotheses advanced in recent papers, not only the coefficients in an exchange rate model are likely to change over time, but the relevant set of fundamentals 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.*, (2012), and Sarno and Valente (2009). Hence in our setting, in addition to allowing for varying degrees of coefficients adaptivity over time, we also entertain the possibility that, potentially, a different predictor may be relevant at each point in time. In this unified framework, we can examine whether models with a certain configuration, characterised by a specific degree of time-variation in coefficients and choice of predictor (fundamental), can forecast well.

Our key contribution in this paper goes entirely beyond establishing whether our model outperforms the typical random walk benchmark. As the evidence on time-varying forecasting performance suggests, the possibility that a model with a specific configuration forecasts 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 analyse the sources of model prediction uncertainty. In this regard, and inspired by Dangl and Halling (2012), we distinguish between (i) model uncertainty due to errors when estimating the coefficients, (ii) model uncertainty originating from time-variation in coefficients, (iii) model uncertainty due to a time-varying set of exogenous predictors, and (iv) model uncertainty due to random or unpredictable fluctuations in the data. Thus, we can 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 apply a Bayesian dynamic model selection and averaging approach of the sort considered in Dangl and Halling (2012) and Koop and Korobilis (2012), among others. The methodology permits to assign posterior probability weights to models 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.

Our predictive regressions employ information sets from Taylor rules and classic fundamentals. Engel and West (2005) use an exchange rate model based on Taylor (1993) rules as an example of models that can be cast within the present-value asset pricing framework. Molodtsova and Papell (2009) and Molodtsova *et al.* (2011) examine the out-of-sample predictive content of different Taylor rule specifications. They find evidence of predictability for most currencies and horizons they consider. Nevertheless, consistent with the hypothesis that the relevant set of predictors may change over time, their results also suggest that the evidence of predictability differs for different specifications across countries and periods. For instance, in Molodtsova and Papell (2009) the strongest support is from Taylor rule specifications with heterogeneous coefficients and interest rate smoothing. In contrast, in Molodtsova *et al.* (2011), the most successful Taylor rules impose equality in coefficients across countries, and do not incorporate interest rate smoothing. Molodtsova and Papell (2012) extend the analysis to incorporate readily available indicators of financial stress, and find evidence of superior forecasting performance of models augmented with these indicators.

In line with these results, our model space encompasses many different Taylor rule specifications, including several augmented with Financial Condition Indexes (FCIs). In our case however, FCIs for some countries in our sample are not readily available. We therefore construct Financial Condition Indexes (FCIs) using the Time-Varying Parameter Factor-Augmented VAR approach of Koop and Korobilis (2014). This approach to constructing FCI's is attractive since it facilitates greater flexibility in capturing turning points in financial conditions. In addition, the approach allows us to purge the effect of past and current output and inflation, such that the resulting FCIs incorporate additional information beyond that already included in the standard Taylor rule; see, Hatzius *et al.* (2010).

In terms of the empirical design, the dataset consists of monthly data spanning 1973M1 - 2013M5 on eight OECD countries' exchange rates relative to the US dollar. We use a direct method to forecast recursively, the period-ahead change in the exchange rate at one-, three-, and twelve-months horizons. The models are compared to the toughest benchmark – the driftless random walk (RW) (Rossi, 2013). We compute the ratio of the Root Mean Squared Forecast Error of the fundamentals-based model relative to that of the RW. 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 procedure proposed in Inoue and Kilian (2005) and implemented, for example, in Rapach and Wohar (2006). An additional measure of relative forecast accuracy is based on predictive likelihoods (see, e.g., Geweke and Amisano, 2010).

Apart from the research on the role of instabilities in model forecasting performance our paper is also related to the literature on forecast combinations. Among articles that study 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 first component, denoted predictive content, captures whether in-sample fit predicts out-of-sample forecasting performance. A second component provides the magnitude of model's in-sample over-fitting which does not translate into out-of-sample predictive power. And a last component captures the relevance of time-variation in forecasting performance. Their results point to a lack of predictive content and time-variation in forecasting performance as the main obstacles to 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 in the latter. Thus, our study complements theirs, as time-variation in parameters is an integral part of our analysis.[1]

Among papers focusing in pooling exchange rate forecasts, we note contributions by Wright (2008), Sarno and Valente (2009), Beckmann and Schuessler (2014), and Li *et al.* (2014). The main difference with our contribution is that the emphasis on these papers is on finding whether combined forecasts from several models with a

[1] 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.

certain configuration are superior to those from a single-variable approach and to the random walk benchmark. Instead, we focus in the same question and extend the analysis to examine the sources of model prediction uncertainty. An additional difference is our use of a data-mining robust bootstrap procedure when evaluating the forecasting performance of the models.[2,3]

To preview our results, we find that models which allow the relevant set of regressors to change over time and with varying degrees of coefficients adaptivity forecast well. These models significantly outperform the benchmark for most currencies at all, but one-month forecast horizon. In particular, at horizons greater than one month, predictive regressions with a high degree of time-variation in coefficients dominate regressions with constant and moderately time-varying coefficients. However, at the one-month forecast horizon our models do better for one quarter of the exchange rates considered. When examining what obstructs models' predictive ability over time, we identify uncertainty in the estimation of the coefficients and uncertainty regarding the correct level of time-variation in coefficients as the main sources of time-varying forecasting performance. When the models successfully embed these sources of uncertainty, they yield a satisfactory out-of-sample forecasting performance. Thus, our findings are 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 models' predictive ability.

The rest of the paper proceeds as follows. The next Section lays out the econometric methodology. Section 3 covers data description and forecast mechanics. 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 literature on 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 Ω_t a set of exchange rate fundamentals. Then, we consider

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

[3] There are also differences in how the predictors are constructed. In most of the mentioned papers, the predictors are constituted by each of the variable that defines macroeconomic models of exchange rate determination (e.g., money supply, inflation, interest rates, among others). In our setting, the predictors are the fundamentals that originate from the macroeconomic models of exchange rate determination (e.g., fundamentals from Taylor rules and the Monetary Model). A more subtle difference is while in our setting the random walk is excluded, in the majority of the related studies the random walk spans the model space.

[4] See, for example, Mark (1995), Cheung *et al.* (2005), Engel *et al.* (2008), Molodtsova and Papell (2009), and Rossi (2013).

predictive regressions of the following form:

$$\Delta e_{t+h} = X_t'\theta_t + v_{t+h}, \quad v_{t+h} \sim N(0, V), \text{ (observation equation)}; \tag{1}$$

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

where,

$$X_t = [1, z_t], \text{ and } \theta_t = [\theta_{0t}; \theta_{1t}]; \tag{3}$$

$$z_t \equiv \Omega_t - e_t. \tag{4}$$

As identity (4) 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 Subsection we discuss what spans our set of fundamentals contained in Ω_t. In this Section we note that the predictive regression given by the system of equations (1) and (2) allows the coefficient linked to the disequilibrium term z_t, and to the constant to change over time. In fact, as equation (2) suggests, we assume a random walk process for parameter θ_t, following Wolff (1987), Rossi (2006), Mumtaz and Sunder-Plassmann (2013), among others. We further assume that the disturbance terms, v_{t+h} and ϖ_t, are uncorrelated and normally distributed with mean zero and variance matrices V and W_t, respectively.[5]

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-parameter 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, Dangl and Halling (2012) suggest imposing some structure on W_t. We define this structure together with the description of the estimation methodology below.

We use Bayesian methods in the spirit of Dangl and Halling (2012) and Koop and Korobilis (2012) to estimate the parameters of the predictive regression. The methods described in these papers involve a full conjugate Bayesian analysis. That is, when prior information on the unknown parameters is combined with the likelihood function, results in 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 come from an inverse-gamma distribution. In a conjugate analysis, the posteriors

[5] Note that the variance of the disturbance term associated with the observation equation V, is constant but unknown. In addition, while we could model this variance as possibly time-varying, we focus in the constant case to isolate the dynamics of time-varying coefficients, from the dynamics of time-varying variance.

are jointly normally/inverse-gamma distributed. In Appendix A.1 we provide details on the updating scheme of the coefficients' vector and the observation equation variance 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., $D_t = [\Delta e_t, \Delta e_{t-h}, ..., X_t, X_{t-h}, ..., Priors_{t=0}]$. For the prior parameters at $t = 0$, we use a natural conjugate g-prior specification:

$$V|D_0 \sim IG\left[\frac{1}{2}, \frac{1}{2}S_0\right], \tag{5}$$

$$\theta_0|D_0, V \sim N\left[0, gS_0(X'X)^{-1}\right], \tag{6}$$

where,

$$S_0 = \frac{1}{N-1}\Delta e'(I - X(X'X)^{-1}X')\Delta e. \tag{7}$$

The prior for the coefficient vector in equation (6) is a diffuse prior centered around the null-hypothesis of no predictability, with g as the scaling factor that determines 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 (Dangl and Halling, 2012). This is consistent with our objective of examining which instabilities are supported by the data. In the empirical results in Section 4, we use a g-prior derived from the entire sample, following Wright (2008) and Dangl and Halling (2012). We also set $g = 10$ for the main results and examine cases of $g = [1, 50,100]$, but find similarities in the results, hence we do not report results based on the other values of g.

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 and V. West and Harrison (1997) 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} (for details, see Appendix A.1):

$$f(\Delta e_{t+h}|D_t) = \mathbf{t}_{n_t}(\Delta e_{t+h}; \widehat{\Delta e}_{t+h}, Q_{t+h}). \tag{8}$$

Using this predictive distribution we can recursively forecast Δe_{t+h}.

Recall that the degree of time-variation in the regressions coefficient is determined by the matrix W_t. Since 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. 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 let W_t be proportional to the estimation variance of the coefficients at time t, following

West and Harrison (1997) and Dangl and Halling (2012):[6]

$$W_t = \frac{1-\delta}{\delta} S_t C_t^*, \quad 0 < \delta \leq 1; \tag{9}$$

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-1}, and δ is a discount factor that controls the degree of time-variation in coefficients.

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 50% within 20 months. Reducing δ to 0.96, translates into 50% increase in 10 months, suggesting very abrupt changes in coefficients. Thus, in our empirical work we consider $\delta = [0.96, 0.97, 0.98, 0.99, 1.00]$ as the possible support points for time-variation in coefficients. We then examine empirically which support point is consistent with the data in a Bayesian model averaging approach, which we discuss in the next section.

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 points out that the set of relevant predictors appears to change over time (see, e.g., 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. Thus, 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 $d.k$.

The range of predictors we consider follows the recent literature that exploits the information content from Taylor (1993) rules.[7] See for example, Engel and West (2005), Engel *et al.* (2008), Mark (2009), Molodtsova *et al.* (2011), Molodtsova and Papell (2009, 2012) and Rossi (2013). The premise is that the home and the foreign central banks conduct monetary policy following the Taylor rule. In line with this rule, the foreign monetary authority, taken as the United States in our empirical section, is concerned with inflation and output deviations from their target values. In addition to these targets and consistent with historical evidence, Engel

[6] See also Raftery et al. (2010), Koop and Korobilis (2012) for a similar approach in modelling the variance of the transition equation.

[7] The Taylor (1993) rule postulates that monetary authorities should set the target for the policy interest rate considering the recent inflation path, inflation deviation from its target, output deviation from its potential level, and the equilibrium real interest rate. Then, it follows that they increase the short-term interest rate when inflation is above the target and/or output is above its potential level. Note that the Taylor principle presupposes an increase in the nominal policy rate more than the rise in inflation rate to stabilize the economy.

and West (2005) assume that the home country also targets the real exchange rate. It is also common, following Clarida *et al.* (1998), to consider that central banks adjust the actual interest rate to eliminate a fraction of the gap between the current interest rate target and its recent past level, known as interest rate smoothing. By subtracting the foreign Taylor rule from the home, the following interest rate differential equation is obtained,

$$\begin{aligned} i_t - i_t^* = \phi_0 + \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* \\ + \phi_3 q_t + \phi_4 i_{t-1} - \phi_4^* i_{t-1}^* + \mu_t, \end{aligned} \tag{10}$$

and from which, using the Uncovered Interest Rate Parity relationship (UIRP), we compute the fundamentals Ω_t, as:

$$\begin{aligned} \Omega_t \equiv \widehat{\phi}_0 + \widehat{\phi}_1 \pi_t - \widehat{\phi}_1^* \pi_t^* + \widehat{\phi}_2 \overline{y}_t - \widehat{\phi}_2^* \overline{y}_t^* \\ + \widehat{\phi}_3 q_t + \widehat{\phi}_4 i_{t-1} - \widehat{\phi}_4^* i_{t-1}^* + e_t, \end{aligned} \tag{11}$$

where, i_t is the short-term nominal interest rate set by the central bank, asterisks indicate foreign (United States) variables; π_t, is inflation; $\overline{y}_t$, denotes the output gap; q_t is the real exchange rate defined as $q_t = e_t + p_t^* - p_t$; p_t, is the log of the price level; ϕ_l for $l = 1, ..., 4$, are regression coefficients, and μ_t is the unexpected disturbance term, which is assumed to be Gaussian. Note that in equation (11), e_t is the log of the exchange rate.

The form of the home and foreign Taylor rules underlying equation (10) is very general. In practice, various specifications can be considered based on a number of assumptions. While the assumptions and the exact form of the rules is provided in Appendix B, here we note that some Taylor rule specifications are augmented with indicators of financial conditions (see also Molodtsova and Papell, 2012). This accord with the recent suggestions that apart from inflation and output gap, central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook (Taylor, 2008; Mishkin, 2010).

To construct the measures of financial conditions, or more precisely Financial Condition Indexes (FCIs), we use the Time-varying Parameter Factor-Augmented Vector Autoregressive model (TVP-FAVAR) of the sort considered in Koop and Korobilis (2014). The details are presented in Appendix B.2. The TVP-FAVAR allows for all the coefficients and the weight attached to the FCI to change over time, providing a framework that can potentially better characterise turning points in financial conditions. Additionally, and in line with Hatzius *et al.* (2010), the approach purges from the FCI the effect of macroeconomic influences, such that the resulting FCI provides extra information beyond that already contained in the original Taylor rule.

In total we consider $k = 23$ potential predictors:

- Fundamentals from 20 Taylor rules specifications, i.e., $TR1, ..., TR20$, each corresponding to a variant of equation (11);[8]

- Fundamentals from the monetary model (MM), $\Omega_{t,MM} \equiv (m_t - m_t^*) - (y_t - y_t^*)$, (where m_t is the log of money supply and y_t is the log of income);[9]
- Fundamentals from the Purchasing Power Parity (PPP) condition, $\Omega_{t,PPP} \equiv (p_t - p_t^*)$; and
- Fundamentals from the Uncovered Interest Rate Parity (UIRP) condition $\Omega_{t,UIRP} \equiv (i_t - i_t^*) + e_t$.

Selecting one specific model characterised by a certain 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 parameters. Then based on the realised likelihood of the model's prediction, the posterior probability of each of the $d.k$ models is updated according to Bayes rule. In Appendix A.2 we provide details on the exact formulae, following Dangl and Halling (2012). Note, however, that we elicit diffuse conditional prior probability for each predictor M_i, and equally, an uninformative prior for the range of support points for the degree of time-variation in coefficients. That is, 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 $k.d$ models. That is, we perform Bayesian Model Averaging (BMA) in a setting with time-varying 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 use 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).[10] 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.

[8] This set of potential specifications encompass the majority of those included in papers employing Taylor rule fundamentals.

[9] Note that we have assumed an income elasticity of one in the monetary model, following Mark (1995) and Engel and West (2005).

[10] In fact, as we show in the empirical section, we can analyze several other cases, depending on specific choices of the model and degree of time-variation, including cases of BMA over time-varying coefficients with single predictors.

2.3 Variance Decomposition and Sources of instability

The variance decomposition method we employ follows directly from the law of total variance, as implemented in Dangl and Halling (2012). That is, we decompose the variance of the random variable Δe, into its constituent parts. Starting with the decomposition with respect to different values of δ, we have,

$$Var(\Delta e) = E_\delta(Var(\Delta e|\delta)) + Var_\delta(E(\Delta e|\delta)), \tag{12}$$

where, E_δ and Var_δ indicate the expected value and the variance with regards to δ. Since the expected value of the variance of δ, is conditional on specific choice of model M, it can be further decomposed as follows:

$$Var(\Delta e|\delta) = E_M(Var(\Delta e|M,\delta)) + Var_M(E(\Delta e|M,\delta)). \tag{13}$$

Substituting back equation (13) for the expression of the expected value of the variance of δ in equation (12), and using the expressions of these variances detailed in appendixes A.1 and A.2, we obtain:

$$\begin{aligned} Var(\Delta e_{t+h}) &= \sum_j \left[\sum_i (S_t|M_i,\delta_j,D_t)P(M_i|\delta_j,D_t)\right] P(\delta_j|D_t) \\ &+ \sum_j \left[\sum_i (X_t' R_t X_t|M_i,\delta_j,D_t)P(M_i|\delta_j,D_t)\right] P(\delta_j|D_t) \\ &+ \sum_j \left[\sum_i (\widehat{\Delta e}^j_{t+h,i} - \widehat{\Delta e}^j_{t+h})^2 P(M_i|\delta_j,D_t)\right] P(\delta_j|D_t) \\ &+ \sum_j (\widehat{\Delta e}^j_{t+h} - \widehat{\Delta e}_{t+h})^2 P(\delta_j|D_t). \end{aligned} \tag{14}$$

The four individual terms in equation (14) 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, 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 characterises model uncertainty with respect to the choice of the predictor. The last term also characterises model uncertainty, but with respect to time-variability of the coefficients. Hence, both, the third and fourth terms capture model uncertainty.

3 Data, Forecast Mechanics, and Evaluation Methods

3.1 Data

We use monthly data spanning 1973M1:2013M5, for eight OECD countries: Canada, Denmark, Japan, Korea, Norway, Sweden, Switzerland and the United Kingdom. The foreign country is taken as the United States. The main data source is the IMF's International Financial Statistics (IFS), supplemented by national central banks. Exchange rates are end-of-month values of the national currencies, relative to the US dollar. The money supply is measured by the aggregate M1.[11]

To estimate Taylor rules we need the short-run central bank nominal interest rate, the inflation rate, and the output gap. We use the central bank's policy rate when available for the entire sample period; alternatively the discount rate or the money market rate. The price level consists of the consumer price index (CPI) and the inflation rate is defined as the (log) CPI monthly change. The proxy for the output is monthly industrial production (IP). 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. However, to correct for the uncertainty about these estimates at our recursive sample end-points, we follow Watson's (2007) method. In this regard, we estimate 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 models are then used to forecast and backcast three years of monthly data-points of IP, and the HP filter is applied to the resulting extended series.[12] The data on money supply, IP, and CPI were seasonally adjusted by taking the mean over twelve months following Engel *et al.* (2012).[13]

To construct the FCIs we select the most common variables used in this literature (see, e.g., Hatzius *et al.*, 2010). Although the specific set considered for each country in our sample differ, in general we include measures of stock market performance, long-term and short-term interest rate spreads, exchange rate indexes, house price indexes (when available), and survey indicators of financial conditions. These data were obtained from the IMF's IFS, OECD Main Economic Indicators, Datastream and country-specific central banks. Refer to Appendix C for extra details on country-specific variables, sources and data transformation.

3.2 Forecast Mechanics and Evaluation Methods

After preliminary data transformations our effective sample runs from 1974M1 to 2013M5. We use the period from 1974M1 to 1978M12 to initialise the recursions in

[11] In cases where the M1 aggregate is unavailable, we use a broader aggregate. This is M3 for Sweden and Belgium; and M4 for the UK. For extra details on Data, see Appendix C.

[12] 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. Note that we use the standard HP smoothing parameter for monthly data frequency (i.e., 14400).

[13] Data limitations prevent us from using real-time data for the countries we consider.

the estimation of Taylor rule fundamentals, as well as in the forecasting regression that we consider in the next section when examining other competing models. Thus, our forecasting window begins in 1978M12+h for most models.[14]

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. According to Wright (2008) both methods yield a similar forecasting performance. The forecasting exercise is based on a recursive approach using data available up to the time the forecast is made. Thus, the output gap, the FCIs, the Taylor fundamentals are also computed without using future information. For example, a three-month ahead forecast of the change in exchange rate for 1995M1 is made using data available up to 1994M10.

The forecasts of our models are compared to those of the driftless random walk (RW). According to Rossi (2013) this is the toughest benchmark. We compute the ratio of the Root Mean Squared Forecast Error (RMSFE) of the fundamentals-based exchange rate model relative to RMSFE of the RW, known as the Theil's U-statistic. Hence, models that perform better than the RW benchmark have a value of Theil's U less than one.

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, 2012). 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 (fundamental-based) 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 minimum mean squared forecast error. Additionally, Rogoff and Stavrakeva (2008) recall that the asymptotics of the CW test are well-defined when forecasting in a rolling, rather than recursive framework. Thus, we construct bootstrapped p-values of the DMW test-statistic in the spirit of Kilian (1999) and Rogoff and Stavrakeva (2008). In light of our search over several predictors however, we employ the bootstrap in a context of a data-mining environment as proposed by Inoue and Kilian (2005) and applied, for instance, in Rapach and Wohar (2006). See Appendix D for full details on the procedure. Bootstrapping also accounts for the fact that for $h > 1$, the forecast errors are likely to be serially correlated.[15,16]

[14] The exception occurs for one of the competing forecast combination method we consider in the next Section, which is based on a discount mean squared prediction error, and therefore requires a holdout out-of-sample period.

[15] While to account for this serial correlation a common practice is to use Newey and West (1987) Heteroskedasticity and Serial Correlation (HAC) robust estimator, Nelson and Kim (1993) point out that there is a strong tendency for the resulting test-statistic to increase with the forecast horizon. Therefore, inference based on bootstrapping is preferred relative to standard asymptotic distribution when testing the null of no predictability. Note than in our bootstrap we use HAC standard errors, with a lag truncation parameter of $int\{Sample^{0.25}\}$, following Rossi (2013).

[16] The Diebold and Mariano (1995) and West (1996) test is computed as: $DMW = \overline{f}\sqrt{P}/$[sample

4 Empirical Results

We begin by examining the out-of-sample forecasting performance of the model that allows predictors and coefficients to change over time, as well as restricted versions of it. Specifically, we report results from the following fundamental-based predictive models:

- BMA including Time-varying Coefficients (BMA incl. TVar-Coeff.): This model constitutes the Bayesian model average over all individual models and with varying degrees of coefficient evolution.

- BMA excluding Time-varying Coefficients (BMA excl. TVar-Coeff.): This is a restricted version of the above, as it represents the Bayesian model average over all individual models, but it excludes time-variation in coefficients. This corresponds to conventional Bayesian averaging as implemented, for example, in Wright (2008).

- BMS including Time-varying Coefficients (BMS incl. TVar-Coeff.): This model 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-Coeff.): This specification is nested in the BMS incl. TVar-Coeff 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 with most posterior probability (Single Predictor incl. TVar-Coeffs.): These models consider only a single predictor at a time, and the degree of time-variation in coefficient with the highest posterior probability among the range of all degrees of time-variation considered.

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

While the above models are based on Bayesian methods, we also consider combination forecast methods based on frequentist approaches. In this case the combination forecast of Δe_{t+h} made at time t, is a weighted average of the k individual

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, and $\widehat{fe}_{2,t+h}$ the corresponding forecast error of the FM. Note that $\overline{f}$ is the mean of $\widehat{f}_{t+h}$.

models' forecast based on OLS estimate of equation (1), excluding time-varying coefficients:

$$\widehat{\Delta e}^{c}_{t+h} = \sum_{i=1}^{k} \omega_{i,t} \widehat{\Delta e}^{i}_{t+h}, \tag{15}$$

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 forecasts combination methods:

- Mean Combination: The combined forecasts are obtained by setting $\omega_{i,t} = 1/k$, for $i = 1, ..., k$ in equation (15). Thus, the weights are constant over time.
- Median Combination: The median combination forecasts is the median of $\{\widehat{\Delta e}^{i}_{t+h}\}_{i=1}^{k}$.
- 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 (15). As in the median combination and the DMSPE combination method below, the weights change over time.
- DMSPE Combination: In this method, the combining weights are related to the historical forecasting performance of the individual models in the 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 $\Phi_{i,t}^{-1} = \sum_{so}^{t-1} \vartheta^{t-1-so} (\Delta e_{so+h} - \widehat{\Delta e}^{i}_{so+h})^2$ 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.[17] 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.

The motivation to consider the Bayesian restricted versions of the model that allows predictors and coefficients to change over time is to differentiate the influence of repeated updating of dynamic coefficients within a model, from the influence of the BMA that changes weights between models depending on past performance (Dangl and Halling, 2012). In contrast to Bayesian model averaging, Bayesian model selection fix a specific choice of predictors and degree of time-variation in coefficients. In addition, examining combined forecasts from models which exclude time-variation in coefficients, allows us to further check the sources of differences in

[17] 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.

Table 1: Forecast Evaluation: Multiple Predictor Models

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

	BMA incl. TVar-Coeff.			BMS incl. TVar-Coeff.		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.009	0.967*	0.843*	1.013	0.970	0.842*
Denmark	1.000	0.959**	0.912*	1.001	0.952**	0.907*
UK	0.996	0.960*	0.894*	0.991**	0.962*	0.886*
Japan	1.005	0.968*	0.859**	1.003	0.966	0.839**
Korea	0.995**	0.978	0.932	0.995**	0.976*	0.927
Norway	1.008	0.985	0.924	1.006	0.988	0.920
Sweden	1.014	0.960*	0.861**	1.012	0.964*	0.858**
Switzerland	1.004	0.992	0.985	1.001	0.990	0.989

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

	BMA excl. TVar-Coeff.			BMS excl. TVar-Coeff.		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.000	0.997	1.035	1.000	0.996	1.035
Denmark	0.998	0.975**	0.977	1.000	0.978*	0.979
UK	0.991***	1.004	0.962	0.990***	1.005	0.960
Japan	1.005	0.988	0.985	1.005	0.988	0.983
Korea	0.997**	0.991	0.936	0.997**	0.992	0.939
Norway	1.007	1.000	1.046	1.007	0.999	1.048
Sweden	1.003	0.987	0.925	1.004	0.987	0.925
Switzerland	1.003	1.014	1.041	1.003	1.012	1.045

Panel C: Combined forecasts, excl. TVar-Coeffs.

	OLS-Mean Combination			OLS-Median Combination		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.002	0.999	0.991	1.002	1.000	1.000
Denmark	1.003	1.001	0.931	1.005	1.005	0.935*
UK	1.005	1.007	1.000	1.005	1.009	1.003
Japan	1.003	1.000	0.974	1.005	1.003	0.990
Korea	0.998**	0.992	0.995	1.000*	0.996	1.003
Norway	1.008	1.013	1.053	1.007	1.011	1.051
Sweden	1.006	1.010	1.022	1.006	1.011	1.029
Switzerland	1.003	1.001	0.957	1.001	0.999	0.953
	OLS-Trimmed Mean			OLS-DMSPE Combination		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.001	0.999	0.993	1.002	0.999	0.979
Denmark	1.003	1.001	0.931*	0.999*	0.991	0.920
UK	1.005	1.009	1.005	1.003	1.003	0.988
Japan	1.004	1.002	0.980	1.005	1.004	0.979
Korea	0.998**	0.993	1.003	0.998*	0.990	0.958
Norway	1.008	1.013	1.056	1.003	1.002	1.004
Sweden	1.005	1.010	1.027	1.000	0.992	0.956
Switzerland	1.003	1.002	0.961	1.001	0.996	0.974

Notes: RMSFE of the fundamentals-based models (FM) - detailed at the beginning of this Section, relative to the RMSFE of the driftless Random Walk (RW). Values less than one indicate that the FM generates a lower RMSFE than RW. The Table also reports the DMW test-statistic, with p-values based on a data-mining robust semi-parametric bootstrap. Thus, asterisks (*10%, **5%, ***1%) denote the level of significance at which the null hypothesis of equal RMSFE is rejected, favouring the alternative that the FM has a lower RMSFE. The forecast evaluation period begins in 1978M12+h in all, but the OLS-DMSPE Combination case (1983M12+h).

forecasting performance and better understand the importance of time-variation in coefficients.

After establishing the main results in terms of the forecasting performance of the competing models, we then proceed and study in detail the characteristics of the BMA including time-varying coefficients model. In this respect, we analyse the sources of prediction uncertainty, the degree of time-variation in coefficients consistent with the empirical findings, and which macroeconomic fundamentals exhibit higher posterior probabilities when predicting exchange rates.

4.1 Out-of-Sample Forecast Evaluation

Table 1 summarises in three panels the results from the predictive regressions that allow predictors and coefficients to change over time, and all the restricted versions that take into account multiple predictors. Comparing the panels, for most countries the regressions that allow predictors and coefficients to change over time significantly outperform the RW benchmark at all forecast horizons, except at h=1 month. In fact, as the forecast horizon increases over one month, there is a notable reduction in the RMSFE. For example, in the case of the BMA including time-varying coefficients model in Panel A, the improvement in terms of the reduction in the RMSFE is of at least 0.8% and maximum of 4.1% at h=3, and a minimum of 1.5% and maximum of 15.7% at h=12. However, at h=1 these regressions still generate a lower RMSFE than RW for two out of eight exchange rates; but in this case the performance is similar to the models that allow only predictors to change over time, excluding time-variation in coefficients, in Panel B. Thus, at h=1 the gains from allowing predictors and coefficients to change over time are modest relative to models that only allow predictors to change. At the same time, these gains are small but greater than the best performing model within the class of models that combine forecasts from all predictors excluding time-varying coefficients, the OLS-DMSPE combination in Panel C.[18]

Results in Table 1 also reveal that allowing only predictors to change over time yields a smaller RMSFE than RW for over half of the currencies and horizons greater than one month. This is the case for the BMA and BMS models excluding time-varying coefficients in Panel B. However, the differences in RMSFEs are usually not significant at the levels of significance we compute using the data-mining robust semi-parametric bootstrap. Additionally, the reduction in the RMSFE is typically smaller than obtained with the more flexible models in Panel A. For instance, the reduction in RMSFE never exceeds 2.5% at h=3 and 7.5% at h=12. Thus for these models, our results are consistent with the findings of Wright (2008).[19]

[18] We also experimented other forecast horizons over one month (i.e., h=6, 24, and 36 months) and found that models with time-varying coefficients still do better than the RW for most exchange rates. However, we do not report these results to save space.

[19] Wright (2008) finds that in a setting of tighter priors and shrinkage towards the null of no predictability, the BMA model excl. 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, the

Table 2: Forecast Evaluation: Single Predictor and BMS incl. TVar-Coeffs

Predictor	Can.	Den.	U.K.	Jap.	Kor.	Nor.	Swe.	Swi.
					h = 1			
TR1	1.016	1.009	1.005	1.002	0.998**	1.014	1.002	1.003
TR2	1.004	1.011	1.006	1.002	1.006*	1.013	0.997**	1.000
TR3	1.015	1.000	1.002	1.001	0.993***	1.010	1.002	1.003
TR4	1.003	1.000*	1.011	1.008	1.001*	1.002	1.007	1.011
TR5	1.001	0.995**	1.013	1.001	1.007	1.001	1.012	0.999**
TR6	1.012	1.010	1.014	1.002	0.999**	1.003	1.014	1.020
TR7	1.008	1.010	1.015	1.009	1.001*	1.018	1.006	1.007
TR8	1.000	1.006	1.010	1.000*	0.998**	1.014	0.999**	0.999*
TR9	1.001	1.010	1.015	1.009	1.000*	1.003	0.998**	1.021
TR10	1.007	1.002	1.004	1.009	1.001*	1.023	1.009	1.001
TR11	1.014	1.000*	1.008	1.001	1.003	1.002	1.015	1.002
TR12	1.008	1.013	1.010	1.003	1.000**	1.018	1.009	1.021
TR13	1.000*	1.007	1.000	1.007	1.004	1.002	1.005	1.018
TR14	1.001	1.009	1.015	1.003	1.000*	1.018	1.007	1.019
TR15	1.000*	1.009	1.014	1.002	0.999**	1.016	1.008	1.019
TR16	1.015	1.011	1.014	1.004	0.991**	1.006	1.000*	1.018
TR17	1.012	0.999*	1.002	1.013	0.995**	1.002	0.997***	1.002
TR18	1.000*	1.009	1.011	1.003	1.011	1.007	1.002	1.000*
TR19	1.002	1.000*	1.013	1.009	0.983***	1.004	1.000*	1.007
TR20	1.001	1.000*	1.011	1.004	1.019	1.007	1.005	0.999**
MM	1.014	1.008	1.010	1.000*	1.027	1.011	0.998**	1.020
PPP	1.010	0.998**	1.012	1.006	0.986***	1.010	1.003	1.015
UIRP	1.008	1.002	0.993**	0.995**	0.985***	1.008	0.972***	0.999*
					h=3			
TR1	0.990	0.977	0.982	0.969**	1.000	0.979	0.976	0.993
TR2	0.957**	0.988	0.992	0.988	0.986	0.998	0.980	0.987
TR3	0.990	0.975	0.997	0.983	0.978*	1.000	0.982	0.980
TR4	0.979	0.982	1.020	0.984	0.966*	1.000	0.978	0.982
TR5	0.992	0.982	0.978	0.989	0.974*	0.980	0.983	0.992
TR6	0.991	0.971**	0.989	0.978	0.989	0.975*	0.999	0.983
TR7	0.983	0.982	1.017	0.984	0.964	0.996	0.978	0.985
TR8	0.985	0.972	1.003	0.979	0.985	0.994	0.978	0.977
TR9	0.980**	0.980	0.995	0.987	0.980	0.998	0.987	0.985
TR10	0.953***	0.981	0.982	0.989	0.982	0.993	0.980	0.984*
TR11	0.988	0.976	0.977*	0.983	0.985**	0.997	0.977	0.986
TR12	0.980	0.978	0.999	0.989	0.969	0.998	0.980	0.987
TR13	0.967*	0.969**	0.997	0.981	0.983	1.001	0.976	0.987
TR14	0.976	0.977	1.000	0.985	0.981	0.994	0.976	0.979*
TR15	0.985	0.979	0.974*	0.988	0.990	0.996	0.995*	0.990
TR16	0.996	0.972*	0.980	0.981	0.991	0.991	0.972**	0.992
TR17	0.999	0.978	1.000	0.985	0.980	0.998	0.977*	0.996**
TR18	0.978	0.966**	0.988	0.975	0.974**	0.997	0.977	0.978
TR19	0.968**	0.977	0.984	0.986	0.980	0.998	0.974*	0.986
TR20	0.970**	0.981	0.987	0.988	0.966**	1.000	0.977*	0.986
MM	0.995	0.972*	0.991	0.981	0.959	0.976**	0.946***	0.981
PPP	0.978**	0.970***	0.964**	0.971**	0.883*	0.954***	0.994***	0.972***
UIRP	0.982	0.989	0.95**	0.965*	0.986	0.980	0.962**	0.981

Notes: RMSFE of the fundamentals-based model (FM), relative to the RMSFE of the driftless Random Walk (RW). Values less than one indicate that the FM generates a lower RMSFE than the RW. Asterisks (*10%, **5%, ***1%) denote the level of significance of the DMW test based on a semi-parametric bootstrap. Here the FM model is based on a Single Predictor incl. TVar-Coeffs. TR1 to TR20 correspond to different Taylor rule specifications as in Appendix B; MM-fundamentals from the Monetary Model, PPP - Purchasing Power Parity; and UIRP- Uncovered Interest Rate Parity. The forecast evaluation period begins in 1978M12+h.

Table 3: Forecast Evaluation: Single Predictor excl. TVar-Coeffs

Predictor	Can.	Den.	U.K.	Jap.	Kor.	Nor.	Swe.	Swi.
				h = 1				
TR1	1.003	1.002	1.004	1.001	0.999*	1.006	1.002	1.004
TR2	1.002	1.000*	1.004	1.003	0.993**	1.006	1.000*	1.002
TR3	1.003	1.000*	1.002	1.000	0.998**	1.005	1.002	1.002
TR4	1.002	1.000*	1.006	1.005	0.992**	1.005	1.000	1.002
TR5	1.001	1.002	1.006	1.002	0.998**	1.001	1.003	1.000*
TR6	1.004	1.003	1.004	1.001	0.999*	1.005	1.003	1.004
TR7	1.002	1.000*	1.006	1.004	0.992**	1.006	1.000*	1.002
TR8	1.000	1.000*	1.006	1.000*	0.997**	1.006	1.001	1.002
TR9	1.002	1.000*	1.006	1.005	0.992**	1.005	0.999*	1.002
TR10	1.003	1.003	1.004	1.001	0.997**	1.006	1.003	1.001
TR11	1.004	1.002	1.004	1.001	0.998**	1.003	1.002	1.004
TR12	1.002	1.000*	1.005	1.005	0.992**	1.006	1.000	1.002
TR13	1.004	0.997**	1.000	1.002	0.998*	1.005	1.001	1.002
TR14	1.002	0.999*	1.006	1.005	0.992**	1.006	1.000*	1.002
TR15	1.003	1.002	1.006	1.002	0.998**	1.006	1.001	1.003
TR16	1.004	1.001	1.006	1.002	0.997***	1.006	1.000*	1.004
TR17	1.002	0.999*	1.002	1.005	0.993**	1.006	0.999**	1.002
TR18	1.000*	0.996**	1.006	1.003	0.998*	1.007	0.999*	1.002
TR19	1.002	1.000*	1.006	1.006	0.992**	1.006	0.998*	1.002
TR20	1.002	1.001	1.005	1.002	0.996***	1.006	1.001	1.002
MM	1.007	1.007	1.008	1.001	1.002	1.007	0.997**	1.005
PPP	1.003	1.003	1.004	1.002	0.999*	1.004	0.999*	1.003
UIRP	1.004	1.003	1.002	0.997*	1.000	1.004	1.018	1.003
				h = 3				
TR1	1.004	1.001	1.004	0.995	1.000	1.005	1.000	1.001
TR2	1.000	0.986	1.006	1.005	0.987	1.005	0.994	0.992
TR3	1.003	0.994	1.003	0.994	0.992*	1.004	0.998	0.993
TR4	1.001	0.985*	1.006	0.999	0.980*	1.006	0.995	0.993
TR5	0.999	0.997	1.006	0.999	0.989**	1.005	1.003	1.000
TR6	1.006	0.996	0.998	0.995	0.987**	1.003	0.999	0.997
TR7	1.001	0.986*	1.008	1.002	0.981*	1.006	0.995	0.991*
TR8	1.008	0.989	1.006	0.996	0.988**	1.006	0.995	0.993
TR9	1.002	0.983*	1.005	1.002	0.982*	1.005	0.993	0.992
TR10	0.994*	0.998	0.996	0.999	0.986**	1.006	1.003	0.994*
TR11	1.005	0.994	1.006	0.994	0.991*	1.004	1.002	0.998
TR12	1.000	0.984*	1.004	1.003	0.981*	1.005	0.993	0.993
TR13	0.998	0.984*	1.004	0.996	0.993	1.006	0.997	0.994
TR14	1.000	0.984*	1.004	1.001	0.981*	1.000	0.994	0.992
TR15	1.003	0.995	1.005	0.999	0.993	1.003	0.997*	0.998
TR16	1.005	0.989*	1.005	0.996	0.996	1.004	0.995	1.000
TR17	1.000	0.983*	1.003	1.001	0.987	1.005	0.991**	0.996**
TR18	0.992	0.978**	1.005	0.996	0.990**	1.006	0.989*	0.991
TR19	1.001	0.985*	1.005	1.002	0.987	1.005	0.992	0.994
TR20	1.001	0.993	0.999	0.999	0.986**	1.001	0.996	0.998
MM	1.005	1.005	1.005	1.001	1.001	1.007	0.982*	1.004
PPP	1.000	0.998	0.984*	0.989**	0.977***	0.999	0.994***	0.994
UIRP	1.007	1.006	1.000	0.979*	0.996	0.998	1.022	0.997

Notes: As in Table 2, except that here the focus is on the Single Predictor excl. TVar Coeffs model.

The combined forecasts based on models that exclude time-variation in coefficients are more accurate than the RW mostly at h=12. For example, the predictions of the best performing combination method among the several methods in Panel C, the OLS - DMSPE combination, are more precise for seven of the eight exchange rates considered at h=12. Nonetheless, it is again the case that this improvement is not superior to the BMA and BMS including time-varying coefficients models, and in many cases the differences in the RMSFE are not statistically significant.

Tables 2 and 3 present results for individual predictors including and excluding time-varying coefficients, respectively. To save space here, and in all the subsequent analyses we focus on h=1 and h=3, as the results for h=12 are qualitatively similar to those of h=3. In general, the relatively good performance of the regressions that allow predictors and coefficients to change over time over horizons greater than one month is supported by the individual models with time-varying coefficients. In Table 2, virtually all the individual models with time-varying coefficients generate a smaller RMSFE than the RW for all countries - though not always significant. By contrast, in Table 3 not all the individual models excluding time-varying coefficients improve upon the RW, and this appears to be impacting upon the average forecasting performance of the model that only allow predictors to change over time. The evidence suggests overall that allowing for varying degrees of time-variation in coefficients yields an improvement in the out-of-sample forecasting performance of the fundamentals-based exchange rate models.

4.2 Characterization of the BMA Model Including Time-Varying Coefficients

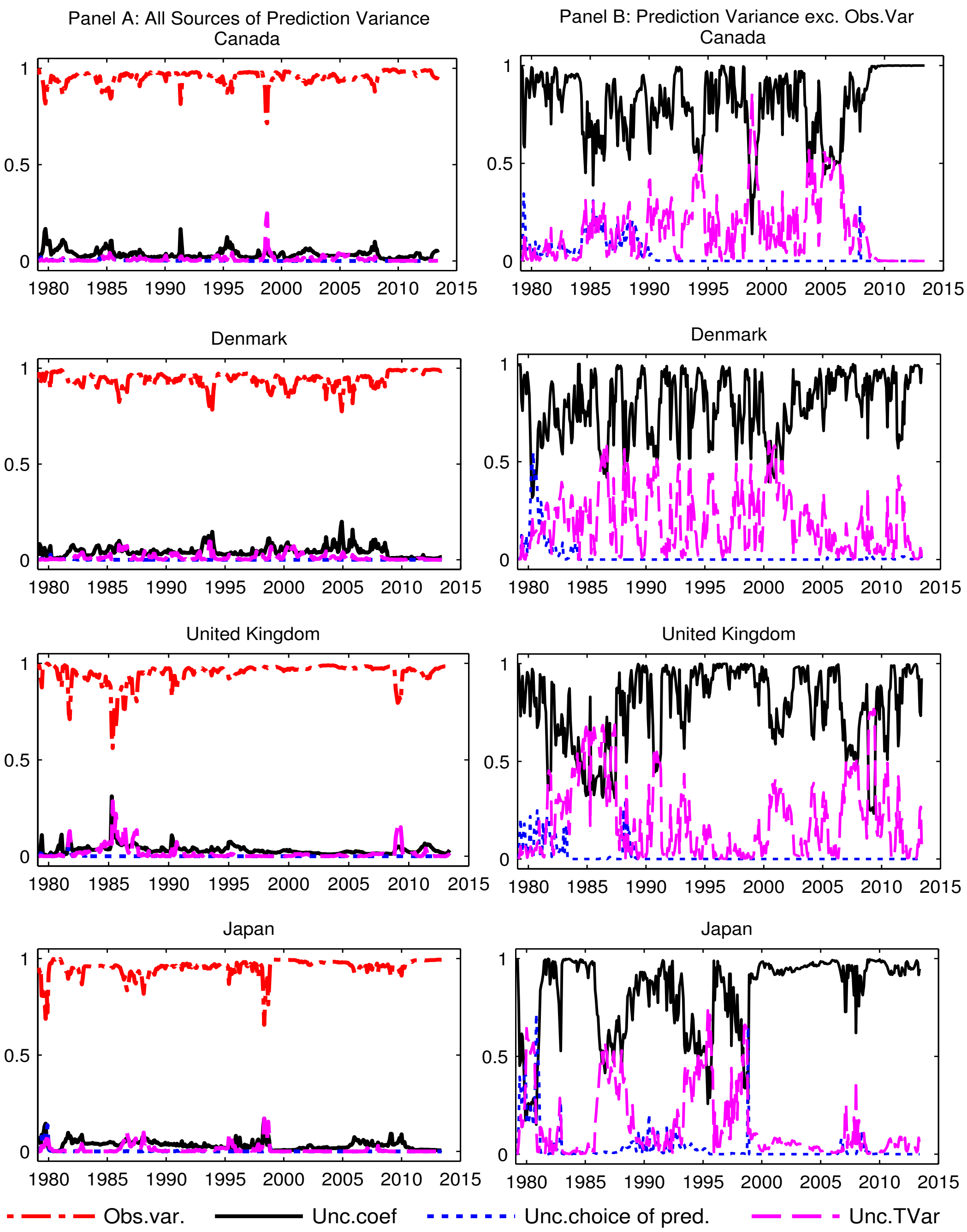
The results in the previous Subsection imply that Bayesian averaging or selection over individual models with varying degrees of coefficients evolution improves relative forecasting performance, particularly at horizons over one month. Since this model emanates from a complex combination of many individual models, understanding its characteristics is useful in explaining the sources of difference in forecasting performance relative to other competing models and across forecast 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 (see, e.g., Dangl and Halling, 2012). As we noted in Subsection 2.3, the total variance can be decomposed 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 due to model uncertainty with respect to the choice of degree of time-variation in coefficients.

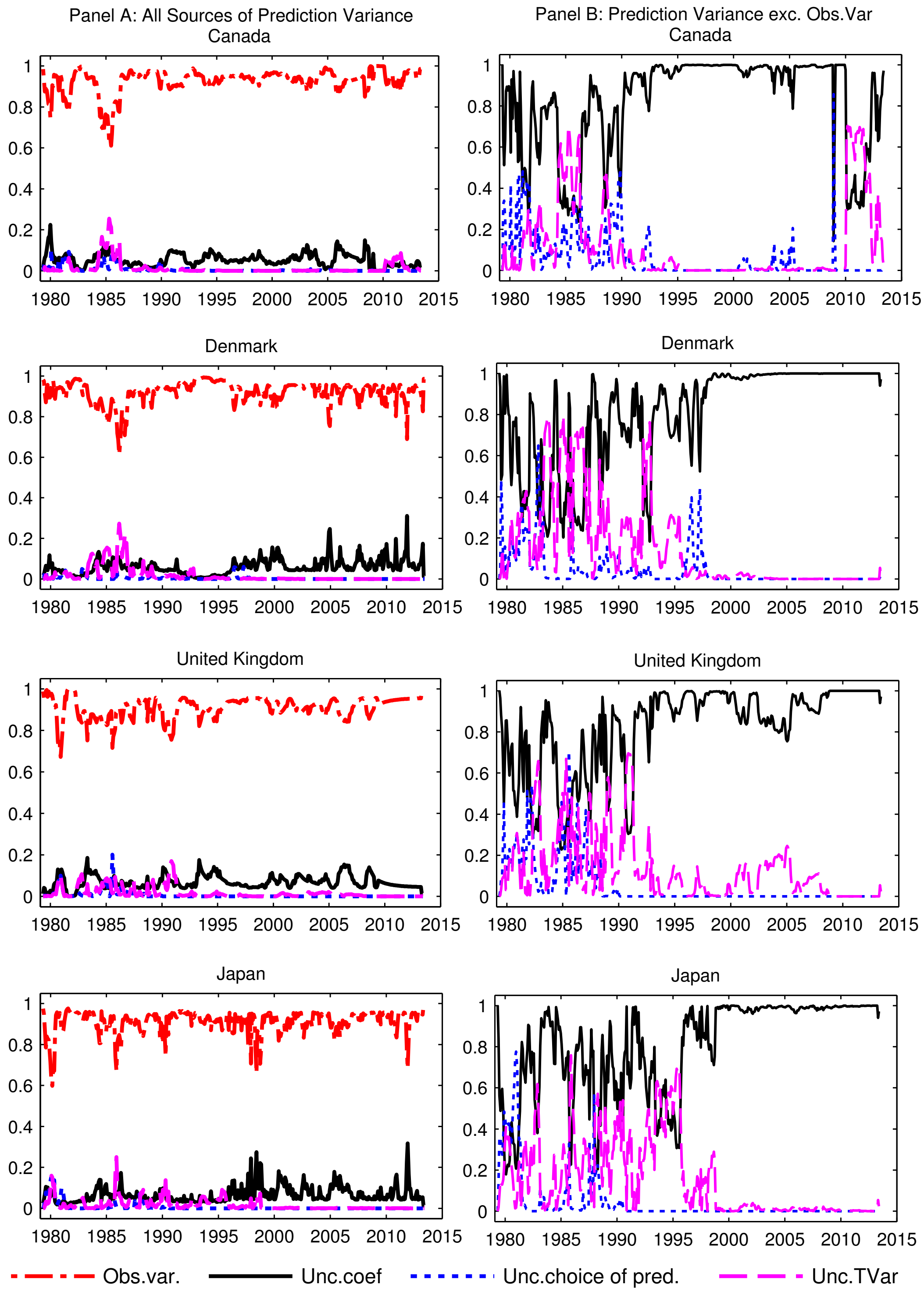
model fails to improve upon the RW.

Figure 1: Sources of Prediction Variance, h=1



Notes: Decomposition of the prediction variance into its constituent parts for h=1 month. 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 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, h=3



Notes: As in Figure 1, except that here h=3 months.

Focusing on a representative selection of four countries, Figures 1 and 2 depict this decomposition for h=1 and h=3, respectively. In both figures, Panel A illustrates the relative weight of each of the components of prediction variance in the total variance. In all cases, the predominant source of uncertainty is observational variance. As Dangl and Halling (2012) point out 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 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 sources of prediction uncertainty. At both forecast horizons, the variance from errors in the estimation of the coefficients is the dominant source of prediction uncertainty. Between the two forecast horizons however, there are differences with respect to uncertainty regarding the choice of the correct degree of time-variation in coefficients. At the one-month forecast horizon, the uncertainty originating from this source is detectable throughout the forecast sample. For example, in the case of Canada and in various periods, it represents over one fifth of the total variance excluding observational variance, peaking in periods around financial stresses, such as the 2008 financial crisis. In contrast, at the three-month horizon the uncertainty regarding the choice of the correct degree of time-variation in coefficients is clustered at the initial out-of-sample period which corresponds to the initial data-points in the expanding window of the forecasting procedure. As more evidence is accumulated, the uncertainty remains mostly low. Finally, the variance due to model uncertainty with respect to the choice of the predictor is for the most part low, except also for the aforementioned initial data-points in the expanding window.

We interpret these findings are suggesting that although the estimation uncertainty is substantial at both horizons, at the one-month forecast horizon the model fails to strongly 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. That is, there is no certainty about the exact degree of time-variation in coefficients necessary to offset the loss in forecast performance emanating from estimation uncertainty. By contrast, at the three-month forecast horizon the model successfully embeds the level of time-variation in coefficients present in the data. Hence, it consistently outperforms the RW by counterbalancing the loss in the precision in coefficient estimation, with increased variability in the coefficients. This signifies that both, estimation uncertainty and coefficient instability obstruct model forecasting performance, and our model adapts to the pattern in the data for horizons over one month.[20]

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 rationalises the Meese and

[20] Recall that we model the variance of the measurement equation error term Wt, as proportional to the estimation variance of the coefficient vector $\theta_t|D_t$, see equation (12).

Rogoff (1983) result of exchange rate unpredictability. They find that estimation uncertainty is the main factor that hinders model's forecasting performance and not time-variation in coefficients. However, Giannone (2010) disputes these findings arguing that both, estimation uncertainty and parameter instability are relevant in explaining 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 characterised by a high forecast error with the first few observations. 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. However, in the third stage, further increasing the window deteriorates model-based forecasting performance relative to the RW, since gains from reduced estimation error are compensated by losses due to the presence of structural instabilities. Thus, the recursive ratio of the relative RMSFE exhibits a U shape. Due to space constraints we show the figures of our recursive ratio of the relative RMSFE in an Additional Result Appendix. Here we note that the figures confirm Giannone's (2010) observations for the first and second stages, but not the third stage. In our case, further increasing the window does not significantly deteriorate model-based forecasting performance relative to the RW. In fact, for the most part of the forecast period and horizons greater than one month, the relative RMSFE is below one, favouring the fundamentals-based model. When read in conjunction with the main sources of instabilities we detect, this result reinforces the finding that the BMA model including time-varying coefficients successfully captures the degree of time-variation in parameters necessary to offset the loss in forecast accuracy due to estimation uncertainty.

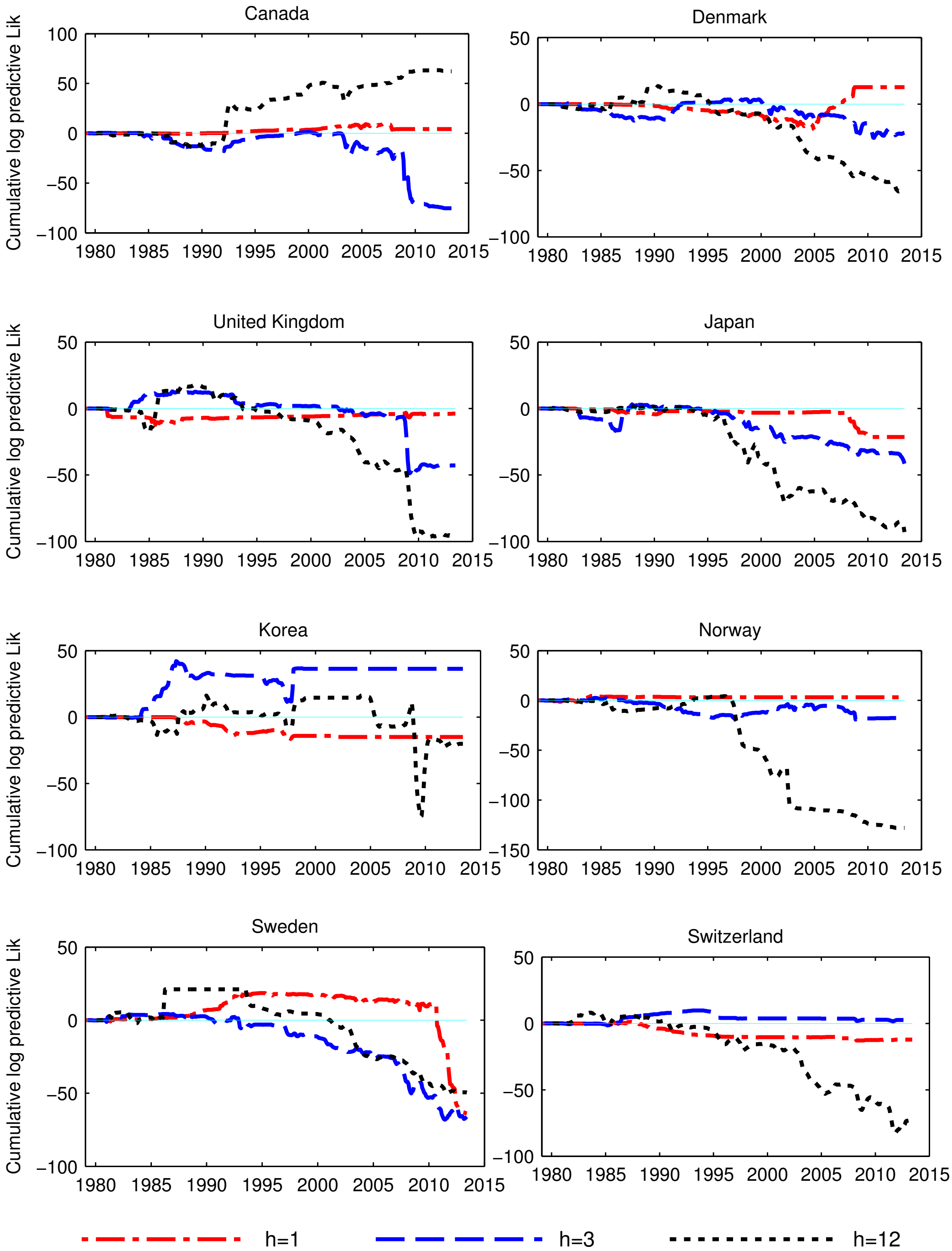
To further shed light on these findings we focus in another measure of forecast accuracy that underlies our Bayesian approach, namely the predictive likelihood, see Geweke and Amisano (2010). Figure 3 depicts the cumulative log predictive likelihood of the models with constant-coefficients relative to models 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.

Three main results are apparent in Figure 3. First, it confirms that models with time-varying coefficients are empirically plausible, especially at h=3 and h=12. At these horizons, with the exception of three out of eight countries, the cumulative log predictive likelihoods become negative after a number of out-of-sample data-points have been accumulated.[21,22] Second, for h=3 and h=12 and occasionally excluding Canada, Korea, and Switzerland, the cumulative log predictive likelihoods shows a downward trend. This is consistent with additive evidence favouring the model with

[21] The three exceptions are Canada, Korea, and Switzerland.

[22] 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. However, as they emphasize, after a number of observations have been accumulated the results become invariant to substantial changes in the prior density distribution.

Figure 3: Cumulative Log Predictive Likelihood: BMS excl. TVar-Coeff. relative to BMS incl. TVar-Coeff



Notes: Cumulative log predictive likelihood of the BMS excl. TVar-Coeff. model relative to the BMS incl. TVar-Coeff. model. 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.

time-varying coefficients. Third, the Figure shows that observations in the more recent turbulent-periods, around 2005-2010, contribute highly to the evidence in favor of the Time-varying coefficients model. Overall, our findings remain invariant to this measure of forecast accuracy.

4.2.2 Analysis of the Degree of Time-Variation in Coefficients

As the preceding results indicate, at the one-month horizon the uncertainty regarding the degree of time-variation in parameters is not trivial, but at longer forecast horizons the uncertainty become low as more evidence is gathered. However, the empirically estimated amount of time-variation was not evoked. Figure 4 provides this information. It depicts the total posterior probability of each of the support points for time-variation in coefficients, δ. At h=1 in Panel A, none of the support points consistently accumulates higher weight than the others for an extensive period. For example in the case of Canada and the United Kingdom, most weight is attributed to $\delta = 1$, consistent with constant-coefficients. Nonetheless, there are still shifts in these weights, dropping substantially around the 2008 financial crisis. For the other countries, the shifts in the support points are even more pronounced, impacting upon the uncertainty about the correct expected degree of coefficient variability.

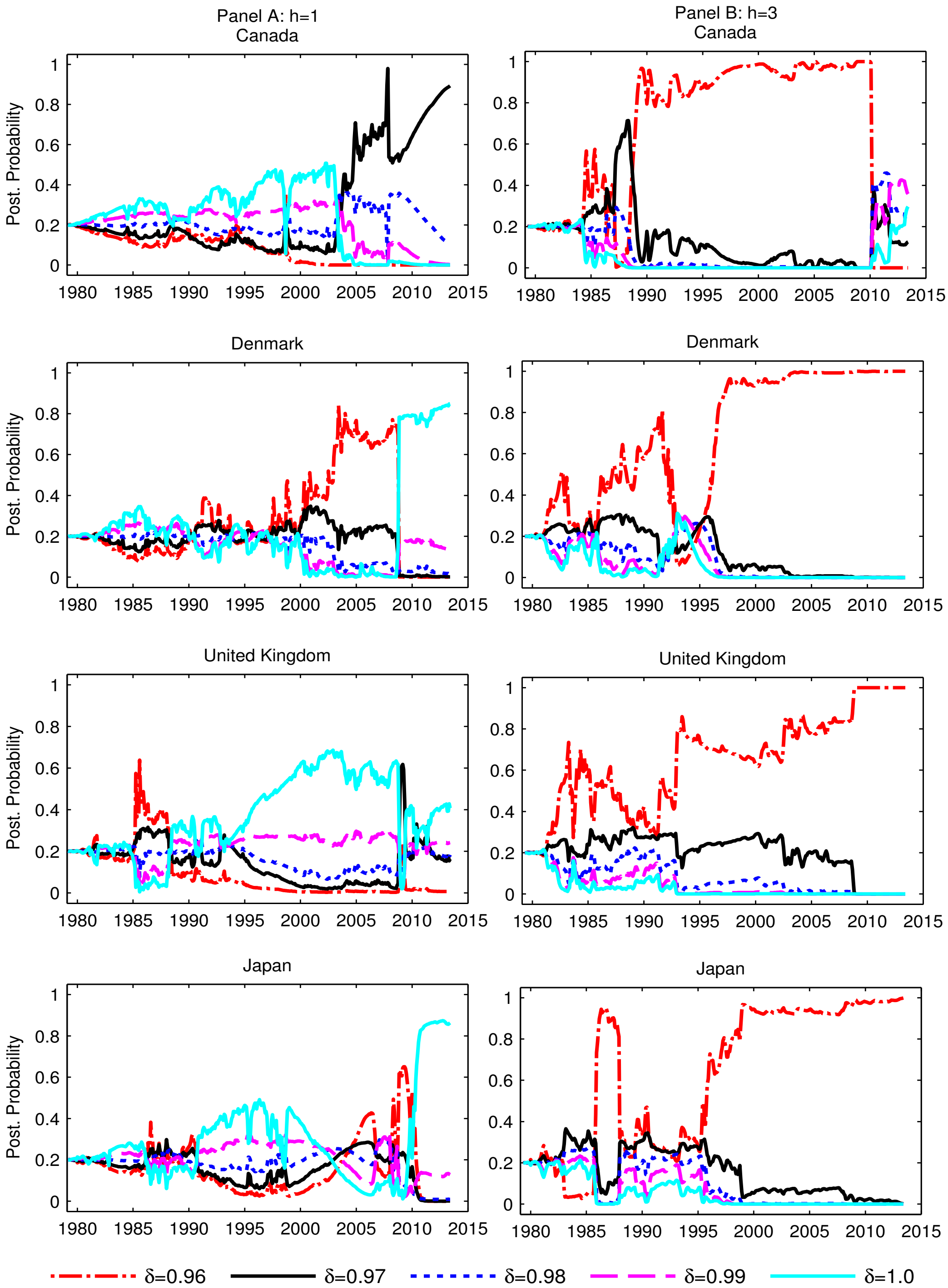
In contrast, at h=3 in Panel B, very dynamic models with $\delta = 0.96$, exhibit the highest posterior probability, while constant-coefficient models typically receive the lowest support. Interestingly, Giannone (2010) finds that to match the pattern of exchange rate unpredictability present in the data, a significant amount of time-variation in coefficients were necessary in his simulations. Beckmann and Schuessler (2014) 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 supported. In the present case, the preponderance of $\delta = 0.96$ over the other support points, is reflected in the low uncertainty with respect to the degree of time-variation in coefficients.[23]

4.2.3 Analysis of the Importance of Individual Predictors

Another characteristic to explore in the BMA including time-varying coefficients model is the importance of individual predictors. Figure 5 shows which predictors accumulate the highest probability at each point in time. Overall and in most cases, after the initial data-points in the sample, there is less variability in the predictor that exhibits the largest posterior probability. This is consistent with the low uncertainty with respect to the choice of the predictor that we found. However, the predictors that receive the highest weight differ over forecast horizons and countries. In the case of the United Kingdom for example, while fundamentals from the Taylor

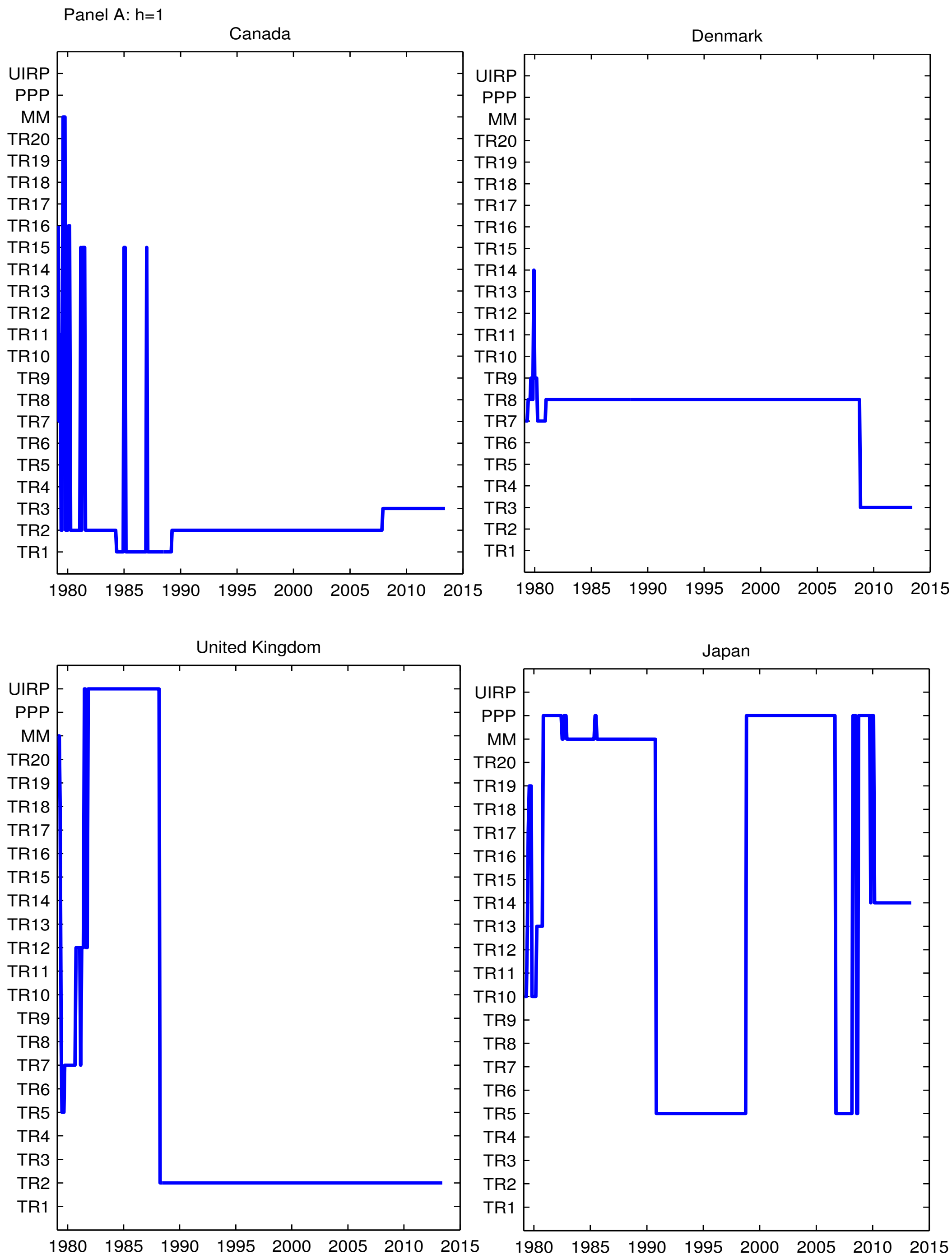
[23] In an Additional Results Appendix we zoom in Figure 4 to illustrate the values of δ with the highest probability at each period.

Figure 4: Posterior Probabilities of Degrees of Time-variation in Coefficients, δ



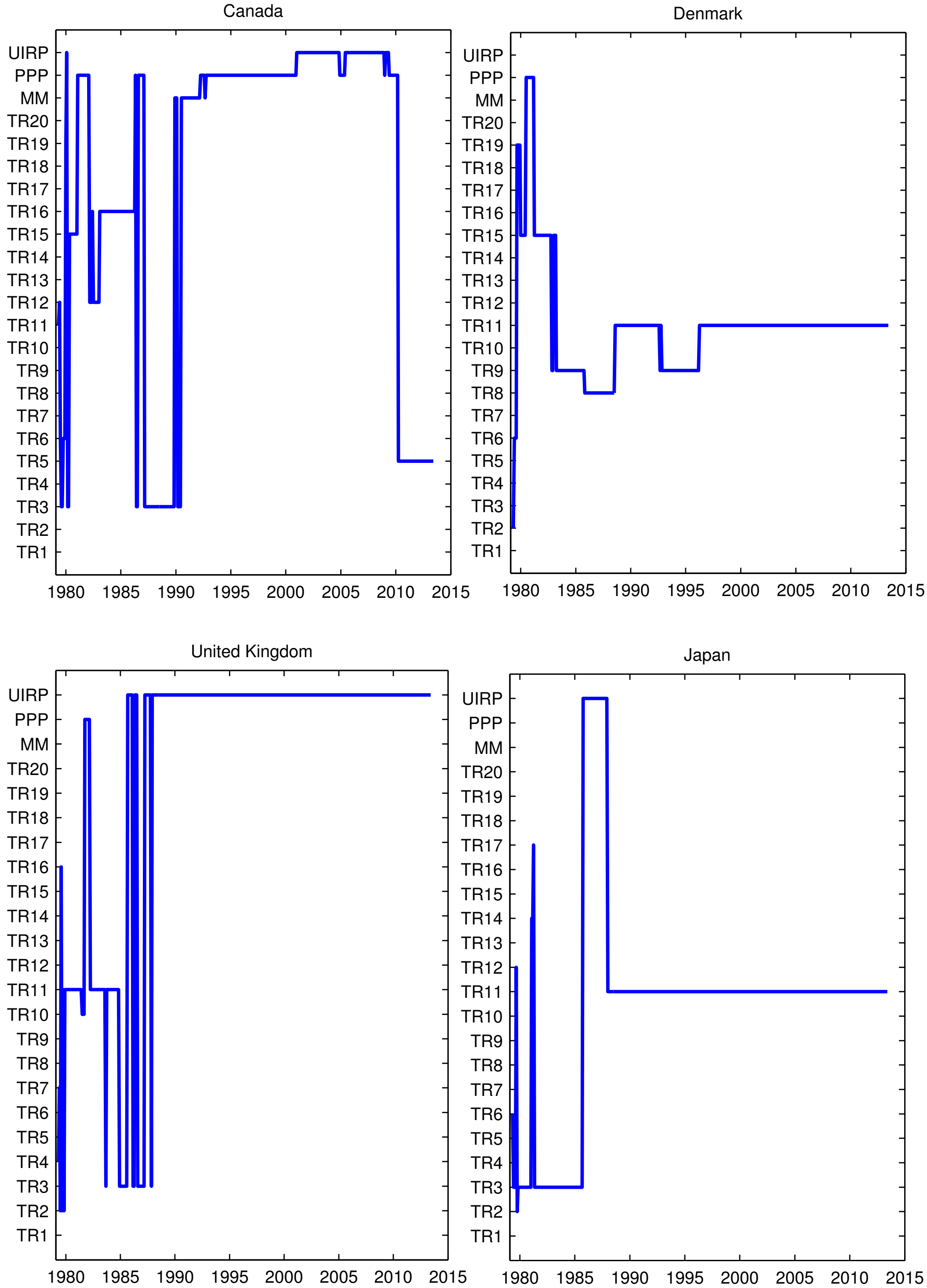
Notes: Posterior probabilities of values of δ (support points for time-variation in coefficients), for a representative selection of countries. Panel A is for h=1 and Panel B for h=3. These are the weights employed to produce the average forecasts in the BMA incl. TVar-Coeff. model. (For all countries, see the Additional Results Appendix).

Figure 5: 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 h=1 and Panel B for h=12. The forecasts from the BMS incl. TVar-Coeff. model are based on these predictors. (For all countries, see the Additional Results Appendix).

Panel B: h=3



Canada
Denmark
United Kingdom
Japan
UIRP
PPP
MM
TR20
TR19
TR18
TR17
TR16
TR15
TR14
TR13
TR12
TR11
TR10
TR9
TR8
TR7
TR6
TR5
TR4
TR3
TR2
TR1
1980
1985
1990
1995
2000
2005
2010
2015

rule with homogeneous coefficients and interest rate smoothing (TR2) are dominant at h=1, at h=3 fundamentals based on UIRP receive the highest weight.

Some of the fundamentals in our set are constituted by Taylor rules augmented with FCIs (TR5, TR10, TR15 and TR20). In the selection of countries represented in Figure 5, one of the specification (TR5) dominate other fundamentals for some period in the case of Japan at h=1. In the extensive results reported in an Additional Results Appendix, other specifications also predominate for the most part of the sample (TR10 for Switzerland at h=3; TR15 for Canada (h=12); and TR20 for Korea at h=1).

5 Robustness Checks

We verified the robustness of the empirical findings in the previous Section in three dimensions. First, while plots of cumulative log predictive likelihoods allowed us to examine the forecast accuracy over the entire path of the forecast window; we also experimented with changing the beginning of the forecast window.[24,25] In particular, we considered an evaluation period starting in 1989M12+h, as in Wright (2008), or in more recent periods, 2005M12+h. We focused in the BMA models including/excluding time-varying coefficients. Overall, in both forecast windows, results in Table 4-Panel A are still supportive of models with time-varying coefficients for horizons greater than one month. For example, at h=3 and h=12, the relative RMSFE is less than one for at least seven, out of eight countries in the case of the BMA model including time-varying coefficients. On the contrary, the BMA model excluding time-varying coefficients achieves small reductions in the RMSFE for no more than five currencies at the same horizons; once again, confirming Wright's (2008) results.

Second, we changed the base numeraire to the UK Pound in place of the US dollar, following Chen *et al.* (2010). Focusing on the BMA model including/excluding time-varying coefficients, results in Table 4-Panel B remain largely unaffected. The corresponding analysis of prediction variance, based on the example of Japan in Figure 6, also reveals the pattern we documented: estimation uncertainty and uncertainty with regards to the exact degree of time-variation are the main obstacles to model's forecasting performance.

Third, we estimated directly the degree of time-variation in coefficients following the approach in Koop and Korobilis (2013), instead of inferring from model's posterior probability. In this case the estimated δ is: $\widehat{\delta} = \delta_{Min} + (1 - \delta_{Min})exp(L_g.v_{t+h}^2)$, where δ_{Min} is the minimum value of support points for time-variation in coefficients

[24] Giacomini and Rossi (2010) formalize the issue of forecast robustness over different windows in presence of instabilities by developing appropriate test statistics. However, their tests require a use of a rolling or fixed estimation window approach when producing the forecasts, rather than our recursive scheme.

[25] Due to computational cost of implementing the bootstrap for the *d.k.* models and each sensitivity analysis we consider, in this Section we evaluate our forecasts solely on the basis of the relative RMSFE.

Table 4: Out-of-Sample Forecast Evaluation: Robustness Checks

Panel A: Sensitivity to change in the forecast window

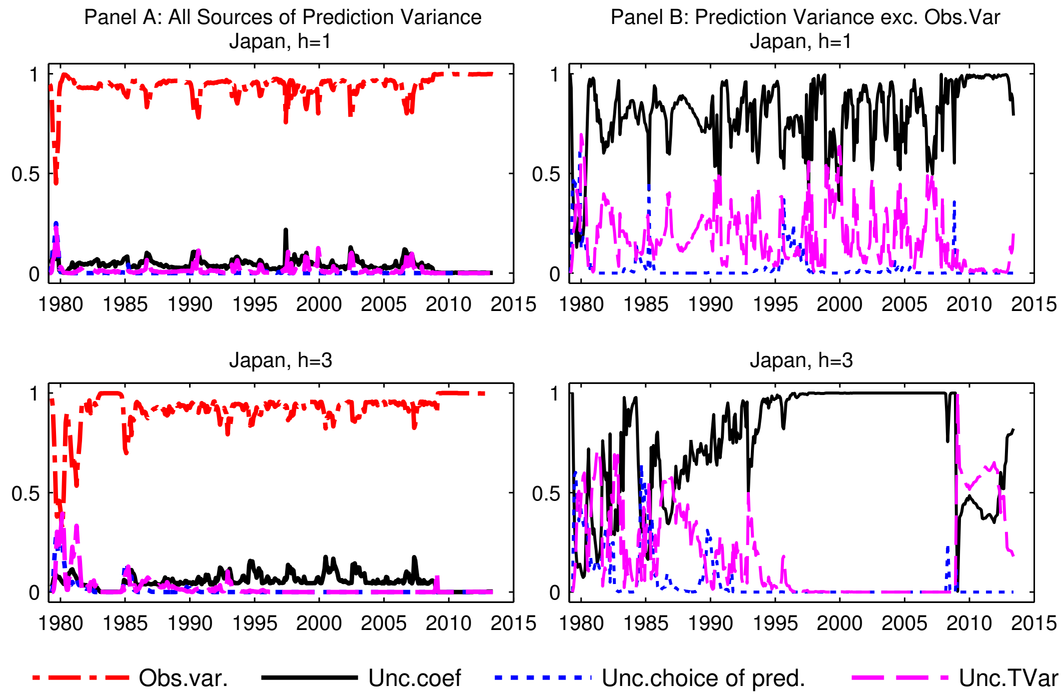
	BMA incl. TVar-Coeff.			BMA excl. TVar-Coeff.		
	1989M12+h			1989M12+h		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.012	0.975	0.857	1.000	1.000	1.051
Denmark	1.011	0.981	0.929	1.005	1.001	1.009
UK	1.014	0.964	0.942	1.005	1.020	1.008
Japan	1.008	0.996	0.837	1.003	1.009	0.995
Korea	0.998	0.983	0.984	1.001	0.996	0.974
Norway	1.009	0.978	0.975	1.007	1.006	1.076
Sweden	1.020	0.965	0.922	1.006	0.997	0.984
Switzerland	1.013	1.009	0.910	1.005	1.009	0.985
	2005M12+h			2005M12+h		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.015	0.991	0.920	0.999	0.992	1.024
Denmark	1.002	0.991	0.952	1.001	0.998	0.999
UK	1.019	0.899	0.934	1.004	1.008	1.002
Japan	1.006	0.980	0.816	1.001	0.996	0.927
Korea	1.000	0.998	0.976	1.000	0.997	0.973
Norway	1.005	0.958	1.007	1.003	1.007	1.008
Sweden	1.023	0.972	0.982	1.004	1.033	0.988
Switzerland	1.017	1.002	0.923	1.000	1.003	0.985

Panel B: Sensitivity to change in base currency to Pound Sterling

	BMA incl. TVar-Coeff.			BMA excl. TVar-Coeff.		
	h=1	h=3	h=12	h=1	h=3	h=12
Canada	1.022	1.012	0.806	1.004	1.003	0.947
Denmark	1.009	0.995	0.854	1.001	1.003	0.941
U.S	1.002	0.967	0.901	0.992	1.006	0.965
Japan	1.000	0.962	0.804	0.996	0.985	0.955
Korea	1.017	1.001	0.912	1.007	0.996	0.984
Norway	1.010	1.002	0.911	1.005	1.002	0.996
Sweden	1.021	0.988	0.957	1.002	0.997	0.974
Switzerland	1.009	0.988	0.808	1.003	0.995	0.933

Notes: Robustness analysis of the results in the Empirical Section. The entries in the Table are the RMSFE of the fundamentals-based models (FM) - BMA incl/exc. TVar-Coeff, relative to the RMSFE of the driftless Random Walk (RW). Values less than one indicate that the FM generates a lower RMSFE than RW. Panel A focus on changing the beginning of the forecast evaluation period to 1989M12+h and 2005M12+h. Panel B reports the sensitivity to change in the base currency to the U.K. Pound.

Figure 6: Prediction Variance: Sensitivity to Change in base Currency to the Pound



Notes: As in Figures 1 and 2, except that here the exchange rate is defined relative to the Pound Sterling, and hence all the data employed in the predictive regression are redefined relative to the U.K.

Table 5: Estimated Degree of Time-Variation in Coefficients

Horizon	Quartiles	Can.	Den.	U.K.	Jap.	Kor.	Nor.	Swe.	Swi.
	First	0.980	0.970	0.970	0.970	0.980	0.970	0.970	0.960
h =1 month	Second	0.990	0.980	0.990	0.980	0.990	0.980	0.980	0.980
	Third	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	First	0.970	0.960	0.960	0.960	0.970	0.960	0.960	0.960
h =3 months	Second	0.980	0.970	0.970	0.970	0.980	0.970	0.970	0.960
	Third	0.990	0.980	0.990	0.980	0.990	0.990	0.990	0.980
	First	0.960	0.960	0.960	0.960	0.960	0.960	0.960	0.960
h =12 months	Second	0.970	0.960	0.960	0.960	0.960	0.960	0.960	0.960
	Third	0.990	0.970	0.970	0.970	0.970	0.960	0.960	0.960

Notes: Estimates of the degree of time-variation in coefficients, δ. For each country, we first estimate $\hat{\delta}$ as we run the predictive model excluding time-variation in coefficients for each t. We then average the estimates over predictors and obtain a series of $\hat{\delta}$. From this series we compute the first, second and third quartiles. The forecast evaluation period begins in 1978M12+h.

considered (we set $\delta_{Min} = 0.96$), exp is the exponential function, L_g is a constant scaling parameter and v_{t+h} is the predictive regression's error. Overall, results are consistent with the degree of time-variation in coefficients we obtained in the empirical section above. As shown in Table 5, at h=1 the median (2^{nd} quartile) $\widehat{\delta}$ is between 0.98 and 0.99, and for over 3/4 of the out-of-sample data-points its value is above 0.97. However, as the forecast horizon increases, the value of $\widehat{\delta}$ is consistent with more time-variation in coefficients, with median $\widehat{\delta} = 0.97$ for h=3 and most countries, and $\widehat{\delta} = 0.96$ for h=12 and virtually all countries.

6 Conclusion

The literature on exchange rate forecasting points out that the out-of-sample predictive power of fundamental-based 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 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 take into account 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; see, for example, Bacchetta and van Wincoop (2004, 2013), Berge (2013), and Sarno and Valente (2009). Inspired by recent advances in Bayesian methods (e.g., Dangl and Halling, 2012; Koop and Korobilis; 2012), we further employ our systematic framework to investigate all sources of uncertainty in the predictive models, through a variance decomposition procedure.

In our findings, fundamentals-based models significantly outperform the driftless random walk benchmark for most currencies at all the forecast horizons we consider, except at the one-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 greater than one 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 the one-month forecast horizon our predictive regressions fail to successfully capture the suitable time-varying weights associated with the regressors; yielding poor model's forecasting performance. Thus, in line with the prevailing view in the literature, we affirm that it is hard to beat the no-change benchmark at the one-month forecast horizon (Rossi, 2013).

We then proceed and track the sources of uncertainty in the regressions. In this regard we find that the uncertainty in the estimation of the models' coefficients, and the uncertainty regarding the correct level of time-variation in coefficients, are the main factors hindering models' predictive ability. When the uncertainty emanating from these 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 model, we find that while the relevant set of predictors generally differs between forecast horizons and between countries; the uncertainty about selecting any individual predictor is generally low within a specific country-horizon. We view our results as providing a direct evidence towards the prevalent conjectures or simulation based evidence 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. (2013). 'On the unstable relationship between exchange rates and macroeconomic fundamentals', *Journal of International Economics*, vol. 91(1), pp. 18–26.

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

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

Beckmann, J. and Schuessler, R. (2014). 'Forecasting exchange rates under model and parameter uncertainty', Center for Quantitative Economics, Münster.

Berge, T.J. (2013). 'Forecasting disconnected exchange rates', *Journal of Applied Econometrics*, vol. 29(5), pp. 713–735.

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

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*, vol. 24(7), pp. 1150–1175.

Clarida, R., Galı, J. and Gertler, M. (1998). 'Monetary policy rules in practice: some international evidence', *European Economic Review*, vol. 42(6), pp. 1033–1067.

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*, vol. 135(1), pp. 155–186.

Clark, T.E. and West, K.D. (2007). 'Approximately normal tests for equal predictive accuracy in nested models', *Journal of econometrics*, vol. 138(1), pp. 291–311.

Dangl, T. and Halling, M. (2012). 'Predictive regressions with time-varying coefficients', *Journal of Financial Economics*, vol. 106(1), pp. 157 – 181.

Diebold, F.X. (2012). 'Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of diebold-mariano tests', National Bureau of Economic Research.

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

Engel, C., Mark, N.C. and West, K.D. (2012). 'Factor model forecasts of exchange rates', National Bureau of Economic Research.

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

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

Fratzscher, M., Sarno, L. and Zinna, G. (2012). 'The scapegoat theory of exchange rates: the first tests', European Central Bank.

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

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

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

Hatzius, J., Hooper, P., Mishkin, F.S., Schoenholtz, K.L. and Watson, M.W. (2010). 'Financial conditions indexes: A fresh look after the financial crisis', National Bureau of Economic Research.

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

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

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

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

Koop, G. and Korobilis, D. (2013). 'Large time-varying parameter vars', *Journal of Econometrics*, vol. 177(2), pp. 185–198.

Koop, G. and Korobilis, D. (2014). 'A new index of financial conditions', *European Economic Review*, vol. 71(0), pp. 101–116.

Li, J., Tsiakas, I. and Wang, W. (2014). 'Predicting exchange rates out of sample: Can economic fundamentals beat the random walk?', *Journal of Financial Econometrics*, vol. 0(0), pp. 1–49.

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

Mark, N.C. (2009). 'Changing monetary policy rules, learning, and real exchange rate dynamics', *Journal of Money, Credit and Banking*, vol. 41(6), pp. 1047–1070.

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

Mishkin, F.S. (2010). 'Monetary policy flexibility, risk management, and financial disruptions', *Journal of Asian Economics*, vol. 21(3), pp. 242–246.

Molodtsova, T., Nikolsko-Rzhevskyy, A. and Papell, D.H. (2011). 'Taylor rules and the euro', *Journal of Money, Credit and Banking*, vol. 43(2-3), pp. 535–552.

Molodtsova, T. and Papell, D.H. (2009). 'Out-of-sample exchange rate predictability with taylor rule fundamentals', *Journal of International Economics*, vol. 77(2), pp. 167–180.

Molodtsova, T. and Papell, D.H. (2012). *Taylor Rule Exchange Rate Forecasting during the Financial Crisis*, NBER International Seminar on Macroeconomics 2012, University of Chicago Press, pp. 55–97.

Mumtaz, H. and Sunder-Plassmann, L. (2013). 'Time-varying dynamics of the real exchange rate: An empirical analysis', *Journal of Applied Econometrics*, vol. 28(3), pp. 498–525.

Nelson, C.R. and Kim, M.J. (1993). 'Predictable stock returns: The role of small sample bias', *The Journal of Finance*, vol. 48(2), pp. 641–661.

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

Raftery, A.E., Kárnỳ, M. and Ettler, P. (2010). 'Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill', *Technometrics*, vol. 52(1), pp. 52–66.

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*, vol. 23(2), pp. 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*, vol. 13(2), pp. 231 – 247.

Rogoff, K. (1996). 'The purchasing power parity puzzle', *Journal of Economic literature*, vol. 34(2), pp. 647–668.

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

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

Rossi, B. (2013). 'Exchange rate predictability', *Journal of Economic Literature*, vol. 51(4), pp. 1063–1119.

Rossi, B. and Sekhposyan, T. (2011). 'Understanding models forecasting performance', *Journal of Econometrics*, vol. 164(1), pp. 158–172.

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

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

Taylor, J.B. (1993). 'Discretion versus policy rules in practice', *Carnegie-Rochester Conference Series on Public Policy*, vol. 39(0), pp. 195 – 214.

Taylor, J.B. (2008). 'Monetary policy and the state of the economy', testimony before the Committee on Financial Services, US House of Representatives.

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

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

West, M. and Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*, New York, NY: 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 and Economic Statistics*, vol. 5(1), pp. 87–97.

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

A Forecasting and Averaging in a Bayesian Framework

This Appendix provides details on the methods used to forecast with the dynamic linear model defined by equations (1) and (2) in the main text, as well as the averaging approach. We follow closely Dangl and Halling (2012) and West and Harrison (1997).

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

For convenience we begin by transcribing the system of equations from Section 2.1 in the main text:

$$\Delta e_{t+h} = X_t' \theta_t + v_{t+h}, \quad v_{t+h} \sim N(0, V), \text{ (observation equation)}; \tag{A.1}$$

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

The essential components of the Bayesian approach we employ are the priors for V and θ_t, along with a method to estimate W_t; the joint or conditional posterior distribution of V 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 $t = 0$:

$$V|D_0 \sim IG\left[\frac{1}{2}, \frac{1}{2}S_0\right], \tag{A.3}$$

$$\theta_0|D_0, V \sim N[0, gS_0(X'X)^{-1}], \tag{A.4}$$

where,

$$S_0 = \frac{1}{N-1}\Delta e'(I - X(X'X)^{-1}X')\Delta e, \tag{A.5}$$

and D_0 indicates the conditioning information at $t = 0$. In general, at any arbitrary subsequent period, $D_t = [\Delta e_t, \Delta e_{t-h}, ..., X_t, X_{t-h}, ..., Priors_{t=0}]$. That is, D_t contains the exchange rate variations, the predictors, and the prior parameters. At this arbitrary period we can form a posterior belief about the unobservable coefficient $\theta_{t-1}|D_t$, and the variance of the observation equation error term (observational variance $V|D_t$). 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:

$$V|D_t \sim IG\left[\frac{n_t}{2}, \frac{n_t S_t}{2}\right], \tag{A.6}$$

$$\theta_{t-1}|D_t, V \sim N(m_t, VC_t^*), \tag{A.7}$$

where, S_t is the mean of the estimate of the observational variance V, at time t, with n_t as the corresponding number of degrees-of-freedom; m_t is the estimate of coefficient vector θ_{t-1} conditional on D_t and V; and C_t^* corresponds to the conditional variance matrix of θ_{t-1}, normalised by the by the observational variance. Integrating out the distribution given by (A.7) with respect to V, yields a multivariate $t-$distribution for the coefficients' posterior:

$$\theta_{t-1}|D_t \sim T_{n_t}(m_t, S_t C_t^*). \tag{A.8}$$

The form of the transition equation given by (A.2) suggests that, when updating the coefficients vector, the posterior distribution of $\theta_{t-1}|D_t$ represented by (A.8) does not necessarily become the prior for $\theta_t|D_t$. The equation indicates that the transition process is exposed to normally distributed random shocks, which widens the variance but maintains the mean:

$$\theta_t|D_t \sim T_{n_t}(m_t, S_t C_t^* + W_t). \tag{A.9}$$

The predictive density of the h-step-ahead change in the exchange rate Δe_{t+h}, is obtained by integrating the conditional density of Δe_{t+h} over the space spanned by θ and V. To derive it, let $\varphi(x; \mu, \sigma^2)$ be the density of a normal distribution evaluated at x and $ig(V; a, b)$ be the density of an $IG[a, b]$ distributed variable, evaluated at V. Then, the predictive density is:

$$\begin{aligned} f(\Delta e_{t+h}|D_t) &= \int_0^\infty \left[\int_\theta \varphi(\Delta e; X_t'\theta, V)\varphi(\theta; m_t, VC_t^* + W_t)d\theta\right] \\ &\quad \times ig\left(\frac{n_t}{2}, \frac{n_t S_t}{2}\right) dV \\ &= \int_0^\infty \varphi(\Delta e; X_t' m_t, X_t'(VC_t^* + W_t)X_t + V) \\ &\quad \times ig\left(\frac{n_t}{2}, \frac{n_t S_t}{2}\right) dV \\ &= \mathbf{t}_{n_t}(\Delta e_{t+h}; \widehat{\Delta} e_{t+h}, Q_{t+h}), \end{aligned} \tag{A.10}$$

where, $\mathbf{t}(\Delta e_{t+h}; \widehat{\Delta e}_{t+h}, Q_{t+h})$ denotes the density of a $t-$distribution 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:

$$\widehat{\Delta} e_{t+h} = X_t' m_t, \tag{A.11}$$

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

$$Q_{t+1} = X_t' R_t X_t + S_t, \tag{A.12}$$

with

$$R_t = S_t C_t^* + W_t, \tag{A.13}$$

where, R_t is the unconditional variance of the coefficient vector θ_t at time t. The first term in equation (A.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.

After observing the exchange rate change at $t + h$, the priors on θ_t and V are updated as described in equations (A.14)-(A.19). The first element, is the prediction error:

$$\varepsilon_{t+h} = \Delta e_{t+h} - \widehat{\Delta e}_{t+h}, \text{ (prediction error)}. \tag{A.14}$$

Insofar as the prediction error equals zero, no updating occurs in the coefficient vector, since the forecast matches the value observed in the data.

$$n_{t+1} = n_t + 1, \text{ (degrees-of-freedom)}, \tag{A.15}$$

$$S_{t+h} = S_t + \frac{S_t}{n_t}\left(\frac{\varepsilon_{t+h}^2}{Q_{t+h}} - 1\right), \text{ (estimator of observational variance)}. \tag{A.16}$$

Given that the total variance of the forecast is expressed by Q_{t+h}, then $E(\varepsilon_{t+h}^2) = Q_{t+h}$. When the prediction error matches its expectation, i.e., $\varepsilon_{t+h}^2 = Q_{t+h}$, the estimate of the observational variance remains unchanged ($S_{t+h} = S_t$). When the prediction error is lower (higher) than the expected error, the estimated observational variance reduces (increases).

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

$$A_{t+h} = \frac{R_t X_t}{Q_{t+h}}, \text{ (adaptive vector)}. \tag{A.17}$$

It characterises the degree to which the posterior of the coefficient vector θ_t changes to new observation. The numerator of equation (A.17) can be recognised as the information content of the current observation, and the denominator as the measure of the precision of the estimated coefficients. With the above elements, we are now in position to update the coefficients' point estimate m_t and the covariance matrix C_t^*:

$$m_{t+h} = m_t + A_{t+h}\varepsilon_{t+h}, \text{ (expected coeff. vector estimator)}, \tag{A.18}$$

$$C_{t+h}^* = \frac{1}{S_t}\left(R_t - A_{t+h}A_{t+h}' Q_{t+h}\right), \text{ (variance of the coeff. vector estimator)}. \tag{A.19}$$

The exposition so far does not include a method to estimate W_t. However, as we noted in Section 2.1, 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 $\theta_t | D_t$ at time t. That is:

$$W_t = \frac{1-\delta}{\delta} S_t C_t^*, \quad \delta \in \{\delta_1, \delta_2, ..., \delta_d\}, \; 0 < \delta_j \leq 1. \tag{A.20}$$

Therefore, the variance of the predicted coefficient vector expressed in equation (A.13) simplifies to:

$$R_t = S_t C_t^* + \frac{1-\delta}{\delta} S_t C_t^* = \frac{1}{\delta} S_t C_t^*. \tag{A.21}$$

This completes the requisites for forecasting with one model. However, the approach we pursue allows for k candidate predictors and d possible support points for time-variation in coefficients, and therefore, $k.d$ models. Hence, part A.2 of this appendix extends the analysis to deal with these possibilities in a Bayesian model selection and averaging approach.

A.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 $\{\delta_1, \delta_2, ..., \delta_d\}$. Clearly, the mean of the predictive distribution of the forecasts we computed above (see equation (A.11)) is influenced by these specific choices. Hence, the point estimate of Δe_{t+h} is now also conditional on M_i and δ_j:

$$\widehat{\Delta e}_{t+h,i}^{j} = E(\Delta e_{t+h} | M_i, \delta_j, D_t) = X_t' m_t | M_i, \delta_j, D_t. \tag{A.22}$$

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:

$$P(M_i | \delta_j, D_0) = 1/k, \tag{A.23}$$

$$P(\delta_j | D_0) = 1/d. \tag{A.24}$$

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:

$$P(M_i | \delta_j, D_t) = \frac{f(\Delta e_t | M_i, \delta_j, D_{t-h}) P(M_i | \delta_j, D_{t-h})}{f(\Delta e_t | \delta_j, D_{t-h})}, \tag{A.25}$$

where,

$$f(\Delta e_t|\delta_j, D_{t-h}) = \sum_M f(\Delta e_t|M_i, \delta_j, D_{t-h})P(M_i|\delta_j, D_{t-h}). \tag{A.26}$$

The key ingredient is the conditional density:

$$f(\Delta e_t|M_i, \delta_j, D_{t-h}) \sim \frac{1}{\sqrt{Q^j_{t,i}}}\mathbf{t}_{n_{t-1}}\left(\frac{\Delta e_t - \Delta e^j_{t,i}}{\sqrt{Q^j_{t,i}}}\right), \tag{A.27}$$

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 (A.10)). The prediction of the average model for each of the specific value of $\delta = \delta_j$ is given by:

$$\widehat{\Delta e}^j_{t+h} = \sum_{i=1}^{k} P(M_i|\delta_j, D_t)\widehat{\Delta e}^j_{t+h,i}. \tag{A.28}$$

Essentially, for each specific δ, it is the sum of the exchange rate prediction of each of the k models weighted by their posterior probability. If there was only one support point for time-variation in coefficients, such that $d = 1$, then equation (A.28) would complete the averaging approach. However, we are considering several possibilities for δ, hence we also perform Bayesian averaging over these values.

Starting with the prior probability in equation (A.24), the posterior probability of a specific δ is:

$$P(\delta_j|D_t) = \frac{f(\Delta e_t|\delta_j, D_{t-h})P(\delta_j|D_{t-h})}{\sum_\delta f(\Delta e_t|\delta, D_{t-h})P(\delta|D_{t-h})}. \tag{A.29}$$

We note that using this probability, we can infer the degree of time-variation in coefficients supported by the data (see also equation (14) in the main text).

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

$$P(M_i, \delta_j|D_t) = P(M_i|\delta_j, D_t)P(\delta_j|D_t), \tag{A.30}$$

and the unconditional average prediction of the average model,

$$\widehat{\Delta e}_{t+h} = \sum_{j=1}^{d} P(\delta_j|D_t)\widehat{\Delta e}^j_{t+h}. \tag{A.31}$$

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

B Model Space

This Appendix outlines the fundamentals implied by Taylor rules under various assumptions. Since some specifications incorporate Financial Condition Indexes (FCIs), the Appendix also describes the methodology employed in their construction.

B.1 Taylor Rules Specifications and Implied Interest Rate Differentials

Taylor (1993) suggested the following rule for monetary policy:

$$i_t^T = \pi_t + \tau_1(\pi_t - \pi^T) + \tau_2\overline{y}_t + r^T, \tag{B.1}$$

where, i_t^T,is the target for the nominal short-term interest rate set by the central bank; π_t is the inflation rate; π^T, is the target inflation rate; $\overline{y}_t = (y_t - y_t^p)$, is the output gap measured as deviation of actual GDP level (y_t) from its potential (y_t^p); and r^T,is the equilibrium real interest rate. In equation (B.1) the central bank rises the short-term interest rate when inflation is above the target and/or output is above its potential level. In Taylor's (1993) formulation, $\tau_1 = 0.5$, $\pi^T = 2\%$, $\tau_2 = 0.5$, and $r^T = 2\%$.

Rearranging equation (B.1) by aggregating the constant terms and the inflation terms yields:

$$i_t^T = \lambda_0 + \lambda_1\pi_t + \lambda_2\overline{y}_t, \tag{B.2}$$

where, $\lambda_0 = r^T - \tau_1\pi^T$; $\lambda_1 = (1+\tau_1)$; and $\lambda_2 = \tau_2$. Note that, since $\lambda_1 = (1+\tau_1)$, a raise in inflation by 1%, induces the central bank to increase the short-term nominal interest rate by a magnitude superior to 1%, satisfying the Taylor rule principle. If the central bank targets the real exchange rate q_t, as suggested in Clarida *et al.* (1998), it follows the following rule:

$$i_t^T = \lambda_0 + \lambda_1\pi_t + \lambda_2\overline{y}_t + \lambda_3 q_t, \tag{B.3}$$

where, $q_t = e_t + p_t^* - p_t$; e_t is the log exchange rate (home price of foreign currency); p_t is the log price level; and asterisk (*) denotes a foreign country's variable. Including q_t in the rule suggests that the monetary authority is concerned with exchange rate deviation from the level implied by Purchasing Power Parity (PPP), with an increase in q_t signalling a rise in i_t^T (see Engel and West, 2005).

The empirical evidence in Clarida *et al.* (1998) also suggests that central banks limit volatility in the interest rate by adjusting the current interest rate in a gradual fashion:

$$i_t = (1-\rho)i_t^T + \rho i_{t-1} + \zeta_t. \tag{B.4}$$

Substituting in equation (B.3) and rearranging yields:

$$i_t = \phi_c + \phi_1 \pi_t + \phi_2 \overline{y}_t + \phi_3 q_t + \phi_4 i_{t-1} + \zeta_t \tag{B.5}$$

where, $\phi_c = (1-\rho)\lambda_0$; $\phi_1 = (1-\rho)\lambda_1$; $\phi_2 = (1-\rho)\lambda_2$; $\phi_3 = (1-\rho)\lambda_3$; $\phi_4 = \rho$.

Denote (B.5) as the home country's Taylor rule. The foreign country is the US, and following Clarida *et al.* (1998) and Engel and West (2005), it is assumed that the Federal Reserve Bank does not target the real exchange rate. Hence, its Taylor rule is similar to equation (B.5), except that the real exchange rate is excluded. Subtracting this US Taylor rule from the home country's rule yields the following interest rate differentials:

$$\begin{aligned} i_t - i_t^* = \phi_0 &+ (\phi_1 \pi_t + \phi_2 \overline{y}_t + \phi_3 q_t + \phi_4 i_{t-1}) \\ &- (\phi_1^* \pi_t^* + \phi_2^* \overline{y}_t^* + \phi_4^* i_{t-1}^*) + \mu_t, \end{aligned} \tag{B.6}$$

where the term ϕ_0, is obtained by collecting the constant terms from both, the home and foreign country (ϕ_c and ϕ_c^*); and hence, it constitutes the sum of the terms capturing the equilibrium real interest rates and inflation targets. Consequently, if the equilibrium real interest rates and the inflation targets are equal between the two countries, the constant term is zero.

The essential mechanism that links the policy actions based on Taylor rules to exchange rates is UIRP. Under UIRP and rational expectations, a home country's policy action characterised by an increase (decrease) in its policy rate relative to the foreign, translates into an expected depreciation (appreciation) of the home country's currency, relative to the foreign country. As the rule suggests, the central bank reacts when for example inflation is above the target, or the output gap is widening, or the real exchange rate is above the target. However, the empirical evidence suggests that UIRP fails to hold and this is known as the forward premium puzzle (see, e.g., Rogoff, 1996). Hence, in linking the interest rate differential equation (B.6) to the forecasting regression, whilst we can substitute out the interest rate differentials by the expected change in the exchange rate, we impose no restrictions on the effect of monetary policy.[26]

The Taylor rule outlined in equation (B.6) is generic. In practice, various specifications can be considered based on a number of specific assumptions. We summarise the exact assumptions in Table (B.1). Here, we outline the general assumptions:

- First, we can assume that the equilibrium real interest rate and the inflation target of the home and foreign country are identical, such that their aggregate value captured by the constant term ϕ_0 in equation (10), is zero. Thus, we can derive sets of fundamentals from Taylor rules with or without the constant term.

[26] See also Molodtsova and Papell (2009) and Engel and West (2005) for more discussion on the links between monetary policy and exchange rates.

- Second, it can be further imposed that the coefficients on variables commonly targeted by both countries are identical. In equation (B.6), for instance, this implies that $\phi_1 = \phi_1^*;\phi_2 = \phi_2^*;\ \phi_4 = \phi_4^*$. Hence, apart from the specifications derived under the first assumption we can also have sets of fundamentals from Taylor rules with homogeneous or heterogeneous coefficients.

- Third, the monetary authority can disregard the partial adjustment of the interest rate towards its target. In this case, the Taylor rule do not allow for interest rate smoothing, consequently, lagged interest rates from both countries are excluded from the specification in (B.6). An additional set of fundamentals can then be constructed with the above characteristics, but also with smoothing or non-smoothing.

- Fourth, the home central bank might not be concerned with exchange rate deviation from its equilibrium level. In this setting, the home and foreign central bank are symmetric in the sense that both countries exhibit similar targets. As a consequence, we can obtain Taylor rule implied fundamentals from asymmetric or symmetric rules.

- Fifth, following the recent financial turmoil, it has been suggested that apart from inflation and output gap, central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook (Taylor, 2008; Mishkin, 2010). Hence, we can exclude the real exchange rate and the lagged interest rate in equation (B.6), and augment with indicators of financial conditions. This provides us with specifications excluding or including financial condition indexes (FCIs).[27]

[27] The methods we use to construct these FCIs are described in part B.2 of this Appendix.

Table B.1: Interest Rate Differentials Implied by Taylor Rules

Assumption	**Interest rate differentials specification**
TR1. Homogeneous rule, symmetric, and without interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii) Central banks do not smooth interest rate; and (v) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \mu_t$
TR2. Homogeneous rule, symmetric with smoothing (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii) Central banks smooth interest rates at an identical magnitude; and (v) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\alpha_3 = \phi_4 = \phi_4^*$.	$i_t - i_t^* = \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_3(i_{t-1} - i_{t-1}^*) + \mu_t$
TR3. Homogeneous rule, asymmetric and without interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii) Central banks do not smooth interest rate; (v) The home central bank targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \phi_3 q_t + \mu_t$
TR4. Homogeneous rule, asymmetric and with interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation, the output gap and the interest rate smoothing are equal between home and foreign country; and (iii) The home central bank targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\alpha_3 = \phi_4 = \phi_4^*$.	$i_t - i_t^* = \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_3(i_{t-1} - i_{t-1}^*) + \phi_3 q_t + \mu_t$

Table B.1: Interest Rate Differentials Implied by Taylor Rules

Assumption	**Interest rate differentials specification**
TR5. Homogeneous rule, symmetric, with FCI and no constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii) Central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook. In eq. (B.6): $\phi_0 = 0$; $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_5(fci_t - fci_t^*) + \mu_t$
TR6. Heterogeneous rule, symmetric and without interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are allowed to differ between home and foreign country; (iii) Central banks do not smooth interest rates; and (iv) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_1\pi_t - \phi_1^*\pi_t^* + \phi_2\overline{y}_t - \phi_2^*\overline{y}_t^* + \mu_t$
TR7. Heterogeneous rule, symmetric and with interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are allowed to differ between home and foreign country; (iii) Central banks smooth interest rates; and (iv) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\phi_3 = 0$.	$i_t - i_t^* = \phi_1\pi_t - \phi_1^*\pi_t^* + \phi_2\overline{y}_t - \phi_2^*\overline{y}_t^* + \phi_4 i_{t-1} - \phi_4^* i_{t-1}^* + \mu_t$
TR8. Heterogeneous rule, asymmetric and without interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are allowed to differ between home and foreign country; (iii) Central banks do not smooth interest rate; and (iv) The home central bank targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_1\pi_t - \phi_1^*\pi_t^* + \phi_2\overline{y}_t - \phi_2^*\overline{y}_t^* + \phi_3 q_t + \mu_t$

Table B.1: Interest Rate Differentials Implied by Taylor Rules

Assumption	**Interest rate differentials specification**
TR9. Heterogeneous rule, asymmetric and with interest rate smoothing. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are allowed to differ between home and foreign country; (iii) Central banks smooth interest rates; and (iv) The home central bank targets the real exchange rate. In eq. (B.6): $\phi_0 = 0$*;*	$i_t - i_t^* = \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_4 i_{t-1} - \phi_4^* i_{t-1}^* + \phi_3 q_t + \mu_t$
TR10. Heterogeneous rule, symmetric, with FCI and no constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are identical, hence their difference is zero; (ii) The coefficients on inflation and the output gap are allowed to differ between home and foreign country; (iii) Central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook. In eq. (B.6): $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_5 fci_t - \phi_5^* fci_t^* + \mu_t$
TR11. Homogeneous rule, symmetric, without interest rate smoothing and includes a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii); Central banks do not smooth interest rates; and (v) None of the two central banks targets the real exchange rate. In eq. (B.6): $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \mu_t$
TR12. Homogeneous rule, symmetric with interest rate smoothing and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii) Central banks smooth interest rates at an identical magnitude; and (v) None of the two central banks targets the real exchange rate. In eq. (B.6): $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\alpha_3 = \phi_4 = \phi_4^*$.	$i_t - i_t^* = \phi_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_3(i_{t-1} - i_{t-1}^*) + \mu_t$

Table B.1: Interest Rate Differentials Implied by Taylor Rules

Assumption	**Interest rate differentials specification**
TR13. Homogeneous rule, asymmetric, without interest rate smoothing, and includes a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii); Central banks do not smooth interest rates; (v) The home central bank targets the real exchange rate. In eq. (B.6): $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \phi_3 q_t + \mu_t$
TR14. Homogeneous rule, asymmetric, with interest rate smoothing and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation, the output gap and the interest rate smoothing are equal between home and foreign country; (iii) Central banks smooth interest rates at an identical magnitude; and (iv) The home central bank targets the real exchange rate. In eq. (B.6): $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\alpha_3 = \phi_4 = \phi_4^*$.	$i_t - i_t^* = \phi_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_3(i_{t-1} - i_{t-1}^*) + \phi_3 q_t + \mu_t$
TR15. Homogeneous rule, symmetric, with FCI and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are equal between home and foreign country; (iii); Central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook. In eq. (B.6): $\alpha_1 = \phi_1 = \phi_1^*$; $\alpha_2 = \phi_2 = \phi_2^*$; $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_0 + \alpha_1(\pi_t - \pi_t^*) + \alpha_2(\overline{y}_t - \overline{y}_t^*) + \alpha_5(fci_t - fci_t^*) + \mu_t$
TR16. Heterogeneous rule, symmetric and without interest rate smoothing, and includes a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are also allowed to differ between home and foreign country; (iii) Central banks do not smooth interest rates; and (iv) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$	$i_t - i_t^* = \phi_0 + \phi_1\pi_t - \phi_1^*\pi_t^* + \phi_2\overline{y}_t - \phi_2^*\overline{y}_t^* + \mu_t$

Table B.1: Interest Rate Differentials Implied by Taylor Rules

Assumption	**Interest rate differentials specification**
TR17. Heterogeneous rule, symmetric and with interest rate smoothing and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are also allowed to differ between home and foreign country; (iii) Central banks smooth interest rates; and (iv) None of the two central banks targets the real exchange rate. In eq. (B.6): $\phi_3 = 0$.	$i_t - i_t^* = \phi_0 + \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_4 i_{t-1} - \phi_4^* i_{t-1}^* + \mu_t$
TR18. Heterogeneous rule, asymmetric and without interest rate smoothing and includes a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are also allowed to differ between home and foreign country; (iii) Central banks do not smooth interest rates; and (iv) The home central bank targets the real exchange rate. In eq. (B.6): $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_0 + \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_3 q_t + \mu_t$
TR19. Heterogeneous rule, asymmetric and with interest rate smoothing and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are also allowed to differ between home and foreign country; (iii) Central banks smooth interest rates; and (iv) The home central bank targets the real exchange rate.	$i_t - i_t^* = \phi_1 + \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_4 i_{t-1} - \phi_4^* i_{t-1}^* + \phi_3 q_t + \mu_t$
TR20. Heterogeneous rule, symmetric, with FCI and a constant. (i) The equilibrium real interest rate and the inflation target of the home and foreign country are allowed to differ; (ii) The coefficients on inflation and the output gap are also allowed to differ between home and foreign country; (iii) Central banks react to financial market conditions, insofar as they signal deterioration in the economic outlook. In eq. (B.6): $\phi_3 = 0$; $\phi_4 = \phi_4^* = 0$.	$i_t - i_t^* = \phi_0 + \phi_1 \pi_t - \phi_1^* \pi_t^* + \phi_2 \overline{y}_t - \phi_2^* \overline{y}_t^* + \phi_5 fci_t - \phi_5^* fci_t^* + \mu_t$

B.2 Construction of Financial Condition Indexes

To construct the Financial Condition Indexes (FCIs), we use the Time-varying Parameters Factor-Augmented VAR (TVP-FAVAR) model as in Koop and Korobilis (2014). More precisely, let $x_t$denote the set of variables to be used in constructing the FCI and Y_ta set of macroeconomic variables for which we want to purge their effect from the FCI. Then, the TVP-FAVAR has the following compact representation:

$$\begin{aligned} x_t &= \digamma_t \gamma_t + \varrho_t, \qquad \varrho_t \sim N(0, V_t^{\varrho}), & \text{(B.7)} \\ \digamma_t &= \digamma_{t-1} \beta_t + \xi_t, \qquad \xi_t \sim N(0, Q_t^{\xi}), & \text{(B.8)} \\ \gamma_t &= \gamma_{t-1} + \iota_t, \qquad \iota_t \sim N(0, W_t^{\iota}), & \text{(B.9)} \\ \beta_t &= \beta_{t-1} + \eta_t, \qquad \eta_t \sim N(0, R_t^{\eta}), & \text{(B.10)} \end{aligned}$$

where, $\digamma_t = \begin{bmatrix} Y_t \\ f_t \end{bmatrix}$; $\gamma_t = \left(\left(\gamma_t^Y\right)', \left(\gamma_t^f\right)' \right)'$; $\beta_t = (B_{0t}', vec(B_t)')$.[28] This notation suggests that γ_t includes the time-varying regression coefficients γ_t^Y and factor loadings γ_t^f, with the factors that constitute our FCI denoted by f_t; and β_t contains the time-varying VAR coefficients and intercepts. Equations (B.9) and (B.10) indicate that these coefficients follow a random walk process. The error terms are all independent and identically distributed, with mean zero and time-varying variance-covariance matrices.

Allowing parameters to change over time renders more flexibility to the model, and Koop and Korobilis (2014) discuss the importance of such flexibility in constructing FCIs. One of their results is that FCIs constructed with models of this sort, can potentially better characterise turning points in financial conditions.

The inclusion of $Y_t$in the above model has also given way to our intention of constructing FCIs that capture true financial shocks as suggested by Hatzius *et al.* (2010). In the cases we consider, $Y_t$contains the change in the log of the industrial production index, inflation and the short-term interest rate. Since the FCIs we construct are used to augment Taylor rules which already incorporate the information from the three variables above, purging these variables allows us to interpret the FCIs as providing marginal information, i.e., beyond that already embodied in the standard Taylor rule.

Estimation is implemented through a two-step algorithm. In the first step, conditional on the principal components estimate of f_t, denoted $\widetilde{f}_t$, the parameters of the TVP-FAVAR are estimated. Thus in this step we have $\widetilde{\digamma}_t = \begin{bmatrix} Y_t \\ \widetilde{f}_t \end{bmatrix}$. In the second step, conditional on the estimated TVP-FAVAR parameters, the Kalman filter is employed to obtain the estimate of f_t(i.e., the FCI). We describe these steps

[28] We can equivalently write equations (B.7) and (B.8) as: $x_t = \gamma_t^Y Y_t + \gamma_t^f f_t + \varrho_t$ and $\begin{bmatrix} Y_t \\ f_t \end{bmatrix} = B_{0t} + B_t \begin{bmatrix} Y_{t-1} \\ f_{t-1} \end{bmatrix} + \xi_t$.

next.[29]

The first step begins by defining the prior hyperparameters and initial conditions for the unknown parameters before observing the data. Thus, at $t = 0$we define:

$$\begin{aligned} f_0 &\sim N(0, 10), \\ \gamma_0 &= N(0, 10), \\ \beta_0 &= N(0, V_{MIN}), \\ V_0^{\varrho} &= 1 \times I, \\ Q_0^{\xi} &= 1 \times I, \end{aligned} \tag{B.11}$$

where, V_{MIN} is a Minnesota type prior which shrinks coefficients associated with the more distant lags according to:

$$V_{MIN} = \begin{Bmatrix} 4, \text{ for intercepts} \\ 4/\rho^2, \text{ for the coefficient on lag } \rho \end{Bmatrix}$$

and $\rho = 1, ..., 12$ lags. Note that the approach in Koop and Korobilis (2014) involves using Exponentially Weighted Moving Average (EWMA) estimators for the V_t^{ϱ} and Q_t^{ξ} matrices, while W_t^{ι} and R_t^{η} are proportional to their covariance matrices obtained from the Kalman filter.

The step proceeds with extraction of the principal components estimate of the factor, $\widetilde{f}_t$ to obtain $\widetilde{F}_t = \begin{bmatrix} Y_t \\ \widetilde{f}_t \end{bmatrix}$. Conditional on this information set, and given the initial conditions, estimate γ_t , β_t , V_t^{ϱ} and Q_t^{ξ}, using the Kalman filter for $t = 1, ..., T$. That is, predict the state vectors γ_t, β_t, and their associated covariance matrices at time t, conditional on information up to $t-1$,(i.e., D_{t-1}):

$$\gamma_t | D_{t-1} \sim N\left(\gamma_t|_{t-1}, \sum\nolimits_{t|t-1}^{\gamma}\right), \tag{B.12}$$

$$\beta_t | D_{t-1} \sim N\left(\beta_t|_{t-1}, \sum\nolimits_{t|t-1}^{\beta}\right), \tag{B.13}$$

where from Kalman filter predicting step,

$$\gamma_t|_{t-1} = \gamma_{t-1}|_{t-1}, \text{ and } \sum\nolimits_{t|t-1}^{\gamma} = \sum\nolimits_{t-1|t-1}^{\gamma} + \widehat{W}_t^{\iota}; \tag{B.14}$$

$$\beta_t|_{t-1} = \beta_{t-1}|_{t-1}, \text{ and } \sum\nolimits_{t|t-1}^{\beta} = \sum\nolimits_{t-1|t-1}^{\beta} + \widehat{R}_t^{\eta}. \tag{B.15}$$

The error covariance are estimated using a forgetting factor approach:

$$\widehat{W}_t^{\iota} = (1 - L_3^{-1}) \sum\nolimits_{t-1|t-1}^{\gamma}, \tag{B.16}$$

$$\widehat{R}_t^{\eta} = (1 - L_4^{-1}) \sum\nolimits_{t-1|t-1}^{\beta}. \tag{B.17}$$

with $L_3 = L_4 = 0.99$. According to the specifications in (B.16) and (B.17), shocks to the coefficients at time t are proportional to the $t-1$ variance, by a factor of

[29] In estimating the VARs we impose a maximum lag of 12 months.

$(1-L_3^{-1})$ and $(1-L_4^{-1})$, respectively. Hence, the value we set for L_3 and L_4, implies a slow time-variation in parameters.

To refine inference about the estimate of the state variables in the Kalman filter updating step, first obtain the prediction errors:

$$\widehat{\varrho}_{b,t} = x_{b,t} - \widetilde{F}_t \gamma_{b,t}|_{t-1}, \text{ and} \tag{B.18}$$
$$\widehat{\xi}_t = \widetilde{F}_t - \widetilde{F}_{t-1}\beta_t|_{t-1}. \tag{B.19}$$

where b, denotes each of the variable used to estimate the factors (i.e., contained in x_t). Then, estimate V_t^{ϱ} and Q_t^{ξ} using EWMA estimators:

$$\widehat{V}_{b,t}^{\varrho} = L_1 V_{b,t-1|t-1}^{\varrho} + (1-L_1)\widehat{\varrho}_{b,t}\widehat{\varrho}_{b,t}', \tag{B.20}$$
$$\widehat{Q}_t^{\xi} = L_2 Q_{t-1|t-1}^{\xi} + (1-L_2)\widehat{\xi}_t\widehat{\xi}_t'. \tag{B.21}$$

with $L_1 = L_2 = 0.96$. With these elements we can now refine inference about γ_t and β_t at time t. More precisely, for each element in x_t we have for γ_t,

$$\gamma_{b,t}|D_t \sim N\left(\gamma_{b,t|t}, \sum\nolimits_{bb,t|t}^{\gamma}\right), \tag{B.22}$$

where, $\gamma_{b,t|t} = \gamma_{b,t|t-1} + \sum_{bb,t|t-1}^{\gamma}\widetilde{F}_t'\left(\widehat{V}_{bb,t}^{\varrho} + \widetilde{F}_t\sum_{bb,t|t-1}^{\gamma}\widetilde{F}_t'\right)^{-1}(x_t - \widetilde{F}_t\gamma_t|_{t-1})$, and $\sum_{bb,t|t}^{\gamma} = \sum_{bb,t|t-1}^{\gamma} - \sum_{bb,t|t-1}^{\gamma}\widetilde{F}_t'\left(\widehat{V}_{bb,t}^{\varrho} + \widetilde{F}_t\sum_{bb,t|t-1}^{\gamma}\widetilde{F}_t'\right)^{-1}\widetilde{F}_t\sum_{bb,t|t-1}^{\gamma}$. And similarly for β_t,

$$\beta_t|D_t \sim N\left(\beta_{t|t}, \sum\nolimits_{t|t}^{\beta}\right), \tag{B.23}$$

where, $\beta_{t|t} = \beta_{t|t-1} + \sum_{t|t-1}^{\beta}\widetilde{F}_{t-1}'\left(\widehat{Q}_t^{\xi} + \widetilde{F}_{t-1}\sum_{t|t-1}^{\beta}\widetilde{F}_{t-1}'\right)^{-1}(\widetilde{F}_t - \widetilde{F}_{t-1}\widehat{\beta}_{t|t-1})$, and $\sum_{t|t}^{\beta} = \sum_{t|t-1}^{\beta} - \sum_{t|t-1}^{\beta}\widetilde{F}_{t-1}'\left(\widehat{Q}_t^{\xi} + \widetilde{F}_{t-1}\sum_{t|t-1}^{\beta}\widetilde{F}_{t-1}'\right)^{-1}\widetilde{F}_{t-1}\sum_{t|t-1}^{\beta}$. Also update V_t^{ϱ} and Q_t^{ξ} at time t:

$$V_{b,t|t}^{\varrho} = L_1 V_{b,t-1|t-1}^{\varrho} + (1-L_1)\widehat{\varrho}_{b,t|t}\widehat{\varrho}_{b,t|t}', \tag{B.24}$$
$$Q_{t|t}^{\xi} = L_2 Q_{t-1|t-1}^{\xi} + (1-L_2)\widehat{\xi}_{t|t}\widehat{\xi}_{t|t}'. \tag{B.25}$$

where, $\widehat{\varrho}_{b,t|t} = x_{b,t} - \widetilde{F}_t\gamma_{b,t|t}$ and $\widehat{\xi}_{t|t} = \widetilde{F}_t - \widetilde{F}_{t-1}\beta_{t|t}$.

The second step involves obtaining smoothed estimates of γ_t, β_t, V_t^{ϱ}, and Q_t^{ξ}, for $t = T-1, ..., 1$. For γ_t and β_t, the updating scheme uses the fixed interval smoother. That is, given information at $t+1$, each $\gamma_{b,t}$ is updated using:

$$\gamma_{b,t}|D_T \sim N\left(\gamma_{b,t|t+1}, \sum\nolimits_{bb,t|t+1}^{\gamma}\right), \tag{B.26}$$

where D_T indicates that we are conditioning on information up to time T, and

$$\gamma_{b,t|t+1} = \gamma_{b,t|t} + G_t^{\gamma}\left(\gamma_{b,t+1|t+1} - \gamma_{b,t+1|t}\right), \tag{B.27}$$

$$\sum\nolimits_{bb,t|t+1}^{\gamma} = \sum\nolimits_{bb,t|t}^{\gamma} + G_t^{\gamma}\left(\sum\nolimits_{bb,t+1|t+1}^{\gamma} - \sum\nolimits_{bb,t+1|t}^{\gamma}\right) G_t^{\gamma\prime}, \tag{B.28}$$

$$G_t^{\gamma} = \sum\nolimits_{bb,t|t}^{\gamma}\left(\sum\nolimits_{bb,t+1|t}^{\gamma}\right)^{-1}. \tag{B.29}$$

Similarly, β_t is updated using:

$$\beta_t|D_T \sim N\left(\beta_{t|t+1}, \sum\nolimits_{t|t+1}^{\beta}\right), \tag{B.30}$$

where,

$$\beta_{t|t+1} = \beta_{t|t} + G_t^{\beta}\left(\beta_{t+1|t+1} - \beta_{t+1|t}\right), \tag{B.31}$$

$$\sum\nolimits_{t|t+1}^{\beta} = \sum\nolimits_{t|t}^{\beta} + G_t^{\beta}\left(\sum\nolimits_{t+1|t+1}^{\beta} - \sum\nolimits_{t+1|t}^{\beta}\right) G_t^{\beta\prime}, \tag{B.32}$$

$$G_t^{\beta} = \sum\nolimits_{t|t}^{\beta}\left(\sum\nolimits_{t+1|t}^{\beta}\right)^{-1}. \tag{B.33}$$

The smoothed estimates of V_t^{ϱ}, and Q_t^{ξ} given information at $t+1$ are updated using:

$$\left(V_{t|t+1}^{\varrho}\right)^{-1} = L_1\left(V_{t|t}^{\varrho}\right)^{-1} + (1 - L_1)\left(V_{t+1|t+1}^{\varrho}\right)^{-1}, \tag{B.34}$$

$$\left(Q_{t|t+1}^{\xi}\right)^{-1} = L_2\left(Q_{t|t}^{\xi}\right)^{-1} + (1 - L_2)\left(Q_{t+1|t+1}^{\xi}\right)^{-1}. \tag{B.35}$$

Finally, conditional on the estimates of γ_t, β_t, V_t^{ϱ}, and Q_t^{ξ} as described above, the f_t is estimated using the standard Kalman filter and smoother.

C Data Appendix

This Appendix describes the data used in the empirical section, including for constructing the Financial Condition Indexes (FCIs). The sample period is 1973M1:2013M5, for nine countries. Table C.1 describes the exchange rate data, as well as the data used to compute the various sets of fundamentals. For each country in the first column, the Table indicates the source of information for each variable in the subsequent columns.

Table C.2 specifies the data used to construct the FCIs. For each country in the first column, the following columns describe the variables used, the transformation code applied to each series (1 = level, 2 = first difference, 5 = first difference of logarithm), the data source, and the sample size.

Table C.1: Data Used to Forecast

Country	**Nominal exchange rate (National currency/USD)**	**Industrial prod. index, NSA, 2005=100**	**Money supply, NSA, National currency (10ˆ9)**	**Short-term nominal interest rate (%)**	**Consumer price index NSA, 2005=100**
Canada	IFS, 156..AE.ZF	IFS, 15666..CZF	M1, OECD, MEI	IFS, 15660B..ZF	IFS, 15664...ZF
Denmark	IFS, 128..AE.ZF	IFS, 12866..BZF	M1, OECD, MEI	IFS, 12860...ZF	IFS, 12864...ZF
UK	IFS, 112..AE.ZF	IFS, 11266..CZF	M4, Bank of England	IFS, 11260...ZF	IFS, 11264B..ZF
Japan	IFS, 158..AE.ZF	IFS, 15866..CZF	M1, OECD, MEI	IFS, 15860B..ZF	IFS, 15864...ZF
Korea	IFS, 542..AE.ZF	IFS, 54266..CZF	M1, OECD, MEI	IFS, 54260B..ZF	IFS, 54264...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
U.S.		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 national currency per U.S. dollar. 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.

Table C.2: Data Used to Construct FCIs

Country	Description	Code	Source	Sample
Canada	S&P/TSX Composite Index	5	DataStream (TTOCOMP)	1973M1-2013M5
	Exchange Rate Index 2005=100	5	IFS (..AHXZF)	1973M1-2013M5
	3m LIBOR, Canadian Dollar/3m TBill Rate	1	IFS (60C..ZF)/FRED (CAD3MTD156N)	1990M5-2013M5
	Prime Corp Paper 90DA/3m TBill Rate Spread	1	IFS(60BC.ZF)	1973M1-2013M5
	5y Mortgage Rate/5y Gov. Bond Spread	1	Bank of Canada/IFS(61A..ZF)	1973M1-2013M5
	New Housing Price Index	5	Statistics Canada	1981M1-2013M5
	Opinion on Overall Business - Lending Conditions	1	Bank of Canada's Loan Survey	1999M6-2013M5
Denmark	OMX Copenhagen (OMXC20) Price Index	5	DataStream (DKKFXIN)	1989M12-2013M5
	Exchange Rate Index 2005=100	5	IFS (..AHXZF)	1973M1-2013M5
	3m CIBOR/CD Interest Rate Spread	1	Danish Central Bank	1992M4-2013M5
	20y Mortgage Yield/10y Gov. Bond Spread	1	IFS (61A..ZF)	1973M1-2013M5
	New Loans for Mortgage	5	Statistics Denmark	1994M1-2013M5
	Credit Standards - Competition: Present Quarter	1	Danish Central Bank (Lending Survey)	2008M12-2013M5
UK	FTSE all Share Price Index	5	DataStream (FTALLSH)	1973M1-2013M5
	FTSE 100 Price Index	5	DataStream (FTSE100)	1978M1-2013M5
	Average Effective Exchange Rate Index (2005 = 100)	5	Bank of England (XUMABK67)	1980M1-2013M5
	Mix-adjusted Average House Prices-1st Time Buyer	5	UK Office of National Statistics	2002M2-2013M5
	3m Libor/3m TBill Rate Spread	1	IFS (60C..ZF)/IFS (60EA.ZF)	1978M1-2013M5
	5y Mortgage Rate/5y Gov. Bond Spread	1	Bank of England/IFS (61A..ZF)	1995M1-2013M5
	10y Gov. Bond Spread/5y Gov. Bond Spread	1	IFS (61...ZF)	1973M1-2013M5
	How have overall unsecured lending spreads changed? Past three months	1	Bank of England	2007M4-2013M5
	How have spreads on loans to medium PNFCs changed? Past three months	1	Bank of England	2007M4-2013M5
Japan	NIKKEI 225 Stock Price Index	5	DataStream (JAPDOWA)	1973M1-2013M5
	TOPIX Price Index	5	DataStream (TOKYOSE)	1973M1-2013M5
	Exchange Rate Index 2005=100	5	IFS (..RECZF)	1980M1-2013M5
	Lending Rate/3 Month Bill rate	1	IFS (60P..ZF)/IFS (60CS.ZF)	1973M1-2013M5
	Rate on Certificate of Deposit/3m TBill Rate Spread	1	IFS (60LA.ZF)	1979M5-2013M5
	3m Libor Yen/3m TBill Rate Spread	1	IFS (60EA.ZF)	1978M9-2013M5

Table C.2: Data Used to Construct FCIs

Country	Description	Code	Source	Sample
	10y Gov. Bond/3y Gov. Bond Spread	1	FRED(IRLTLT01JPM156N)/IFS(61..ZF)	1989M1-2013M5
	How has demand from households for housing loans changed?	1	Bank of Japan	2001M3-2013M5
	How has demand from households for consumer loans changed	1	Bank of Japan	2001M3-2013M5
	Change in bank's credit standards for approving applications from large and medium firms	1	Bank of Japan	2001M3-2013M5
Korea	KOREA SE Composite Price Index	5	DataStream (KORCOMP)	1973M1-2013M5
	Exchange Rate Index 2005=100	5	IFS (..RELZF)	1973M1-2013M5
	Loans of CBs & SBs	5	Bank of Korea	1973M1-2013M5
	Yields on CD 91 days/Uncollateralized Rate Spread	1	Bank of Korea	1991M3-2013M5
	Yields on Comm. Paper 91 days/Uncollateralized Call Rate Spread	1	Bank of Korea	1994M9-2013M5
	Yields of National Housing 5y Bonds/5y Gov. Bonds Spread	1	Bank of Korea	1995M5-2013M5
	Yields of National Housing 5y Bonds/3y Corporate Bonds, AA Spread	1	Bank of Korea	1987M1-2013M5
	Yields on 10y Gov. Bonds/Yields on 5y Gov. Bonds Spread	1	Bank of Korea	2000M10-2013M5
	Overall Credit Risk	1	Bank of Korea	2002M3-2013M5
	Business Condition BSI All Industries Index	5	Bank of Korea	2003M1-2013M5
	Economic Sentiment Index	5	Bank of Korea	2003M1-2013M5
	Consumer Survey: Expectations of Domestic Economic Situation Index	5	Bank of Korea	1995M9-2013M5
	Consumer Survey: Durable Goods Expense Plan Index	5	Bank of Korea	2004M9-2013M5
Norway	10y Gov. Bond/3m TBill rate Spread	1	Norge Bank	1973M1-2013M5
	OSLO SE OBX Price Index	5	DataStream (OSLOOBX)	1987M1-2013M5
	Exchange Rate Index 2005=100	5	IFS (..AHXZF)	1975M1-2013M5
	Default assessment in the past three months	1	Norge Bank	2007M10-2013M5
Sweden	10y Gov. Bond/3m TBill rate Spread	1	IFS (61ZF)/IFS (60C..ZF)	1973M1-2013M5

Table C.2: Data Used to Construct FCIs

Country	Description	Code	Source	Sample
	OMXS30 Stockholm Price Index	5	Data Stream(SWEDOMX)	1986M1-2013M5
	Exchange Rate Index 2005=100	5	IFS (..AHXZF)	1973M1-2013M5
	3m Libor Swedish Krona/3m TBill Rate Spread	1	IFS (60CZF)/FRED (SEK3MTD156N)	2006M2-2013M5
	Money Supply M3	5	OECD-MEI	1973M1-2013M5
Switzerland	Mortgage Rate/10y Gov. Bond Spread	1	DataStream (S03801)/IFS (61ZF)	1973M1-2013M5
	SMI Price Index	5	DataStream (SWISSMI)	1988M7-2013M5
	Exchange Rate Index 2005=100	5	IFS (..AHXZF)	1973M1-2013M5
	Lending Rate/3m TBill rate Spread	1	IFS (60P..ZF)/IFS (60C..ZF)	1981M1-2013M5
U.S.	S&P 500 Composite Price Index	5	DataStream (S&PCOMP)	1973M1-2013M5
	Nominal Major Currencies Dollar Index	1	FRED-TWEXMMTH	1973M1-2013M5
	Household Sector: Liabilities: Household Credit Market Debt Outstanding	5	FRED-CMDEBT	1973M1-2013M5
	Mortgage rate/10y Gov. Bond Yield Spread	1	IFS (60PA.ZF)/IFS (61...ZF)	1973M1-2013M5
	Issuers of Asset-Backed Securities; Total Liabilities	1	FRED (ABSITCMDODFS)	1983M4-2013M5
	Finance Rate on Consumer Instalment Loans, New Autos 48 Month Loan-2YB	1	FRED (TERMCBAUTO48NS)	1976M6-2013M5
	Libor USD/3m Tbill Rate Spread	1	FRED(USD3MTD156N)/IFS (60CS.ZF)	1986M1-2013M5
	10y Treasury Constant Maturity Rate/2y Treasury Constant Maturity Rate Spread	1	IFS(61ZF)/FRED (GS2)	1976M6-2013M5
	2y Treasury Constant Maturity Rate/3m Tbill Rate Spread	1	FRED (GS2)/IFS (60CS.ZF)	1976M6-2013M5
	3m Commercial Paper/ 3m Tbill Rate Spread	1	IFS (60BC.ZF)	1973M1-2013M5
	BofA Merrill Lynch US High/Moody's BAA Spread	1	FRED (BAMLH0A0HYM2EY), (BAA)	1996M12-2013M5
	Wilshire 5000 Total Market Index	5	FRED (WILL5000IND)	1973M1-2013M5
	CBOE Volatility Index: VIX	1	FRED (VIXCLS)	2004M1-2013M5
	Total Consumer Credit Owned and Securitized, Outstanding	5	FRED (TOTALNS)	1973M1-2013M5
	Michigan Survey: Expected Change in Financial Condition	1	Surveys of Consumers, Univ. of Michigan	1978M1-2013M5

D The Bootstrap

The bootstrap is primarily based on Kilian (1999) and Rogoff and Stavrakeva (2008); but it also includes procedures to account for data-mining as proposed in Inoue and Kilian (2005). More precisely, 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) we also assume cointegration between the exchange rate and fundamentals. For each country we postulate the following DGP under the null of no predictability (the country subscript i, is omitted for simplicity):

$$\Delta e_t = v_t^e, \tag{D.1}$$

$$\Delta z_t = c_0 + t + \Upsilon z_{t-1} + \sum_{\ell=1}^{\ell e} B_\ell^e \Delta e_{t-\ell} + \sum_{\ell=1}^{\ell z} B_\ell^z z_{t-\ell} + v_t^z, \tag{D.2}$$

where, $\Delta e_t = e_t - e_{t-1}$; $\Delta z_t = z_t - z_{t-1}$; c_0 is a constant, t is a trend, and v_t^e and v_t^z are $i.i.d$ error terms error terms with covariance matrix Σ. We first estimate equations (D.1) and (D.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.[30] Subsequently we re-sample with replacement the residuals matrix (v_t^e,v_t^z) in tandem to preserve the contemporaneous correlation in the original sample. We then use the re-sampled residuals to recursively generate pseudo-samples of e_t and z_t. The first 100 observations are discarded to avoid potential bias due to using the sample averages as initial values for the recursions. We then use each of the predictive model (Single Predictor incl./excl. TVar-Coeff., 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 analysed in isolation, but our forecasting approach allows for many potential predictors. Even in the cases where a single predictor enters the regression, we are actually considering and searching over many predictors. To take into account concerns about data-mining we implement the bootstrap in the context of a data-mining environment, following Inoue and Kilian (2005). (See also Rapach and Wohar, 2006 for an application to stock returns). The procedure involves assuming that under the null of no predictability the DGP comprises:

[30] In equation (D.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).

$$
\begin{aligned}
\Delta e_t &= v_t^e, \\
\Delta z_{1,t} &= c_{1,0} + t + \Upsilon_1 z_{1,t-1} + \sum_{\ell=1}^{\ell e} B_{1,\ell}^e \Delta e_{t-\ell} + \sum_{\ell=1}^{\ell z} B_{1,\ell}^z z_{1,t-\ell} + v_{1,t}^z \\
&\vdots \\
\Delta z_{k,t} &= c_{k,0} + t + \Upsilon_k z_{k,t-1} + \sum_{\ell=1}^{\ell e} B_{k,\ell}^e \Delta e_{t-\ell} + \sum_{\ell=1}^{\ell z} B_{k,\ell}^z z_{k,t-\ell} + v_{k,t}^z
\end{aligned}
$$

where, $v_t^e, v_{1,t}^z$,..., $v_{k,t}^z$, are $i.i.d$ error terms with covariance matrix Σ. As the system of equations suggests, we now consider k candidate predictors. Each of the equation is also estimated via OLS. We then follow the same steps as in the bootstrap procedure above, except that for each bootstrap we store the maximal DMW statistic, providing us with an empirical distribution of the maximal statistic. After ordering the empirical distribution for each maximal statistic the 900th, 950th, and 990th values constitute the 10%, 5% and 1% critical values, respectively. The forecast performance of the BMA incl or excl. TVar-Coeff and the BMS incl or excl. TVar-Coeff. are evaluated on the basis of these critical values.[31]

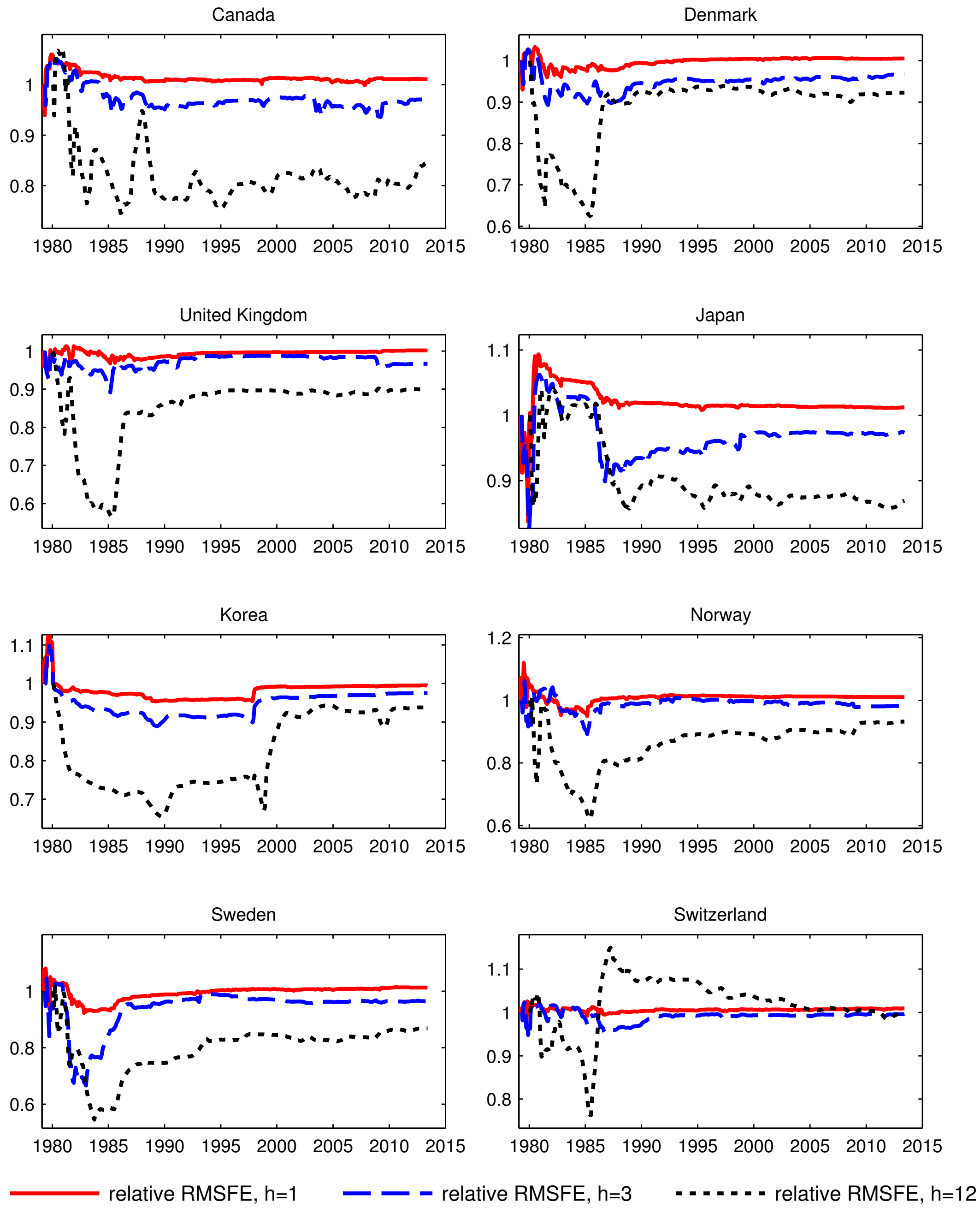
[31] Note that the same procedure can also be applied to construct bootstrapped critical values and p-values for the Clark and West (2006, 2007) test statistic, and the Theil's U-statistic.

E Additional Results Appendix (Not-for-Publication)

In the Paper, we concentrate on a representative selection of countries and omit some figures to save space. This Appendix contains all additional figures. These are:

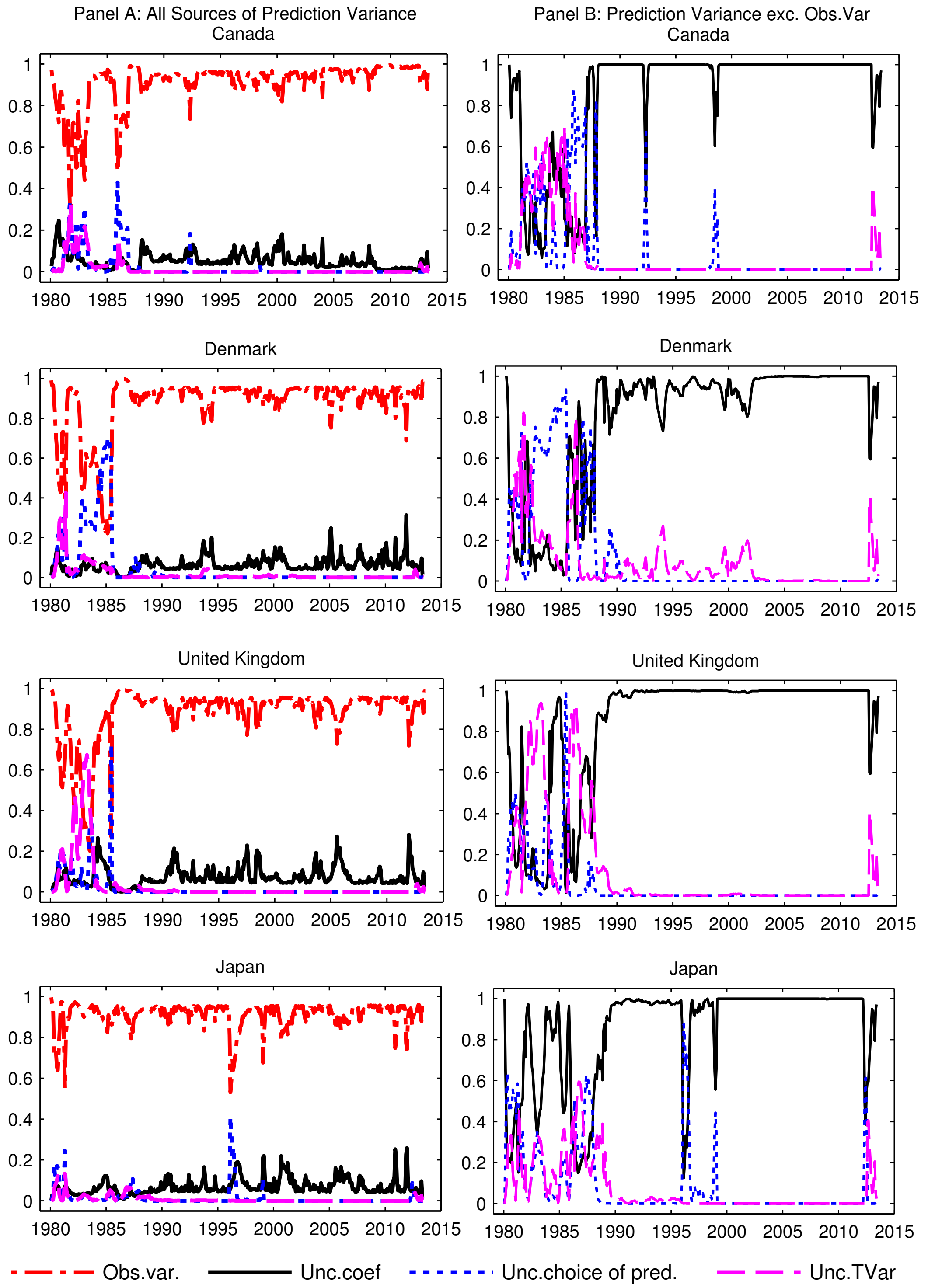
- Figure E.1: RMSFE of the BMA incl. TVar-Coeff. model relative to the RMSFE of the Random Walk without drift
- Figure E.2: Variance Decomposition, h=12
- Figure E.3: Posterior Probabilities of Values of δ, h=1
- Figure E.4: Posterior Probabilities of Values of δ, h=3
- Figure E.5: Posterior Probabilities of Values of δ, h=12
- Figure E.6: Values of δ with the Largest Posterior Probability at Each Period, h=1
- Figure E.7: Values of δ with the Largest Posterior Probability at Each Period, h=3
- Figure E.8: Values of δ with the Largest Posterior Probability at Each Period, h=12
- Figure E.9: Predictors with the Largest Posterior Probability at Each Period, h=1
- Figure E.10: Predictors with the Largest Posterior Probability at Each Period, h=3
- Figure E.11: Predictors with the Largest Posterior Probability at Each Period, h=12
- Figure E.12: Financial Conditions Indexes - Monthly Frequency

Figure E.1: RMSFE of the BMA incl. TVar-Coeff. Model relative to the RMSFE of the Random Walk



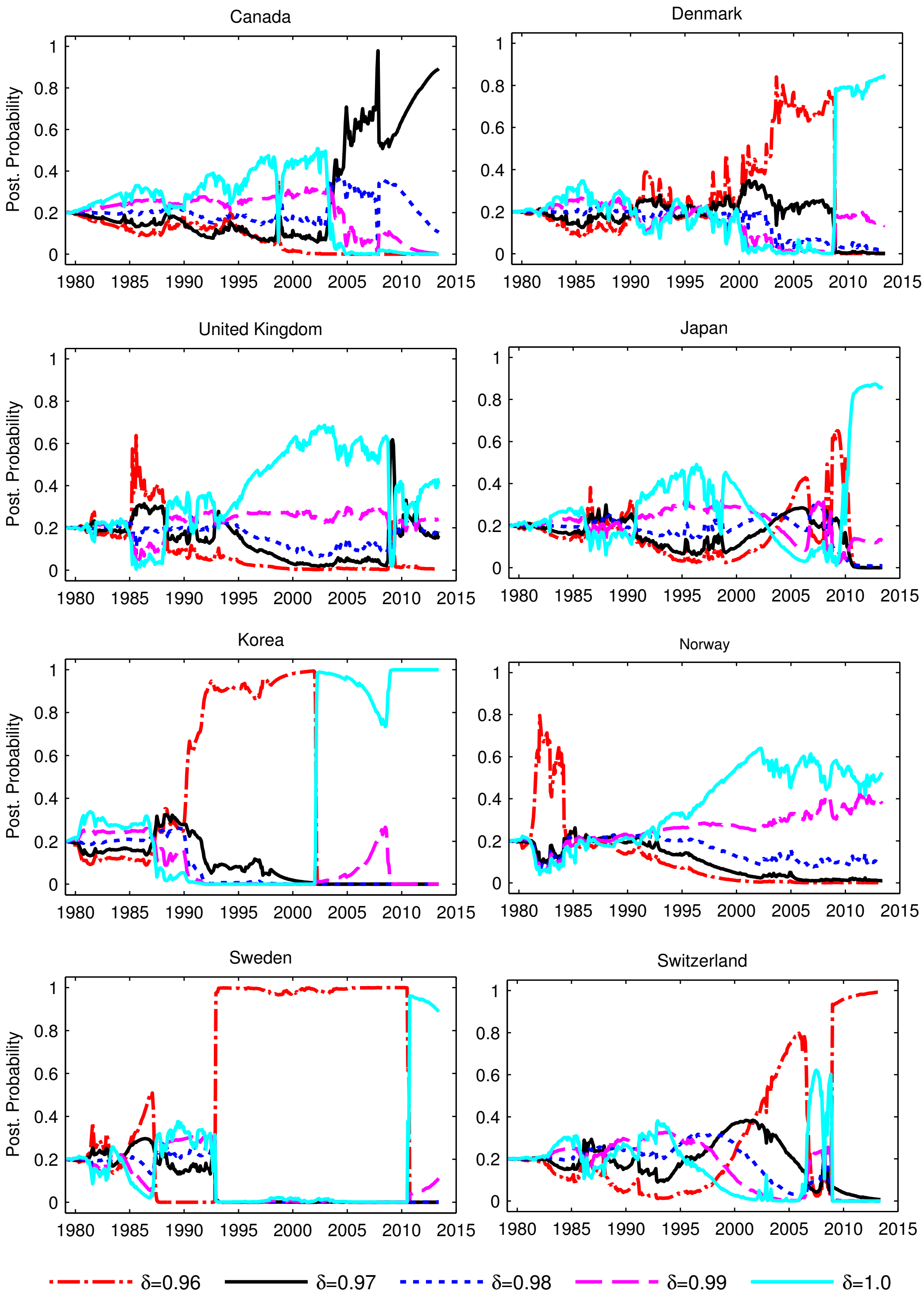
Notes: Cumulative ratio of the Root Mean Squared Forecast Error (RMSFE) of the models that allow for predictors and coefficients to change over time (BMA incl. TVar-Coeff.) relative to the RMSFE of the driftless random walk (RW) benchmark. Values less than one are consistent with a better forecasting performance of the BMA incl. TVar-Coeff. model relative to the RW.

Figure E.2: Sources of Prediction Variance, h=12



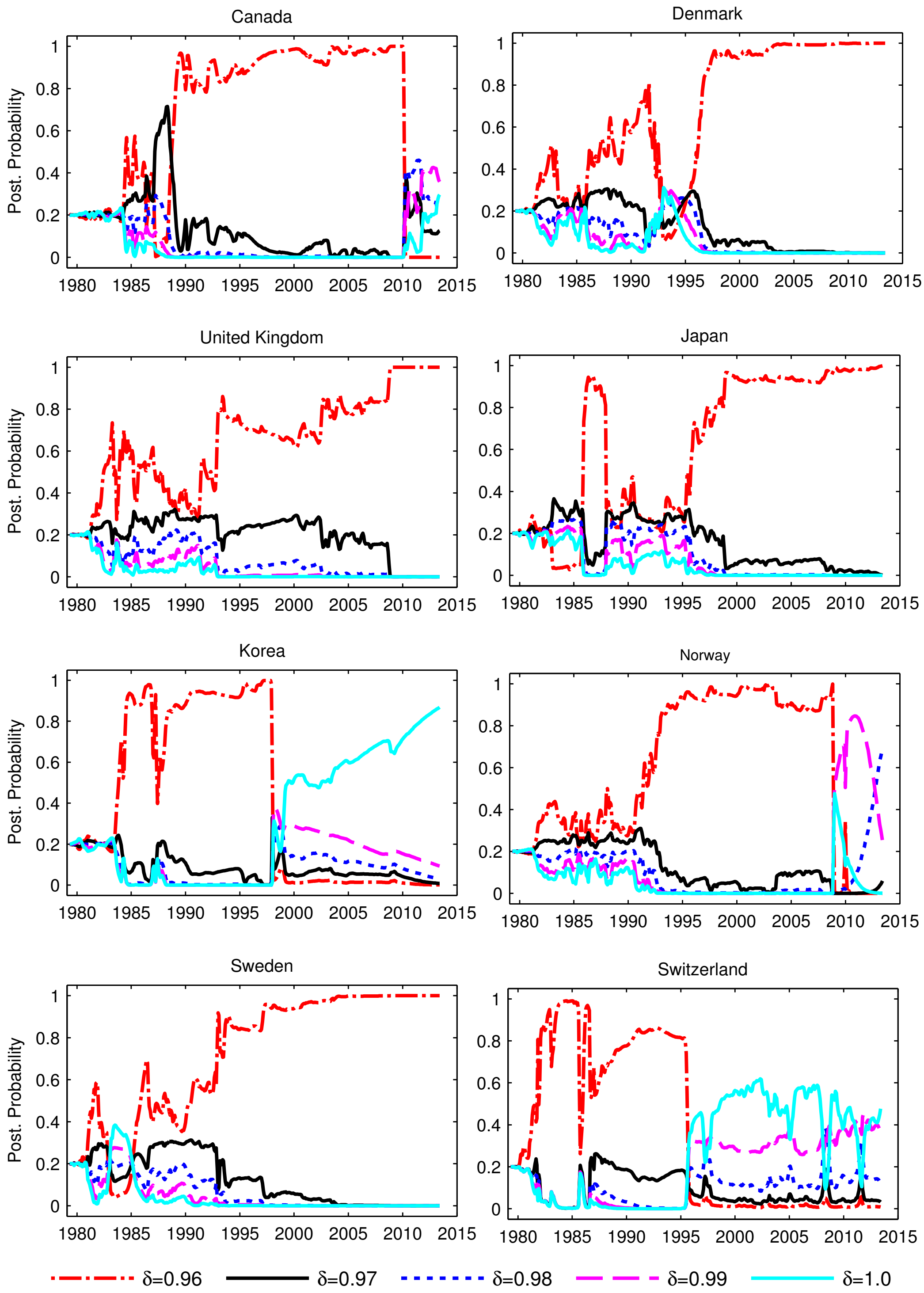
Notes: Decomposition of the prediction variance into its constituent parts for h=12 months. Panel A shows all sources of prediction variance: (i) the variance caused by random fluctuations in the data (Obs.var.); (ii) variance due to errors in the estimation of the coefficients (Unc.coef); (iii) variance 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 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 E.3: Posterior Probabilities of Values of δ, h=1



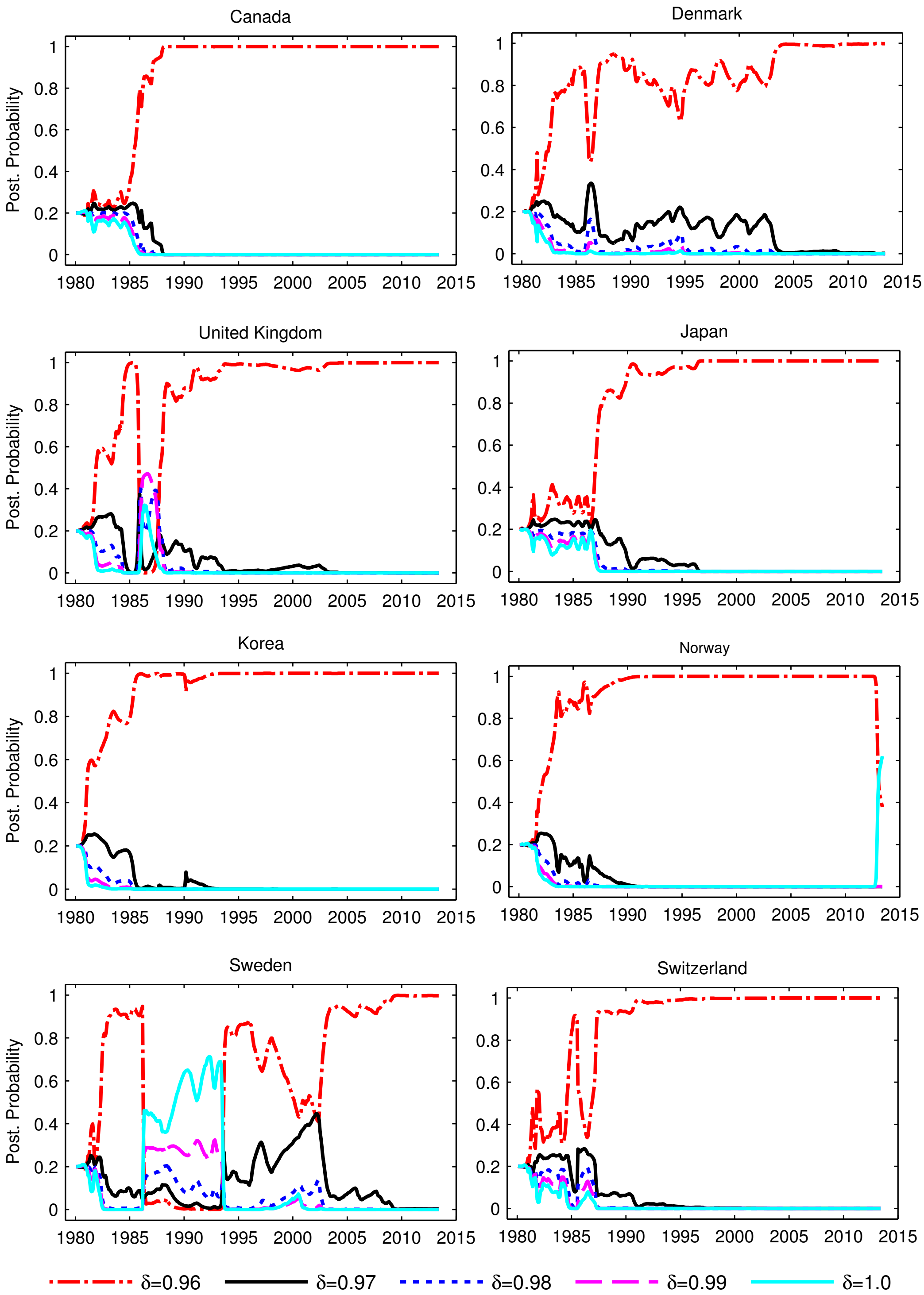
Notes: Posterior probabilities of specific values of degrees of time variation in coefficients (δ). These are the weights employed to produce the average forecasts in the BMA incl. TVar-Coeff. model.

Figure E.4: Posterior Probabilities of Values of δ, h=3



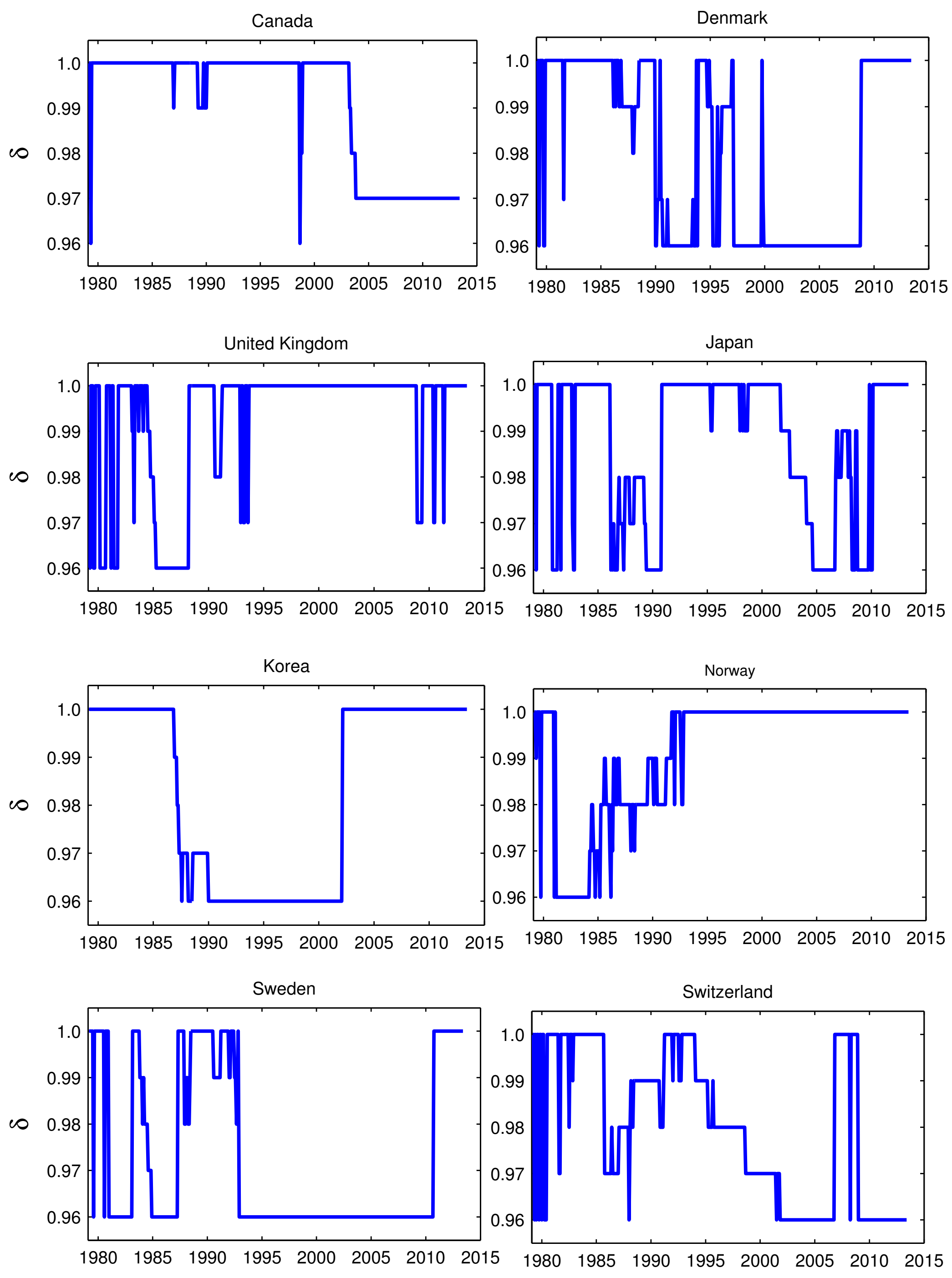
Notes: As in Figure E.3, except that here h=3 months.

Figure E.5: Posterior Probabilities of Values of δ, h=12



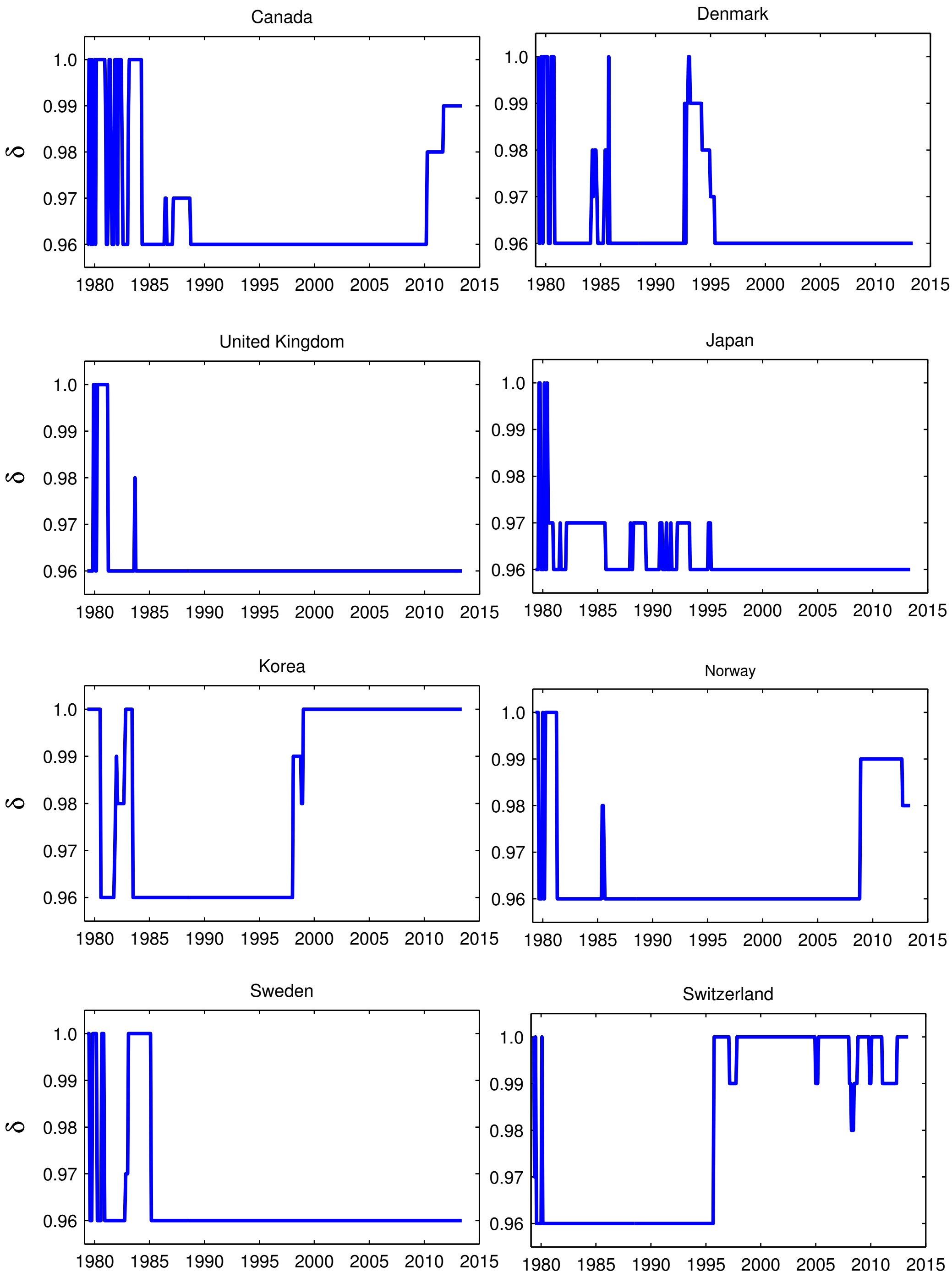
Notes: As in Figure E.3, except that here h=12 months.

Figure E.6: Values of δ with the Largest Posterior Probability at Each Period, h=1



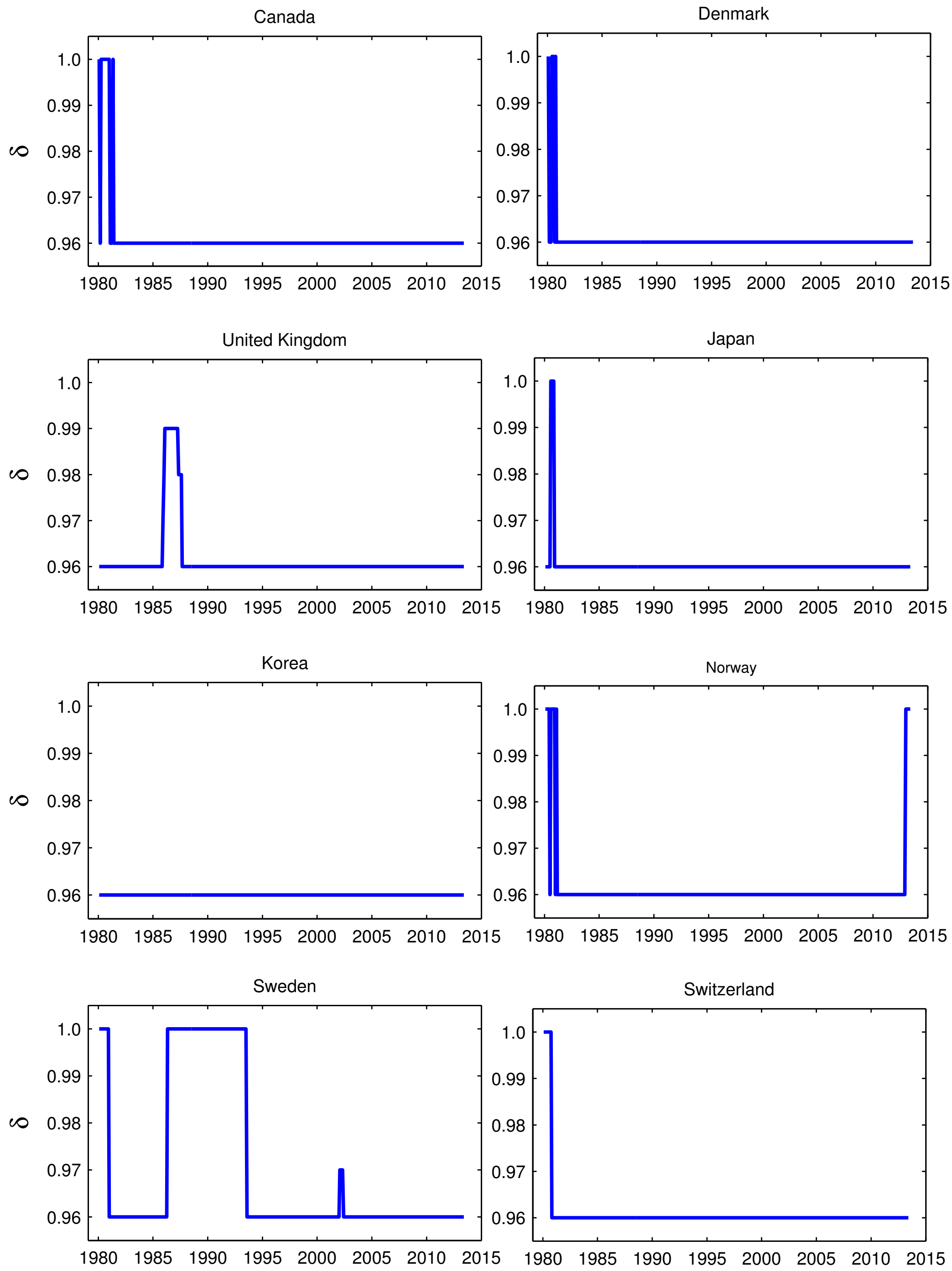
Notes: Values of δ (support points for time-variation in coefficients) with the highest probability at each point in time for h=1 month. These are the probabilities used to select the best single model at each period, and hence determine the BMS incl. TVar-Coeff model.

Figure E.7: Values of δ with the Largest Posterior Probability at Each Period, h=3



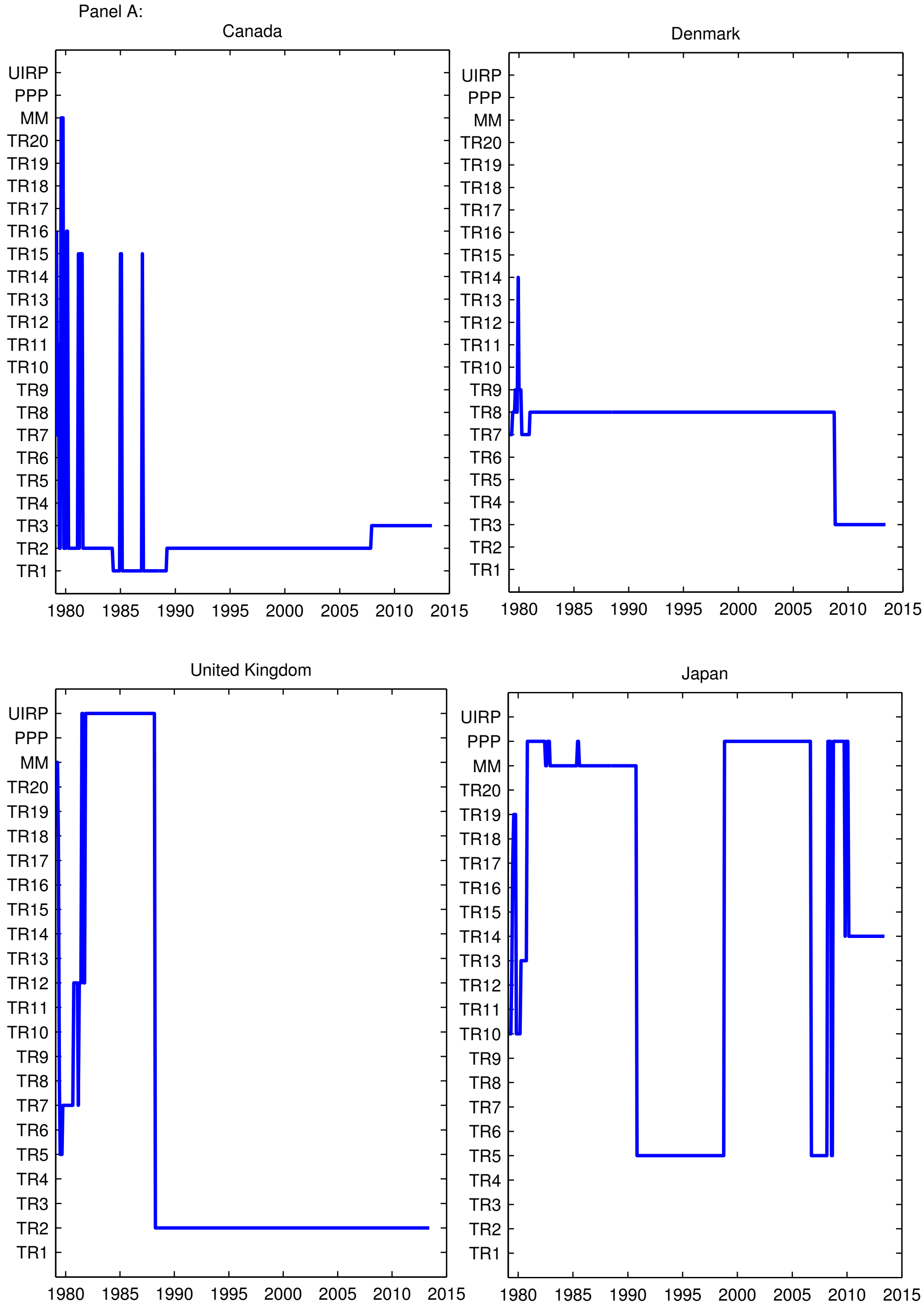
Notes: As in Figure E.6, except that here h=3 months.

Figure E.8: Values of δ with the Largest Posterior Probability at Each Period, h=12



Notes: As in Figure E.6, except that here h=12 months.

Figure E.9: Predictors with the Largest Posterior Probability at Each Period, h=1



Notes: Predictors with the highest probability at each point in time for h=1 month (Panel A and B). The forecasts from the single best model at each period are based on these predictors (i.e., from the forecasts from the BMS incl. TVar-Coeff.).

Panel B:

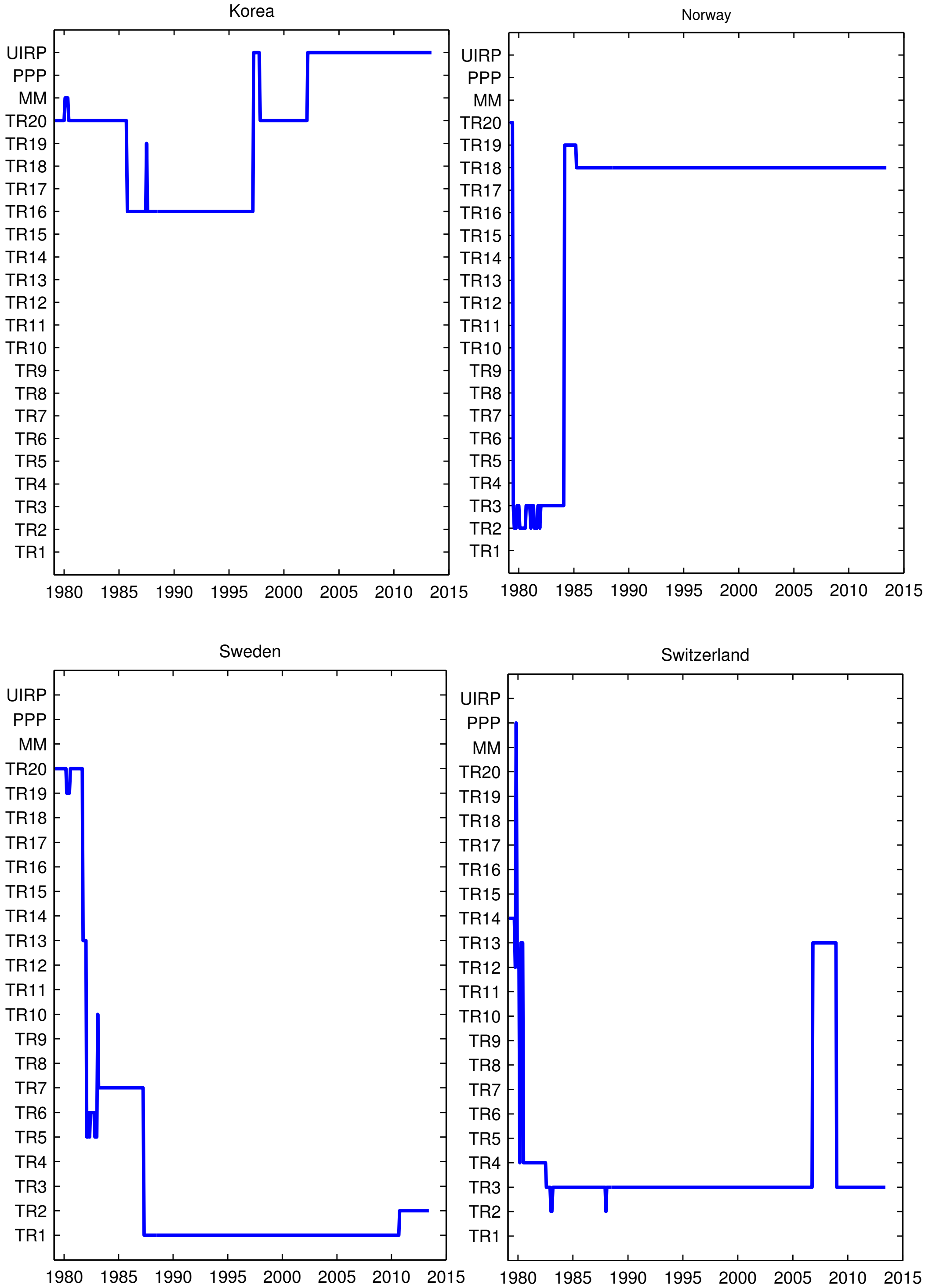
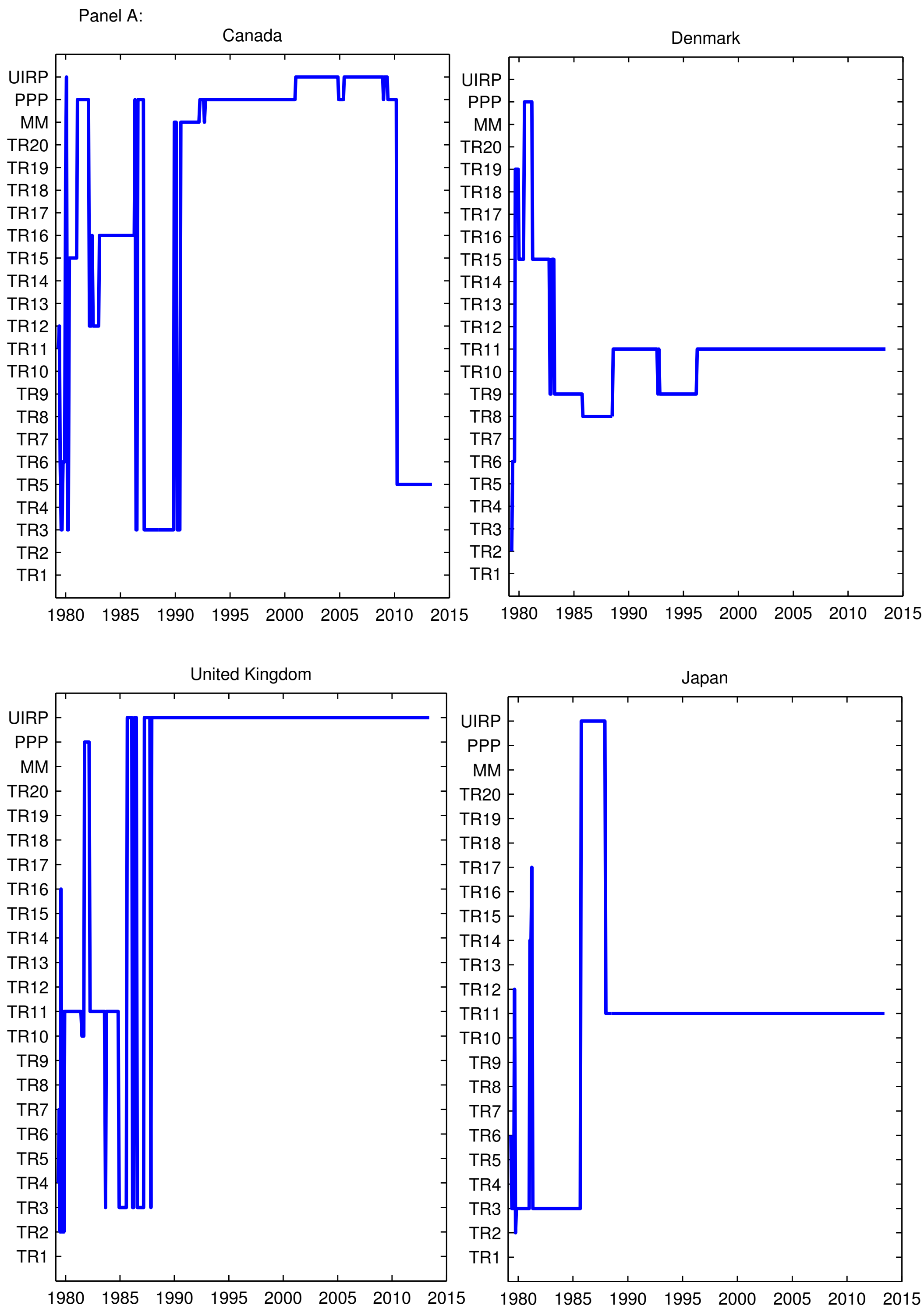

Korea
UIRP
PPP
MM
TR20
TR19
TR18
TR17
TR16
TR15
TR14
TR13
TR12
TR11
TR10
TR9
TR8
TR7
TR6
TR5
TR4
TR3
TR2
TR1
1980
1985
1990
1995
2000
2005
2010
2015
Norway
Sweden
Switzerland

Figure E.10: Predictors with the Largest Posterior Probability at Each Period, h=3



Notes: As in Figure E.9, except that here h=3 months.

Panel B:

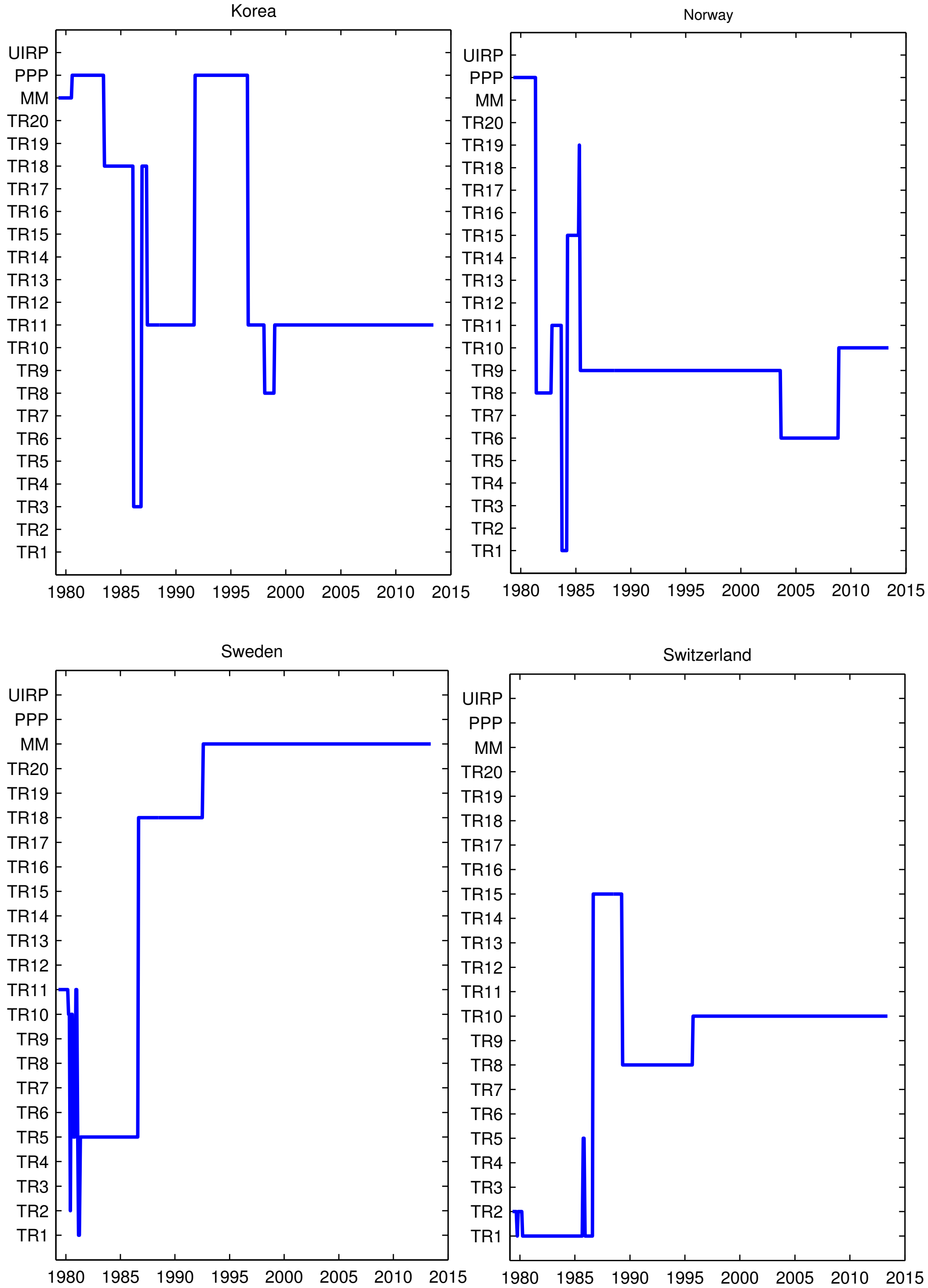
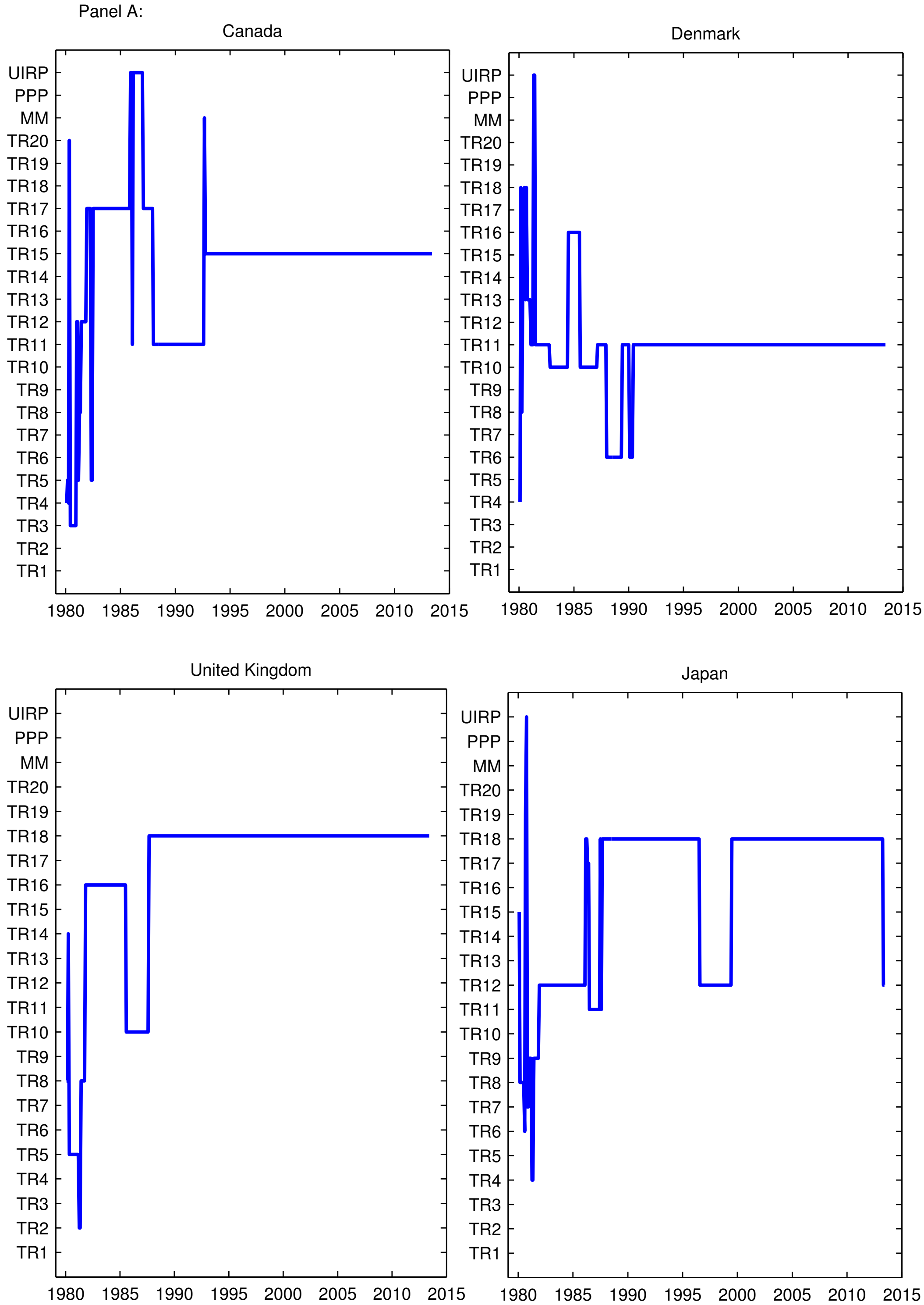

Korea
Norway
Sweden
Switzerland
UIRP
PPP
MM
TR20
TR19
TR18
TR17
TR16
TR15
TR14
TR13
TR12
TR11
TR10
TR9
TR8
TR7
TR6
TR5
TR4
TR3
TR2
TR1
1980
1985
1990
1995
2000
2005
2010
2015

Figure E.11: Predictors with the Largest Posterior Probability at Each Period, h=12



Notes: As in Figure E.9, except that here h=12 months.

Panel B:

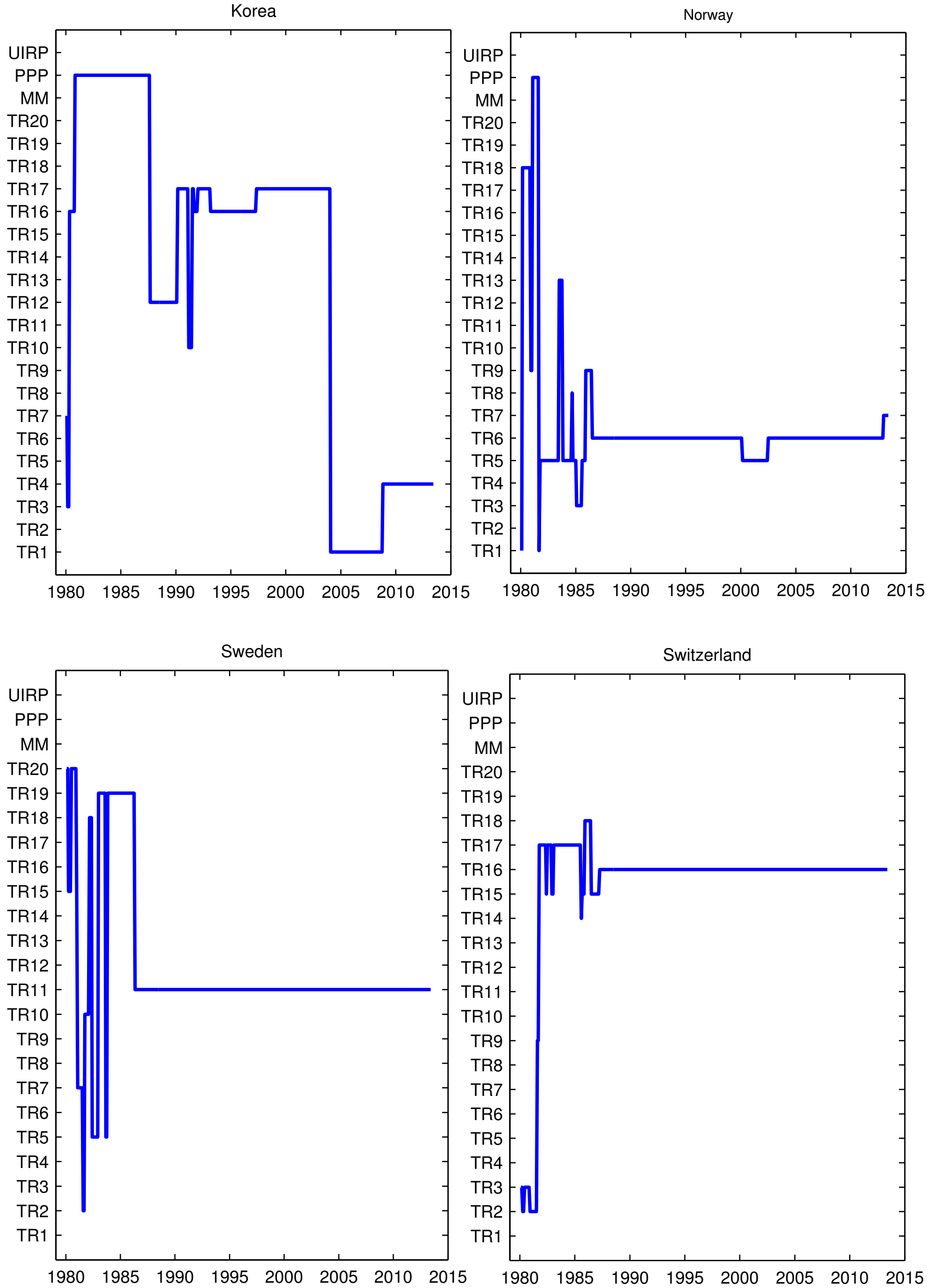
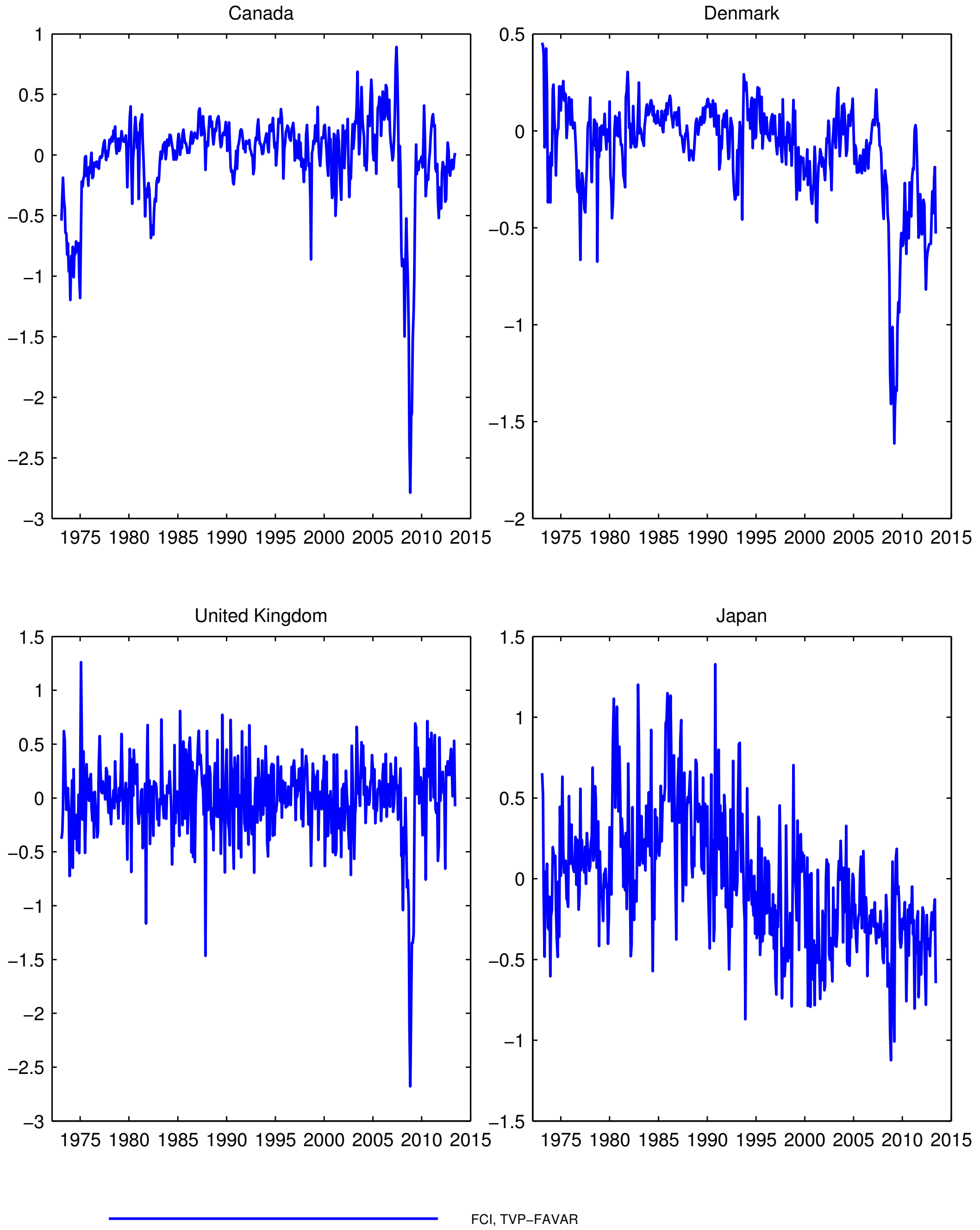
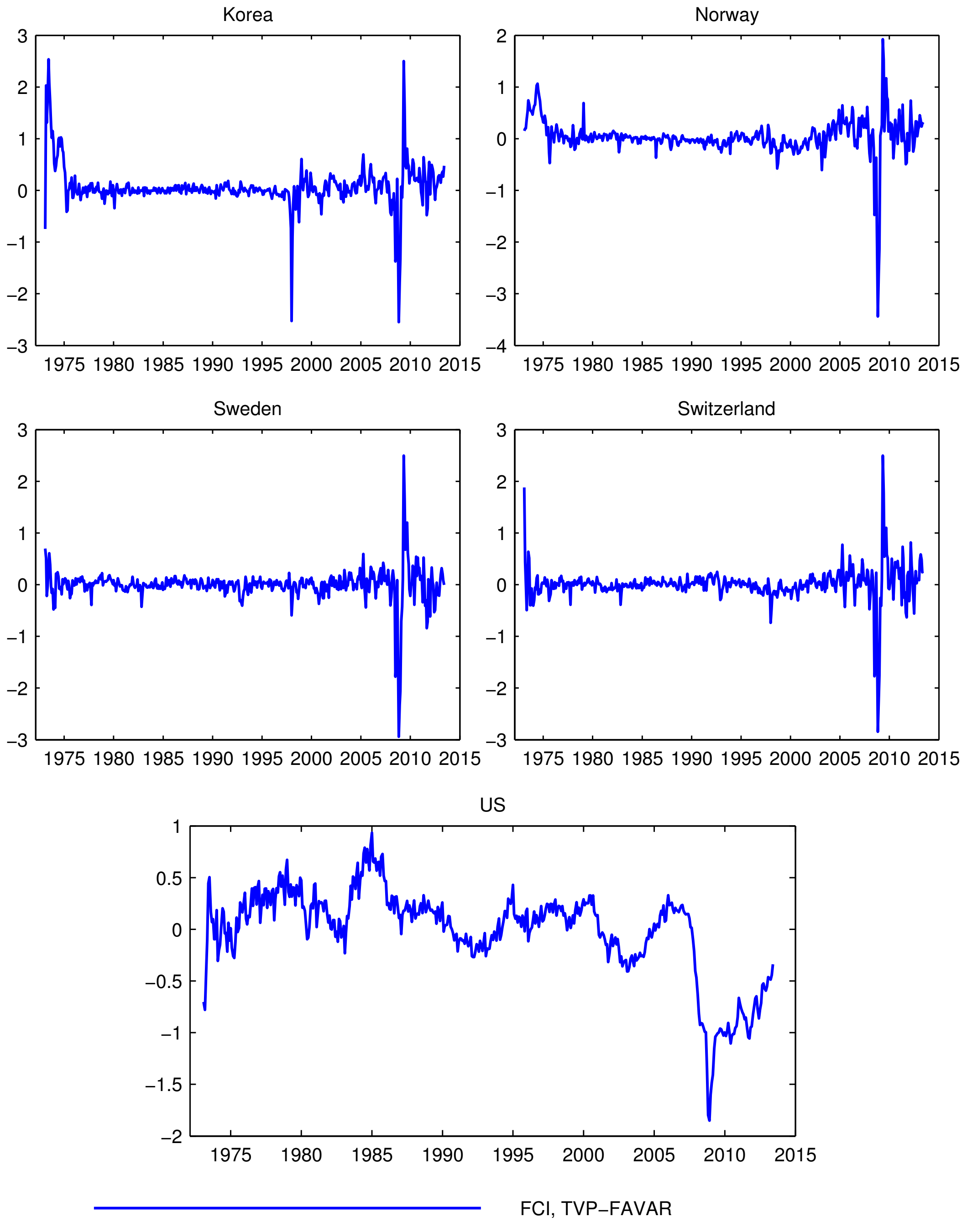

Korea
Norway
Sweden
Switzerland
UIRP
PPP
MM
TR20
TR19
TR18
TR17
TR16
TR15
TR14
TR13
TR12
TR11
TR10
TR9
TR8
TR7
TR6
TR5
TR4
TR3
TR2
TR1
1980
1985
1990
1995
2000
2005
2010
2015

Figure E.12: Financial Conditions Indexes - Monthly Frequency



Notes: Financial Conditions Indexes estimated with the TVP-FAVAR model. Positive values indicate financial conditions that are looser than on average, and negative values indicate financial conditions that are tighter than on average.


Korea
Norway
Sweden
Switzerland
US
1975 1980 1985 1990 1995 2000 2005 2010 2015
FCI, TVP−FAVAR