2019

Real-time inflation forecast combination for time-varying coefficient models

Bo Zhang
University of Wollongong, bzhang@uow.edu.au

Publication Details
Zhang, B. (2019). Real-time inflation forecast combination for time-varying coefficient models. Journal of Forecasting, 38 (3), 175-191.

Real-time inflation forecast combination for time-varying coefficient models

Abstract
We use real-time macroeconomic variables and combination forecasts with both time-varying weights and equal weights to forecast inflation in the USA. The combination forecasts compare three sets of commonly used time-varying coefficient autoregressive models: Gaussian distributed errors, errors with stochastic volatility, and errors with moving average stochastic volatility. Both point forecasts and density forecasts suggest that models combined by equal weights do not produce worse forecasts than those with time-varying weights. We also find that variable selection, the allowance of time-varying lag length choice, and the stochastic volatility specification significantly improve forecast performance over standard benchmarks. Finally, when compared with the Survey of Professional Forecasters, the results of the best combination model are found to be highly competitive during the 2007/08 financial crisis.

Keywords
forecast, combination, real-time, time-varying, models, coefficient, inflation

Disciplines
Business

Publication Details
Zhang, B. (2019). Real-time inflation forecast combination for time-varying coefficient models. Journal of Forecasting, 38 (3), 175-191.

Real-Time Inflation Forecast Combination for Time-Varying Coefficient Models

Bo Zhang

Research School of Economics,
The Australian National University
Canberra, ACT 2601 Australia
Email: bozhangyc@gmail.com
Phone: +61 2 612 52651
Fax: +61 2 612 55124

Abstract

We use real-time macroeconomic variables and combination forecasts with both time-varying weights and equal weights to forecast inflation in the United States. The combination forecasts compare three sets of commonly used time-varying coefficient autoregressive models: Gaussian distributed errors, errors with stochastic volatility, and errors with moving average stochastic volatility. Both point forecasts and density forecasts suggest that models combined by equal weights do not produce worse forecasts than those with time-varying weights. We also find that variable selection, the allowance of time-varying lag length choice and the stochastic volatility specification significantly improve forecast performance over standard benchmarks. Finally, when compared with the Survey of Professional Forecasters, the results of the best combination model are found to be highly competitive during the 2007/08 financial crisis.

Keywords: inflation forecasts, real-time data, time-varying coefficient model, forecast combination

JEL classification codes: C11, C52, C53, C55

1 Introduction

Inflation is a core macroeconomic indicator that is closely monitored by both central bankers and macroeconomic researchers. Many researchers have consequently investigated the time series properties of inflation. The consensus from these studies is that the underlying trend and volatility of inflation have changed considerably over time; however, there is still no agreement on the best way to forecast inflation dynamics (e.g., Stock and Watson, 2007; Cogley and Sbordone, 2008; Koop and Korobilis, 2012; Chan, 2013).

Since no single best model exists, forecast combination methods have proven to be a useful way for improving inflation forecast performance (e.g., Bunn, 1975; Faria and Mubwandarikwa, 2008; Kascha and Ravazzolo, 2010; Raftery, Madigan, and Hoeting, 1997; Tibiletti, 1994). Despite the existence of many sophisticated combination methods, e.g., the linear combination method, the geometric combination method, the logarithmic combination method, Bayesian model averaging, and combinations of experts' forecasts (the Survey of Professional Forecasters), the empirical evidence suggests that forecast accuracy is often best when simply averaging the forecasts across the set of models (Timmerman, 2003). Two commonly used types of averages are time-varying and equal weight methods. However, there is no consensus on whether time-varying or equally weighted combinations work better over a wide variety of models (e.g., Stock and Watson, 2004; Clark and McCracken, 2009; Jore, Mitchell, and Vahey, 2010).

With this in mind, our main objective in this paper is to compare the forecast performance of both time-varying and equally weighted combinations of a wide variety of time series models with time-varying parameters and various error structures. The main result is that models combined by equal weights do not have worse forecast performance than those with time-varying weights. We also find that both combination strategies tend to provide better forecast performance than univariate models in both point forecasts and density forecasts. Finally, the forecast results of our proposed models are also highly competitive when comparing with the quarterly reports from the SPF during the financial crisis.

Another key contribution is that we investigate the temporal relationship between inflation and other explanatory variables when conducting combination forecasts. This is done by considering models with different predictors and lag structures. The present paper combines forecasts based on one inflation predictor, which reduces the number of models significantly. Furthermore, Koop and Korobilis (2012) introduced dynamic model averaging (DMA) and dynamic model selection (DMS), which use a forgetting factor strategy to update time-varying coefficients and averaging models with a set of explanatory variables and different lag lengths. However, the computational burden must be considered when the number of lags is greater than two with multiple explanatory variables. For instance, eight explanatory variables with three lags could produce more than 400 million candidate models. Another interesting aspect of the forecasting results from DMA and DMS is that high weights are given to parsimonious models or parsimonious models that rarely have more than two predictors selected. Moreover, quarterly data are generally

widely used for inflation forecasts, and most models use up to four lags (e.g., Cogley and Sargent, 2005; Stock and Watson, 2007; Clark and Ravazzolo, 2015). In the present paper, the lag length can increase to four without a heavy computational burden.

Finally, the present paper also compares the forecasting results with the quarterly reports from the SPF. The results of the proposed combination forecasts models are highly competitive during the 2007/08 financial crisis. The SPF collects forecasts from professional forecasters and their forecasts are generally quite close to the actual values and difficult to beat (Tibiletti, 1994; Smith and Vahey, 2016).

The remainder of the chapter proceeds as follows. Section 2 describes the specifications of time-varying coefficients models and the component models for inflation forecast combination. Section 3 provides a brief introduction to the competing models for forecasts and the forecast combination methodology. Section 4 discusses the forecasting results of the model combination for point and density inflation forecasts. Section 5 concludes.

2 Component Models

We consider three broad classes of time-varying coefficient models with different specifications of error terms: (i) models with constant variance (TVC); (ii) models with stochastic volatility (TVC-SV); and (iii) models with moving average stochastic volatility (TVC-SVMA). Within each class of model, parameters are estimated by different measures of economic activities, then they are employed in the forecasting exercises. In addition, for each inflation predictor, various lag structures are considered, such as from one to four single lags, and up to four more lags. By combining all lag structures with the eight predictors, this model's averaging approach is conducted for TVC, TVC-SV and TVC-SVMA. In the following subsections, the specifications of TVC, TVC-SV and TVC-SVMA are described first, followed by the eight inflation predictor candidates.

2.1 Time-Varying Coefficient Models

2.1.1 Constant Variance

A generic TVC model can be described as a generalized Phillips curve with time-varying coefficients:

$$y_{t+k} = \beta_{1,t} + \Sigma_{j=0}^{n} \beta_{2+j,t} x_{t-j} + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \sigma_y^2), \tag{1}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t^{\beta}, \quad \boldsymbol{\varepsilon}_t^{\beta} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{2}$$

$$\boldsymbol{\beta}_t = \begin{pmatrix} \beta_{1,t} & \beta_{2,t} & \cdots & \beta_{n,t} \end{pmatrix}', \mathbf{Q}_0 = \begin{pmatrix} \sigma_{0\beta_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{0\beta_n}^2 \end{pmatrix}, \mathbf{Q} = \begin{pmatrix} \sigma_{\beta_1}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\beta_n}^2 \end{pmatrix},$$

where k is the forecasting horizon. Note that the subscripts in the specification represent the order of the coefficients and time points (e.g., $\boldsymbol{\beta}_{1,t}$ to $\boldsymbol{\beta}_{n,t}$), while the superscripts indicate the underlying relationship between variables (e.g., $\mathbf{y}$ and $\boldsymbol{\varepsilon}^y$). x_t is a vector of covariates that may include lagged values of inflation or inflation predictors. The model above incorporates both a time-varying intercept and regression coefficients.

In Equation (2), the intercept and coefficients are assumed to follow independent random walks (e.g., Clark, 2011; Clark and Ravazzolo, 2015). By allowing the coefficients to evolve gradually over time, this specification accommodates a slowly changing relationship between inflation and the explanatory variables. Atkeson and Ohanian (2001) criticize the ability of the Phillips curve models to forecast inflation compared with random walk forecasts. However, the Phillips curve models they adopt all have constant coefficients, and they do not perform as well as random walk naive forecasts in some historical periods. It can be expected that a time-varying Phillips curve performs better than one with constant coefficients.

The covariance matrices $\mathbf{Q}_0$ and $\mathbf{Q}$ of $\boldsymbol{\beta}_0$ and $\boldsymbol{\beta}_t$ respectively are assumed to be diagonal matrices, where $\mathbf{Q}_0$ and $\boldsymbol{\beta}_0$ are initial values of $\mathbf{Q}$ and $\boldsymbol{\beta}_t$, respectively. It indicates that $\boldsymbol{\beta}_{1,t} \cdots \boldsymbol{\beta}_{n,t}$ have individual independent white noise disturbances with zero means and variances $\varepsilon_t^{\boldsymbol{\beta}_1} \cdots \varepsilon_t^{\boldsymbol{\beta}_n}$, respectively. Both ε_t^y and $\boldsymbol{\varepsilon}_t^\beta$ have constant variances.

2.1.2 Stochastic Volatility

Next, we extend (1)-(2) to allow for stochastic volatility (e.g., Groen, Paap, and Ravazzolo, 2013):

$$y_{t+k} = \beta_{1,t} + \Sigma_{j=0}^{n} \beta_{2+j,t} x_{t-j} + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \mathrm{e}^{h_t}), \tag{3}$$

$$h_t = h_{t-1} + \boldsymbol{\varepsilon}_t^h, \quad \boldsymbol{\varepsilon}_t^h \sim \mathcal{N}(0, \sigma_h^2), \tag{4}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t^\beta, \quad \boldsymbol{\varepsilon}_t^\beta \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{5}$$

where $\boldsymbol{\beta}_t$ and $\boldsymbol{\varepsilon}_t^\beta$ are the same as those in TVC, but the variance of ε_t^y is time-varying. The variance of ε_t^y is controlled by log stochastic volatility h_t, which changes over time. Equation (4) presents the evolution of the stochastic volatility parameter, which is an instantaneous volatility component of the model. The log-volatilities in Equation (4) are initialized by $h_1 \sim \mathcal{N}(0, \sigma_{0h}^2)$ with variance given in advance, and h_t follows random walk innovation.

2.1.3 Moving Average Stochastic Volatility

We further extend the SV model (3)-(5) using the framework in Chan (2013). The TVC-SVMA model is given below:

$$y_{t+k} = \beta_{1,t} + \Sigma_{j=0}^{n} \beta_{2+j,t} x_{t-j} + \varepsilon_t^y, \tag{6}$$

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\varepsilon}_t^{\beta}, \quad \boldsymbol{\beta}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_0), \quad \boldsymbol{\varepsilon}_t^{\beta} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \tag{7}$$

$$\varepsilon_t^y = \omega_t + \psi_1 \omega_{t-1} + \cdots + \psi_q \omega_{t-q}, \quad \omega_t \sim \mathcal{N}(0, \mathrm{e}^{h_t}), \tag{8}$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \tag{9}$$

where $\boldsymbol{\beta}_t$ and $\boldsymbol{\varepsilon}_t^{\beta}$ are the same as those in TVC. Equation (8) presents the moving average error feature of the specification. It can be rewritten as a polynomial of the lag operator L:

$$\varepsilon_t = \psi(L) u_t,$$

where $\psi(L) = 1 + \psi_1 L + \cdots + \psi_q L^q$. For identification, we assume that all roots of $\psi(L)$ are outside the unit circle. We assume $q = 1$ for simplicity.

2.2 Explanatory Variables and Real Time Data

In the present study, we use real-time data in the forecasting exercise instead of heavily revised data. There has already been much Bayesian work studying real-time macroeconomic variable forecasting, such as forecasts using Bayesian vector autoregressive models (Clark, 2011), forecasts of inflation and the output gap (Garratt et al., 2011), UK monetary aggregates (Garratt et al., 2009), inflation forecasts by Bayesian model averaging (Groen et al., 2013) and forecasts macroeconomic variables by a copula model with asymmetric margins (Smith and Vahey, 2016). The present study follows these pioneer studies and employs both Bayesian estimation and real-time data to study inflation.

Although inflation has become less responsive to oil price shocks since 1970s, the share of energy in aggregate output and the effects of oil price to inflation cannot be ignored (Mishkin, 2007). Meanwhile, oil prices are a component of many Phillips curve models (Hooker, 2002). Thus, oil prices are also introduced in the forecasting models. For simplicity, we consider contemporary oil prices (op_t) for each component forecasting model.

Motivated by the predictive performance of inflation predictors in Groen et al. (2013), there are a total of eight inflation predictors being considered for each time-varying coefficient model (TVC, TVC-SV, and TVC-SVMA). Among these eight predictors, there are seven real economy activity predictors and one nominal predictor, M2. The real-time inflation predictors are from the Real-Time Data Set for Macroeconomists (RTDSM) database at the Federal Reserve Bank of Philadelphia (Croushore and Stark, 2001, 2003). The inflation predictors are as follows:

1. Real unemployment rate (UR),
2. Real capacity utilization rate in manufacturing (CUR),
3. Housing starts (HSTS),

4. Real imports of goods and services (IMP),
5. M2 growth rate (M2),
6. Real durable consumption growth (RCON),
7. Real residential investment (RINV),
8. Real output growth (ROUT).

The transformations of the variables are listed below:

1. The quarterly inflation rate is calculated as the first difference of the logged inflation deflator.
2. The oil prices are the West Texas Intermediate crude oil spot prices with first difference logarithm, and they are obtained from Federal Reserve Bank of St. Louis.
3. The quarterly UR is the number of unemployed as a percentage of the labor force. The original data vintages of UR are used for both estimation and forecasting studies.
4. To obtain CUR, the original monthly data are first transformed into quarterly average data, and then the first differences are taken on the logarithm to measure quarterly changes.
5. HSTS are also monthly vintages, and they are transformed by taking the first difference logarithm.
6. The monthly M2 growth rate vintages are transformed in the same way as the CUR; however, vintages 1981Q1 and 1981Q2 of M2 are incomplete, so we replace them with vintage 1981Q3.
7. For IMP, the original quarterly data vintages are available for imports, so we simply take the natural logarithm of the raw data to construct the quarterly frequency of import price inflation.
8. For RCON, the quarterly data vintages are transformed by the first difference of logarithm.
9. For RINV and ROUT, the original quarterly data vintages are all available, so we follow the same transformation as for IMP to construct the growth rates.

The last vintages of the time series after transformation are plotted in Figure 1. The full sample estimation employs the last vintage of data. Most component models present decreasing intercepts and coefficients of inflation predictors over time. The decreasing intercept indicates inflation remains a low level recently (Stock and Watson, 2007), while

the decreasing coefficients are consistent with the flattening of the Phillips curve (Lakova, 2007).

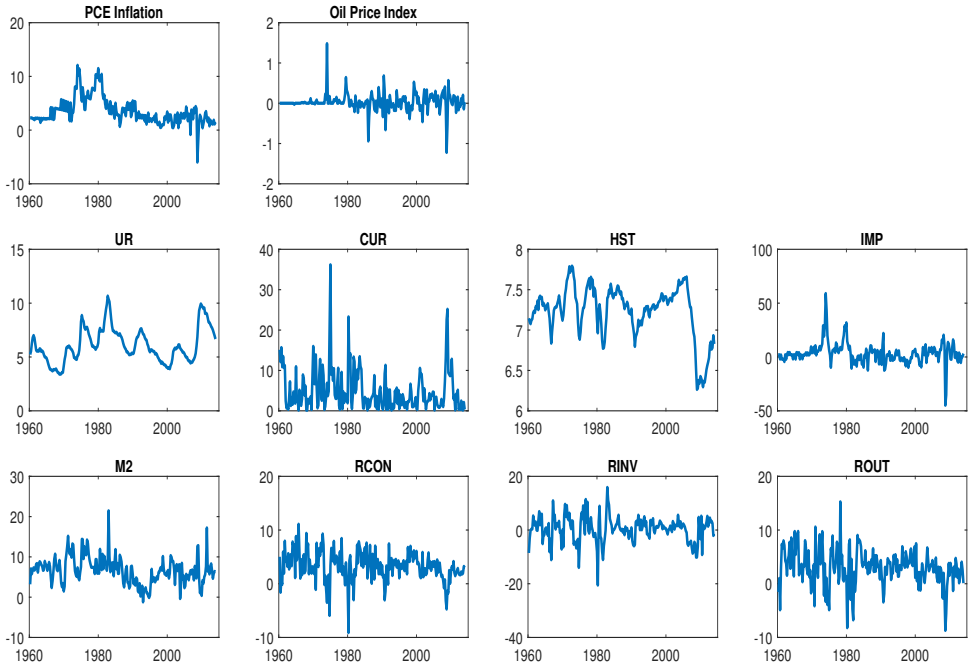

Figure 1: PCE inflation, the oil price index and other explanatory variables.

For forecasting exercises, we use the second available estimates as actuals because Corradi et al. (2009) provide tests showing that the second revision error is concentrated around zero and is better than the first revision error, which is normally distributed. For example, considering that we use vintage 2000Q2 data to make one-step-ahead forecasts for 2000Q2, the second available estimate for 2000Q2 is only in the time series 1959Q1-2000Q3 in vintage 2000Q4.

The starting point of modeling estimation is 1960Q2, and the time period of 1959Q2 to 1960Q1 is trimmed as lags. We use the personal consumption expenditures (PCE) deflator as the measure of US inflation rather than the customer price index (CPI) since the first available vintage of CPI is vintage 1994Q3, which is too late for a comparison with PCE, for which we can adopt a vintage as early as 1985Q3. Thus, the data from 1960Q2 to 1985Q2 in vintage 1985Q3 is considered the first evaluation time period, and the last vintage is 2014Q2; thus, the forecasting exercises are complete when the forecasting results of 2014Q1 are obtained. The estimation window is a rolling window that retains the same window length from 1985Q3 by adding the newest data and deleting the oldest data. The forecasting horizons are one quarter, one year, two years and four years ($k = 1, 4, 8$ and 16). In the 1-, 4-, 8- and 16-step-ahead forecasts, $y_{t+1}, \cdots, y_{t+16}$ are calculated by Equation (10) to (18).

2.3 Forecasting Models

For each inflation predictor, different forms of lag structure are considered: coincident, or as a leading inflation predictor (one quarter ahead, half-year ahead, three quarters ahead, and one year ahead). We also include different numbers of lags. Explicitly, the list of models is as follows:

$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \beta_{3,t} x_t + \varepsilon_t^y, \tag{10}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \beta_{3,t} x_{t-1} + \varepsilon_t^y, \tag{11}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \beta_{3,t} x_{t-2} + \varepsilon_t^y, \tag{12}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \beta_{3,t} x_{t-3} + \varepsilon_t^y, \tag{13}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \beta_{3,t} x_{t-4} + \varepsilon_t^y, \tag{14}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \Sigma_{j=0}^{1} \beta_{3+j,t} x_{t-j} + \varepsilon_t^y, \tag{15}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \Sigma_{j=0}^{2} \beta_{3+j,t} x_{t-j} + \varepsilon_t^y, \tag{16}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \Sigma_{j=0}^{3} \beta_{3+j,t} x_{t-j} + \varepsilon_t^y, \tag{17}$$
$$y_{t+k} = \beta_{1,t} + \beta_{2,t} op_t + \Sigma_{j=0}^{4} \beta_{3+j,t} x_{t-j} + \varepsilon_t^y, \tag{18}$$

where k is the same as that in Section 2.1.1. In total, there are nine component models for each predictor. The time-varying coefficients in each component model and time-varying weights of component models make it possible to update the information set at any time.

2.4 The Priors

Each component model is estimated using Bayesian methods that incorporate parameter uncertainty. The priors of the parameter initial values are set as follows. We assume that the priors of the intercept and coefficient initial values are normal: $\beta_{1,1} \sim \mathcal{N}(\beta_{1,0}, V_{\beta_1})$, $\beta_{2,1} \sim \mathcal{N}(\beta_{2,0}, V_{\beta_2}), \cdots, \beta_{j,1} \sim \mathcal{N}(\beta_{j,0}, V_{\beta_j}), \cdots$, and $\beta_{n,1} \sim \mathcal{N}(\beta_{n,0}, V_{\beta_n})$ for $j = 3, \cdots, n-1$. We set $\beta_{1,0} = 5, \beta_{2,0} = -0.2, \beta_{j,0} = -0.1, V_{\beta_1} = 2, V_{\beta_2} = 0.2$ and $V_{\beta_j} = 0.1$. The prior means of $\beta_{j,1}$ are set to small values, which reflects the belief that the covariates are weakly informative about the inflation initial conditions.

Next, the variances $\sigma_{\beta_1}^2, \sigma_{\beta_2}^2, \cdots, \sigma_{\beta_n}^2$ and the elements of the covariance matrix Q for $\boldsymbol{\beta}$ are assumed to have independent inverse-gamma priors: $\sigma_{\beta_1}^2 \sim \mathcal{IG}(\nu_{\beta_1}, S_{\beta_1})$, $\sigma_{\beta_2}^2 \sim \mathcal{IG}(\nu_{\beta_2}, S_{\beta_2})$, and $\sigma_{\beta_j}^2 \sim \mathcal{IG}(\nu_{\beta_j}, S_{\beta_j})$. To have vague priors for the variances, we choose large prior variances. Specifically, we set small values for the degree of freedom parameters: $\nu_{\beta_1} = \nu_{\beta_2} = \nu_{\beta_j} = 10$. We then set the scale parameters $S_{\beta_1} = 2$,$S_{\beta_2} = 0.2, S_{\beta_j} = 0.1$, so that the prior means are $E(\sigma_{\beta_1}^2) = 0.22, E(\sigma_{\beta_2}^2) = 0.02$, and $E(\sigma_{\beta_j}^2) = 0.01$. The prior means indicate that the parameters transit in the desired smoothness from one state to another.

The prior of the stochastic volatility parameter initial value h_1 is also supposed to be normal: $h_1 \sim \mathcal{N}(h_0, V_h)$ and $\sigma_h^2 \sim \mathcal{IG}(\nu_{\sigma_h}, S_{\sigma_h})$, where $h_0 = 0, V_h = 0.05, \nu_h = 10$ and $S_h = 0.45$, so that the prior mean of $\sigma_h^2 = 0.05$.

Finally, the prior of the MA (1) coefficient ψ is assumed to be a truncated normal prior following that of Chan (2013): $\psi \sim \mathcal{N}(\psi_0, V_\psi)\mathbb{1}(|\psi| < 1)$, where $\psi_0 = 0.9$ and $V_\psi = 1$.

A detailed description of the Bayesian estimation methods is provided in Appendix 3.A.

3 Real-Time Forecasts of US Inflation

We compute forecast combinations by weighting the forecasts from different models. Specifically, recursive weights based on historical forecast performance are used for model averaging forecasts (e.g., Jore, Mitchell, and Vahey, 2010; Garratt, Mitchell, Vahey, and Wakerly, 2011), so that the model averaging weights are evaluated repeatedly by the following vintages.

Model averaging based on three time-varying coefficient models with specifications on error terms are considered as competing models in the following forecasting exercise sections. Various lag forms for the eight macroeconomic variables are the component models for combining forecasts, as discussed above (see Equation (10) to (18)). Instead of using Bayesian model selection criteria to determine the optimal forecast model, such as the Bayesian information criterion or deviance information criterion, model averaging does not need to select a single model, which may not forecast well all the time, but instead weights models with better forecast performance more heavily.

3.1 List of Competing Models

As mentioned, we consider three classes of models: TVC, TVC-SV, and TVC-SVMA. In the inflation forecasting literature, stochastic volatility is found to be an important component in inflation forecasting models, and we would like to determine whether the moving average stochastic volatility can improve forecasting results.

An autoregressive model with homoscedatic variance (AR(m)) is used as the benchmark. The lag length m is selected as three via the Bayesian information criteria (BIC). Thus, AR(3) is used as the benchmark model in forecasting exercises.

The competing models are as follows:

1. AR model with the lag length three (AR(3)):

$$y_t = \beta_1 + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2).$$

2. Time-varying coefficient model without stochastic volatility (TVC), and $\beta_{2,t}x_t$ can be replaced according to Equation (10) to (18):

$$\begin{aligned} y_{t+k} &= \beta_{1,t} + \beta_{2,t}x_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \sigma_y^2), \\ \beta_t &= \beta_{t-1} + \varepsilon_t^\beta, \qquad \beta_1 \sim \mathcal{N}(\mathbf{0}, Q_0), \quad \varepsilon_t^\beta \sim \mathcal{N}(\mathbf{0}, Q). \end{aligned}$$

3. Time-varying coefficient model with stochastic volatility (TVC-SV), and $\beta_{2,t}x_t$ can be replaced according to Equation (10) to (18):

$$\begin{aligned} y_{t+k} &= \beta_{1,t} + \beta_{2,t}x_t + \varepsilon_t^y, \quad \varepsilon_t^y \sim \mathcal{N}(0, \mathrm{e}^{h_t}) \\ \beta_t &= \beta_{t-1} + \varepsilon_t^\beta, \quad \beta_1 \sim \mathcal{N}(\mathbf{0}, Q_0), \varepsilon_t^\beta \sim \mathcal{N}(\mathbf{0}, Q), \\ h_t &= h_{t-1} + \varepsilon_t^h, \quad h_1 \sim \mathcal{N}(0, \sigma_{0h}^2), \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2). \end{aligned}$$

4. Time-varying coefficient model with moving average stochastic volatility (TVC-SVMA) with lag of ψ equal to one for simplicity as in Chan (2013), and $\beta_{2,t}x_t$ can be replaced according to Equation (10) to (18):

$$\begin{aligned} y_{t+k} &= \beta_{1,t} + \beta_{2,t}x_t + \varepsilon_t^y, \\ \beta_t &= \beta_{t-1} + \varepsilon_t^\beta, \quad \beta_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_0), \varepsilon_t^\beta \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}), \\ \varepsilon_t^y &= \omega_t + \psi\omega_{t-1}, \quad \omega_t \sim \mathcal{N}(0, \mathrm{e}^{h_t}), \\ h_t &= h_{t-1} + \varepsilon_t^h, \quad h_1 \sim \mathcal{N}(0, \sigma_{0h}^2), \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \end{aligned}$$

3.2 Forecast Metrics

In this section, we discuss the metrics for evaluating the forecasts. Both the mean absolute forecast errors (MAFE) and the continuous ranked probability score (CRPS) proposed by Gneiting and Raftery (2007) are employed to compare forecasts from different models. For easier comparison, the MAFE and CRPS values of each model relative to the benchmark AR(3) are reported in Table 1. For the relative MAFE, a score under 1.00 indicates a better forecast than the benchmark, whereas a score over 1.00 presents a worse forecast performance. For the relative average CRPS, a negative score of the relative CRPS represents a better forecast performance, whereas a positive one indicates a worse forecast performance than the benchmark.

The MAFE uses absolute values of the difference between the actual values and the forecasts to evaluate the forecast performance. It therefore measures the average magnitude of the forecasting errors without considering the sign of the forecasting results. The result is that it provides equal weight to forecasting errors, which is simple but useful for measuring the accuracy of point forecasts. The rolling window MAFE is defined as:

$$\mathrm{MAFE}_{k,i} = \frac{1}{T - T_0 - k + 1} \sum_{t=1}^{T-T_0-k+1} |\widehat{y}_{T_0+t+k-1,i} - y^0_{T_0+t+k-1}|,$$

where $\widehat{y}_{T_0+t+k-1,i}$ is the k-step ahead point forecast for model i, and $y^0_{T_0+t+k-1}$ is the realized inflation at time $T_0+t+k-1$.

For density forecasts, the CRPS provides a density forecast evaluation for the quality of the forecast performance. It defines the out-of-sample forecasting prediction performance directly in terms of predictive likelihood functions (Gneiting and Raftery, 2007); thus, it provides a straightforward way for probabilistic forecasts. In practice, a smaller value of CRPS implies a better density forecast. The CRPS is defined as:

$$\begin{aligned}\text{CRPS}_{\text{Model}_i} =& \frac{1}{T-T_0-k+1}\sum_{t=1}^{T-T_0-k+1} p(|\widehat{y}_{1,(T_0+t+k-1,\text{Model}_i)} - y_{(T_0+t+k-1)}| \\ & - 0.5\times|\widehat{y}_{1,(T_0+t+k-1,\text{Model}_i)} - \widehat{y}_{2,(T_0+t+k-1,\text{Model}_i)}| \,|\, y_{(1:T_0+t)}),\end{aligned} \tag{19}$$

3.3 Forecast Combination

For point forecasts, both equal weights and time-varying weights are considered. Similar to the combination weights of mean square forecast error in Stock and Watson (2004), we use a window of the previous 40 periods to calculate MAFE so that for each forecast horizon k and for each type of specifications (TVC, TVC-SV, and TVC-SVMA), the MAFE at time T_0+t $\text{MAFE}_{T_0+t,T_0+t+k,i} = \sum_{\tau=T_0+t-40}^{T_0+t-1}|y_\tau - \widehat{y}_{\tau,i}|$, where i stands for a specific inflation predictor or one model in model combination. The results of window 20 and 60 are also presented for forecast sensitivity analysis. The time-varying weights, based on the inverse absolute forecast errors for point forecast model averaging, are calculated by:

$$\widehat{w}^{\text{MAFE}}_{T_0+t,T_0+t+k,i} = \frac{1}{\text{MAFE}_{T_0+t,T_0+t+k,i}} / \sum_{j=1}^{N}(1/\text{MAFE}_{T_0+t,T_0+t+k,i}), \tag{20}$$

where N is the total number of inflation predictors. The weights $w^{\text{MAFE}}_{T_0+t,T_0+t+k,i}$ are all nonnegative and can be summed to unity. This may vary with recursive forecasts in the entire forecasting evaluation time period.

Thus, the forecast combination of each type of time-varying coefficient specifications (TVC, TVC-SV and TVC-SVMA) for k-step ahead forecasts is:

$$\widehat{y}^{\text{comb-MAFE}}_{T_0+t,T_0+t+k} = \sum_{i=1}^{N}(\widehat{w}^{\text{MAFE}}_{T_0+t,T_0+t+k,i}\cdot\widehat{y}_{T_0+t,T_0+t+k,i}).$$

For density forecast combinations, in addition to using equal weights, we also consider an approach that is based on forecast performance. Specifically, the density combination weights use the CRPS with a rolling window of 40 periods (Ravazzolo and Vahey, 2014).

The forecast combination of each type of time-varying coefficient specifications (TVC, TVC-SV and TVC-SVMA) for the k-step ahead forecasts is:

$$\widehat{y}^{\text{comb-CRPS}}_{T_0+t,T_0+t+k} = \sum_{i=1}^{N} (\widehat{w}^{\text{CRPS}}_{T_0+t,T_0+t+k,i} \cdot \widehat{y}_{T_0+t,T_0+t+k,i}).$$

4 Forecasting Results

Both the point and density forecasting results of the PCE deflator inflation are summarized in Table 1. The forecasting results are listed in three blocks relating to the benchmark model, model combination with time-varying weight (TVW), and model combination with equal weight (EW). An asymptotic test of Diebold and Mariano (1995) is used for the null hypothesis of equal forecast accuracy between competing models and the benchmark. The results with significance levels of 0.05 and 0.1 are reported in Table 1.

Table 1: Real-time forecasts for PCE inflation.

	Relative MAFE				Relative CRPS			
	$k=1$	$k=4$	$k=8$	$k=16$	$k=1$	$k=4$	$k=8$	$k=16$
AR(3)	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
TVW								
TVC	1.01**	1.10**	1.11*	0.94	0.07	0.06	0.04	0.01
TVC-SV	0.94**	0.98*	0.99*	0.83	-0.02**	-0.03*	-0.05*	-0.08
TVC-SVMA	0.90**	1.03**	1.06*	0.89	-0.08**	-0.07**	-0.08**	-0.10
EW								
TVC	1.02**	1.11**	1.12*	0.95	0.06	0.05	0.03	0.00
TVC-SV	0.94**	0.98*	1.00*	0.83	-0.02**	-0.03*	-0.05*	-0.08
TVC-SVMA	0.90**	1.03**	1.06*	0.89	-0.08**	-0.08**	-0.09**	-0.11

Note: ** and * indicate superior forecast performance relative to AR(3) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

In Table 1, the real-time forecasting results suggest that the forecast combinations based on time-varying coefficient models outperform the benchmark for both point forecasts and density forecasts. In particular, the forecast combination of TVC-SV with time-varying coefficients improves forecast performance relative to the univariate model AR(3) all the time in Table 1. It supports the conclusion of Koop and Korobilis (2012) that allowing for changes of models over time is important for inflation forecasts.

The results of both the point density forecasts suggest that the specification of TVC-SVMA has the best forecast performance with both time-varying weights and equal weights combinations over all forecast horizons except TVC-SV in the point forecasts

in horizon one. For all three specifications (TVC, TVC-SV, and TVC-SVMA), the results from both equal weight and time-varying weight combinations have similar forecast performance. These results are in line with the findings that equal weights may not have worse forecast performance than time-varying weights (e.g., Stock and Watson, 2004; Clark and McCracken, 2009; Jore, Mitchell, and Vahey, 2010).

Comparing the forecast performance of TVC-SV and TVC-SVMA with that of TVC, it appears that adding either SV or SVMA specifications improves both point and density forecasts. Without SV, the forecast results of TVC are no better than the benchmark, while TVC-SV and TVC-SVMA have better forecast performance than the benchmark in most of the exercises.

Although the specifications of both TVC-SV and TVC-SVMA have competitive forecast performance in time-varying averaging and equal weight blocks, their performance does not consist of the point and density forecasts. The point forecasts of TVC-SVMA are not as good as those of TVC-SV on longer forecast horizons, while the density forecasts of TVC-SVMA have better performance than those of TVC-SV. This finding also suggests that when forecast combination is allowed, the specification of moving average does not significantly assist in point forecasts. For point forecast performance improvement, a model combination with various covariates is more important than various sepecifications of SV.

4.1 Weights from the top 10 models

Figure 2 to 4 report the sum of time-varying weights from the overall top 10 models on horizon one for TVC, TVC-SV and TVC-SVMA, respectively. Most of the figures have similar trends on the time-varying weights: the weights of the top 10 models become heavy until 2002, drop sharply in 2009, and then remain at low levels until 2014. This suggests that the top 10 models work particularly well approximately 2002 but are less suitable after the financial crisis. The results show that no single model can forecast inflation very well over the entire period and suggest that forecast combinations could provide a superior alternative to single model predictions.

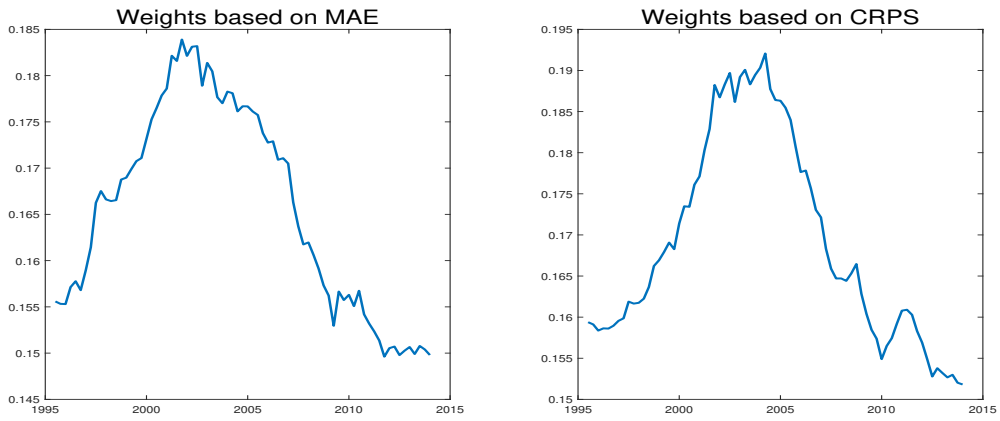

Figure 2: Time-varying weights from the top 10 models based on MAE and CRPS for TVC, forecasting horizon one.

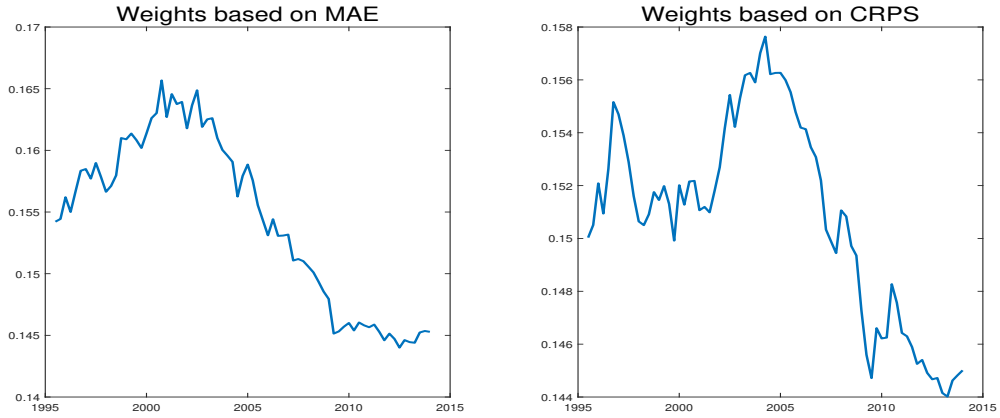

Figure 3: Time-varying weights from the top 10 models based on MAE and CRPS for TVC-SV, forecasting horizon one.

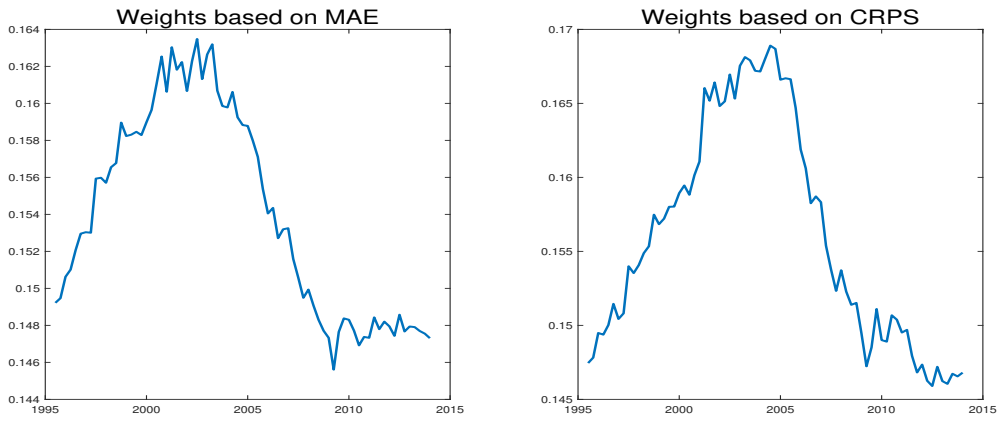

Figure 4: Time-varying weights from the top 10 models based on MAE and CRPS for TVC-SVMA, forecasting horizon one.

4.2 Weights of Inflation Predictors Grouped by Inflation Predictors

In this section, we present the time-varying weights computed for forecast combinations in the previous sections. For one-step-ahead point forecasts, the weights of 72 models with TVC-SVMA specification are grouped by eight inflation predictors (see Section 2.2) and are plotted in Figure 5. Similar plots for other forecast horizons are given in Appendix B.

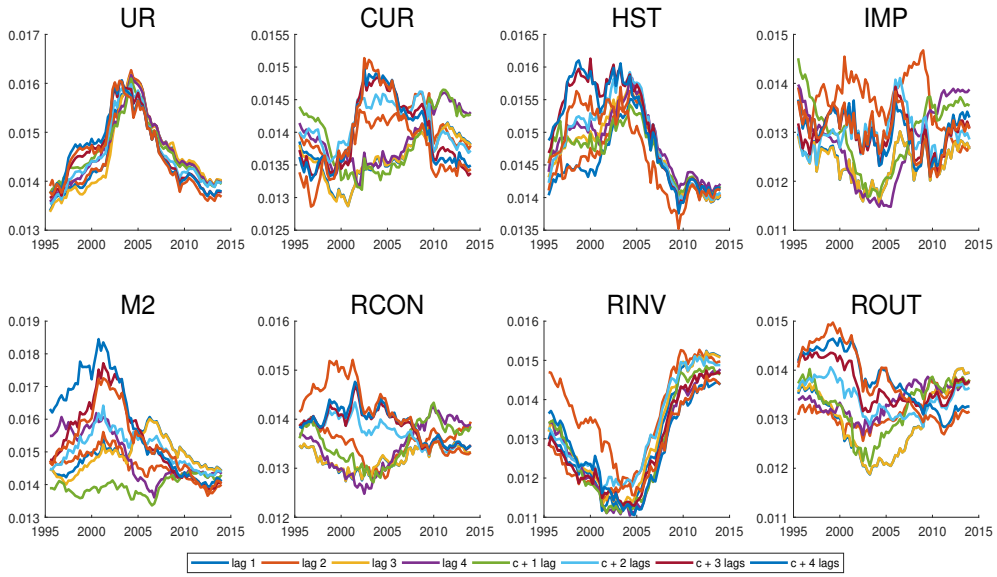

Figure 5: Time-varying weights based on MAE for TVC-SVMA specification, grouped by inflation predictors, forecasting horizon one.

The results in Figure 5 show that the weights of models with UR and HST decrease until 2005 and then begin to increase. In contrast, the weights of RINV increase until 2005 and then drop sharply until the end of the sample. The remaining regressors do not have such clear and consistent trends. This suggests that the choice of inflation predictor is important and that the best predictor changes over time. Moreover, no single inflation predictor dominates any other predictor across the examined time period. This result highlights the importance of forecast combinations.

4.3 Forecast Densities from Individual Models

Figure 6 and 7 plot the forecast densities for PCE inflation from the individual models TVC-SVMA in the first and third quarters of 2009, respectively. The vertical lines in Figure 6 and 7 stand for the actual values. US inflation drops very low in 2009Q1, while it grows up in 2009Q3. In 2009Q1, the medians of the individual forecast densities are away and larger than the actual value, especially for longer forecast horizons. However, the forecast medians keep moving back to the actual values as it shows in 2009Q3. The forecast densities of short forecast horizons are more diffuse than those of longer forecast horizons in 2009Q1 and 2009Q3, which indicates that there is more uncertainty from short forecast horizons than that from longer forecast horizons from individual models in 2009Q1 and 2009Q3.

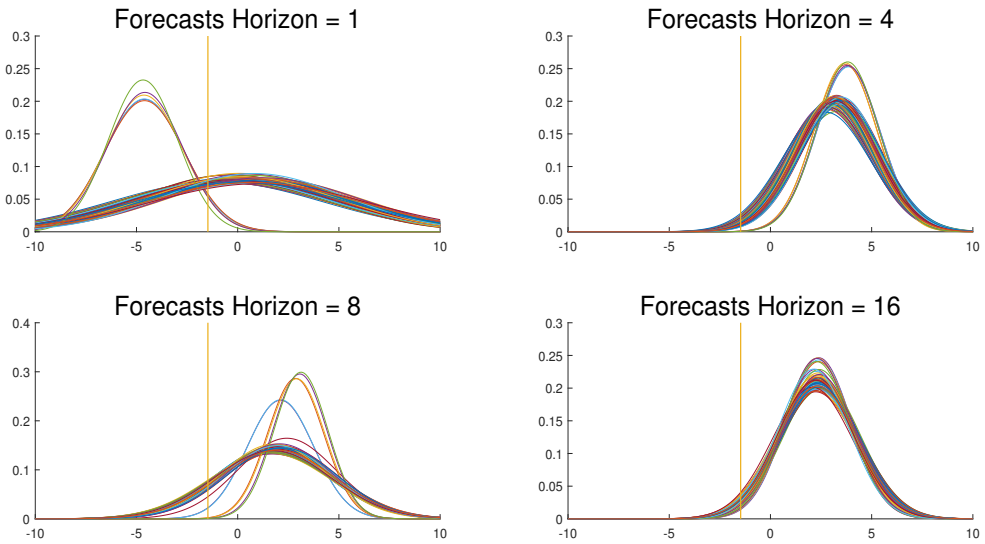

Figure 6: Forecast densities for TVC-SVMA specification from the combinations in 2009Q1.

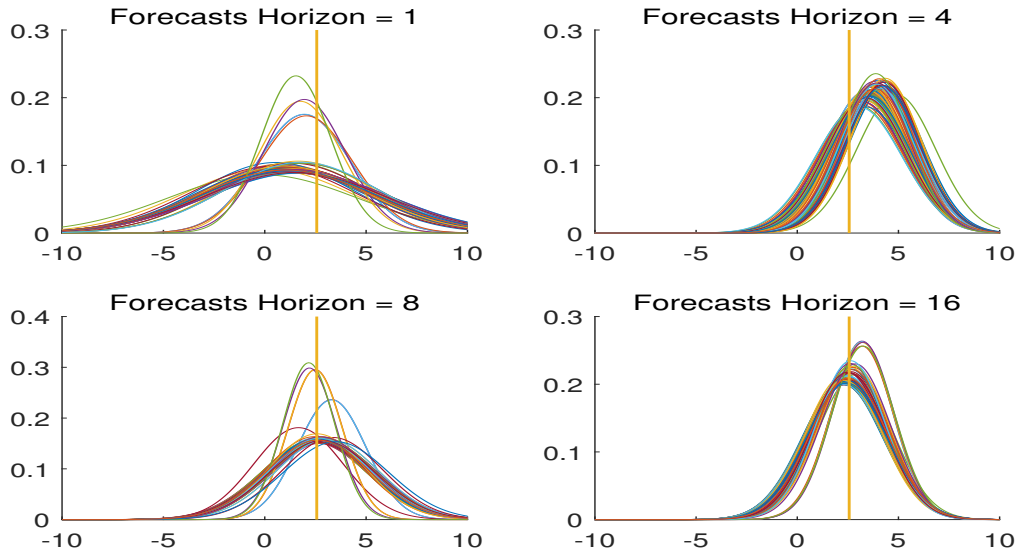

Figure 7: Forecast densities for TVC-SVMA specification from the individual models in 2009Q3.

4.4 Comparison with SPF in Financial Crisis

In this section, we compare the forecast performance of competing models with the forecasts produced by SPF for PCE inflation in 2009. In the SPF reports, only the forecast means of PCE inflation forecasts are available. Since these values are also the mean of the predictions, we list the point forecast results from the competing models in Table 2. To maintain consistency with the SPF forecasts, the forecast horizons are one quarter ahead, one year ahead and two years ahead. The actual values are real-time data. Models with closer forecast results to the actual values than those from the SPF are in bold.

The primary result is that the forecasts from the SPF are highly competitive, especially during the crucial time period of the financial crisis. Generally, the professional forecasters can rapidly adjust their views on the inflation forecasts, and their forecasts are quite close to the actual values, especially in the short horizons. In Table 2, the one-quarter-ahead forecast results show that the SPF is difficult for the competing models to beat. The SPF forecasts gave forecasts of -1 for 2009Q1, which is very close to the actual value -1.47, while all the results from the competing models are over 0 except TVC-SVMA. This is also true for the one-year and two-years-ahead forecasts for the first half year in 2009. However, the rapid adjustment of the SPF forecasts does not work well for longer forecast horizons. One year ago, the forecasts made by the SPF were over 2, which was an optimistic estimate of inflation similar to that from the competing models. Once the economy began to recover from the crisis, the forecasting models show better forecast performance for longer horizons in the second half of the year in 2009.

Table 2: Forecast results comparing with SPF in 2009.

2009Q1					*2009Q2*				
		$k=1$	$k=4$	$k=8$			$k=1$	$k=4$	$k=8$
RT Data	**-1.47**				RT Data	**1.1**			
SPF		-1	2.25	1.7	**SPF**		1.3	1.5	1.6
AR(3)		2.63	3.35	1.86	**AR(3)**		-0.11	3.21	2.88
TVW					*TVW*				
TVC		0.53	3.41	1.96	**TVC**		-1.43	3.23	2.69
TVC-SV		1.67	2.94	1.97	**TVC-SV**		0.55	2.86	2.37
TVC-SVMA		-0.22	3.18	1.9	**TVC-SVMA**		0.15	3.05	2.49
EW					*EW*				
TVC		0.48	3.42	1.96	**TVC**		-1.42	3.24	2.7
TVC-SV		1.65	2.94	1.97	**TVC-SV**		0.55	2.87	2.37
TVC-SVMA		-0.24	3.19	1.91	**TVC-SVMA**		0.14	3.06	2.5
2009Q3					*2009Q4*				
		$k=1$	$k=4$	$k=8$			$k=1$	$k=4$	$k=8$
RT Data	**2.57**				RT Data	**2.55**			
SPF		2.8	1.48	1.68	**SPF**		2.7	1.18	1.41
AR(3)		1.68	3.84	1.72	**AR(3)**		1.16	**2.64**	**2.64**
TVW					*TVW*				
TVC		0.92	4.02	**2.82**	**TVC**		**2.57**	0.53	3.15
TVC-SV		1.55	**3.34**	**2.45**	**TVC-SV**		2.75	**1.68**	**2.72**
TVC-SVMA		1.27	3.72	**2.68**	**TVC-SVMA**		2.96	**1.78**	**2.94**
EW					*EW*				
TVC		0.93	4.04	**2.82**	**TVC**		**2.6**	0.48	3.16
TVC-SV		1.55	**3.35**	**2.45**	**TVC-SV**		2.74	**1.65**	**2.72**
TVC-SVMA		1.28	3.73	**2.69**	**TVC-SVMA**		2.96	**1.74**	**2.95**

Note: ** and * indicate superior forecast performance relative to AR(3) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

4.5 Forecast Sensitivity Analysis

In this section, we conduct a sensitivity analysis of the forecast results by changing the window length to 20 and 60. Since the time-varying weight of each model is determined by the previous forecast performance in a certain window length, it is necessary to check whether shorter or longer window lengths could change the forecast results.

Table 3 and 4 report the forecast results of TVC, TVC-SV and TVC-SVMA with windows of 20 and 60 for forecast combination, respectively. While the short horizon forecast results between the 20 and 40 quarter windows are comparable, the shorter window tends to produce a worse forecast performance at the longest forecast horizon 16. This indicates that the length of 20 is too short and is insufficient for long horizon forecasts. In contrast, the results from the 60-quarter window are no better than those from the 40-quarter window at both the short and long forecast horizons. This suggests that too much information from a certain window is not helpful for combination forecasts.

Table 3: Real-time forecasts for PCE inflation, combination window = 20.

	Relative MAFE				Relative CRPS			
	$k=1$	$k=4$	$k=8$	$k=16$	$k=1$	$k=4$	$k=8$	$k=16$
AR(3)	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
TVW								
TVC	1.06**	1.13**	1.15*	1.07*	0.12**	0.10**	0.08*	0.05*
TVC-SV	0.94**	0.96**	0.98**	0.91*	-0.02**	-0.02**	-0.04**	-0.07
TVC-SVMA	0.92**	1.04**	1.08**	1.00*	-0.05**	-0.04**	-0.06**	-0.08
EW								
TVC	1.07**	1.14**	1.15*	1.08*	0.10**	0.08**	0.06*	0.03*
TVC-SV	0.94**	0.96**	0.98**	0.92*	-0.03**	-0.02**	-0.04**	-0.07*
TVC-SVMA	0.93**	1.05**	1.08**	1.01*	-0.05**	-0.05**	-0.06**	-0.09*

Note: ** and * indicate superior forecast performance relative to AR(3) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 4: Real-time forecasts for PCE inflation, combination window = 60.

	Relative MAFE				Relative CRPS			
	$k=1$	$k=4$	$k=8$	$k=16$	$k=1$	$k=4$	$k=8$	$k=16$
AR(3)	1.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00
TVW								
TVC	1.10**	1.17*	1.12	1.06	0.15**	0.14**	0.14	0.14
TVC-SV	1.05**	0.96**	1.03	0.95*	0.09**	-0.01**	0.09*	0.08
TVC-SVMA	1.02**	1.11**	1.08*	1.01	-0.08**	-0.07**	-0.08*	-0.10
EW								
TVC	1.10**	1.17**	1.12	1.06	0.15**	0.14**	0.13*	0.13
TVC-SV	1.05*	0.96**	1.04	0.96*	0.09**	-0.02**	0.08*	0.08
TVC-SVMA	1.02**	1.12**	1.08*	1.01	0.04**	0.02**	0.03*	0.04

Note: ** and * indicate superior forecast performance relative to AR(3) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

5 Concluding Remarks

In this study, forecast combinations using both time-varying and equal weights of 72 component models have been investigated. Three specifications of time-varying coefficient models were employed for PCE inflation forecasts in the US. In the real-time forecasting exercise, forecast combinations tend to improve the univariate model AR(3) with only inflation, which suggests that adding variables significantly improves forecast performance. Both point forecasts and density forecasts suggest that equal weights and time-varying weights have similar forecast performance.

Our results show that the specification of TVC-SVMA has the best forecast performance

in both point and density forecasts most of the time. The analysis of weights among inflation predictors and lag lengths for TVC further suggests that an inflation predictor choice is more important than the choice of the lag length.

Comparing with the benchmark, both TVC-SV and TVC-SVMA have better forecast performances in both point and density forecasts. It suggests that allowing stochastic volatility in forecast combinations is helpful for improving forecast performance. Moreover, the moving average does not seem to improve forecast performance over stochastic volatility models in point forecasts but works well in density forecasts.

For the special time period of the 2007/08 financial crisis, the forecasting results of the proposed combination models are highly competitive compared with the quarterly reports from the Survey of Professional Forecasters, which is gerenally hard to beat.

Acknowledgements

I am thankful to my supervisor, Joshua C.C. Chan, for his helpful comments. I am also grateful to Jamie Cross, Timothy Kam, Chenghan Hou and two anonymous referees for their detailed suggestions which helped the paper evolve considerably.

Appendix A: Bayesian Estimation Method: MCMC Algorithm

The model is estimated by Bayesian methods using the priors specified in Section 2.4. Specifically, a Markov chain Monte Carlo (MCMC) algorithm is derived to sample from the joint posterior distribution. In particular, Gibbs sampling, the Metropolis-Hastings algorithm, and precision-based algorithm are employed for simulation.

Here, we take the estimation of parameters in TVC-SVMA as an example. There are five blocks in the MCMC algorithm. we draw $\boldsymbol{\beta}$, stochastic volatility parameter $\mathbf{h}$, the covariance matrix $\mathbf{Q}$ of $\boldsymbol{\beta}$, and the moving average coefficient ψ sequentially by conditioning on the inflation and inflation predictors:

1. $p(\boldsymbol{\beta} \,|\, \mathbf{y}, \mathbf{u}, \psi, \mathbf{h}, \mathbf{Q})$;
2. $p(\mathbf{h} \,|\, \mathbf{y}, \mathbf{u}, \psi, \sigma_h^2)$;
3. $p(\mathbf{Q} \,|\, \mathbf{y}, \mathbf{u}, \boldsymbol{\beta})$;
4. $p(\psi \,|\, \mathbf{y}, \mathbf{u}, \boldsymbol{\beta}, \mathbf{h})$;
5. $p(\sigma_h^2 \,|\, \mathbf{h})$.

In the first step, we draw vectors of the intercept and coefficients $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_j$ together by rearranging them into one vector $\boldsymbol{\beta} = (\beta_{1,1}, \beta_{2,1}, \beta_{j,1}, \ldots, \beta_{1,T}, \beta_{2,T}, \beta_{j,T})'$. Equation (6) and (8) can be stacked over t

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}^y, \tag{21}$$

$$\boldsymbol{\varepsilon}^y = \mathbf{H}_\psi \mathbf{u}. \tag{22}$$

Hence,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{H}_\psi \mathbf{u}. \tag{23}$$

This is a standard normal model. Normal conjugate priors are assumed for the variables and the intercept and coefficients $\boldsymbol{\beta}$, so that the posterior is also normal. In fact, they are:

$$(\boldsymbol{\beta} \,|\, \mathbf{y}, \mathbf{u}, \psi, \mathbf{h}, \mathbf{Q}) \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}, \mathbf{D}_{\boldsymbol{\beta}}).$$

The TVC-SVMA model has a similar model structure as the state space model used by Chan (2013); the precision sampler can be used (Chan and Jeliazkov, 2009) to draw the posterior of $\boldsymbol{\beta}$ efficiently.

The second step draws the stochastic volatility $\mathbf{h}$ conditioning on the observed variables, and other parameters. This is done using the auxiliary mixture sampler of Kim et al. (1998). This method provides an accessible and efficient approximation of a nonlinear stochastic volatility model using a mixture of linear Gaussian state space models.

The third step draws the covariance matrix of $\boldsymbol{\beta}$. Stacking (7) over t:

$$\mathbf{H}\boldsymbol{\beta} = \boldsymbol{\varepsilon}^{\beta}, \quad \boldsymbol{\varepsilon}^{\beta} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}), \tag{24}$$

where

$$\boldsymbol{\Omega} = diag(\mathbf{Q}_0, \mathbf{Q}, \cdots, \mathbf{Q}).$$

As $\mathbf{Q}$ is assumed to be diagonal, $\boldsymbol{\Omega}$ is also diagonal. Thus, σ_β^2 in step 3 and σ_h^2 in step 5 are both conditionally independent. Both of them have conjugate inverse-gamma priors, and their posteriors can be estimated by the standard method discussed in Koop (2003).

Below, the details of each step are presented for deriving the conditional posterior distributions of parameters in the proposed model.

Step 1: Sampling for the Coefficient Parameter $\boldsymbol{\beta}$

To derive the conditional posterior distribution $p(\boldsymbol{\beta} \,|\, \mathbf{y}, \mathbf{u}, \psi, \mathbf{h}, \mathbf{Q})$, we first have:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{H}_\psi \mathbf{u},$$
$$\boldsymbol{\beta} = \mathbf{H}^{-1}\boldsymbol{\varepsilon}_{\boldsymbol{\beta}},$$

where

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ 0 & \cdots & -1 & 1 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & -1 & 1 \end{pmatrix}, \quad \mathbf{H}_\psi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \psi_1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \psi_p & \cdots & \psi_1 & 1 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \psi_p & \cdots & \psi_1 & 1 \end{pmatrix}.$$

Then we pre-multiply both sides of the above two equations by $\mathbf{H}_\psi^{-1}$:

$$\widetilde{\mathbf{y}} = \mathbf{X}\widetilde{\boldsymbol{\beta}} + \mathbf{u},$$
$$\widetilde{\boldsymbol{\beta}} = \mathbf{H}_\psi^{-1}\mathbf{H}^{-1}\boldsymbol{\varepsilon}_{\boldsymbol{\beta}},$$

where $\widetilde{\mathbf{y}} = \mathbf{H}_\psi^{-1}\mathbf{y}$ and $\widetilde{\boldsymbol{\tau}} = \mathbf{H}_\psi^{-1}\boldsymbol{\tau}$.

Thus, the log posterior density for $\widetilde{\boldsymbol{\beta}}$ is:

$$\log p(\widetilde{\boldsymbol{\beta}} \,|\, \widetilde{\mathbf{y}}, \mathbf{h}, \psi, \sigma_\beta^2) \propto \log p(\widetilde{\boldsymbol{\beta}} \,|\, \sigma_\beta^2) + \log p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\beta}}, \mathbf{h}, \psi), \tag{25}$$

where $p(\widetilde{\boldsymbol{\beta}} \,|\, \sigma_\beta^2)$ is the prior for $\widetilde{\boldsymbol{\beta}}$ and $p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\beta}}, \mathbf{h}, \psi)$ is the likelihood for $\widetilde{\mathbf{y}}$. So that the log-likelihood of $\widetilde{\mathbf{y}}$ and the prior of $\widetilde{\boldsymbol{\beta}}$ are:

$$\log p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\beta}}, \mathbf{h}, \phi, \psi) \propto -\frac{1}{2}\sum_{t=1}^{T} h_t - \frac{1}{2}(\widetilde{\mathbf{y}} - \mathbf{X}\widetilde{\boldsymbol{\beta}})'\boldsymbol{\Omega}_u^{-1}(\widetilde{\mathbf{y}} - \mathbf{X}\widetilde{\boldsymbol{\beta}}), \tag{26}$$

$$\log p(\widetilde{\boldsymbol{\beta}} \,|\, \sigma_\tau^2) \propto -\frac{T-1}{2}\log \sigma_\beta^2 - \frac{1}{2}\widetilde{\boldsymbol{\beta}}'\mathbf{X}'\mathbf{H}_\psi'\mathbf{H}'\mathbf{Q}^{-1}\mathbf{H}\mathbf{H}_\psi\mathbf{X}\widetilde{\boldsymbol{\beta}}, \tag{27}$$

where $\boldsymbol{\Omega}_u = diag(\mathrm{e}^{h_1}, \cdots, \mathrm{e}^{h_t})$.

Putting (27) and (26) into (25), we have:

$$\begin{aligned}
\log p(\widetilde{\boldsymbol{\beta}} \,|\, \widetilde{\mathbf{y}}, \mathbf{h}, \phi, \psi, \sigma_\tau^2) &\propto -\frac{1}{2}\widetilde{\boldsymbol{\beta}}'\mathbf{H}_\psi'\mathbf{H}'\mathbf{Q}^{-1}\mathbf{H}\mathbf{H}_\psi\widetilde{\boldsymbol{\beta}} - \frac{1}{2}(\widetilde{\mathbf{y}} - \mathbf{X}\widetilde{\boldsymbol{\beta}})'\boldsymbol{\Omega}_u^{-1}(\widetilde{\mathbf{y}} - \mathbf{X}\widetilde{\boldsymbol{\beta}}) \\
&\propto -\frac{1}{2}(\widetilde{\boldsymbol{\beta}}'(\mathbf{H}_\psi'\mathbf{H}'\mathbf{Q}^{-1}\mathbf{H}\mathbf{H}_\psi \ + \mathbf{X}'\boldsymbol{\Omega}_u^{-1}\mathbf{X})\widetilde{\boldsymbol{\beta}} - 2\mathbf{X}'\widetilde{\boldsymbol{\beta}}'\boldsymbol{\Omega}_u^{-1}\widetilde{\mathbf{y}}) \\
&\propto -\frac{1}{2}(\widetilde{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}})'\mathbf{D}_{\widetilde{\beta}}^{-1}(\widetilde{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}),
\end{aligned}$$

where $\mathbf{D}_{\widetilde{\beta}} = (\mathbf{H}_\psi'\mathbf{H}'\mathbf{Q}^{-1}\mathbf{H}\mathbf{H}_\psi + \mathbf{X}'\boldsymbol{\Omega}_u^{-1}\mathbf{X})^{-1}$ is a sparse matrix and $\widehat{\boldsymbol{\beta}} = \mathbf{D}_{\widetilde{\beta}}\mathbf{X}'\boldsymbol{\Omega}_u^{-1}\widetilde{\mathbf{y}}$. So that:

$$(\widetilde{\boldsymbol{\beta}} \,|\, \widetilde{\mathbf{y}}, \mathbf{h}, \phi, \psi, \sigma_\tau^2) \sim \mathcal{N}(\widehat{\boldsymbol{\beta}}, \mathbf{D}_{\widetilde{\beta}}).$$

By Cholesky decomposition and forward and backward substitution (details in Chan (2013)), we finally sample $\boldsymbol{\beta}$ by $\boldsymbol{\beta} = \mathbf{H}_\psi\widetilde{\boldsymbol{\beta}}$.

Step 2: Sampling for h

We follow Kim et al. (1998) to sample the SV component $\mathbf{h}$, which uses an auxiliary mixture of seven normal distributions to draw $\mathbf{h}$ efficiently. In practice, we adopt the algorithm used in Chan (2013), which is a precision-based sampler, instead of the forward-backward smoothing algorithm used in Kim et al. (1998). To apply this method, we first define:

$$\mathbf{y}^* = \mathbf{H}_\psi^{-1}(\mathbf{y} - \boldsymbol{\beta}\mathbf{X})),$$

So that $\mathbf{y}^* = \mathbf{u}, \mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_y)$, where $\mathbf{S}_y = \mathrm{diag}(e^{h_1}, \cdots, e^{h_T})$. The detailed sampling procedure can also be obtained from Koop and Korobilis (2009).

Step 3: Sampling for σ_h^2 and Q

We assume that σ_h^2 and the diagonal elements of $\mathbf{Q}$ $(\sigma_{\beta_1}^2, \cdots, \sigma_{\beta_k}^2)$ are conditionally independent, so that the derivations of their posteriors can follow the standard method discussed in Koop (2003). Thus, their posteriors can be obtained after a simple transformation. Given a conjugate inverse-gamma prior $\sigma_h^2 \sim \mathcal{IG}(\nu_h, S_h)$, we can derive an

inverse-gamma posterior for $(\sigma_h^2 \,|\, \mathbf{h})$,

$$\begin{aligned} p(\sigma_h^2 \,|\, \mathbf{h}) &\propto p(\mathbf{h} \,|\, \sigma_h^2) + p(\sigma_h^2) \\ &= (\sigma_h^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma_h^2} \sum_{t=2}^{T} (h_t - h_{t-1})^2 \right) \cdot (\sigma_h^2)^{-(\nu_{0h}-1)} \exp(-\frac{S_h}{\sigma_h^2}) \\ &\propto (\sigma_h^2)^{-((\frac{T}{2}+\nu_0)-1)} \exp\left(-\frac{1}{\sigma_h^2} (\sum_{t=2}^{T} (h_t - h_{t-1})^2/2 + S_h) \right). \end{aligned}$$

That is:

$$(\sigma_h^2 \,|\, \mathbf{h}) \sim \mathcal{IG}\left(T/2 + \nu_h, \quad \sum_{t=2}^{T} (h_t - h_{t-1})^2/2 + S_h \right).$$

Similarly, the posterior density of each diagonal element in $\mathbf{Q}$ can be derived as above, and we can stack them into matrix form in empirical simulation:

$$(\mathbf{Q} \,|\, \boldsymbol{\beta}) \sim \mathcal{IG}\left(T/2 + \nu_\beta, \quad \sum_{t=2}^{T} (\mathbf{H}\boldsymbol{\beta})^2/2 + S_\beta \right).$$

Step 4: Sampling for ψ

Note that given $\mathbf{y}, \boldsymbol{\beta}$ and $\mathbf{h}, \psi$ is conditionally independent from σ_h^2 and $\mathbf{Q}$; thus, we can draw σ_h^2, $\mathbf{Q}$ and ψ sequentially in the simulation. We first stack (6) and (8) into matrix form:

$$\mathbf{H}_\phi(\mathbf{y} - \boldsymbol{\tau}) = \mathbf{H}_\psi \mathbf{u}. \tag{28}$$

Then, the log-likelihood of posterior $(\psi \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h})$ is:

$$\begin{aligned} \log p(\psi \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}) &\propto \log p(\mathbf{y} \,|\, \psi, \boldsymbol{\tau}, \mathbf{h}) + \log p(\psi) \\ &\propto \log p(\psi) - \frac{1}{2} (\mathbf{H}_\phi(\mathbf{y} - \boldsymbol{\tau}))' (\mathbf{H}_\psi' \boldsymbol{\Omega}_u \mathbf{H}_\psi)^{-1} \mathbf{H}_\phi(\mathbf{y} - \boldsymbol{\tau}). \end{aligned}$$

Unlike $\mathbf{y}, \boldsymbol{\beta}, \sigma_h^2$ or σ_τ^2, the distributions of moving average parameters ψ are unknown, and ψ is typically low dimensional in empirical economic studies (Chan, 2013), so it can be evaluated numerically by maximizing $\log p(\psi|\mathbf{y}, \boldsymbol{\beta}, \mathbf{h})$ to obtain the mode and the negative Hessian evaluated at the mode. Then we use the Metropolis-Hastings algorithm detailed in Chan (2013) to sample ψ, which is widely used to simulate the posterior of a multivariate model. The proposed density for ψ is multi-normal distribution $q(\psi)$, and the updated ψ^c is accepted with the probability:

$$\min\{1, \quad \frac{p(\psi^c \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h})}{p(\psi \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h})} \cdot \frac{q(\psi)}{q(\psi^c)} \}.$$

Appendix B: Figures of Weights for Each Inflation Predictor and Lag Forms

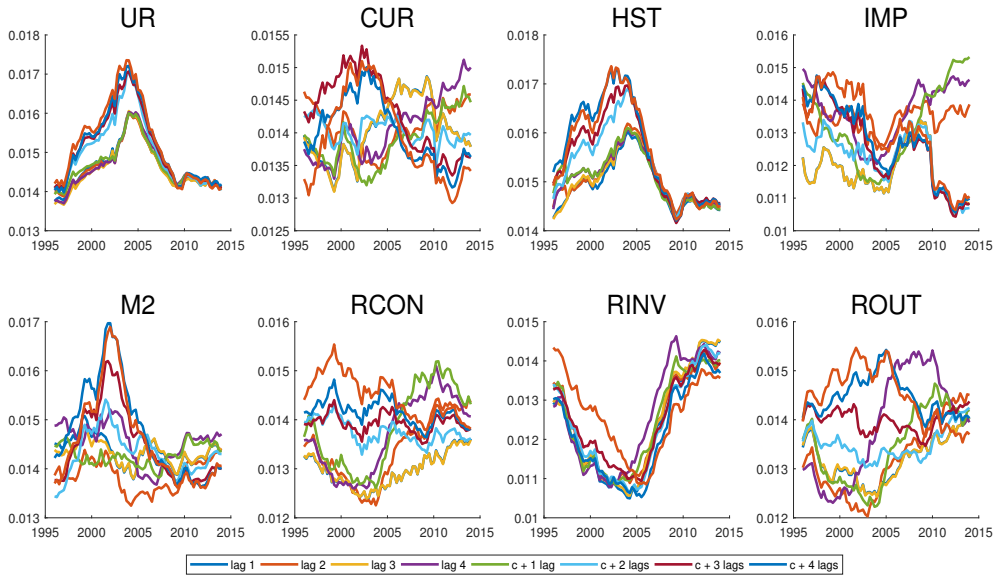

Figure 8: Time-varying weights based on MAE for TVC-SVMA specification, grouped by inflation predictors, forecast horizon four.

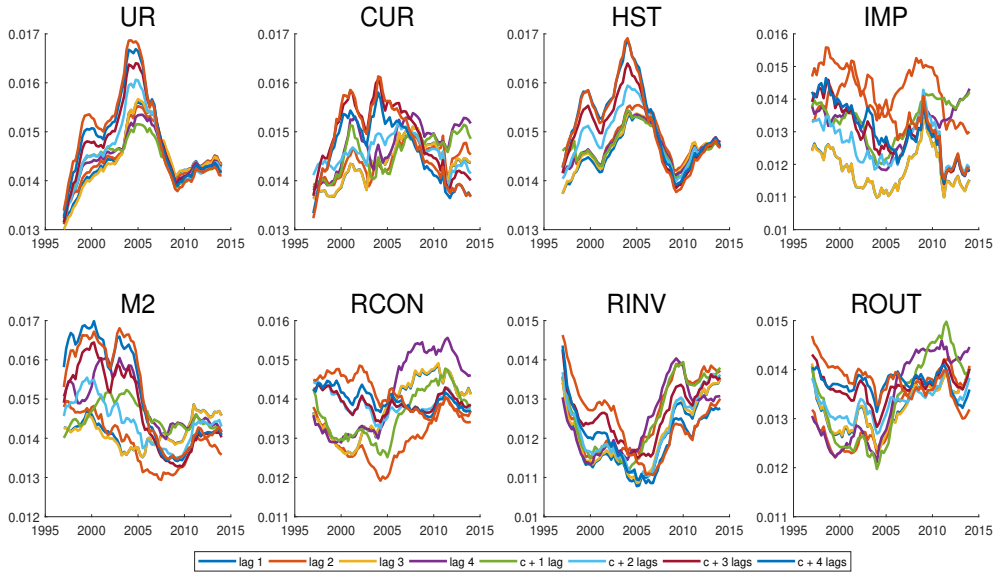

Figure 9: Time-varying weights based on MAE for TVC-SVMA specification, grouped by inflation predictors, forecast horizon eight.

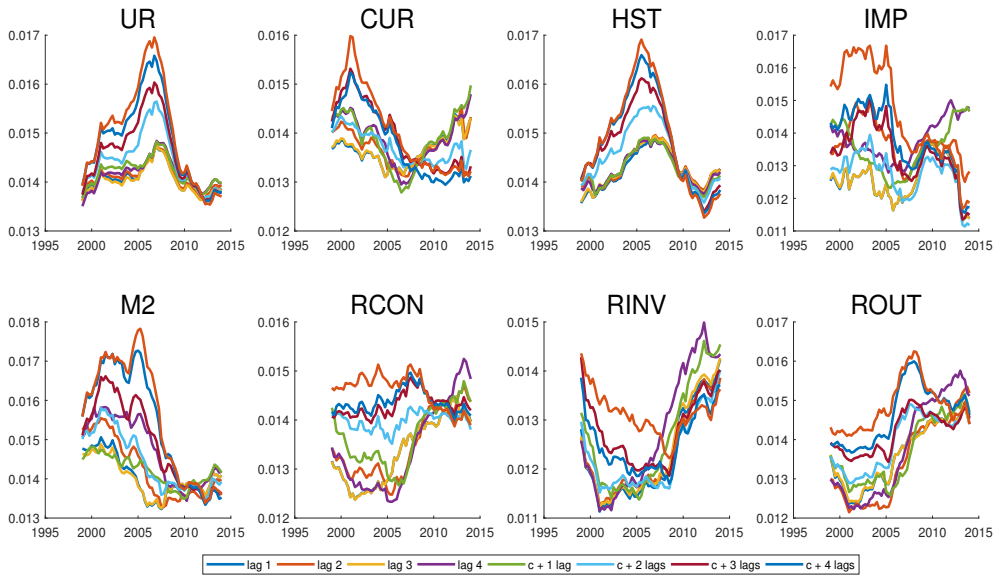

Figure 10: Time-varying weights based on MAE for TVC-SVMA specification, grouped by inflation predictors, forecast horizon sixteen.

References

A. Atkeson and L. E. Ohanian. Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1):2–11, 2001.

D. W. Bunn. A Bayesian approach to the linear combination of forecasts. *Operational Research Quarterly*, 26(2):325–329, 1975.

J. C. C. Chan. Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, 176(2):162–172, 2013.

J. C. C. Chan and I. Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1:101–120, 2009.

T. E. Clark. Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. *Journal of Business and Economic statistics*, 29(3):327–341, 2011.

T. E. Clark and M. W. McCracken. Averaging forecasts from VARs with uncertain instabilities. *Journal of Applied Econometrics*, 25(1):5–29, 2009.

T. E. Clark and F. Ravazzolo. Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575, 2015.

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

T. Cogley and A. M. Sbordone. Trend inflation and inflation persistence in the new Keynesian Phillips curve. *American Economic Review*, 98(5):2101–2126, 2008.

V. Corradi, A. Fernandez, and N. R. Swanson. Information in the revision process of real-time datasets. *Journal of Business and Economic Statistics*, 27:455–467, 2009.

D. Croushore and T. Stark. A real-time data set for macroeconomists. *Journal of Econometrics*, 105(1):111–130, 2001.

D. Croushore and T. Stark. A real-time data set for macroeconomists: Does the data vintage matter? *The Review of Economics and Statistics*, 85(3):605–617, 2003.

F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Econometric Statistics*, 13(3):134–144, 1995.

A. E. Faria and E. Mubwandarikwa. The geometric combination of Bayesian forecasting models. *Journal of Forecasting*, 27(6):519–535, 2008.

A. Garratt, G. Koop, E. Mise, and S. P. Vahey. Real-time prediction with U.K. monetary aggregates in the presence of model uncertainty. *Journal of Business and Economic Statistics*, 27(4):480–491, 2009.

A. Garratt, J. Mitchell, S. P. Vahey, and E. C. Wakerly. Real-time inflation forecast densities from ensemble Phillips curve. *The North American Journal of Economics and Finance*, 22(1):77–87, 2011.

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

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

M. A. Hooker. Are oil shocks inflationary? Asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit and Banking*, 34(2):540–561, 2002.

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

C. Kascha and F. Ravazzolo. Combining inflation density forecasts. *Journal of Forecasting*, 29:231–250, 2010.

S. Kim, N. Shepherd, and S. Chib. Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65:361–393, 1998.

G. Koop. *Bayesian Econometrics*. John Wiley and Sons, New York, 2003.

G. Koop and D. Korobilis. Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends in Econometrics*, 3(4):267–358, 2009.

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

D. M. Lakova. Flattening of the Phillips curve; implications for monetary policy. *IMF Working Papers*, 07/76, 2007.

F. S. Mishkin. Inflation dynamics. *International Finance*, 10(3):317–334, 2007.

A. E. Raftery, D. Madigan, and J. A. Hoeting. Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92:179–191, 1997.

F. Ravazzolo and S. P. Vahey. Forecast densities for economic aggregates from disaggregate ensembles. *Studies in Nonlinear Dynamics and Econometrics*, 18(4):1–15, 2014.

M. S. Smith and S. P. Vahey. Asymmetric forecast densities for U.S. macroeconomic variables from a Gaussian copula model of cross-sectional and serial dependence. *Journal of Business and Economic Statisitics*, 34(3):416–434, 2016.

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

J. H. Stock and M. W. Watson. Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, 39:3–33, 2007.

L. Tibiletti. A non-linear combination of experts' forecasts: a Bayesian approach. *Journal of Forecasting*, 13(1):21–27, 1994.

A. Timmerman. Forecast combination. In G. Elliott, C. W. J. Granger, & A. Timmerman (Eds.) Handbook of Economic Forecasting. 1:135–196, 2003.