Stochastic Volatility Models with ARMA Innovations: An Application to G7 Inflation Forecasts

Bo Zhang
University of Wollongong

Joshua C. C. Chan
Purdue University, UTS and CAMA

Jamie L. Cross
BI Norwegian Business School

February 2019

Abstract

We introduce a new class of stochastic volatility models with autoregressive moving average (ARMA) innovations. The conditional mean process has a flexible form that can accommodate both a state space representation and a conventional dynamic regression. The ARMA component introduces serial dependence which renders standard Kalman filter techniques not directly applicable. To overcome this hurdle we develop an efficient posterior simulator that builds on recently developed precision-based algorithms. We assess the usefulness of these new models in an inflation forecasting exercise across all G7 economies. We find that the new models generally provide competitive point and density forecasts compared to standard benchmarks, and are especially useful for Canada, France, Italy and the US.

Keywords: autoregressive moving average errors, stochastic volatility, inflation forecast, state space models, unobserved components model.

JEL classification codes: C11, C52, C53, E37

1 Introduction

Following the seminal work of Box and Jenkins (1970), autoregressive moving average (ARMA) models have become the standard tool for modeling and forecasting univariate time series. More recently, coefficient instability in macroeconomic time series models has been widely acknowledged (see, e.g., Stock and Watson, 1996; Ludbergh et al., 2003; Marcellino, 2004; Stock and Watson, 2007; Cross and Poon, 2016). For example, Stock and Watson (2007) show that US CPI inflation is best modeled by an unobserved components model in which both the transitory and trend equations allow for time-varying volatility. It is therefore necessary to extend the ARMA framework to allow for the possibility of heteroscedasticity.

There are two popular approaches to achieve this objective: autoregressive conditional heteroscedasticity models of Engle (1982)—or their generalized counterparts introduced in Bollerslev (1986) called GARCH models—and stochastic volatility (SV) models (Taylor, 1994). In a recent paper, Clark and Ravazzolo (2015) put these two classes of models head-to-head in a forecasting exercise involving a few key US macroeconomic time series. They find that stochastic volatility models generally provide superior point and density forecasts across all variables. Thus, SV seems to be a more appropriate specification, at least for macroeconomic forecasting. In addition, given the historical success of ARMA models and their more recent time-varying extensions, one might expect a more flexible class of dynamic models with ARMA and SV errors could further improve the forecast performance. This idea is partially investigated by Chan (2013), who shows that MA-SV errors are useful in forecasting US inflation. Nonetheless, it remains to be seen whether more general ARMA-SV errors can further enhance forecast accuracy.[1]

With this idea in mind, our objective in this paper is to investigate the forecast performance of a new class of ARMA-SV models. By allowing the conditional mean process to have a flexible state space representation, our general framework is able to accommodate numerous popular specifications, such as unobserved components and time varying parameter models, as special cases. Thus, from a methodological perspective, our models can be viewed as an extension of the unobserved components model in Stock and Watson (2007) to follow for ARMA errors. In addition, we also extend the recent work of Chan (2013) to allow for a more flexible error structure.

[1]Nakatsuma (2000) develops a posterior simulator for estimating ARMA-GARCH models. Thus, our paper can also be viewed as filling an important gap in completing the econometricians toolbox of possible ARMA error models.

A second contribution of our paper is to develop an efficient posterior simulator to estimate this new class of models. A major computational hurdle is that the ARMA component introduces serial dependence into the measurement equation, which makes standard Kalman filter techniques not directly applicable. To overcome this issue, one may seek a suitable transformation of the data to make the errors serially independent (Chib and Greenberg, 1994). Here we take a different route and build upon recent advances in precision-based algorithms for state space models, which have been shown to be computationally more efficient than Kalman filter based methods (Chan and Jeliazkov, 2009; McCausland et al., 2011). The key to our efficiency gain is that despite having a full covariance matrix implied by the ARMA structure, we can work with only band matrices, which substantially speed up the computations. In this way, we are able to overcome the computational challenge of having a full covariance matrix and maintain the same type of advantages obtained by Chan (2013).

A third contribution of our paper is that we provide a substantive forecasting exercise involving two commonly used inflation measures: CPI and the GDP Deflator, across the G7 economies. Since inflation plays a key role in modern monetary policy, any forecast improvement over the standard set of benchmark models will have substantive practical significance. Our primary result is that the proposed ARMA-SV models generally provide superior out-of-sample point and density forecasts, and are especially useful for Canada, France, Italy and the US. From an empirical perspective, given that our study includes state-of-the-art models such as UC-SV and UC-MA-SV, our analysis can be viewed as an extension of the results presented in Stock and Watson (2007) and Chan (2013).

The rest of the paper is structured as follows. Section 2 presents the general ARMA-SV framework, discusses how this embeds a variety of popular model specifications, and develops an efficient posterior simulator to estimate this new class of models. In Section 3, we discuss our application of forecasting CPI and GDP Deflator inflation measures in each of the G7 countries. We further provide a comparison with the Survey of Professional Forecasters for US data in Section 4. Finally, we conclude in Section 5.

2 Stochastic Volatility Models with ARMA Errors

In this section we first introduce a class of stochastic volatility models with ARMA errors and provide some theoretical motivation. We then outline our computational approach to estimate this class of models efficiently using recent advances in precision-

based algorithms.

First let y_t denote a variable of interest at date t, where $t = 1, \ldots, T$. Then, the general state space representation of the ARMA(p, q)-SV framework is given by:

$$y_t = \mu_t + \varepsilon_t^y, \tag{1}$$

$$\varepsilon_t^y = \phi_1 \varepsilon_{t-1}^y + \cdots + \phi_p \varepsilon_{t-p}^y + u_t + \psi_1 u_{t-1} + \cdots + \psi_q u_{t-q}, \quad u_t \sim \mathcal{N}(0, \mathrm{e}^{h_t}), \tag{2}$$

$$h_t = h_{t-1} + \varepsilon_t^h, \quad \varepsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \tag{3}$$

where the error terms ε_t^τ and u_t are independent across all leads and lags. We assume for simplicity that the initial conditions $\varepsilon_0 = \cdots = \varepsilon_{-p} = u_0 = \cdots = u_{-q} = 0$. One can treat these initial innovations as parameters to be estimated if desired, and the estimation procedures discussed in the next section can be straightforwardly extended to allow for this possibility. For typical situations where T is much larger than p and q, whether these initial conditions are modeled explicitly or not makes little difference in practice.

In the measurement equation (1) above, we leave the time-varying conditional mean μ_t unspecified. By choosing a suitable time-varying process for μ_t, our framework can accommodate a variety of popular model specifications. Two such examples are:

1. The autoregressive model:

$$\mu_t = \rho_0 + \sum_{i=1}^{m} \rho_i y_{t-i} \tag{4}$$

2. The unobserved components model:

$$\tau_t = \tau_{t-1} + \varepsilon_t^\tau, \quad \varepsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \tag{5}$$

Of course, many other models can be considered as well, such as the standard linear regression model with constant or time-varying coefficients.

Equation (2) specifies an ARMA(p, q) error structure, where the variance $\exp(h_t)$ follows a stochastic volatility process — i.e., the log-volatility h_t evolves according to the random walk in Equation (3). Moreover, the ARMA(p, q) process is assumed to be stationary and invertible. Specifically, rewrite (2) in terms of two polynomials:

$$\phi(L)\varepsilon_t^y = \psi(L)u_t,$$

where $\phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$, $\psi(L) = 1 + \psi_1 L + \cdots + \psi_q L^q$ and L is the lag operator. Then, we assume that all roots of $\phi(L)$ and $\psi(L)$ are outside of the unit circle for stationarity and invertibility, respectively (Chib and Greenberg, 1994).

Under our framework, y_t is decomposed into a (typically) non-stationary process μ_t and a stationary error process ε_t^y, which we model as ARMA(p, q). The theoretical justification of the latter choice is the well-known Wold decomposition theorem, which states that any zero mean covariance-stationary time series has an infinite moving average representation. An implication of this theorem is that any covariance-stationary process can be approximated arbitrarily well by a sufficiently high order ARMA model.

A second motivation of our framework in (1)–(3) is that it can include many state-of-the-art forecasting models for inflation as special cases. In other words, we can embed many seemingly different models within a unifying framework. One prominent example is the UCSV model of Stock and Watson (2007), which has become the benchmark for forecasting inflation. In our framework it amounts to assuming the conditional mean μ_t to follow a random walk process with SV, and turning off both the AR and MA components—i.e., setting $\phi_i = \psi_j = 0$ for all $i = 1, \ldots, p$ and $j = 1, \ldots, q$. One seemingly restrictive assumption in the UCSV model is that the transitory component of inflation or the inflation gap is in fact serially independent. More recent papers have allowed for some form of serial dependence in the transitory component and showed that these extensions better forecast inflation. For example, Chan (2013) proposes a variant where the transitory component has an MA structure. By contrast, Clark and Doh (2014) considers a version in which the inflation gap follows an AR(1) process. Both of these variants of the UCSV model can naturally be included within the proposed framework.

2.1 Estimation

We estimate the model with Bayesian methods and simulate the joint distributions of interest through an efficient Metropolis-within-Gibbs sampler that builds upon recent developments in precision-based algorithms. The main challenge is that due to the ARMA error structure, the covariance matrix of the joint distribution for $\mathbf{y} = (y_1, \ldots, y_T)'$ is a full matrix. Therefore, in order to use the conventional Kalman filter, the original data need to be transformed so that the transformed errors are serially independent. Here, however, we extend the results in Chan (2013) and employ a direct approach using precision-based algorithms. Specifically, for the case of MA errors, Chan (2013) exploits the fact that

the implied inverse covariance matrix or the *precision matrix* is banded—i.e., it has few non-zero elements and they are arranged along the main diagonal. Consequently, fast band matrix routines can be used to speed up computations. For the general ARMA case, unfortunately, both the covariance and precision matrices are full, and we cannot use the previous results directly. However, by a careful manipulation of the ARMA error structure, we show in the following sections that one can work with only banded matrices and achieve the associated efficiency gains.

2.1.1 Likelihood Evaluation

We first investigate the likelihood function of our model and present a fast and simple way to evaluate it for both Bayesian and maximum likelihood estimation. Since the likelihood function is the joint distribution of the data, we seek so stack the system of equations in (1)–(3) over $t = 1, \dots, T$. To this end, note that (2) can be written as:

$$\mathbf{H}_\phi \boldsymbol{\varepsilon}^y = \mathbf{H}_\psi \mathbf{u}, \quad \mathbf{u} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_u), \tag{6}$$

where $\boldsymbol{\varepsilon}^y = (\varepsilon_1^y, \dots, \varepsilon_T^y)'$, $\mathbf{u} = (u_1, \dots, u_T)'$, $\boldsymbol{\Omega}_u = \text{diag}(\text{e}^{h_1}, \dots, \text{e}^{h_T})$, $\mathbf{H}_\phi$ is a $T \times T$ difference matrix and $\mathbf{H}_\psi$ is a $T \times T$ lower triangular matrix with ones along the main diagonal and ψ_j on the j-th lower diagonal, $j = 1, \dots, q$. For example, if we have an ARMA(2,2) error structure then $\mathbf{H}_\phi$ and $\mathbf{H}_\psi$ are defined to be:

$$\mathbf{H}_\phi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\phi_1 & 1 & 0 & 0 & \cdots & 0 \\ -\phi_2 & -\phi_1 & 1 & 0 & \cdots & 0 \\ 0 & -\phi_2 & -\phi_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -\phi_2 & -\phi_1 & 1 \end{pmatrix}, \quad \mathbf{H}_\psi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \psi_1 & 1 & 0 & 0 & \cdots & 0 \\ \psi_2 & \psi_1 & 1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \psi_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_2 & \psi_1 & 1 \end{pmatrix}.$$

Since $\mathbf{H}_\phi$ is a lower triangular matrix with ones along the main diagonal, $|\mathbf{H}_\phi| = 1$ for any $\boldsymbol{\phi} = (\phi_1, \dots, \phi_p)'$. Thus, $\mathbf{H}_\phi$ is invertible and (6) can be written as:

$$\boldsymbol{\varepsilon}^y = \mathbf{H}_\phi^{-1} \mathbf{H}_\psi \mathbf{u}. \tag{7}$$

Finally, stacking (1) over all dates and substituting (7) gives:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{H}_{\phi}^{-1}\mathbf{H}_{\psi}\mathbf{u}, \tag{8}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_T)'$. By a change of variable, it follows that:

$$(\mathbf{y} \,|\, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\mu}, \mathbf{h}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Omega}_y),$$

where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_p)'$, $\mathbf{h} = (h_1, \ldots, h_T)'$ and $\boldsymbol{\Omega}_y = \mathbf{H}_{\phi}^{-1}\,\mathbf{H}_{\psi}\,\boldsymbol{\Omega}_u\,(\mathbf{H}_{\phi}^{-1}\,\mathbf{H}_{\psi})'$. Thus, the log-likelihood function is given by:

$$\log p(\mathbf{y} \,|\, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\mu}, \mathbf{h}) = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} h_t - \frac{1}{2}(\mathbf{y} - \boldsymbol{\mu})'\boldsymbol{\Omega}_y^{-1}(\mathbf{y} - \boldsymbol{\mu}). \tag{9}$$

Evaluation of the log-likelihood function requires the computation of the $T \times T$ inverse of the matrix $\boldsymbol{\Omega}_y$. In general, this is an intensive procedure, requiring $\mathcal{O}(T^3)$ operations. For the case of only MA errors, Chan (2013) realizes that $\boldsymbol{\Omega}_y^{-1}$ is a band matrix, and computations involving band matrices are much faster. For example, the Cholesky decomposition of a $T \times T$ band matrix takes $\mathcal{O}(T)$ operations, which is substantially less than the $\mathcal{O}(T^3)$ operations required for the same operation on a full matrix of the same size.[2]

For our general ARMA case, however, both $\boldsymbol{\Omega}_y^{-1}$ and $\boldsymbol{\Omega}_y$ are full. The key to overcoming the computational hurdle is to show that the matrices $\mathbf{H}_{\phi}^{-1}$ and $\mathbf{H}_{\psi}$ commute, i.e., $\mathbf{H}_{\phi}^{-1}\mathbf{H}_{\psi} = \mathbf{H}_{\psi}\mathbf{H}_{\phi}^{-1}$. We prove this claim in Appendix A. Using this result, we show that we can evaluate the log-likelihood in (9) using only band matrix operations, thus preserving the computational gains. More specifically, rewrite (8) as:

$$\widetilde{\mathbf{y}} = \widetilde{\boldsymbol{\mu}} + \mathbf{H}_{\phi}^{-1}\mathbf{u}, \tag{10}$$

where $\widetilde{\mathbf{y}} = \mathbf{H}_{\psi}^{-1}\mathbf{y}$ and $\widetilde{\boldsymbol{\mu}} = \mathbf{H}_{\psi}^{-1}\boldsymbol{\mu}$. Thus, by a change of variable:

$$(\widetilde{\mathbf{y}} \,|\, \boldsymbol{\phi}, \boldsymbol{\psi}, \boldsymbol{\mu}, \mathbf{h}) \sim \mathcal{N}\left(\boldsymbol{\mu}, \mathbf{S}_{\widetilde{y}}\right),$$

[2]For a textbook treatment see Chapter 4 in Golub and Van Loan (2013).

where $\mathbf{S}_{\widetilde{y}}^{-1} = \mathbf{H}_{\phi}'\mathbf{\Omega}_u^{-1}\mathbf{H}_{\phi}$ and the log-likelihood for $\widetilde{\mathbf{y}}$ is:

$$\log p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\mu}}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}) \propto -\frac{1}{2}\sum_{t=1}^{T} h_t - \frac{1}{2}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})'\mathbf{S}_{\widetilde{y}}^{-1}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}}). \tag{11}$$

The fact that $\mathbf{S}_{\widetilde{y}}$ is invertible stems from noting that $|\mathbf{H}_{\phi}| = 1$ for any vector $\boldsymbol{\phi}$ and that $|\mathbf{\Omega}_u| = \mathrm{e}^{\sum_{t=1}^{T} h_t} > 0$ for any $h_1, \ldots, h_T$. Moreover, since $\mathbf{\Omega}_u$ is a diagonal matrix, its inverse is simply given by taking the reciprocal of the diagonal elements, i.e., $\mathbf{\Omega}_u^{-1} = \text{diag}\left(\mathrm{e}^{-h_1}, \ldots, \mathrm{e}^{-h_T}\right)$.

To summarize, we can employ a simple 3-step procedure to evaluate the log-likelihood in (9) using the equivalent representation in (11). First, we obtain the Cholesky decomposition $\mathbf{C}_{\widetilde{y}}$ of the band matrix $\mathbf{S}_{\widetilde{y}}$, which involves $\mathcal{O}(T)$ operations. Second, we implement forward substitution and backward substitution to get:

$$\mathbf{A} = \mathbf{C}_{\widetilde{y}}' \backslash (\mathbf{C}_{\widetilde{y}} \backslash (\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})),$$

which, by definition, is equivalent to $\mathbf{A} = \mathbf{C}_{\widetilde{y}}^{-1'}(\mathbf{C}_{\widetilde{y}}^{-1}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})) = \mathbf{S}_{\widetilde{y}}^{-1}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})$.[3] Finally, we compute:

$$\mathbf{B} = -\frac{1}{2}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})'\mathbf{A} = -\frac{1}{2}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})'\mathbf{S}_{\widetilde{y}}^{-1}(\widetilde{\mathbf{y}} - \widetilde{\boldsymbol{\mu}})$$

Thus, conditional on $\boldsymbol{\mu}, \boldsymbol{\phi}, \boldsymbol{\psi}$ and $\mathbf{h}$, the log-likelihood function (11) can be efficiently evaluated without implementing the Kalman filter.

2.1.2 Posterior Simulation

After discussing an efficient way to evaluate the likelihood function, we now outline an efficient posterior sampler for estimating the ARMA-SV model presented in (1)–(3). In our empirical application we will consider the forecast performance of various nested specifications where the conditional mean μ_t follows a non-stationary random walk process as in (5) (UC-ARMA-SV) or an stationary AR process (AR-ARMA-SV). In this section we restrict our focus to the UC-ARMA-SV model, noting that the estimation for all alternative models considered in this paper can be obtained in a similar manner.

The UC-ARMA-SV model is given by (1)–(3) and (5). As before, the main computational hurdle is that the ARMA structure implies a full error covariance matrix $\mathbf{\Sigma}_y$. Unlike the

[3]Note that since $\mathbf{H}_{\psi}$ is a lower triangular band matrix, $\widetilde{\mathbf{y}}$ can be obtained by forward substitution in $\mathcal{O}(T)$ operations. Similarly for $\widetilde{\boldsymbol{\mu}}$.

pure MA case analyzed in Chan (2013), the precision matrix $\boldsymbol{\Sigma}_y^{-1}$ in our case is full as well. Consequently, sampling both the trend $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_T)'$ and log-volatilities $\mathbf{h}$ are difficult. However, by a careful manipulation of the ARMA error structure, we show below how we can sample the latent states $\boldsymbol{\tau}$ and $\mathbf{h}$ using fast band matrix routines.

To complete the model specification, we initialize the transition equations for the trend and log-volatilities with $\tau_1 \sim \mathcal{N}(\tau_0, \sigma_{0\tau}^2)$ and $h_1 \sim \mathcal{N}(h_0, \sigma_{0h}^2)$, where $\tau_0, \sigma_{0\tau}^2, h_0$, and σ_{0h}^2 are known constants. The priors for $\boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2$ and σ_h^2 are assumed to be independent. In particular, we set:

$$\sigma_\tau^2 \sim \mathcal{IG}(\nu_\tau, S_\tau), \quad \sigma_h^2 \sim \mathcal{IG}(\nu_h, S_h),$$

$$\boldsymbol{\phi} \sim \mathcal{N}(\boldsymbol{\phi}_0, \mathbf{V}_\phi)\mathbb{1}(\boldsymbol{\phi} \in \mathbf{A}_\phi), \quad \boldsymbol{\psi} \sim \mathcal{N}(\boldsymbol{\psi}_0, \mathbf{V}_\psi)\mathbb{1}(\boldsymbol{\psi} \in \mathbf{A}_\psi),$$

where $\mathcal{IG}$ denotes the inverse-gamma distribution, $\mathbb{1}(\cdot)$ is the indicator function that takes the value of one if the argument is true and the value of 0 otherwise, $\mathbf{A}_\phi$ and $\mathbf{A}_\psi$ are the stationary and invertible regions, respectively. Then, posterior draws can be obtained by sequentially sampling from:

1. $p(\boldsymbol{\tau} \,|\, \mathbf{y}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2)$;
2. $p(\mathbf{h} \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2)$;
3. $p(\sigma_\tau^2, \sigma_h^2 \,|\, \mathbf{y}, \mathbf{h}, \boldsymbol{\tau}, \boldsymbol{\phi}, \boldsymbol{\psi}) = p(\sigma_h^2 \,|\, \mathbf{h})p(\sigma_\tau^2 \,|\, \boldsymbol{\tau})$;
4. $p(\boldsymbol{\psi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \boldsymbol{\phi}, \sigma_h^2, \sigma_\tau^2)$;
5. $p(\boldsymbol{\phi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \boldsymbol{\psi}, \sigma_h^2, \sigma_\tau^2)$.

In what follows we discuss an efficient way to implement Step 1. The remaining steps are discussed in detail in Appendix B. As mentioned earlier, simulating directly from $p(\boldsymbol{\tau} \,|\, \mathbf{y}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2)$ is cumbersome due to the ARMA error structure. It will prove more efficient to work with $\widetilde{\boldsymbol{\tau}} = \mathbf{H}_\psi^{-1}\boldsymbol{\tau}$; once a draw for $\widetilde{\boldsymbol{\tau}}$ is obtained, we can simply compute the implied $\boldsymbol{\tau} = \mathbf{H}_\psi\widetilde{\boldsymbol{\tau}}$. Now, we derive the log posterior density of $\widetilde{\boldsymbol{\tau}}$:

$$\log p(\widetilde{\boldsymbol{\tau}} \,|\, \widetilde{\mathbf{y}}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2) \propto \log p(\widetilde{\boldsymbol{\tau}} \,|\, \boldsymbol{\psi}, \sigma_\tau^2) + \log p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\tau}}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}), \tag{12}$$

where $p(\widetilde{\mathbf{y}} \,|\, \widetilde{\boldsymbol{\tau}}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi})$ is the likelihood for the transformed data $\widetilde{\mathbf{y}} = \mathbf{H}_\psi^{-1}\mathbf{y}$ obtained by setting $\boldsymbol{\mu} = \boldsymbol{\tau}$ in (10). Next we derive an explicit expression for $p(\widetilde{\boldsymbol{\tau}} \,|\, \boldsymbol{\psi}, \sigma_\tau^2)$, the prior density of $\widetilde{\boldsymbol{\tau}}$.

This can be done by first noting that (5) can be written as:

$$\mathbf{H}\boldsymbol{\tau} = \boldsymbol{\varepsilon}_{\boldsymbol{\tau}}, \quad \boldsymbol{\varepsilon}_{\boldsymbol{\tau}} \sim \mathcal{N}(0, \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_\tau}), \tag{13}$$

where $\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}_\tau} = \text{diag}(\sigma_{0\tau}^2, \sigma_\tau^2, \ldots, \sigma_\tau^2)$ and $\mathbf{H}$ is a first-difference matrix of size T. Since $|\mathbf{H}| = 1$, $\mathbf{H}$ is invertible and we can write (13)as:

$$\boldsymbol{\tau} = \mathbf{H}^{-1}\boldsymbol{\varepsilon}_\tau. \tag{14}$$

Thus $(\boldsymbol{\tau} \,|\, \sigma_\tau^2) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_\tau)$, where $\boldsymbol{\Omega}_\tau^{-1} = \mathbf{H}'\boldsymbol{\Omega}_{\varepsilon_\tau}^{-1}\mathbf{H}$. Pre-multiplying both sides of (14) by $\mathbf{H}_\psi^{-1}$ gives

$$\widetilde{\boldsymbol{\tau}} = \mathbf{H}_\psi^{-1}\mathbf{H}^{-1}\boldsymbol{\varepsilon}_\tau, \tag{15}$$

so that $(\widetilde{\boldsymbol{\tau}} \,|\, \boldsymbol{\psi}, \sigma_\tau^2) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_{\widetilde{\tau}})$ with $\boldsymbol{\Omega}_{\widetilde{\tau}}^{-1} = \mathbf{H}_\psi'\boldsymbol{\Omega}_\tau^{-1}\mathbf{H}_\psi$. The log prior density for $\widetilde{\boldsymbol{\tau}}$ is therefore given by:

$$\log p(\widetilde{\boldsymbol{\tau}} \,|\, \boldsymbol{\psi}, \sigma_\tau^2) \propto -\frac{T-1}{2}\log\sigma_\tau^2 - \frac{1}{2}\widetilde{\boldsymbol{\tau}}'\mathbf{H}_\psi'\boldsymbol{\Omega}_\tau^{-1}\mathbf{H}_\psi\widetilde{\boldsymbol{\tau}}, \tag{16}$$

Finally, substituting (11) and (16) into (12), gives:

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

where $\mathbf{K}_{\widetilde{\tau}} = \mathbf{H}_\psi'\boldsymbol{\Omega}_\tau^{-1}\mathbf{H}_\psi + \mathbf{H}_\phi'\boldsymbol{\Omega}_u^{-1}\mathbf{H}_\phi$ and $\widehat{\boldsymbol{\tau}} = \mathbf{K}_{\widetilde{\boldsymbol{\tau}}}^{-1}\mathbf{H}_\phi'\boldsymbol{\Omega}_u^{-1}\mathbf{H}_\phi\widetilde{\mathbf{y}}$. The conditional posterior distribution is therefore Gaussian:

$$(\widetilde{\boldsymbol{\tau}} \,|\, \widetilde{\mathbf{y}}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2) \sim \mathcal{N}(\widehat{\boldsymbol{\tau}}, \mathbf{K}_{\widetilde{\tau}}^{-1}).$$

Since $\mathbf{H}_\psi, \mathbf{H}_\phi$ and $\mathbf{H}$ are all band matrices, so is the precision matrix $\mathbf{K}_{\widetilde{\tau}}$. Sampling from this Gaussian distribution can be efficiently conducted via the precision-based algorithm (see, e.g., Chan and Jeliazkov, 2009). In particular, since $\mathbf{K}_{\widetilde{\tau}}$ is a band matrix, its Cholesky decomposition $\mathbf{C}_{\widetilde{\tau}}$ can be quickly obtained. Then, forward and backward substitutions give:

$$\widehat{\boldsymbol{\tau}} = \mathbf{C}_{\widetilde{\tau}}' \backslash (\mathbf{C}_{\widetilde{\tau}} \backslash (\mathbf{H}_\phi'\boldsymbol{\Omega}_u^{-1}\mathbf{H}_\phi\widetilde{\mathbf{y}})).$$

A draw of $\widetilde{\boldsymbol{\tau}} \sim \mathcal{N}(\widehat{\boldsymbol{\tau}}, \mathbf{K}_{\widetilde{\tau}}^{-1})$ can be obtained by:

$$\widetilde{\boldsymbol{\tau}} = \widehat{\boldsymbol{\tau}} + \mathbf{C}_{\widetilde{\tau}}' \backslash \mathbf{Z},$$

where $\mathbf{Z}$ is a $T \times 1$ vector of standard normal random variables, i.e., $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_T)$. Finally, we return a draw of $\boldsymbol{\tau}$ by the transformation $\boldsymbol{\tau} = \mathbf{H}_\psi \widetilde{\boldsymbol{\tau}}$.

3 Application to Inflation Forecasting

In this section we assess the proposed ARMA-SV models to forecast two commonly used inflation measures, namely, CPI and the GDP deflator, in each of the G7 countries. In the inflation forecasting exercise, we compute both point and density forecasts from the proposed models and compare them to a few popular benchmarks.

3.1 Data and Preliminary Analysis

For each of the inflation index, z_t, we compute the annualized inflation rate as $y_t = 400 \log(z_t/z_{t-1})$. The resulting series are plotted in Figures 1 and 2. All our time series end in 2016Q4. But due to data availability, they have different starting dates. In particular, while most CPI series start in 1960Q2, Canada's is only available from 1961Q1 and Germany's from 1970Q2.

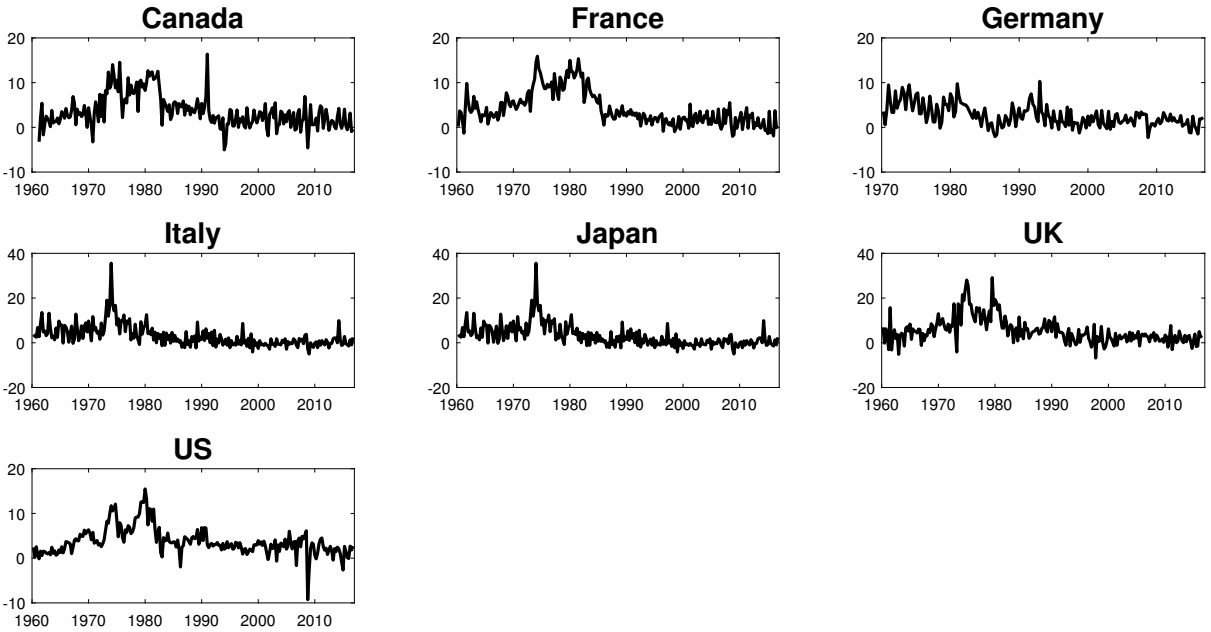

Figure 1: CPI inflation across G7 countries.

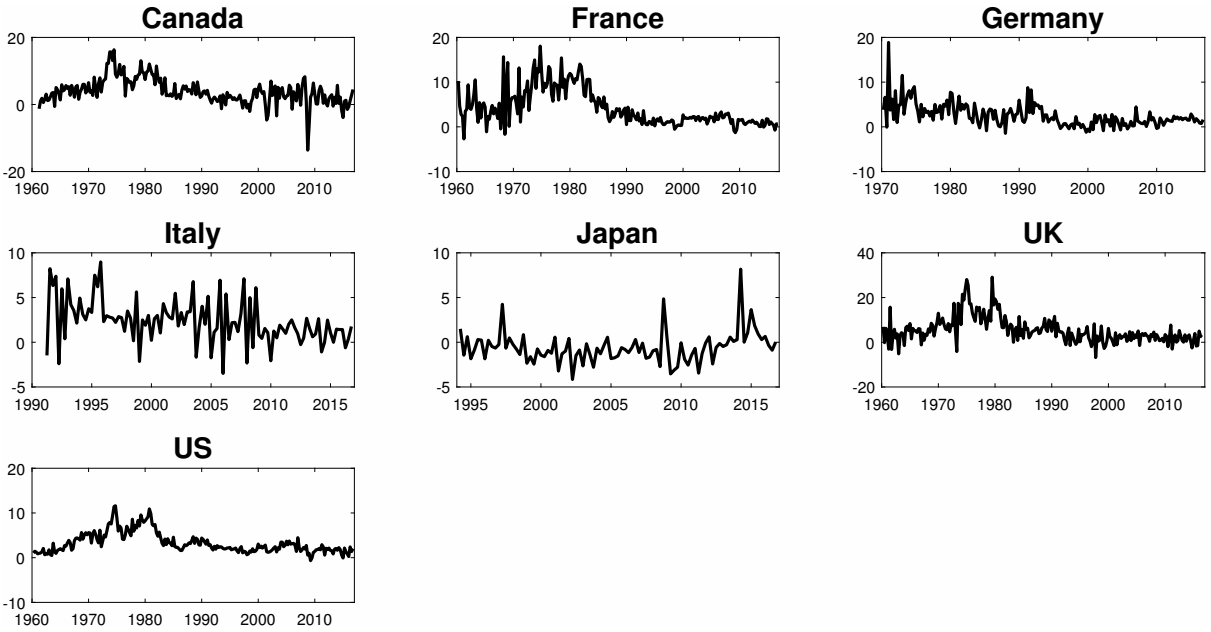

Figure 2: GDP Deflator inflation across G7 countries.

Before reporting the forecasting results, we first perform a preliminary analysis to understand when a specification is expected to work better for an inflation series. In particular, we focus on the expected differences between the proposed ARMA-SV models and the MA-SV models in Chan (2013). Since the SV component is common in both classes

of models, expected differences in forecast performance boil down to cases in which the ARMA error structure is preferable to the simpler case with MA errors only.

One way to gain insights into these differences is by comparing the conditional autocovariance under the two classes of models. To that end, let $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_T)'$ and note that the MA(q)-SV is a simplified version of the ARMA(p,q)-SV model in equations (1)-(3) in which $\phi_i = 0$, $i = 1, \ldots, p$ in equation (2). The conditional autocovariance of the MA(q)-SV model is given by

$$\gamma(j) = \begin{cases} \sum_{k=0}^{q} \psi_k^2 \mathrm{e}^{h_{t-k}}, & \text{for } j = 0, \\ \sum_{k=0}^{q-j} \psi_{k+j}\psi_k \mathrm{e}^{h_{t-k}}, & \text{for } j = 1, \ldots, q, \\ 0, & \text{for } j > q, \end{cases} \tag{17}$$

where $\psi_0 = 1$. The conditional autocovariance for the ARMA(p,q)-SV model is more complicated, and we compute them recursively using

$$\gamma(j) - \sum_{i=1}^{p} \gamma(j-i)\phi_i = \begin{cases} \sum_{k=0}^{q} \psi_k\theta_k \mathrm{e}^{h_{t-k}}, & \text{for } j = 0, \\ \sum_{k=0}^{q-j} \psi_{k+j}\theta_k \mathrm{e}^{h_{t-k}}, & \text{for } j \in [1, \max(p, q+1)), \\ 0, & \text{for } j \geqslant \max(p, q+1), \end{cases} \tag{18}$$

where $\psi_0 = 1$ and θ_k denotes the k-th element of the lag polynomial $\theta(L) = \frac{\psi(L)}{\phi(L)}$. Notice that if we have no AR terms, i.e. $\phi_i = 0$, $i = 1, \ldots, p$, then $\theta_k = \psi_k$ and the autocovariance function in equation (18) is equivalent to that in equation (17). This equivalence is in spite of the fact that both models possess stochastic volatility. Thus any differences in covariance structure is entirely due to the presence of AR terms.

The above comparison suggests that we can gain some practical insights as to whether models with the ARMA-SV or MA-SV errors may provide relatively superior inflation forecasts in some countries, but not others, by investigating the autocorrelation function (ACF) of each series. Specifically, equation (18) suggests that if the ACF exhibits a high degree of persistence then we would expect models with ARMA errors to provide relatively superior forecasts due to the presence of AR terms. Conversely, if the ACF goes to zero relatively quickly, then we would expect MA type models to provide superior forecasts.

In light of these insights, Figure 3 presents the sample autocorrelations for CPI inflation in each of the G7 nations. It is immediately obvious that CPI inflation in Canada, France,

Italy and the US exhibit very strong persistence. This suggests that models with ARMA errors may be particularly useful in forecasting inflation in these countries. The ACFs for inflation in Japan and the UK also exhibit some persistence, albeit to a lesser extent than the aforementioned countries. We would therefore not be surprised if a model with an ARMA error structure helped improve the forecast performance in these countries too, however our expectations are certainly not as strong. Finally, inflation in Germany is the least persistent of all G7 economies. For this reason, we would not be surprised if the simpler MA error model outperforms those with ARMA errors.

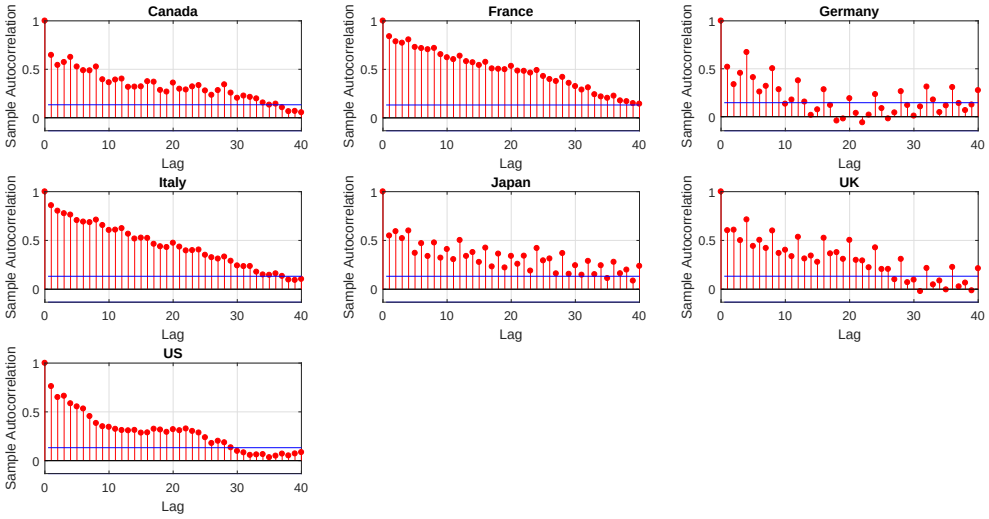

Figure 3: Sample ACF for CPI inflation across G7 countries.

3.2 Competing Models

The primary benchmark model is taken to be a stationary AR(m) model—as in (4)—with homoscedatic variance. The reason for using this benchmark is two-fold. First, it is still a competitive model among both univariate and multivariate models. Second, given the parsimonious structure of the model, any finding that it provides competitive forecasts has practical significance.

In addition to the homoscedastic AR(m) model, we also consider a more general class of heteroscedastic autoregressive (AR) and unobserved components (UC) models. These include specifications with and without SV, MA-SV and ARMA-SV errors. In each country, the lag length is selected via the Bayesian information criteria (BIC). The results from these tests are summarized in Table 1.

Table 1: Lag length selection for AR(m) models using BIC.

Country	CPI Inflation	GDP Deflator Inflation
Canada	4	5
France	5	6
Germany	4	4
Italy	4	1
Japan	8	1
UK	5	3
US	3	2

In each case we limit the ARMA-SV component to the observation equation. While the extension to ARMA-SV errors in the state equation is conceptually straight forward to implement, this specification allows for a more direct comparison with the broader literature (e.g., Stock and Watson (2007), Chan (2013) and Chan et al. (2013)). In total, this constitutes 10 models, each of which is summarized in Table 2.

Table 2: A list of competing models.

Model	Description
AR(m)	Autoregressive model with homoscedastic errors
AR(m)-SV	Autoregressive model with SV errors
AR(m)-MA-SV	Autoregressive model with MA-SV errors
AR(m)-ARMA-SV	Autoregressive model with ARMA-SV errors
AR(m)-ARMA	Autoregressive model with ARMA errors
UC	Unobserved components model with homoscedastic errors
UC-SV	Unobserved components model with SV errors
UC-MA-SV	Unobserved components model with MA-SV errors
UC-ARMA-SV	Unobserved components model with ARMA-SV errors
UC-ARMA	Unobserved components model with ARMA errors

3.3 Priors and Initial Conditions

In each of the UC models we set the initial value of τ_t as $\tau_1 \sim \mathcal{N}(\tau_0, \sigma_{0\tau}^2)$, where $\tau_0, h_0, \sigma_{0\tau}^2$. In particular, we set $\tau_0 = 0$ and $\sigma_{0\tau}^2 = 5$. Similarly, we initialize UC models with SV with $h_1 \sim \mathcal{N}(h_0, \sigma_{0h}^2)$, where $h_0 = 0$ and $\sigma_{0h}^2 = 5$. Moreover, we set $\nu_\tau = \nu_h = 10$, $S_\tau = 0.18$ and $S_h = 0.45$. These values imply prior means $\mathbb{E}\sigma_\tau^2 = 0.02$ and $\mathbb{E}\sigma_h^2 = 0.05$.

For the AR models, we set an independent truncated prior for the conditional mean coefficients. In particular, the prior mean is the zero vector and the variance is $\mathbf{V}_{\boldsymbol{\rho}} = 5 \times \mathbf{I}_n$, where $\mathbf{I}_n$ denotes the identity matrix of size n. The posterior distribution of

this model is then obtained by following the procedure in Section 2.1.2, where Step 1 is replaced with draws from $p(\boldsymbol{\rho} \,|\, \mathbf{y}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2)$, with $\boldsymbol{\rho} = (\rho_1, \cdots, \rho_m)'$. To this end, note that by following similar steps in Section 2.1.2, the conditional posterior distribution is:

$$(\boldsymbol{\rho} \,|\, \mathbf{y}, \mathbf{h}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2) \sim \mathcal{N}(\widehat{\boldsymbol{\rho}}, \mathbf{D}_\rho) \mathbb{1}(\boldsymbol{\rho} \in \mathbf{A}_\rho), \tag{19}$$

where $\mathbf{D}_\rho = (\mathbf{V}_{\boldsymbol{\rho}} + \mathbf{H}_\phi' \boldsymbol{\Omega}_u^{-1} \mathbf{H}_\phi)^{-1}$, $\widehat{\boldsymbol{\rho}} = \mathbf{D}_{\boldsymbol{\rho}} \mathbf{H}_\phi' \boldsymbol{\Omega}_u^{-1} \mathbf{H}_\phi \widetilde{\mathbf{y}}$, and the truncated region ensures all roots lay outside of the unit circle. Samples from this truncated density are then obtained using the acceptance-rejection method discussed in Section 2.1.2 using the proposal density $\mathcal{N}(\widehat{\boldsymbol{\rho}}, \mathbf{D}_\rho)$.

Finally, following Chan (2013) we set the moving average order in the MA-SV model variants to be one. For consistency, we also set each of the specifications with ARMA(p, q) errors to be ARMA(1,1).

3.4 Forecasting Setup

We conduct a pseudo out-of-sample forecasting exercise in which we consider both point and density forecasts. In each exercise, we divide the data into three sub-samples. The first part contains the first m observations used to initialize the AR(m) models. This guarantees that all AR and UC model variants have the same initial observation. The second part is the *estimation period*, which consists of the next 40 observations. The third part is the *hold-out period*, which contains the remaining observations that are used to assess the forecast performance of the model.

To see how the forecasts are conducted, let $\mathbf{y}_{1:t}$ denote the data from the *estimation period* and $\widehat{\mathbf{y}}_{t+k}$ represent the vector of k-steps-ahead forecasts with k=1,4,8,12 and 16. Density forecasts are obtained by the predictive density: $f(\mathbf{y}_{t+k}|\mathbf{y}_{1:t})$, and point forecasts are taken to be the mean of this density: $\widehat{\mathbf{y}}_{t+k} = \mathbb{E}[\mathbf{y}_{t+k}|\mathbf{y}_{1:t}]$. These forecasts are conducted with predictive simulation. For concreteness, suppose we want to produce a 4-step ahead forecast of US CPI inflation from 1975Q1 to 1976Q1. Then, given the MCMC draws up to 1975Q1 along with the relevant transition equations, we simulate the future states up to time 1975Q4. The conditional expectation of this equation is then taken to be the point forecast and the observed data is used to evaluate implied density to produce a density forecast. The exercise is then repeated using data from 1975Q2 up to the end of the hold-out period, i.e. 2012Q4.

In each period, the parameter estimates are based on 45,000 draws from the posterior simulator discussed in 2.1, after discarding the first 5,000 draws as a burn-in period.

3.4.1 Forecast Metrics

The accuracy of the point and density forecasts are respectively assessed with the mean square forecast error (MSFE) and log-predictive-likelihood (LPL).

To compute the MSFE, we first evaluate the forecast $\widehat{y}_{T_0+t+k-1}$ by averaging all the posterior means $\mathbb{E}(y_{T0+t+k-1} \,|\, \mathbf{y}_{1:T_0+t})$ when it is time $T_0 + t$. Given that the forecasting error is $\mathbf{e}^2_{T_0+t+k-1} = \mathbf{y}^0_{T_0+t+k-1} - \mathbb{E}(y_{T0+t+k-1} \,|\, \mathbf{y}_{1:T_0+t})$, where k denotes a k-step-ahead forecast, the MSFE is defined as:

$$\text{MSFE} = \frac{1}{T - T_0 - k + 1} \sum_{t=1}^{T-T_0-k+1} \mathbf{e}^2_{T_0+t+k-1}$$

Since a smaller forecast error corresponds to a smaller MSFE, a relatively smaller MSFE indicates better forecast performance. As mentioned in the previous section, we use a stationary AR(m) model as the benchmark. To facilitate our discussion when presenting the results we therefore standardize the MSFE of each model to the MSFE of the AR(m) model. Hence the relative MSFE (RMSFE) for the AR(m) model is 1.00. Moreover, if a given model produces a RMSFE less than one then this indicates superior forecast performance to the AR(m), and vice-versa.

To compute the predictive likelihood we simply sum the evaluated likelihood function implied by the forecasts. That is:

$$\text{LPL} = \sum_{t=1}^{T-T_0-k+1} \log p(\widehat{y}_{T_0+t+k-1} = y_{T_0+t+k-1} \,|\, \mathbf{y}_{1:T_0+t})$$

Since a larger likelihood implies better fit, for interpretation purposes, a larger predictive likelihood value implies a better density forecast performance. To facilitate our discussion we also standardize the LPL to the AR(m) benchmark. Since it is in logs, the relative LPL (RLPL) for the AR(m) model is 0.00. Moreover, if a given model produces a positive LPL then this indicates superior forecast performance to the AR(m), and vice-versa.

3.5 Empirical Results

To facilitate the discussion, we separately present the point and density forecast results.

3.5.1 Point Forecasts

The point forecast results for CPI are presented in Tables 3–9. While there is no single best model across all countries, there are some general results. For instance, in line with previous studies (e.g., Stock and Watson, 2007; Chan, 2013; Clark and Ravazzolo, 2015), the AR-SV and UC-SV models tend to dominate their constant volatility counterparts in CPI forecasts. Interestingly, this results does not extend to the GDP Deflator. In that case, the simple AR model often provides better forecasts than its SV counterpart. One exception is the US, in which the SV variants dominate their homoscedastic counterparts.

Table 3: MSFEs relative to AR benchmark: Canada CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	0.99	0.97**	0.95**	0.93**	0.90*
AR-ARMA	0.99	0.96	0.85	0.77	0.70*
AR-MA-SV	0.99	0.97**	0.95**	0.93**	0.89*
AR-ARMA-SV	**0.97**	**0.95**	**0.82**	**0.73**	**0.67***
UC	1.01	0.99	0.95**	0.91**	0.88*
UC-SV	1.05*	1.03	0.96	0.89	0.88
UC-ARMA	1.02	0.96	0.88*	0.80*	0.74*
UC-MA-SV	1.01	1.04	0.96	0.88	0.87*
UC-ARMA-SV	1.02	1.03	0.90	0.80*	0.74*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 4: MSFEs relative to AR benchmark: France CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	0.95**	0.91**	0.86**	0.83**	0.83**
AR-ARMA	0.82**	0.87	0.89**	0.93*	0.95*
AR-MA-SV	0.83**	0.89**	0.85**	0.82**	0.83*
AR-ARMA-SV	**0.81****	0.88	0.90**	0.94*	0.95
UC	0.84**	0.82**	0.79*	0.81	0.82
UC-SV	0.91**	0.93	0.88	0.87	0.85
UC-ARMA	0.85**	0.77**	0.66**	0.63**	0.59*
UC-MA-SV	0.93**	1.02	0.91	0.87	0.84
UC-ARMA-SV	0.84**	**0.73****	**0.57****	**0.52****	**0.46***

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 5: MSFEs relative to AR benchmark: Germany CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	**1.00**	1.00	1.00	1.00	1.00
AR-SV	1.01	**0.98**	**0.96**	0.92	0.88
AR-ARMA	1.01	1.01	1.00	0.97	0.88
AR-MA-SV	1.01	0.99	**0.96**	0.92	0.88
AR-ARMA-SV	1.01	1.01	0.98	0.95	0.89
UC	1.08*	1.10*	1.01	0.95	0.89
UC-SV	1.13**	1.14*	1.00	0.93	0.85
UC-ARMA	1.14**	1.16*	1.00	0.92	0.83
UC-MA-SV	1.14**	1.15*	0.99	**0.91**	**0.83**
UC-ARMA-SV	1.14**	1.15*	0.99	**0.91**	**0.83**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 6: MSFEs relative to AR benchmark: Italy CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	**0.89**	0.67**	0.63**	0.60**	0.57**
AR-ARMA	0.93	0.65**	0.59*	0.55**	0.49
AR-MA-SV	**0.89**	0.67**	0.63**	0.61**	0.58**
AR-ARMA-SV	**0.89**	0.63	0.56	0.50**	0.43**
UC	1.00	0.77**	0.79	0.83	0.86
UC-SV	1.11*	0.89	0.87	0.86	0.84
UC-ARMA	1.04	0.67**	0.70*	0.75	0.81
UC-MA-SV	1.08**	0.95	0.86	0.81	0.75
UC-ARMA-SV	0.99	**0.60****	**0.51****	**0.47****	**0.41****

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 7: MSFEs relative to AR benchmark: Japan CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	1.04	0.83**	**0.72****	0.76**	**0.73***
AR-ARMA	**0.70****	0.79*	0.94	0.95	0.99
AR-MA-SV	1.06**	0.82**	**0.72****	0.75**	**0.73***
AR-ARMA-SV	**0.70****	0.79	0.94	0.95	0.99
UC	0.78**	**0.76****	0.78**	**0.74****	0.75*
UC-SV	0.80**	0.81**	0.83**	0.77**	0.76*
UC-ARMA	0.83**	0.79**	0.87**	0.83**	0.84*
UC-MA-SV	0.80**	0.81**	0.82**	0.77**	0.75*
UC-ARMA-SV	0.73**	0.80**	0.84**	0.76**	0.74*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 8: MSFEs relative to AR benchmark: UK CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	0.97	0.95**	0.77**	0.77**	0.76**
AR-ARMA	**0.75**	0.65	0.85	0.89	0.94
AR-MA-SV	0.86**	0.91**	0.78**	0.77**	0.76**
AR-ARMA-SV	**0.75***	0.65**	0.85**	0.89**	0.94**
UC	0.96	0.74**	0.80	0.79	0.81
UC-SV	0.99	0.81**	0.81**	0.75**	0.73*
UC-ARMA	0.99	0.68**	0.70**	**0.64****	**0.63****
UC-MA-SV	1.02	0.79**	0.78**	0.72**	0.69**
UC-ARMA-SV	0.98	**0.63****	**0.68****	**0.64****	0.64**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 9: MSFEs relative to AR benchmark: US CPI inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	0.99	0.97	0.86*	0.84	0.82
AR-ARMA	1.03	1.01	0.88	0.83	0.75
AR-MA-SV	0.99	0.97	0.86*	0.84	0.80
AR-ARMA-SV	1.00	0.97	**0.82**	**0.73**	**0.63**
UC	0.99	0.97	0.88**	0.91	0.91*
UC-SV	0.96	**0.93**	0.92	0.97	0.88*
UC-ARMA	1.00	0.97	0.83*	0.79	0.70
UC-MA-SV	**0.94**	0.94	0.90	0.92	0.81*
UC-ARMA-SV	1.00	1.07	0.86	0.79	0.70

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 10: MSFEs relative to AR benchmark: Canada GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	1.01	0.97**	0.93**	0.92**	0.91*
AR-ARMA	0.95**	0.96**	0.93	0.83	0.78
AR-MA-SV	0.96	0.97**	0.93**	0.92**	0.91
AR-ARMA-SV	0.96**	0.95**	**0.90**	**0.78**	**0.73**
UC	0.94	0.98	0.96	0.91	0.89
UC-SV	0.95	**0.94**	0.93	0.89	0.86
UC-ARMA	0.92	0.97	0.91	0.82**	0.78**
UC-MA-SV	**0.89****	0.95	0.93	0.89	0.86
UC-ARMA-SV	0.92	1.15	0.98	0.84**	0.81*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 11: MSFEs relative to AR benchmark: France GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	1.00	0.91**	0.89	0.88	0.87
AR-ARMA	0.71**	0.94*	0.96	0.98	0.99
AR-MA-SV	1.18**	0.91**	0.88	0.88	0.87
AR-ARMA-SV	**0.66****	0.82*	0.80	0.81	0.83
UC	1.18*	0.98	0.92	0.89	0.87
UC-SV	1.18**	1.02	0.95	0.90	0.88
UC-ARMA	0.70**	0.78**	0.71**	0.67*	0.64
UC-MA-SV	0.70	1.04	0.96	0.91	0.89
UC-ARMA-SV	0.67**	**0.77**	**0.66***	**0.61***	**0.55**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 12: MSFEs relative to AR benchmark: Germany GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	1.00	**0.98**	**0.95***	**0.95***	0.94
AR-ARMA	0.98	1.06**	1.18**	1.30**	1.42*
AR-MA-SV	0.98	0.99	0.96	**0.95***	0.95
AR-ARMA-SV	**0.96**	1.01	1.06*	1.13	1.15
UC	**0.96**	1.02	0.97	0.96	0.95
UC-SV	**0.96**	1.02	0.97	0.96	0.95
UC-ARMA	1.01	1.04	0.99	0.98	0.96
UC-MA-SV	1.01	1.01	**0.95**	**0.95**	0.94
UC-ARMA-SV	1.00	1.01	**0.95**	**0.95**	**0.93**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 13: MSFEs relative to AR benchmark: Italy GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	1.00**	1.01**	1.01*	1.01	1.01
AR-ARMA	1.13**	0.92*	0.97	1.00	1.00
AR-MA-SV	1.13	1.01*	1.01*	1.01	1.01
AR-ARMA-SV	1.23**	0.98	1.00	1.01	1.01
UC	1.23	0.75**	0.76*	**0.75**	**0.74**
UC-SV	1.23*	0.78**	0.79*	0.81	0.81
UC-ARMA	**0.96****	0.75**	0.76*	0.76	0.75
UC-MA-SV	1.23**	0.76**	0.77	0.77	0.78
UC-ARMA-SV	0.97**	**0.73****	**0.75***	0.76	0.77

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 14: MSFEs relative to AR benchmark: Japan GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	**1.00**
AR-SV	1.00**	1.01	1.00	1.00	1.00
AR-ARMA	**0.93**	**0.92**	**0.93**	**0.93**	1.01
AR-MA-SV	**0.93****	1.01	1.00	1.00	1.00
AR-ARMA-SV	1.04*	1.00*	1.00	1.00	1.00
UC	1.00	0.98	1.03	1.06	1.06
UC-SV	1.00	0.98	1.02	1.04	1.04
UC-ARMA	1.00	0.99	1.04	1.08	1.08
UC-MA-SV	1.00	0.98	1.02	1.05	1.04
UC-ARMA-SV	1.00	0.98	1.02	1.05	1.04

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 15: MSFEs relative to AR benchmark: UK GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	**1.00**	1.00	1.00	1.00	1.00
AR-SV	1.00**	0.96**	0.91**	0.87**	0.83**
AR-ARMA	1.04	1.03	1.00	0.97	0.93
AR-MA-SV	1.04**	0.95**	0.90**	0.86**	0.82**
AR-ARMA-SV	1.02*	1.03	1.01	0.98	0.94
UC	1.00	0.92	0.89	0.91	0.87
UC-SV	1.00	0.98	0.89*	0.84**	0.78*
UC-ARMA	1.02	**0.90****	**0.81****	**0.75***	**0.68***
UC-MA-SV	1.02	0.98	0.88**	0.83**	0.77**
UC-ARMA-SV	1.00**	1.09	0.95	0.83*	0.72*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 16: MSFEs relative to AR benchmark: US GDP Deflator inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	1.00	1.00	1.00	1.00	1.00
AR-SV	0.97**	0.87**	0.81**	0.80**	0.78**
AR-ARMA	0.97**	0.86*	0.73*	0.64*	0.60*
AR-MA-SV	0.96**	0.87**	0.82**	0.82**	0.82*
AR-ARMA-SV	0.95*	**0.84**	**0.67***	**0.54***	**0.49***
UC	**0.93**	0.89	0.87	0.92	0.93
UC-SV	0.99**	0.98	0.98	0.98	0.96
UC-ARMA	0.96*	0.88**	0.83**	0.80**	0.77**
UC-MA-SV	1.05**	1.07	1.02	0.97	0.94
UC-ARMA-SV	0.96	0.88	0.82**	0.77**	0.75*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Since our major innovation is the estimation of ARMA-SV error models, it is useful to compare the relative forecast performance of these models to those made by their nested variants: ARMA, SV, and MA-SV models. In the case of CPI forecasts, models with ARMA-SV errors provide the best point forecasts across all horizons in both Canada, France and Italy. They also do quite well in the UK, and at longer horizons in the US (i.e. $k = 8, 12, 16$). Finally, we highlight the fact that Japan's CPI inflation is quite difficult to forecast.

A slightly weaker result exists when forecasting GDP Deflator inflation. In that case, ARMA-SV models still provide the best forecasts in France, and are only bested by the simple UC model at the one-step-ahead horizon in the US. They also provide good medium term forecasts in Italy and longer term forecasts in both Canada and Germany. Finally, despite having no clearly dominant model in the CPI forecasts, the AR-ARMA model does quite well in forecasting Japan's GDP Deflator inflation.

Taken together, our main conclusion from this exercise is that ARMA-SV error models provide good point forecasts of CPI and GDP deflator inflation measures. In particular, with few exceptions they are able to improve on the point forecast performance gained by the simpler SV model in Stock and Watson (2007) and the MA-SV model in Chan (2013).

3.5.2 Density Forecasts

In this subsection we report the results corresponding to the density forecasts. In particular, the forecast performance of the models are presented in Tables 17–30.

Table 17: Sum of log predictive likelihoods relative to AR benchmark: Canada CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	5.5**	9.6**	5.7**	9.6	3.4*
AR-ARMA	**8.6****	**17.1***	**30.7***	**46.0****	**49.3****
AR-MA-SV	5.4**	9.5**	5.9**	9.6	3.6*
AR-ARMA-SV	4.4**	7.2*	18.2*	29.7	30.2*
UC	-2.5**	1.7**	5.9*	8.6*	9.9**
UC-SV	-1.5**	7.4**	9.4**	15.5*	8.4*
UC-ARMA	-3.3**	4.3**	13.8*	21.7*	26.4*
UC-MA-SV	0.6**	5.1**	9.6**	17.0*	10.2**
UC-ARMA-SV	-0.8**	-11.9**	-4.4*	6.5	4.9**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 18: Sum of log predictive likelihoods relative to AR benchmark: France CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	5.0**	15.9**	25.0**	27.5**	31.2*
AR-ARMA	15.2**	24.1**	9.5**	-13.8**	-33.6*
AR-MA-SV	22.9**	15.5**	23.4**	25.9**	29.6*
AR-ARMA-SV	**24.0****	17.8**	7.5**	-6.1**	-13.5*
UC	18.4**	22.0**	30.4**	28.3**	26.5*
UC-SV	17.3**	27.9**	38.6**	37.6**	40.1**
UC-ARMA	18.6**	21.3**	34.7*	40.9**	47.7*
UC-MA-SV	14.1**	21.6**	36.0**	39.8**	43.7**
UC-ARMA-SV	21.1**	**30.1****	**55.2****	**65.4****	**77.6****

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 19: Sum of log predictive likelihoods relative to AR benchmark: Germany CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	2.0*	-2.4**	-1.4**	4.8**	11.4**
AR-ARMA	**3.8****	**0.3****	-1.2**	-0.8**	0.5**
AR-MA-SV	1.6**	-3.1**	-2.8**	3.3**	9.7**
AR-ARMA-SV	-0.1*	-11.4**	-12.1**	-8.9**	-1.7*
UC	-12.0**	-13.2**	-2.4**	2.3*	6.6
UC-SV	-11.7**	-14.0**	**0.4****	**12.3****	**24.1****
UC-ARMA	-17.1**	-18.4**	-3.1*	5.0	12.9
UC-MA-SV	-12.9**	-17.3**	-1.9**	9.2**	22.3*
UC-ARMA-SV	-12.8**	-16.8**	-2.1**	9.2**	22.0

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 20: Sum of log predictive likelihoods relative to AR benchmark: Italy CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	63.9**	84.7**	90.1**	88.0**	88.5*
AR-ARMA	25.0**	57.2**	69.9**	79.0**	87.3**
AR-MA-SV	**65.2****	84.7**	89.9**	86.8**	86.3*
AR-ARMA-SV	59.9**	80.1**	88.8**	92.8**	100.3*
UC	0.5**	27.0**	29.7**	25.6**	20.0*
UC-SV	38.2**	66.9**	63.6**	56.1**	49.5**
UC-ARMA	0.4	4.4**	0.4**	-4.5**	-9.9**
UC-MA-SV	48.7**	70.1**	77.8**	78.4**	78.6**
UC-ARMA-SV	46.9**	**87.5****	**102.0****	**105.9****	**110.9***

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 21: Sum of log predictive likelihoods relative to AR benchmark: Japan CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	-26.4**	50.1**	**68.7****	60.6**	63.2**
AR-ARMA	47.1**	53.1**	39.1**	33.3**	24.2**
AR-MA-SV	-27.5**	51.4**	67.1**	60.6**	62.0**
AR-ARMA-SV	**47.8****	50.1**	32.3**	19.0*	5.1
UC	11.5**	29.8**	31.5**	35.9**	35.6**
UC-SV	33.4**	55.8**	57.7**	66.2**	**68.9***
UC-ARMA	11.3**	25.4**	22.3**	25.9**	26.1**
UC-MA-SV	32.9**	**56.5****	57.0**	**64.5****	66.8*
UC-ARMA-SV	46.8**	52.2**	41.2**	40.7**	37.4**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 22: Sum of log predictive likelihoods relative to AR benchmark: UK CPI Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	21.1**	17.3**	60.8**	51.7**	51.9*
AR-ARMA	45.3**	69.7**	48.3**	37.9**	26.2**
AR-MA-SV	41.8**	21.1**	59.2**	50.9**	51.3*
AR-ARMA-SV	**57.7****	**72.3****	34.4**	15.0**	-0.9*
UC	-7.6**	29.3**	34.2**	34.8**	31.3**
UC-SV	20.9**	63.0**	**68.7****	**71.2****	**70.7***
UC-ARMA	-4.7**	31.6**	36.0**	39.8**	40.4**
UC-MA-SV	17.4**	62.6**	68.3**	70.3**	69.3*
UC-ARMA-SV	31.2**	59.7**	53.5**	50.5**	46.3*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 23: Sum of log predictive likelihoods relative to AR benchmark: US CPI Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	30.8	18.5	19.4	22.5	35.4
AR-ARMA	11.1	5.1**	27.3**	39.0*	55.0
AR-MA-SV	30.6	17.0	17.8	20.0	34.4
AR-ARMA-SV	28.6	22.6	30.6*	**40.9**	**61.2**
UC	-0.6**	-0.3**	11.5	2.9	9.3
UC-SV	**32.3**	25.9	26.2	25.2	36.9
UC-ARMA	-0.2**	2.8**	17.5*	18.7*	37.8
UC-MA-SV	30.5	**27.4**	**31.7***	34.6	50.8
UC-ARMA-SV	26.3	13.6	26.8	37.4	55.7

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 24: Sum of log predictive likelihoods relative to AR benchmark: Canada GDP Deflator Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	28.5	6.8	1.3*	3.7*	4.0
AR-ARMA	11.9**	3.4**	-2.6**	25.8**	33.0**
AR-MA-SV	33.7	6.1	0.8*	2.9*	3.1
AR-ARMA-SV	35.5**	**15.1****	**13.9****	**28.7****	**35.4****
UC	-67.2**	-68.4**	-60.3**	-57.1**	-46.1**
UC-SV	33.0	14.4	6.9**	9.8**	11.1*
UC-ARMA	14.0**	-0.3**	13.8**	26.6	31.7*
UC-MA-SV	35.5	12.1	6.5**	11.4**	11.2*
UC-ARMA-SV	**37.2**	0.1	2.3**	10.7**	15.1*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 25: Sum of log predictive likelihoods relative to AR benchmark: France GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	66.7**	60.0**	65.1**	57.7**	58.2*
AR-ARMA	19.7**	-2.8**	2.3**	8.2**	14.3**
AR-MA-SV	68.3**	59.8**	64.6**	57.4**	58.4*
AR-ARMA-SV	74.0**	57.6**	**66.2****	**69.4***	**72.3**
UC	-3.8**	4.6**	17.4**	23.1**	25.5**
UC-SV	69.7**	**59.9****	64.3**	59.9**	56.5**
UC-ARMA	3.1**	9.6**	23.5**	31.9**	37.8*
UC-MA-SV	**70.6****	58.0**	62.4**	58.8**	56.1**
UC-ARMA-SV	3.1**	9.6**	23.5**	31.9**	37.8**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 26: Sum of log predictive likelihoods relative to AR benchmark: Germany GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	**15.5****	**15.0****	19.5**	17.3*	16.8*
AR-ARMA	4.3**	-12.1**	-20.9**	-29.4**	-36.0**
AR-MA-SV	15.4**	14.2**	18.0**	15.6**	15.3*
AR-ARMA-SV	11.2**	2.3**	1.8**	-2.8**	-3.7*
UC	-5.2**	-3.0**	4.3**	4.2**	4.5*
UC-SV	13.4**	13.7**	**21.6****	**20.3***	**21.4**
UC-ARMA	-3.1**	-3.2**	4.5**	6.1*	7.9*
UC-MA-SV	13.4**	14.0**	20.9**	19.3**	20.3
UC-ARMA-SV	13.9**	13.9**	19.8**	18.0**	19.1*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 27: Sum of log predictive likelihoods relative to AR benchmark: Italy GDP Deflator Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	-1.8**	-1.6*	-2.0*	-3.6	-3.8
AR-ARMA	6.6**	1.4*	0.3	-0.8	-1.1
AR-MA-SV	-0.5**	-1.6*	-2.1*	-3.8	-3.8
AR-ARMA-SV	3.5**	-0.4*	-1.6*	-3.4	-3.6
UC	7.6**	4.1**	3.8*	3.7	3.4
UC-SV	11.8**	7.5**	5.8	3.5	2.3
UC-ARMA	10.2**	6.9**	6.5*	**6.5**	**6.3**
UC-MA-SV	13.0**	9.5**	7.9	5.2	4.1
UC-ARMA-SV	**11.9****	**10.3****	**8.7**	6.1	4.9

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 28: Sum of log predictive likelihoods relative to AR benchmark: Japan GDP Deflator Inflation

	$k = 1$	$k = 4$	$k = 8$	$k = 12$	$k = 16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	7.9	5.7	7.5	10.6	8.3
AR-ARMA	-9.1	-1.7	-6.0	-2.7	-18.4
AR-MA-SV	8.0	5.9	7.5	10.8	8.4
AR-ARMA-SV	8.4	7.0	8.8	**12.3**	**9.0**
UC	8.1	**9.1**	7.4	6.3	3.6
UC-SV	7.7	7.3	9.3	11.1	9.1
UC-ARMA	-4.9*	-3.9*	-6.0	-9.5	-5.9
UC-MA-SV	**9.0**	7.3	9.1	10.8	8.7
UC-ARMA-SV	8.6	7.4	**9.5**	11.0	8.8

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 29: Sum of log predictive likelihoods relative to AR benchmark: UK GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	21.1**	17.3**	**60.8****	**51.7****	**51.9***
AR-ARMA	45.3**	69.7**	48.3**	37.9**	26.2**
AR-MA-SV	41.8**	21.1**	59.2**	50.9**	51.3*
AR-ARMA-SV	**57.7****	**72.3****	34.4**	15.0**	-0.9*
UC	-10.2**	23.8**	27.5**	25.9*	23.9*
UC-SV	8.0**	42.6**	48.8**	52.3**	55.3*
UC-ARMA	-7.5**	23.8**	27.0**	29.3**	29.9*
UC-MA-SV	6.8**	42.0**	48.1**	52.6**	55.0*
UC-ARMA-SV	8.1**	10.1**	9.7**	12.2**	16.3**

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

Table 30: Sum of log predictive likelihoods relative to AR benchmark: US GDP Deflator Inflation

	$k=1$	$k=4$	$k=8$	$k=12$	$k=16$
AR	0.0	0.0	0.0	0.0	0.0
AR-SV	15.5**	34.6**	38.3**	38.9*	42.0
AR-ARMA	13.2**	35.4**	48.6**	60.9**	68.8**
AR-MA-SV	17.2**	28.1**	29.9**	29.6*	31.6
AR-ARMA-SV	**18.8****	**40.7****	**59.7****	**71.6****	**80.8****
UC	-4.8**	13.8*	17.3	10.5	6.3
UC-SV	17.5**	35.3**	35.0**	28.3*	25.2
UC-ARMA	4.9**	13.5**	17.3**	19.2**	20.6**
UC-MA-SV	14.1**	31.6**	33.3**	31.5**	32.2*
UC-ARMA-SV	18.1**	40.7**	49.6**	54.6**	58.6*

Note: ** and * indicate rejection of equal forecast accuracy relative to AR(m) at significance level 0.05 and 0.1, respectively, when using an asymptotic test in Diebold and Mariano (1995).

In line with the point forecast results, the AR-SV and UC-SV models tend to dominate their constant volatility counterparts in CPI forecasts. However, the result here is stronger in the sense that it also extends to the GDP deflator. More generally, models with ARMA-SV errors tend to provide better forecasts than those with SV, ARMA or homoscedastic errors. This is especially true in forecasting the GDP deflator in the US, or CPI in France. One notable exception is that the AR-ARMA model provides the best forecast of CPI in Canada across all horizons. Another is the one-step-ahead forecast of US CPI, in which the simpler UC-SV model provides the best density forecast. This result is important

because it is in contrast with those in Chan (2013). In that paper, the UC-MA-SV model is found to provide better forecasts of the same variable on an earlier sample. More precisely, Chan (2013) uses CPI inflation from 1947Q1 to 2011Q3, while we extend this up to 2017Q4. Thus, our result suggests that, in the case of US inflation forecasts, the relative performance of the UC-MA-SV model has declined since the end of 2011. That being said, the UC-MA-SV model still provides good forecasts of US CPI inflation at the one- and two-year-ahead forecast horizons.

Thus, our main conclusion from this exercise is that ARMA-SV error models again provide competitive forecasts across the sample of countries considered in this study. This is especially true in France, Canada and the US, where they provide the best forecasts of CPI for France and the GDP Deflator for Canada and the US across all forecast horizons.

4 Comparison with the Survey of Professional Forecasters

Recent research on US inflation dynamics has shown that subjective, non-model based forecasts, may provide superior point forecasts compared to conventional models (Ang et al., 2007; Croushore, 2010; Faust and Wright, 2009, 2013). For instance, Faust and Wright (2013) show that survey based point forecasts are more accurate than a range of state of the art econometric models, including the UC-SV model, over the period 1985Q1 to 2011Q4. In light of these results, it is interesting to see whether our ARMA-SV error specification can improve on these survey based forecasts.

To that end, we compare the forecast performance of the AR-ARMA-SV and UC-ARMA-SV models against those given by the Survey of Professional Forecasters (SPF). Established in 1968, the SPF is the oldest quarterly survey of macroeconomic forecasts in the United States. The current version of the survey contains forecasts for more than 23 key macroeconomic indicators, including both CPI and GDP Deflator based inflation measures.[4] The first survey to include CPI and GDP Deflator measured inflation was 1992 Q1.[5] The survey specifically asks participants to provide forecasts for the seasonally adjusted, annualized rate of both headline CPI inflation and the chain-weighted GDP price index. These survey participants are predominantly from the business sector, all make

[4]Documentation for the SPF is provided by the Federal Reserve Bank of Philadelphia: https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters.

[5]Prior to 1992 participants were asked to forecast the GNP implicit price deflator.

their living via forecasting, and consequently have strong incentives to do it accurately. The survey is consequently more accurate than other commonly used surveys such as the Livingston and Michigan Surveys (Thomas, 1999), making it a highly competitive benchmark.

When comparing the SPF and model based forecasts, an important detail is that the SPF forecasts of the GDP price index are always scaled to the base year that was in effect at the time of the survey. This means that the base year varies over the SPF sample making it incomparable with the most recently available vintage of GDP Deflator data used in our model based forecasts. To make these two series comparable, we rebase the SPF series so that the base year is 2009Q1 as in the most recent vintage of available data.

The Root MSFEs for the one- and four- quarter ahead CPI and GDP Deflator measured inflation forecasts from the SPF and our ARMA-SV error models are presented in Table 31. The samples for the CPI and GDP deflator forecasts are respectively 1993Q4-2016Q4 and 1997Q2-2016Q4. The main message is that the ARMA-SV error models can improve upon the SPF forecasts when forecasting the GDP Deflator measured inflation, but the SPF provides superior CPI inflation forecasts. This suggests that while ARMA-SV error models are useful when forecasting the GDP deflator, the SPF still remains a useful benchmark when forecasting US CPI inflation.

Table 31: Root MSFEs for one- and four-step-ahead CPI and GDP Deflator inflation forecasts.

	CPI		GDP Deflator	
	$k=1$	$k=4$	$k=1$	$k=4$
SPF	1.81	1.93	1.51	1.95
AR-ARMA-SV	2.02	2.19	0.83	0.89
UC-ARMA-SV	2.01	2.42	0.83	0.98

5 Concluding Remarks

We have introduced a new class of dynamic models with ARMA-SV errors, provided details on how to estimate them, and shown that they can be useful in forecasting inflation. The main difficulty in estimating such models is that the ARMA component induces serial dependence in the measurement errors, making the standard Kalman filter not directly applicable. We showed that this could be overcome by carefully designing the order of matrix operations. Moreover, by exploiting the model structure, we were

able to develop an efficient algorithm that avoids the forward and backward recursions in the Kalman filter. To illustrate the usefulness of the models, we assessed their forecast performance of two commonly used inflation measures: CPI and the GDP Deflator, in each of the G7 countries. More specifically, we presented both out-of-sample point and density forecast performance to various nested AR and UC models.

While there was no clearly dominant model across each of the countries, the AR-ARMA-SV model provided highly competitive forecasts of both inflation measures. In particular, they provided the best one-step-ahead point forecasts of CPI in all countries except Germany and the US. The model also dominated the CPI point forecasts at all other horizons in Canada, while the UC-ARMA-SV variant dominated in both France and Italy. This latter result extended to the density forecasts. In the former case, however, the simpler AR-ARMA model produced the best density forecasts across all forecast horizons. Finally, ARMA-SV models dominated the GDP Deflator density forecasts in Canada the US, provided good short-term forecasts in Italy and the UK, and good long-term forecasts in France.

Appendix A: Proof of $\mathbf{H}_\psi^{-1}\mathbf{H}_\phi = \mathbf{H}_\phi\mathbf{H}_\psi^{-1}$

Proposition: Suppose $\mathbf{H}_\phi$ and $\mathbf{H}_\psi$ are the following matrices of size T:

$$\mathbf{H}_\phi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ -\phi_1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ -\phi_p & \cdots & -\phi_1 & 1 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & -\phi_p & \cdots & -\phi_1 & 1 \end{pmatrix}, \quad \mathbf{H}_\psi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \psi_1 & 1 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \psi_q & \cdots & \psi_1 & 1 & \cdots & 0 \\ \vdots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \psi_q & \cdots & \psi_1 & 1 \end{pmatrix}.$$

Then: $\mathbf{H}_\phi^{-1}$ and $\mathbf{H}_\psi$ commute, i.e., $\mathbf{H}_\psi^{-1}\mathbf{H}_\phi = \mathbf{H}_\phi\mathbf{H}_\psi^{-1}$.

Proof: Let $\mathbf{L}_i$ be a $T \times T$ matrix which only has the nonzero elements 1 on the i-th lower diagonal for $i = 0, \ldots, T-1$, i.e.,

$$\mathbf{L}_i = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & & 0 \\ \vdots & 0 & 0 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & \ddots & & & \vdots \\ 1 & \ddots & \ddots & 0 & \cdots & & 0 \\ 0 & \ddots & 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix}.$$

In particular, $\mathbf{L}_0 = \mathbf{I}_T$ (identity matrix). It is easy to check that $\mathbf{L}_i\mathbf{L}_j = \mathbf{L}_{i+j} = \mathbf{L}_j\mathbf{L}_i$, when $i, j \geqslant 0$ and $i + j \leqslant T - 1$. Then we can write $\mathbf{H}_\phi$ and $\mathbf{H}_\psi$ as:

$$\mathbf{H}_\phi = \mathbf{I}_T - \sum_{i=1}^{p} \phi_i \mathbf{L}_i, \quad \mathbf{H}_\psi = \mathbf{I}_T + \sum_{j=1}^{q} \psi_j \mathbf{L}_j.$$

So that:

$$\begin{aligned}
\mathbf{H}_\phi \mathbf{H}_\psi &= \left(\mathbf{I}_T - \sum_{i=1}^{p} \phi_i \mathbf{L}_i\right)\left(\mathbf{I}_T + \sum_{j=1}^{q} \psi_j \mathbf{L}_j\right) \\
&= \mathbf{I}_T - \sum_{i=1}^{p} \phi_i \mathbf{L}_i + \sum_{j=1}^{q} \psi_j \mathbf{L}_j - \sum_{i=1}^{p}\sum_{j=1}^{q} \phi_i \psi_j \mathbf{L}_i \mathbf{L}_j \\
&= \mathbf{I}_T + \sum_{j=1}^{q} \psi_j \mathbf{L}_j - \sum_{i=1}^{p} \phi_i \mathbf{L}_i - \sum_{j=1}^{q}\sum_{i=1}^{p} \psi_j \phi_i \mathbf{L}_j \mathbf{L}_i \\
&= \left(\mathbf{I}_T + \sum_{j=1}^{q} \psi_j \mathbf{L}_j\right)\left(\mathbf{I}_T - \sum_{i=1}^{p} \phi_i \mathbf{L}_i\right) \\
&= \mathbf{H}_\psi \mathbf{H}_\phi.
\end{aligned}$$

Hence:

$$\begin{aligned}
\mathbf{H}_\psi^{-1}(\mathbf{H}_\phi \mathbf{H}_\psi)\mathbf{H}_\psi^{-1} &= \mathbf{H}_\psi^{-1}(\mathbf{H}_\psi \mathbf{H}_\phi)\mathbf{H}_\psi^{-1} \\
\mathbf{H}_\psi^{-1}\mathbf{H}_\phi &= \mathbf{H}_\phi \mathbf{H}_\psi^{-1}.
\end{aligned}$$

Appendix B: Estimation Details

In this appendix we provide the details of the posterior simulator. As mentioned in the main text, we implement a Metropolis-within-Gibbs sampler with five steps. The first one involves sampling the trend $\boldsymbol{\tau}$ from its complete conditional distribution, which is discussed in the main text. We provide the details of other steps below.

Step 2: Sample h

To sample from $p(\mathbf{h} \,|\, \mathbf{y}, \boldsymbol{\tau}, \boldsymbol{\phi}, \boldsymbol{\psi}, \sigma_\tau^2, \sigma_h^2)$, note that (8) can be written as:

$$\mathbf{y}^* = \mathbf{u},$$

where $\mathbf{y}^* = \mathbf{H}_\psi^{-1}\mathbf{H}_\phi(\mathbf{y} - \boldsymbol{\tau})$. Thus, $(\mathbf{y}^* \,|\, \mathbf{h}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_u)$ with $\boldsymbol{\Omega}_u = \text{diag}\left(\text{e}^{h_1}, \ldots, \text{e}^{h_T}\right)$. With this transformation the auxiliary mixture sampler proposed by Kim, Shepherd, and Chib (1998) can be directly applied. The only difference here is that we replace their forward-backward smoothing algorithm with the precision-based sampler discussed in the main text.

Step 3: Sample σ_h^2 and σ_τ^2

Since an inverse-gamma prior is conjugate for the normal likelihood, sampling σ_h^2 and σ_τ^2 is standard. For example, it follows from (5) and the inverse-gamma prior on σ_h^2 that

$$\begin{aligned} p(\sigma_\tau^2 \,|\, \boldsymbol{\tau}) &\propto p(\boldsymbol{\tau} \,|\, \sigma_\tau^2) p(\sigma_\tau^2) \\ &= (\sigma_\tau^2)^{-\frac{T-1}{2}} \text{e}^{-\frac{1}{2\sigma_\tau^2}\sum_{t=2}^{T}(\tau_t - \tau_{t-1})^2} \times (\sigma_\tau^2)^{-(\nu_\tau+1)} \text{e}^{-\frac{S_\tau}{\sigma_\tau^2}}, \\ &\propto (\sigma_\tau^2)^{-(\frac{T-1}{2}+\nu_\tau+1)} \text{e}^{-\frac{1}{\sigma_\tau^2}(\sum_{t=2}^{T}(\tau_t - \tau_{t-1})^2/2 + S_\tau)}. \end{aligned}$$

Hence, we have

$$(\sigma_\tau^2 \,|\, \boldsymbol{\tau}) \sim \mathcal{IG}\left((T-1)/2 + \nu_\tau, \quad \sum_{t=2}^{T}(\tau_t - \tau_{t-1})^2/2 + S_\tau\right).$$

Similarly, the posterior density of σ_h^2 is given by:

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

Step 4: Sample $\boldsymbol{\psi}$

Since the complete conditional distribution of $\boldsymbol{\psi}$ is non-standard, we implement the independence-chain Metropolis-Hastings algorithm (Kroese, Taimre, and Botev, 2011) using a suitable proposal density. Below we first derive an analytical expression of the full conditional density of $\boldsymbol{\psi}$. Stacking (1) and (2) over $t = 1, \ldots, T$ gives:

$$\mathbf{y}^{**} = \mathbf{H}_{\psi}\mathbf{u}, \tag{20}$$

where $\mathbf{y}^{**} = \mathbf{H}_{\phi}(\mathbf{y} - \boldsymbol{\tau})$. By a change of variable, we have $(\mathbf{y}^{**} \,|\, \boldsymbol{\psi}, \mathbf{h}) \sim \mathcal{N}\left(\mathbf{0}, \mathbf{H}_{\psi}\boldsymbol{\Omega}_{u}\mathbf{H}_{\psi}'\right)$. Therefore, given the truncated normal prior on $\boldsymbol{\psi}$, the log conditional posterior of $\boldsymbol{\psi}$ is given by:

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

The above log density can be evaluated using the method discussed in Section 2.1.1. Since the dimension of $\boldsymbol{\psi}$ is typically low, we can use numerical optimization routines to obtain the mode and negative Hessian of $\log p(\boldsymbol{\psi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \boldsymbol{\phi})$ evaluated at the mode, which we denote as $\widehat{\boldsymbol{\psi}}$ and $\mathbf{K}_{\psi}$, respectively. Let $q(\boldsymbol{\psi})$ represent the $\mathcal{N}(\widehat{\boldsymbol{\psi}}, \mathbf{K}_{\psi}^{-1})$ density, and we use $q(\boldsymbol{\psi})$ to generate candidates. Given the current draw $\boldsymbol{\psi}$, a candidate draw $\boldsymbol{\psi}^{c} \sim \mathcal{N}(\widehat{\boldsymbol{\psi}}, \mathbf{K}_{\psi}^{-1})$ is accepted with probability:

$$\min\left\{1, \quad \frac{p(\boldsymbol{\psi}^{c} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \boldsymbol{\phi})}{p(\boldsymbol{\psi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}, \boldsymbol{\phi})} \times \frac{q(\boldsymbol{\psi})}{q(\boldsymbol{\psi}^{c})}\right\};$$

otherwise we return $\boldsymbol{\psi}$.

Step 5: Sample $\boldsymbol{\phi}$

Finally, we sample $\boldsymbol{\phi}$ from its complete conditional distribution. To that end, note that given $\mathbf{y}$ and $\boldsymbol{\tau}$, we can compute $\boldsymbol{\varepsilon}^{y} = \mathbf{y} - \boldsymbol{\tau}$. Then, we can rewrite (20) as:

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

where $\mathbf{X}_{\boldsymbol{\varepsilon}^y}$ is a $T \times p$ matrix of lagged residuals, i.e.,

$$\mathbf{X}_{\boldsymbol{\varepsilon}^y} = \begin{pmatrix} \varepsilon_0^y & \varepsilon_{-1}^y & \cdots & \varepsilon_{1-p}^y \\ \varepsilon_1^y & \varepsilon_0^y & \cdots & \varepsilon_{2-p}^y \\ \vdots & \vdots & & \vdots \\ \varepsilon_{T-1}^y & \varepsilon_{T-2}^y & \cdots & \varepsilon_{T-p}^y \end{pmatrix}$$

Therefore, $\boldsymbol{\phi}$ is equivalent to the coefficients of a linear regression model with MA errors. Given the truncated normal prior, the complete conditional density of $\boldsymbol{\phi}$ is:

$$(\boldsymbol{\phi} \,|\, \mathbf{y}, \boldsymbol{\tau}, \mathbf{h}) \sim \mathcal{N}(\widehat{\boldsymbol{\phi}}, \mathbf{K}_\phi^{-1}) \mathbb{1}(\boldsymbol{\phi} \in \mathbf{A}_\phi),$$

where $\mathbf{K}_\phi = \mathbf{V}_\phi^{-1} + \mathbf{X}'_{\boldsymbol{\varepsilon}^y}(\mathbf{H}_\psi \boldsymbol{\Omega}_u \mathbf{H}'_\psi)^{-1} \mathbf{X}_{\boldsymbol{\varepsilon}^y}$ and $\widehat{\boldsymbol{\phi}} = \mathbf{K}_{\boldsymbol{\phi}}^{-1}(\mathbf{V}_\phi^{-1}\boldsymbol{\phi}_0 + \mathbf{X}'_{\boldsymbol{\varepsilon}^y}(\mathbf{H}_\psi \boldsymbol{\Omega}_u \mathbf{H}'_\psi)^{-1} \boldsymbol{\varepsilon}^y)$. Sampling from this distribution can then be done using a standard acceptance-rejection algorithm (see, e.g. Koop, 2003; Kroese et al., 2011).

References

Andrew Ang, Geert Bekaert, and Min Wei. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics*, 54(4):1163–1212, 2007.

Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

George EP Box and Gwilym M Jenkins. *Time series analysis: Forecasting and control*. Holden-Day, 1970.

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

Joshua C. C. Chan and Ivan Jeliazkov. Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation*, 1(2):101–120, 2009.

Joshua C. C. Chan, Gary Koop, and Simon M. Potter. A new model of trend inflation. *Journal of Business and Economic Statistics*, 31(1):94–106, 2013.

Siddhartha Chib and Edward Greenberg. Bayes inference in regression models with ARMA(p, q) errors. *Journal of Econometrics*, 64:183–206, 1994.

Todd E. Clark and Taeyoung Doh. Evaluating althernative models of trend inflation. *International Journal of Forecasting*, 30(3):426–448, 2014.

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

Jamie Cross and Aubrey Poon. Forecasting structural change and fat-tailed events in Australian macroeconomic variables. *Economic Modelling*, 58:34–51, 2016.

Dean Croushore. An evaluation of inflation forecasts from surveys using real-time data. *The BE Journal of Macroeconomics*, 10(1), 2010.

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

Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

Jon Faust and Jonathan H Wright. Comparing greenbook and reduced form forecasts using a large realtime dataset. *Journal of Business & Economic Statistics*, 27(4):468–479, 2009.

Jon Faust and Jonathan H Wright. Forecasting inflation. In *Handbook of economic forecasting*, volume 2, pages 2–56. Elsevier, 2013.

Gene H. Golub and Charles F. Van Loan. *Matrix computations*, volume 3. The Johns Hopkins University Press, 2013.

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

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

Dirk P. Kroese, Thomas Taimre, and Zdravko I. Botev. *Handbook of Monte Carlo methods*. John Wiley and Sons, New York, 2011.

Stefan Ludbergh, Timo Teräsvirta, and Dick van Dijk. Time-varying smooth transition autoregressive models . *Journal of Business and Economic Statistics*, 21(1):104–121, 2003.

Massimliano Marcellino. Forecasting EMU macroeconomic variables . *International Journal of Forecasting*, 20:359–372, 2004.

William J. McCausland, Shirley Miller, and Denis Pelletier. Simulation smoothing for state-space models: A computational efficiency analysis. *Computational Statistics and Data Analysis*, 55(1):199–212, 2011.

Teruo Nakatsuma. Bayesian analysis of ARMA–GARCH models: A Markov chain sampling approach. *Journal of Econometrics*, 95(1):57–69, 2000.

James H. Stock and Mark W. Watson. Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics*, 14:11–30, 1996.

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

Stephen J. Taylor. Modeling stochastic volatility: A review and comparative study. *Mathematical finance*, 4(2):183–204, 1994.

Lloyd B Thomas. Survey measures of expected US inflation. *Journal of Economic Perspectives*, 13(4):125–144, 1999.