Erasmus University Rotterdam
Erasmus School of Economics
Seminar in Forecasting (FEB23019)

Enhancing Inflation Forecasting Models through Bayesian Hierarchical Approaches and Stochastic Seasonality

Philipp Höcker
564398

Maurits van Altvorst
557923

Matthias Hofstede
612810

Stefan van Diepen
573588

Team	2
Supervisor:	Daan Opschoor
Date:	22nd of April 2024

Abstract

To enhance decision-making, accurate inflation forecasts are of high importance for both governments and the general public. This paper examines whether existing forecasting techniques can be improved upon using hierarchical priors. We argue that inflation is a global phenomenon, which is why a Bayesian hierarchical model with a prior based on the distance between countries is of special interest. Similarly, this paper extends the unobserved components model with stochastic volatility to incorporate stochastic seasonality, and it constructs a multivariate unobserved components model that builds upon a geospatial prior. These models are evaluated based on their point forecasts and we find that adding a seasonal component to the unobserved components model with stochastic volatility leads to the most accurate forecasts both when using headline and core consumer price indices. The results also suggest that a Bayesian hierarchical model with a geospatial prior does not improve inflation forecasting.

1 Introduction

Central banks around the world have clearly defined inflation targets, such as the 2% target of the European Central Bank, U.S. Federal Reserve, and Bank of England, and the 3% inflation target of Columbia. Besides having clear inflation targets, policymakers should make predictions of inflation in order to take early policy measures to keep inflation close to its target. Therefore, forecasting inflation is of relevance to policymakers all around the world. Furthermore, inflation rates affect varying aspects of the economy, ranging from consumer purchasing power to business investment decisions. Thus, accurate inflation forecasts can help the general public to make better-informed decisions when forming their spending patterns. Higher inflation might incentivise consumers to, for example, buy less premium products or make different choices about booking holidays to ensure that they are able to pay their bills during this period of high inflation. Furthermore, as inflation affects real investment returns, accurate inflation forecasts are also relevant within the financial sector. Investment returns should exceed inflation to ensure a positive real return on investment, which leads to higher required returns on investments in high inflationary periods. Finally, high inflation usually leads to an increase in interest rates, as central banks use interest rates as a policy measure to keep inflation close to its target. Therefore, inflation affects interest rate expectations, and thus influences stock and bond markets via this dynamic.

Due to its high societal relevance, inflation forecasting is a widely discussed topic in academic literature. For example, Stock and Watson (2007) show that U.S. inflation has become both easier and harder to forecast over the last decades. By this, they mean that the root mean square forecast error of the models they implement has decreased over time, so in that sense, it is easier to get 'good' inflation forecasts. However, it has become harder to beat an easy benchmark AR(p) model, where p is selected using the Akaike information criterion (AIC). Therefore, in order to have inflation forecasting models that add value over simple benchmark autoregressive models, future research should be undertaken to identify novel and better ways to forecast inflation.

Furthermore, Stock and Watson (2007) describe the unobserved components stochastic volatility model (UC-SV), in which variances of permanent and transitory disturbances evolve randomly over time. As this model was the best performing among the models considered by Stock and Watson (2007), this model is used as a 'tough-to-beat' benchmark model in this paper.

Stock and Watson (2016) show that sector-specific data on inflation can be used to improve the estimate of trend inflation, as the multivariate estimates have a better forecasting performance than the univariate estimates when using a multivariate UC-SV (M-UC-SV) model. These models are Bayesian in the sense that we incorporate priors on the way the unobserved components evolve over time. Stock and Watson (2016) use this model to forecast core and trend inflation. This paper extends the research by Stock and Watson (2007, 2016) by using Bayesian hierarchical models across countries to improve forecasts for headline inflation as measured by the change in the Consumer Price Index (CPI).

Atkeson and Ohanian (2001) introduce the Atkeson-Ohanian random walk model (AO-RW). This is a simple naive model that forecasts inflation in the next four quarters to be equal to that of the previous four quarters. They find that the existing Phillips curve models of the time did not provide more accurate inflation forecasts than this naive model. Therefore, it was used as a benchmark model by Stock and Watson (2007). Due to the simplicity of the model and its proven performance, the model will also be used as a benchmark model in this paper.

Many other contributions have been made to try to improve inflation forecasts using different techniques. These contributions include Groen, Paap, and Ravazzolo (2013), who use Bayesian model averaging. They find that their model provides accurate forecasts, and outperforms their benchmark model, which is the same UC-SV model that this paper takes as a benchmark. Medeiros, Vasconcelos, Veiga, and Zilberman (2021) use machine learning techniques and show they can outperform simpler inflation benchmarks. However, one of the downsides of machine learning techniques is that the interpretation of these models is often not straightforward.

Lastly, an often cited contribution is that of Stock and Watson (1999), where the effect of the Phillips curve on inflation forecasting is examined. They find that Phillips curve based models produce more reliable out-of-sample forecasts than autoregressive and random walk models. The authors also incorporate alternative Phillips curves, defined using measures of economic activity other than unemployment to improve forecasting performance. However, for these models, it is yet unknown how the performance would compare to the more sophisticated UC-SV models. Therefore, the use of unemployment data during inflation forecasting seems relevant, and we include this as one of our exogenous variables for some of our models.

Lis and Porqueddu (2018) argue that inflation is seasonal, even when correcting for food and energy prices. Reasons for this finding include seasonal demand for certain goods and services, such as clothing or vacation. This seasonality feature - largely unexploited for forecasting purposes in the literature - is used in this paper as inspiration for adding stochastic seasonality components to unobserved component models.

As a final remark on the literature regarding inflation forecasting, Ciccarelli and Mojon (2010) argue that 70% of the variance of inflation in OECD countries can be explained by a common global factor, which is the simple average of the inflation in the countries under consideration. This factor is estimated using a static principle component method. They furthermore show that using a model that incorporates this factor outperforms other, simpler, inflation models such as AR models. The model used to estimate inflation incorporating a global inflation component is a one-factor model including lags for inflation and the common factor. We extend upon this, and argue that the data generating processes of inflation across countries are similar to each other. A global surge in oil prices likely has a positive effect on inflation around the globe. Therefore, this paper focuses on the use of Bayesian hierarchical models in inflation forecasting on a global scale, taking into account that inflation rates between countries are highly correlated and that shocks have similar effects across countries.

Summing up, the literature suggests that inflation is both seasonal and similar across countries. Therefore, the focus of this research is to construct models for inflation forecasting that incorporate seasonality and cross-sectional similarities of inflation across countries. Concretely, we investigate the following two research questions:

Research Question 1

Do seasonality features improve the one-quarter ahead forecasting performance of existing quarterly headline consumer price index inflation models?

Research Question 2

Does imposing similarity in inflation patterns across countries improve the one-quarter ahead forecasting performance of quarterly headline consumer price index inflation models?

Therefore, this paper constructs several Bayesian models that we compare against more traditional baseline models. We construct Bayesian hierarchical models, imposing priors on inflation across countries, as well as unobserved component models, in which we incorporate seasonality as separate components. The main aim is to compare pointwise forecast performance, despite that some of these models have the capability to forecast entire densities.

The pointwise prediction performance of these models is assessed using different performance metrics. We find that Bayesian hierarchical models do not outperform the UC-SV, AO-RW, and AR benchmark models. This suggests that it is not possible to improve existing inflation forecast models by using Bayesian hierarchical models that take advantage of the similarity of inflation rates of different countries. However, the UC-SV model with a stochastic seasonality component is shown to outperform all of the benchmark models defined in this paper when forecasting one-quarter ahead, including the UC-SV model. It is thus possible to take advantage of the under-explored seasonality in inflation data to improve existing forecasting models.

The rest of this paper has the following structure. Section 2 introduces the data used in the research, as well as some stylised facts of inflation data. Section 3 introduces the models used in the paper, as well as estimation techniques used for the estimation of the models and performance metrics to evaluate their performance. Section 4 presents and discusses the results of the research. Finally, Section 5 concludes the paper.

2 Data

This section outlines the key data needed to produce inflation forecasts with this paper's methodology and displays the most important patterns in the data that motivate the choice of model construction. Initially, the methodological aspects of the data selection are explored, specifically addressing the relevance and utility of the chosen exogenous covariates for inflation forecasting.

This includes a detailed overview of the data sources, measurement methods employed, and the necessary data transformations, which ensure the data's applicability for timely forecasting. Following the descriptive exposition, a series of stylised facts derived from the datasets is presented. These include comparative analyses of median global inflation against commodity prices, and graphical representations correlating interest rates and inflation. Additionally, we discuss the observed seasonality in inflation data post-2000s and conclude with visual insights into the geographic correlation of inflation rates among proximate countries. The latter insights motivate the choice of models presented in this paper.

2.1 Description of Data Sources

Besides lagged inflation rates, four more exogenous covariates are used to construct inflation forecasts. Those are global commodity prices, country-level unemployment rates, GDP growth, and central bank interest rates. Furthermore, data on geographical distances between countries is used to construct priors on the correlation of their inflation. After describing the inflation dataset, the relevance of each feature is briefly outlined and the data sources are described in detail.

A global inflation database published by the World Bank is used to retrieve quarterly inflation data for the research (Ha, Kose, & Ohnsorge, 2023). In particular, we opt to use headline inflation, as measured by the changes in the Consumer Price Index (CPI). Although core inflation (CPI inflation excluding food and energy prices) would allow us to better distinguish short-term from permanent price movements (Laflèche & Armour, 2006), historical data available for this is limited. Core inflation is only available for 18 countries on a quarterly basis since 1970, whilst headline inflation is available for 92 countries for the same frequency and time period. Hence, we opt to initially construct and evaluate all models using headline inflation and later perform a robustness check on a subset of countries by using the same methods to forecast core inflation.

Next, exogenous features are discussed, starting with commodity prices. Resources such as crude oil or natural gas form the basis of many production processes and also make up the lion's share of many countries' energy mix. Thus, changes in their prices have direct impacts on consumer end-prices. It is therefore not surprising that Boughton and Branson (1988) have shown that historically, changes in commodity price indices have been a leading indicator for G7 inflation. Further, Chen, Turnovsky, and Zivot (2014) have shown that global commodity prices improve out-of-sample CPI inflation forecasts, outperforming a simple AR(1) model. Commodity prices are available along the inflation data published by Ha et al. (2023) at a quarterly frequency. In this paper, a subset of the most important price indicators is used: crude oil prices and four indices representing global price levels for natural gas, agricultural products, metals and minerals, and finally precious metals. Whilst those four indicators generally move in the same direction, occasionally there are spikes in one not followed by increases in the other (e.g., natural gas after Russia's invasion of Ukraine in 2022, see Figure 1). Using multiple indicators allows for modelling differing inflation across countries, as different nations exhibit different levels of dependence on different sorts of energy sources and raw materials.

Further, the central bank interest rates are considered. Existing literature makes this a natural choice. From a theoretical perspective, the well-known Fisher equation prescribes equality between the real rate of return and the difference between the nominal interest rate and the expected inflation rate (Fisher, 2006). Indeed, Fama (1975) presents evidence suggesting that the real rate of return is constant, thus suggesting a clear linear relationship between interest rates and inflation. Indeed, this insight is consistent with the practice of major central banks, such as the European Central Bank or the Federal Reserve, using interest rates as the main policy tool to influence the rate of inflation. Central bank interest rates are available from the Bank for International Settlements (2024). Overall, the number of countries that have quarterly data available is limited, but well above 20 from 1990 onwards.

Similar to interest rates, there is an established theoretically based inverse relationship between inflation and unemployment, in the form of the Phillips curve. Stock and Watson (1999) have indeed shown that Phillips curve based forecasts for inflation rates outperform forecasts both based on interest rates and commodity prices, thus encouraging the inclusion of this macroeconomic covariate in inflation forecasting models. Data on quarterly unemployment rates is available from the International Labour Organisation (2024), from here on abbreviated by ILO. One struggle with this data is the range of countries it is available for. Only from around 1997, there are over 30 countries with quarterly available unemployment data in the ILO database.

As a last covariate, we consider GDP growth rates. The relationship between economic growth and inflation is a lot more debated than the relevance of the previous covariates. In particular, the causal direction is unclear. For example, Koulakiotis, Lyroudi, and Papasyriopoulos (2012) when focusing on European countries, do not only find that GDP growth affects inflation, but also that inflation affects GDP growth. GDP growth data is available from the Organisation for Economic Co-operation and Development (2024). Quarterly data here is readily available for the original member states from 1960 onwards. Figures on over 40 countries, including the OECD's current 38 member states, are only available as of the late 1990s.

Finally, geographical distances between countries are used. In general, one would expect that countries that are subject to similar economic circumstances - a similar energy mix, related consumption patterns, the same trade partners - should react similarly to changes in economic conditions, and thus have correlated rates of inflation. We hence choose to use data on the geographical distance between countries to construct priors on the closeness of their model parameters, i.e., neighbouring countries should have close model parameters. Data on distances between countries are found in the work of Mayer and Zignago (2011). The authors do not only provide a weighted distance measure, taking into account the location of the main population centres for geographical distance calculations, but further provide dummies regarding the contiguity of nations.

One may rightfully argue that geographical distance is an imperfect proxy for the economic resemblance between countries. Its use can however be justified by the lack of readily available alternatives. For instance, international trade flows may be a much better metric, but to the best of our knowledge, no comprehensive database of trade flows - on a country-pair basis - exists. Data is particularly sparse historically prior to the year 2003, the first year for which the World Trade Organization published their annual World Trade Report. Even with complete trade data, its use for this research report would be questionable, as these figures are often published with great delay. Except for GDP growth, all other covariates of choice - commodity prices, interest rates, and unemployment rates - are quarterly and swiftly published for developed economies. GDP growth figures pose a practical problem, as they are often published with a delay of around six weeks (Bańbura & Rünstler, 2011). This limits their usability for short term (e.g., one-quarter ahead) forecasts, but is less of an issue if the forecast horizon far exceeds the publication delay (e.g., one-year ahead).

Near to no transformations or missing value filling are needed for the data presented in this section. Solely for the purpose of principal component analysis, we fill a small number of non-diagonal missing elements of the inflation covariance matrix with the median covariance across country pairs, and make slight adjustments to the diagonal to ensure positive semi-definiteness. Naturally, however, most countries do not have all features available for a long enough period of time. When filtering out countries that lack data for a too large part of the sample period, we end up with 20 countries that we use for the analysis: Australia, Belgium, Canada, Denmark, France, Germany, Italy, Japan, Korea, The Netherlands, New Zealand, Norway, Portugal, South Africa, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The sample ranges from Q1 1977 up to and including Q1 2023.

2.2 Stylised Facts

Now, a high-level overview of the patterns in the data that motivate the choice of models is presented.

First, we illustrate the relationship between commodity prices and inflation already outlined in the previous subsection.

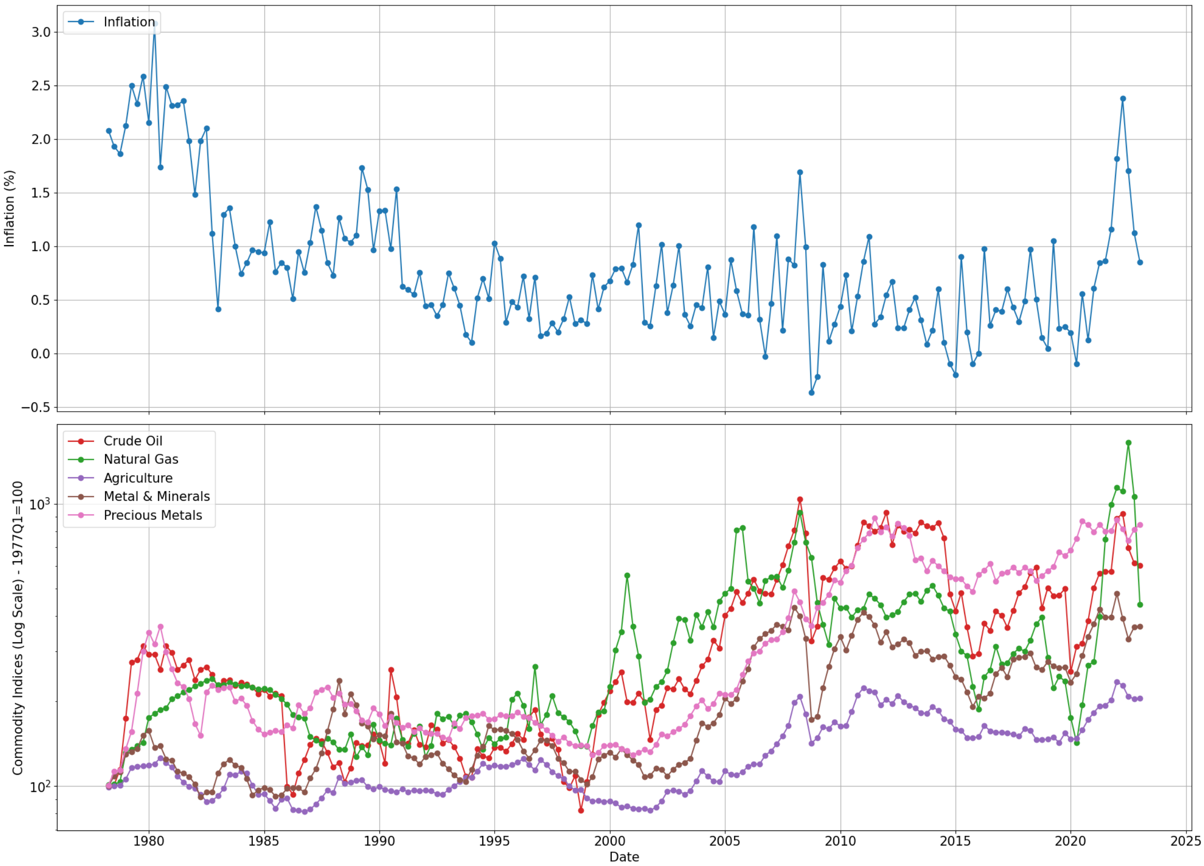


Figure 1: Inflation and various commodity indices over time

In Figure 1, we can see the median inflation of the 20 countries in the sample, and below the commodity indices, over time. We see that inflation has generally been high in the 1970s and 1980s, and relatively low and stable in the 1990s and the 21st century, with the exception of spikes during the 2008 financial crisis and the post-COVID-19 pandemic era. Further, while no clear relationship between commodity prices is apparent at times of low inflation, peaks in inflation are preceded by sharp increases in commodity prices, particularly in oil and gas. This is indeed in line with the results from Boughton and Branson (1988).

Similarly, we illustrate the empirical relationship between interest rates and inflation.

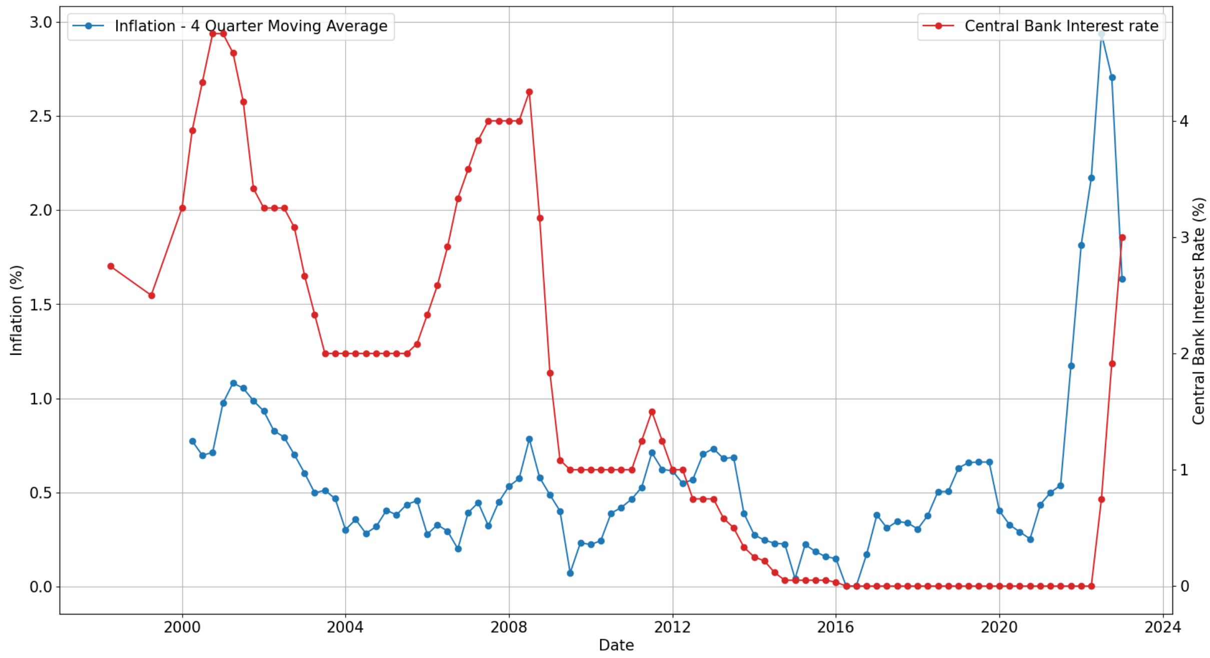


Figure 2: Moving average inflation and central bank interest rates in The Netherlands over time

Figure 2 shows the central bank interest rate and the four-quarter moving average inflation in The Netherlands over time. We can see multiple occasions, but most notably in 2023, where inflation shot up, and subsequently the European Central Bank reacted by raising interest rates, eventually bringing inflation down again, hence suggesting that an increase in interest rates causes a decrease in inflation with lag.

For other covariates, the relationships established by the literature are less apparent from simple visual inspection.

We now move to a feature that is yet under-exploited in the inflation forecasting literature: seasonality.

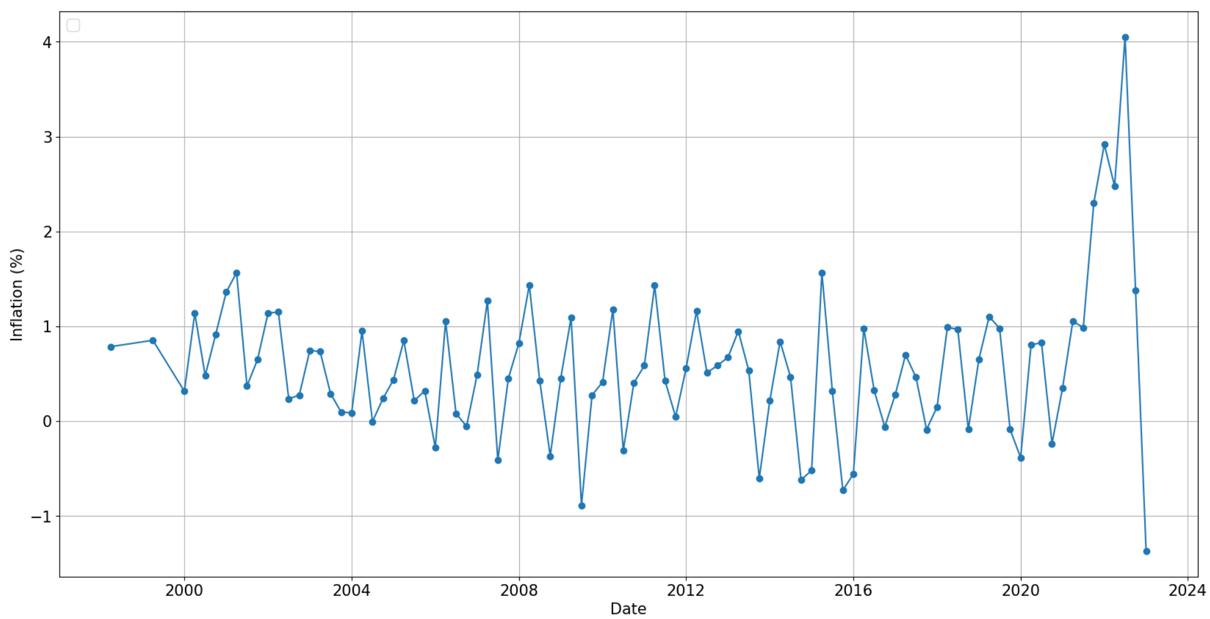


Figure 3: Quarterly inflation in The Netherlands

Figure 3 again shows quarterly inflation in The Netherlands, but now with no moving average smoothing. Upon inspection, one can clearly observe small peaks and lows occurring in a regular pattern, with an annual period. This prompts the additive decomposition of countries' inflation time series into a constant trend component and four-quarterly constant seasonal components. The results of splitting the sample into three roughly equal periods and extracting the seasonal components for each sub-sample are shown in Figure 4.

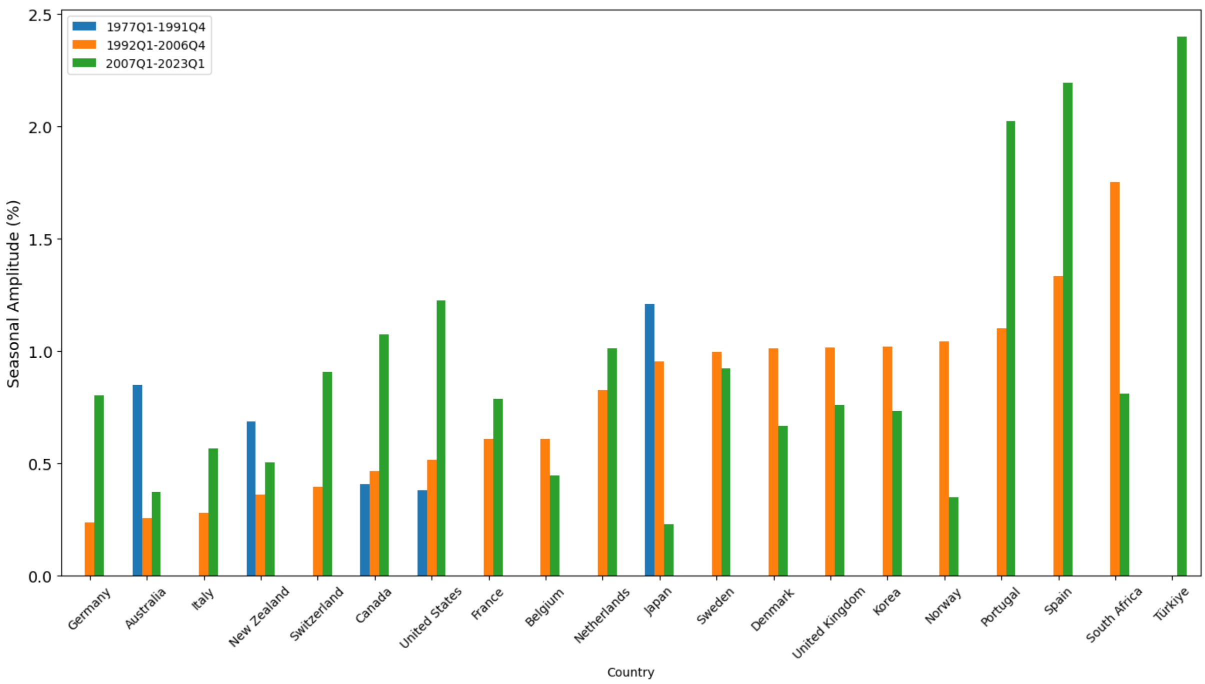


Figure 4: Range of the seasonal components across countries and sub-samples

Figure 4 shows the difference between the maximum and minimum seasonal component in the additive decomposition on a per-country level. One can observe that for the majority of countries, the seasonal effect increased significantly over time. For the United States, for example, it more than doubled from the period of 1992-2006 to 2007-2023, from 0.4% to over 1.2%. There is little literature on the possible causes of inflation seasonality. A part of headline CPI inflation seasonality can be attributed to food and energy prices. However, the pattern persists even with core inflation. The most common explanation for the latter is seasonal demand for holidays or goods like clothing (Lis & Porqueddu, 2018). Overall, the observation of seasonality motivates a model specification allowing for it.

Finally, we illustrate the phenomenon that inflation rates of geographically close countries are more correlated than inflation rates of distant countries.

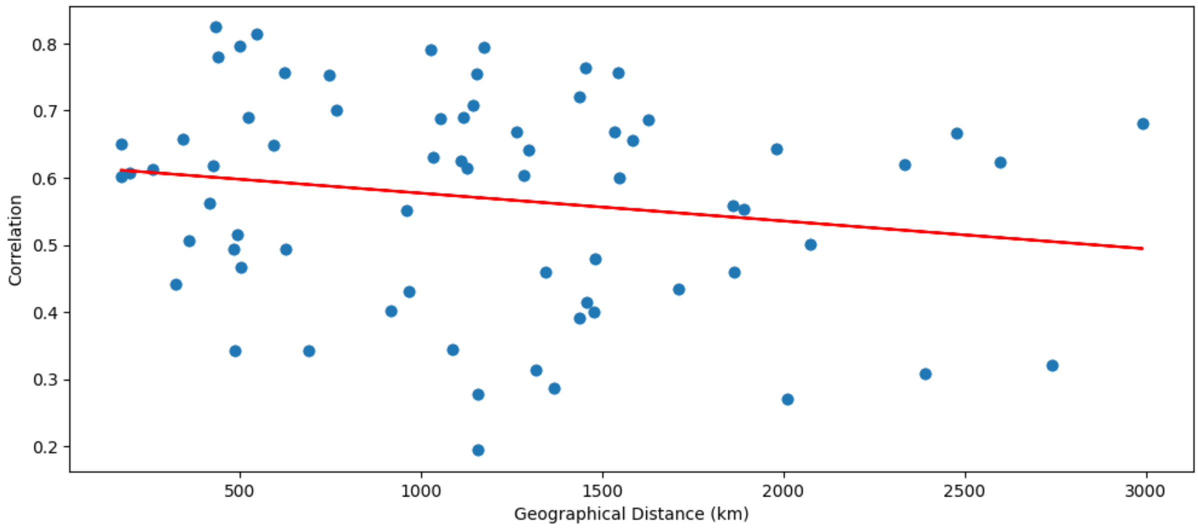


Figure 5: Correlation of countries and geographical distance

Figure 5 shows a scatter plot of all 190 country pair inflation correlations in the sample against their geographical distance. Although a visually less obvious relationship, a linear regression of correlation on distance and a constant shows a statistically significant slope coefficient at the 1% level. This is sufficient evidence to prompt the use of geographical distance to construct priors on the similarity of countries' inflation model parameters.

Further summary statistics on all datasets presented in this section can be found in Tables 4-8 in Appendix A.

3 Methodology

This section introduces the models that are used to forecast inflation in this paper. First, the Bayesian models we consider are introduced in Sections 3.1-3.2. Next, Sections 3.3-3.4 introduce unobserved component models. In both cases cases, we first present univariate models, and then provide an appropriate extension that imposes similarity in inflation patterns across countries. Sections 3.5-3.7 present three simpler baseline models for comparison purposes. Finally, Section 3.8 introduces some evaluation metrics.

For all models, we denote inflation at time t for country i as $y_{t,i}$ for all $i \in \{1, \ldots, N\}$ and $t \in \{1, \ldots, T\}$. The m-th independent variable is denoted by $x_{m,t,i}, \quad \forall m \in \{1, ..., M\}$, where the independent variables are as defined in Section 2.

3.1 Bayesian Hierarchical Models

Bayesian hierarchical models (BHMs) allow for the incorporation of prior distributions across different levels of a hierarchical structure (Lindley & Smith, 1972). This is particularly useful in the context of group-level effects or in the context of nested data. A fundamental advantage of BHMs in this context is that they are able to utilize information across individuals to improve estimates, especially when data is limited.

We introduce three different BHMs to forecast inflation. All three models are based on an autoregressive linear model with exogenous regressors: predictions are a linear function of lagged observed inflation values within a country, lagged macroeconomic variables within a country, and lagged global macroeconomic variables that are shared across countries. While these types of models are usually estimated by using OLS, Ridge or Lasso regression, we use a BHM to estimate these parameters simultaneously for different countries to act as a form of regularisation. Note that when estimating these Bayesian models, all features, lagged variables, and the target are normalised to have a mean of 0 and a standard deviation of 1 which implies that we can compare coefficients between the estimated BHMs.

We start by introducing the random effects Bayesian hierarchical model (RE-BHM). We estimate inflation by constructing a linear model per country, while incorporating a prior that the coefficients of each linear model should be approximately similar across countries. For example, a positive shock in unemployment in The Netherlands has a negative impact on inflation in The Netherlands in the next period, and we expect other countries to have approximately the same effect with the same magnitude. This model allows each country to have their own coefficients but penalizes differing coefficients between countries. This can be interpreted as a form of regularisation when comparing this model to, for example, a standard linear regression per country, because this BHM has a prior structure that a standard linear regression does not have. The BHM considered in this paper is given by

$$y_{t,i} = \alpha_i + \beta_i x_{t-1,i} + \gamma_i w_{t-1} + \epsilon_{t,i}, \tag{1}$$

where α_i is a country-specific constant, $x_{t-1,i}$ are country specific covariates at time $t-1$ and w_{t-1} contains global covariates at time $t-1$. So, we model the inflation in the i-th country as a function of both country-specific and global variables. Next, because as illustrated in Section 2.2, one would expect similar coefficients across countries, we impose the following structure on the coefficients

$$\alpha_i \sim N(\alpha_0, \tau_\alpha^{-1}), \tag{2}$$
$$\beta_i \sim N(\beta_0, \tau_\beta^{-1}), \tag{3}$$
$$\gamma_i \sim N(\gamma_0, \tau_\gamma^{-1}), \tag{4}$$
$$\epsilon_{t,i} \sim N(0, \tau_\epsilon^{-1}), \tag{5}$$
$$\alpha_0 \sim N(0, \tau_{\alpha_0}^{-1}), \tag{6}$$
$$\beta_0 \sim N(0, \tau_{\beta_0}^{-1}), \tag{7}$$
$$\gamma_0 \sim N(0, \tau_{\gamma_0}^{-1}), \tag{8}$$

$$\tau_{\alpha_0}, \tau_{\beta_0}, \tau_{\gamma_0}, \tau_\alpha, \tau_\beta, \tau_\gamma, \tau_\epsilon \sim \Gamma(1,1).$$

For some simple BHMs, it is possible to analytically derive the posterior distribution $\theta \mid x, y$ by making use of e.g. conjugate priors. In this formulation, Equations (2)-(4) are the equations that incorporate the prior that the coefficients per country are centered around a global mean. Because the variances follow inverse gamma distributions, one can derive the posterior distribution analytically. An analytical derivation of a simplified random effects model can be found in Appendix B. However, for more complicated models, it is difficult, and sometimes impossible to analytically derive the posterior distribution. The key intuition that allows us to numerically estimate this distribution is that $P(\theta \mid x, y) = \frac{P(x,y|\theta)\cdot P(\theta)}{P(x,y)} \propto P(x, y \mid \theta) \cdot P(\theta)$. One can easily calculate the numerator, but not the denominator. However, when applying Metropolis-Hastings to obtain the posterior distribution, the denominator cancels out in the ratio of likelihoods. Metropolis-Hastings is a specific form of Markov Chain Monte Carlo (MCMC), which is a more general class of methods that can be used to numerically estimate the posterior distribution by generating a Markov chain of samples when direct sampling is impractical.

3.2 Covariance Priors

With the RE-BHM, we introduce a prior that coefficients should be close to each other when compared across countries. This is modeled by letting each coefficient follow a normal distribution which is centered around the same mean for every country. We can introduce an additional prior by letting coefficients follow a multivariate normal distribution with a non-diagonal covariance matrix. The purpose of this is to correlate coefficient values between countries. As an example, if you expect that the coefficients for The Netherlands and Germany are quite similar, you could set the corresponding elements in the corresponding covariance matrices to positive values to specify this prior for the BHM. We can estimate this type of BHM using MCMC because a closed-form solution of this type of model does not exist.

The second BHM we consider incorporates such a covariance prior. The covariance matrix we use as a prior for this model is the sample covariance matrix of inflations. Hence, for countries with similar inflation over time, we force the coefficient values to be closer to each other. This model is similar to the RE-BHM, in the sense that the linear predictive model per country stays the same. However, rather than sampling independent coefficients and intercepts from a normal distribution, we impose a prior on the covariance between coefficients across countries. More formally, let $\theta_{s,i}$ denote either a coefficient or intercept θ_s for country i, e.g., θ_{si} can denote either the intercept α_i of country i, or either coefficient β_i or γ_i of country i. Then, we impose a prior on β_{si} and β_{sj} for each $i, j \in \{1, \ldots, n\}$ where $i \neq j$:

$$\boldsymbol{\beta}_s \sim \mathcal{N}_n(\mathbf{0}, \hat{\Sigma}), \tag{9}$$

where $\hat{\Sigma}$ corresponds to the sample covariance matrix of historical inflation rates. This model is called the sample covariance Bayesian hierarchical model (COV-BHM). Note that we do not impose any correlation constraints within an individual country.

The third BHM we introduce is a distance based Bayesian hierarchical model (DB-BHM) and this model also incorporates a covariance prior. More specifically, we incorporate a prior that geospatially nearby countries have similar coefficients. We expect that the coefficients between The Netherlands and Belgium are likely more similar than the coefficients between The Netherlands and South Africa, for example. Instead of taking the sample covariance matrix of historical inflation, we consider the covariance matrix $\Sigma_{\text{Matérn kernel,dist}}$. This covariance matrix is generated by the Matérn kernel using the distance function $\text{dist}(\cdot)$, where $\text{dist}(i, j)$ denotes a distance function between country i and j. The Matérn kernel is an often used kernel to convert distances to covariances. We restrict ourselves to the geospatial distance as described in Section 2, but other distance functions could be interesting for future research. We should note that we restrict $\Sigma_{\text{Matérn kernel,dist}} = 0 \iff \text{dist}(i,j) = \infty, \quad \forall i,j$ where i and j are in different continents. The purpose of this is that countries in different continents could behave quite differently despite being geographically close. For example, Greece and Turkey are geographically close but it could be that the coefficients are drastically different across these countries.

3.3 Unobserved Component Stochastic Volatility Model

The unobserved component (UC) model, first introduced by Harvey (1990), splits a time series into underlying components. The version of this model we consider splits inflation into a trend component, $\tau_{t,i}$, and an irregular component, $\eta_{t,i}$, and lets the variances of the disturbances evolve randomly over time. This is the unobserved component stochastic volatility model (UC-SV), as introduced by Stock and Watson (2007). The model is formally defined as

$$y_{t,i} = \tau_{t,i} + \eta_{t,i}, \qquad (10) \qquad\qquad \eta_{t,i} \sim N(0, \sigma^2_{\eta,t,i}), \qquad (14)$$
$$\tau_{t,i} = \tau_{t-1,i} + \epsilon_{t,i}, \qquad (11) \qquad\qquad \epsilon_{t,i} \sim N(0, \sigma^2_{\epsilon,t,i}), \qquad (15)$$
$$\ln(\sigma^2_{\eta,t,i}) = \ln(\sigma^2_{\eta,t-1,i}) + v_{\eta,t,i}, \qquad (12) \qquad\qquad v_{\eta,t,i} \sim N(0, \gamma_i), \qquad (16)$$
$$\ln(\sigma^2_{\epsilon,t,i}) = \ln(\sigma^2_{\epsilon,t-1,i}) + v_{\epsilon,t,i}, \qquad (13) \qquad\qquad v_{\epsilon,t,i} \sim N(0, \gamma_i). \qquad (17)$$

In Equation (10), we split inflation $y_{t,i}$ into a stochastic trend component $\tau_{t,i}$ and a transitory component $\eta_{t,i}$. In Equation (11), we define the stochastic trend component of inflation as a random walk. We then define the variances of the disturbances to be time-varying in Equations (12) and (13). More specifically, the natural logarithms of both variances follow a random walk. Making the variances of the disturbance terms $\eta_{t,i}$ and $\epsilon_{t,i}$ time-invariant, leads back to the UC model with a stochastic trend, which is equivalent to the IMA(1,1) model, as noted by Stock and Watson (2007). The UC-SV model is thus a generalization of the IMA(1,1) model, which is why we do not consider the latter separately. All error terms are independently distributed and follow a normal distribution as shown in Equations (14)-(17). We remark that the UC-SV model considers each country to be independent: there are no common components between countries, nor are errors correlated.

One disadvantage of this model is that it does not consider seasonality as a potential effect. This is understandable because seasonality is not apparent if we consider inflation up until approximately 2000, which is later than the publishing date of Harvey (1990). However, if we consider inflation after approximately 2000, we observe that yearly seasonality starts to play a more prominent role across multiple countries, as we have noted before in Section 2.2. To account for this, we introduce the unobserved component stochastic volatility model with stochastic seasonality (UC-SV-SS). The model is formally defined as

$$y_{t,i} = \tau_{t,i} + \delta_{t,i,1} \cdot seas_1(t) + \ldots + \delta_{t,i,4} \cdot seas_4(t) + \eta_{t,i}, \tag{18}$$

$$\tau_{t,i} = \tau_{t-1,i} + \epsilon_{t,i}, \tag{19}$$

$$\delta_{t,i,j} = \delta_{t-1,i,j} + seas_j(t) \cdot \xi_{t,i}, \quad j = 1, \ldots, 4, \tag{20}$$

$$ln(\sigma^2_{\eta,t,i}) = ln(\sigma^2_{\eta,t-1,i}) + v_{\eta,t,i}, \tag{21}$$

$$ln(\sigma^2_{\epsilon,t,i}) = ln(\sigma^2_{\epsilon,t-1,i}) + v_{\epsilon,t,i}. \tag{22}$$

In this model, j is used to denote the four quarters of the year. We let $seas_j(t) = 1$ if at time t we are in season j and 0 otherwise. $\xi_{t,i}$ follows a normal distribution with zero mean and fixed variance equal to $\Theta = 0.002$ which is a hyperparameter. A robustness check can be found in Appendix F which shows that the model is insensitive to the exact value of Θ. The error terms still follow the same distributions as in the UC-SV model, i.e., Equations (14)-(17) still apply. In Equation (18) we split inflation per country into a trend component, seasonal components and an irregular component. The trend component follows the same random walk as before. The seasonal components also follow a stochastic process similar to a random walk, as displayed in Equation (20). However, we constrain the sum of seasonal components to be zero by subtracting the mean seasonal component at every timestep. The purpose of this is to ensure the trend component captures all of the stochastic trend.

Finally, we introduce another novel model that is based on the original UC-SV model by Stock and Watson (2007). In the two previous models we considered, the unobserved components that underpin the inflation process were assumed to be independent. This is a special case of the multivariate UC-SV-SS model (M-UC-SV-SS) that we introduce, where the innovations in the unobserved component stochastic processes are assumed to be correlated.

Consider the UC-SV-SS model as described in Equations (18)-(22). Let $\boldsymbol{\tau}_t = (\tau_{t,1}, \ldots, \tau_{t,n})$, $\boldsymbol{\eta}_t = (\eta_{t,1}, \ldots, \eta_{t,n})$ and $\boldsymbol{\xi}_t = (\xi_{t,1}, \ldots, \xi_{t,n})$. The UC-SV-SS model can be interpreted as the following:

$$\begin{aligned}
\boldsymbol{\tau}_t &\sim \mathcal{N}_n(\boldsymbol{\tau}_{t-1}, \mathrm{diag}(\sigma^2_{\epsilon,t,1}, \ldots, \sigma^2_{\epsilon,t,n})), \\
\boldsymbol{\eta}_t &\sim \mathcal{N}_n(\mathbf{0}, \mathrm{diag}(\sigma^2_{\eta,t,1}, \ldots, \sigma^2_{\eta,t,n})), \\
\boldsymbol{\xi}_t &\sim \mathcal{N}_n(\mathbf{0}, \mathrm{diag}(\sigma^2_{\xi,t,1}, \ldots, \sigma^2_{\xi,t,n})),
\end{aligned}$$

where

$$\mathrm{diag}(\sigma_1^2, \ldots, \sigma_n^2) = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

is a diagonal matrix with the arguments on the diagonal. Because the covariance matrices are diagonal, each country has independent innovations in the unobserved components.

To generalize this while retaining stochastic volatility, we introduce the covariance matrix $\Sigma_{\boldsymbol{\rho}}(\sigma_1^2, \ldots, \sigma_n^2) = D\boldsymbol{\rho}D$, where $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ such that $\left(\Sigma_{\boldsymbol{\rho}}(\sigma_1^2, \ldots, \sigma_n^2)\right)_{ij} = \boldsymbol{\rho}_{ij}\sigma_i\sigma_j$. With this, we essentially decomposed the time-varying covariance matrix into time-varying individual components and a fixed correlation structure. This fixed correlation matrix is our prior, and the UC-SV-SS model corresponds to the case where $\boldsymbol{\rho} = I_n$. We determine $\boldsymbol{\rho}$ by applying the Matérn kernel function on the geospatial distance as described in Section 2 between each pair of countries. This results in a correlation matrix where each entry is positive and countries that are geospatially closer obtain a higher correlation. In the M-UC-SV-SS model, we then have

$$
\begin{aligned}
\boldsymbol{\tau}_t &\sim \mathcal{N}_n(\boldsymbol{\tau}_{t-1}, \Sigma_{\boldsymbol{\rho}}(\sigma^2_{\epsilon,t,1}, \ldots, \sigma^2_{\epsilon,t,n})), \\
\boldsymbol{\eta}_t &\sim \mathcal{N}_n(\mathbf{0}, \Sigma_{\boldsymbol{\rho}}(\sigma^2_{\eta,t,1}, \ldots, \sigma^2_{\eta,t,n})), \\
\boldsymbol{\xi}_t &\sim \mathcal{N}_n(\mathbf{0}, \Sigma_{\boldsymbol{\rho}}(\sigma^2_{\xi,t,1}, \ldots, \sigma^2_{\xi,t,n})).
\end{aligned}
$$

To estimate (M-)UC-SV(-SS) models, we use particle filters, which we describe in Section 3.4.

3.4 Particle Filters

Particle filtering, also known as sequential Monte Carlo methods, is a set of algorithms used for estimating the state of a hidden Markov model where the hidden state evolves over time. In the context of an (M-)UC-SV(-SS) model, the 'hidden' part refers to the latent trend, the latent stochastic volatility, and the latent stochastic seasonality, which are not directly observable but influence the observed variable: inflation.

The basic idea behind particle filtering is to represent the posterior distribution of the state variables by considering the empirical distribution of a set of particles. Each particle is a possible state, and each one has an associated weight that represents its importance or likelihood. The filter updates these particles and weights over time as new observations become available, allowing for real-time state estimation. Subsequently, these estimates can be used to construct forecasts.

A potential pitfall of a naive sequential importance sampling (SIS) particle filter is that the particles might stray away from the important parts of the distribution. This is referred to as the degeneracy problem, where the weights of the particles become smaller over time. One way to reduce the degeneracy of the weights is to introduce a resampling step, which leads to the sequential importance sampling with resampling (SISR) particle filter. This is the particle filter we use for the (M-)UC-SV(-SS) models.

Particle filters are a generalisation of Kalman filters and extended Kalman filters. However, ordinary Kalman filters assume that the transformations the hidden states undergo are linear and that the noise is Gaussian. In this case, we have non-Gaussian noise due to the stochastic volatility component, which is an effect that is not possible to capture with (extended) Kalman filters.

For the UC-SV and UC-SV-SS model, we use 10,000 particles, as in our experiments this was more than sufficient to obtain convergence. For the M-UC-SV-SS model, we use 100,008 particles, which is divisible by the amount of CPU cores in our PC to make multiprocessing easier. Note that the M-UC-SV-SS model with this amount of particles requires up to 128GB of RAM for the countries we considered.

3.5 Atkeson–Ohanian Random Walk Model

The first benchmark model we consider is the Atkeson–Ohanian random walk model (AO-RW), introduced by Atkeson and Ohanian (2001) and later extended by Stock and Watson (2007). This model forecasts inflation as the average rate of inflation over the previous four quarters. It is included, because Atkeson and Ohanian (2001) show that this naive model often performs better than the more sophisticated models, available at that time, based on the Philips curve.

The AO-RW model (Stock & Watson, 2007) is given by

$$y_{t,i} = \frac{1}{4}(y_{t-1,i} + y_{t-2,i} + y_{t-3,i} + y_{t-4,i}) + \epsilon_{t,i}, \tag{23}$$

where $\epsilon_{t,i}$ is a random variable with zero mean and constant variance.

3.6 ARMA(X) Model

The autoregressive moving average model with exogenous variables (ARMAX) for the inflation of the i-th country is given by

$$y_{t,i} = c_i + \Sigma_{p=1}^{P} \alpha_{p,i} y_{t-p,i} + \Sigma_{q=1}^{Q} \theta_{q,i} \epsilon_{t-q,i} + \Sigma_{m=1}^{M} \beta_{m,i} x_{m,t-1,i} + \epsilon_{t,i}, \tag{24}$$

where the regressors represent a constant, the AR terms, the MA terms, the exogenous regressors and an error term respectively.

This model is estimated for the individual inflation of each country, with P and Q between 0 and 2 using maximum likelihood. The optimal model for each country is then determined based on the AIC. We consider this model because it is a flexible time series model that contains different models as special cases. We apply this model both with and without exogenous regressors to investigate the benefit of adding covariates. In our analysis, we also investigate the performance of an AR(X) model, using the same strategy as for the ARMA(X) but restricting Q to be zero. Note that to enable a comparison of the coefficients of the AR model with the BHM, the features and target are normalised to have a mean of 0 and a standard deviation of 1 when estimating the AR(MA)(X) models, similar to what we do when estimating the BHMs.

3.7 PCA-VAR Model

Next, we built upon the results of Ciccarelli and Mojon (2010), who argue that a common global factor can explain 70% of the variance of inflation in the OECD countries. That is why we also consider a model that can identify the common factors that drive the variance of inflation. This is done by applying principal component analysis (PCA) to the sample covariance matrix of inflation which is determined from the train data. Next, we apply a vector autoregressive model of order one (VAR(1)) to the first k principal components. The VAR(1) model is given by

$$f_t = c + A_1 f_{t-1} + \epsilon_t, \tag{25}$$

where $f_t = (f_{t,1}, ..., f_{t,k})'$, $c = (c_1, ..., c_k)'$, $\epsilon_t = (\epsilon_{t,1}, ..., \epsilon_{t,k})'$ and f_t denotes the principal components at time t. The number of principal components is chosen based on a scree plot. This model is estimated using system OLS and is used to obtain forecasts of principal components. These forecasts are then transformed back to the original data space using the eigenvectors obtained from PCA during training to obtain the desired inflation forecasts. This approach can lead to more parsimonious forecasts compared to a normal VAR model, because of the dimension reduction. This method is applied directly to normalized data with a mean value of 0 and a standard deviation of 1 to avoid principal components that load too much on inflation of a country that is extremely volatile.

Summarising, we introduced three BHMs: the RE-BHM, which does not impose restrictions on covariances across countries, and two covariance prior based models, namely the COV-BHM and the DB-BHM. Additionally, we extend the univariate UC-SV model with stochastic seasonality and we extend this model to the multivariate case. Further, we outlined a series of simpler frequentist approaches. The multivariate extensions allow us to determine whether viewing inflation as a global phenomenon occurring similarly across countries, instead of on national level, improves forecasting performance.

3.8 Evaluation Metrics

To compare the different models introduced in the previous subsections, we look at multiple evaluation metrics. The next subsections outline the evaluation metrics we use to compare the predictive power of the different models.

3.8.1 Mean Squared Prediction Error

A first measure to compare the accuracy of a univariate forecast is the mean squared prediction error (MSPE). The MSPE of the j-th model, for the inflation of the i-th country, is given by

$$MSPE_{i,j} = \frac{\Sigma_{t=1}^{T} (\hat{y}_{t,i,j} - y_{t,i})^2}{T}. \tag{26}$$

A low MSPE indicates that the model's predictions are close to the actual inflation values in the dataset, which implies good accuracy. For the frequentist models, we obtain point forecasts, such that $\hat{y}_{t,i,j}$ is directly obtained. For the models that produce density forecasts, namely (M-)UC-SV(-SS) models and BHMs, we take the median to obtain point forecasts.

3.8.2 Mean Absolute Error

Because MSPE is highly sensitive to a few large errors, periods of high inflation or individual countries leading to a few large forecasting errors might distort the overview of a model's performance based on MSPE. Therefore, similarly to the MSPE, the mean absolute error (MAE) is also used to evaluate the forecasting performance of the models under consideration. The MAE of the j-th model, for the inflation of the i-th country, is given by

$$MAE_{i,j} = \frac{\Sigma_{t=1}^{T} |\hat{y}_{t,i,j} - y_{t,i}|}{T}. \quad (27)$$

Similarly to the MSPE, low values for the MAE indicate that model predictions are close to the observed inflation values.

3.8.3 R^2

As a third way of evaluating the predictive performance of the model, the R^2 of actual observed inflation versus the out-of-sample predicted inflation of country i by model j is calculated. The formula used is as follows:

$$R^2 = 1 - \frac{\Sigma_{t=1}^{T}(y_{t,i} - \hat{y}_{t,i,j})^2}{\Sigma_{t=1}^{T}(y_{t,i} - \bar{y}_i)^2}, \quad (28)$$

where $y_{t,i}$ is the observed inflation of country i at period t, $\hat{y}_{t,i,j}$ is the forecast of inflation of model j for country i at time t, and $\bar{y}_i$ is the average observed inflation in country i. Note that the highest value of R^2 possible is 1, but there is no lower bound because these are out-of-sample predictions. Namely, if $R^2 < 0$, it indicates that the predictions are worse than predicting the sample mean. In general, higher values indicate better forecasting performance than lower values.

3.8.4 Mincer-Zarnowitz Regression

A fourth way of evaluating the point forecasts of a single model is by using the Mincer-Zarnowitz (MZ) regression, as introduced by Mincer and Zarnowitz (1969). The regression is given by

$$y_{t,i} = \alpha_i + \beta_i \hat{y}_{t,i,j} + \epsilon_{t,i}, \quad (29)$$

where $y_{t,i}$ is the true inflation and $\hat{y}_{t,i,j}$ is the forecasted inflation of country i at time t for model j. We test the joint hypothesis of $\alpha_i = 0, \beta_i = 1$ using a Wald test, where rejecting the null hypothesis indicates a biased and inefficient forecast. The Wald statistic follows a $F(2, T-2)$ distribution where T denotes the number of timesteps in the MZ regression. Note that the MZ regression is performed for each model separately.

4 Results

As introduced in Section 2, data on both inflation and covariates is available between Q1 1977 and Q1 2023 for a total of twenty countries. In this paper, inflation is forecasted one-quarter ahead using an iterative expanding window approach. In the first iteration, we estimate the models using data up until Q4 1999 to forecast inflation one-quarter ahead for the next four years, i.e., until Q4 2003. This means that we use all available historical data to forecast inflation, but not to estimate the model. For example, to predict the inflation in Q4 2003, we use data up until Q3 2003, but a model that is estimated on data until Q4 1999. In the next iteration of this approach, the models are re-estimated on data until Q4 2003 and used to forecast inflation until Q4 2007. In reality, models are often re-fitted after a period of time to improve forecasting performance, so this expanding window approach more closely mimics reality than taking a fixed 20-year test set without re-fitting. These iterations are performed until we exhaust all the data.

For the evaluation of the PCA-VAR model, Figure 10 in Appendix C shows a scree plot of the fraction of variance explained per principal component. Using the elbow method to define the number of principal components, we select three principal components when estimating the model.

4.1 Point Forecasts

This section discusses the results of the point forecasts of the different models by first looking at the model-specific results, followed by the country-specific performance as measured by MSPE.

4.1.1 Model Specific Performance

In order to evaluate the performance of the different models introduced in Section 3, the out-of-sample MSPE, MAE, R^2, and Mincer-Zarnowitz statistics, as discussed in Section 3.8, are presented in Table 1. All statistics shown in Table 1 are averaged across countries and the MSPE and MAE are reported relative to the benchmark UC-SV model make these metrics more interpretable. The p-value of the Wald test is given in brackets in the last column.

Table 1 clearly shows that the UC-SV-SS model has the lowest MAE and MSPE, as well as the highest R^2 value. Additionally, the Mincer-Zarnowitz R^2 is the highest, which indicates that these predictions are closest in line with the actual values. As this model outperforms even the benchmark UC-SV model, we conclude that the stochastic seasonality component of the UC-SV-SS model adds forecasting power in this sample. This is interesting to see, as this seasonality component seems to be underexplored in existing literature. Therefore, we argue that existing forecasting models can be improved by adding a seasonal component in the model, thereby capturing the seasonality in inflation data, as evidenced in Section 2.2.

Table 1: Average MSPE (normalized with UC-SV = 1), MAE (normalized with UC-SV = 1), R^2, and Mincer-Zarnowitz (MZ) statistics (p-value in between brackets), one-quarter ahead, out-of-sample, 2000-2023

	MSPE	MAE	R^2	MZ-intercept	MZ-slope	MZ-R^2	MZ-Wald-Stat
UC-SV-SS	0.7837	0.8480	0.3530	0.0010	0.8880	0.3600	9.96(0.00)
PCA-VAR	0.8456	0.9636	0.2360	0.0020	0.7330	0.2720	41.519(0.0)
AO-RW	0.9263	0.9693	0.2380	0.0020	0.7550	0.2670	33.09(0.00)
UC-SV	1.0000	1.0000	0.1740	0.0010	0.7980	0.1870	13.55(0.00)
AR	1.0007	1.0004	0.2010	0.0010	0.7910	0.2170	14.78(0.00)
RE-BHM	1.2641	1.1217	-0.0480	0.0030	0.4030	0.0380	73.75(0.00)
COV-BHM	1.3038	1.1390	-0.0810	0.0040	0.3550	0.0350	99.00(0.00)
DB-BHM	1.3061	1.1396	-0.0830	0.0040	0.3590	0.0370	102.76(0.00)
RE-BHM-X	1.3069	1.1299	-0.0840	0.0040	0.3870	0.0430	109.25(0.00)
DB-BHM-X	1.5956	1.2729	-0.3230	0.0050	0.2410	0.0340	304.32(0.00)
COV-BHM-X	1.6139	1.2742	-0.3380	0.0050	0.2380	0.0330	316.22(0.00)
ARMA	3.9908	1.3409	-2.1880	0.0050	0.1420	0.0620	1727.32(0.00)
M-UC-SV-SS	4.5571	2.1596	-2.7650	0.0060	0.0600	0.0110	2328.06(0.00)
ARX	13.8756	2.2090	-10.0840	0.0060	0.0600	0.0400	7601.77(0.00)
ARMAX	369.8425	9.0717	-294.4480	0.0060	0.0060	0.0120	214774.19(0.00)

Another model that outperforms the UC-SV model is the PCA-VAR model with three principal components. This supports the argument that inflation across countries is similar and that there exist common global factors that explain inflation, as argued by Ciccarelli and Mojon (2010). Moreover, it shows that forecasting performance can be improved by dimension reduction techniques.

Looking further, we see that the baseline UC-SV model performs well, based on the presented metrics. However, it does not beat the AO-RW model in this sample. It turns out to be difficult to beat a relatively simple, slightly adjusted, random walk model with more sophisticated models, which was already shown by Stock and Watson (2007).

The M-UC-SV-SS model does not perform as well as its univariate counterpart. One possible reason for this is that the dimensionality of the particle filter becomes too large in a multivariate setting to accurately capture the dynamics of the joint density of the unobserved components. If there are 7 dimensions for a single UC-SV-SS model, an M-UC-SV-SS model has $7N$ dimensions, with N being the number of countries. This high-dimensional space is difficult to fill with the amount of points that can fit on modern computers. To test this theory, we also applied the M-UC-SV-SS model on a subset of countries: The Netherlands, Germany, the United States and Belgium. This leads to a lower dimensionality, which should make the particle filter more accurate. However, this model still underperformed the UC-SV-SS model for these specific countries in all relevant metrics, as can be seen in Appendix E. This could suggest that a geospatial correlation prior is not a good prior of unobserved component correlation in the context of forecasting inflation, but more research is required to definitively show this.

Furthermore, the proposed Bayesian hierarchical models, the RE-BHM, COV-BHM, and

DB-BHM, underperform the AR model, all UC-SV(-SS) models, as well as the AO-RW model. We conclude from this that using Bayesian priors on a linear data generating process for inflation does not improve forecasting performance. Potential causes could include the presence of countries with very high inflation rates due to detrimental monetary and fiscal policy, such as Turkey. Because of the structure of the BHMs, the fitted models cannot cope well with the large differences in inflation rates between countries with 'sound' monetary policy and countries with exceptionally high inflation rates, as the coefficients of these models are imposed to be relatively similar.

However, when looking at the BHMs, we clearly see that the RE-BHM outperforms the DB-BHM and COV-BHM. We conclude from this that the approach of using geospatial distances as a prior of the correlation between inflation processes of different countries does not improve forecasting performance, nor does imposing a sample covariance prior improve performance. Although this seems surprising at first, as we expect inflation rates across geographically close countries to be correlated more closely than inflation rates between countries that are geographically far apart, issues could be related to the used proxy for distance, or the imposed prior could be too strong. The COV-BHM performs almost identical to the DB-BHM and does not outperform the RE-BHM. This indicates that using the sample covariance matrix to proxy the similarity between inflation in countries does not add predictive power for one-quarter ahead forecasts.

The worst-performing models are the ARX and ARMA(X) models. Modelling inflation as an autoregressive or autoregressive moving average model with exogenous variables results in remarkably poor forecasts. Even though the AR model has a relatively good performance in this sample, there is a simpler, even better performing alternative (the AO-RW model). Therefore, even though these models are relatively simple, they should be avoided in this setting. An explanation for this poor performance of the ARMAX model could be overfitting. The ARMAX model has many parameters to fit without any form of regularisation like the Bayesian models have. This can, and in this case does, result in poor out-of-sample performance. It is also interesting to note that adding exogenous variables to the data reduces model performance, which can be seen from the fact that all 'X' models have worse forecasting performance than their counterparts without exogenous variables. Note that this difference is especially large for the AR and ARMA models. Although we initially expected the exogenous variables to add predictive power to the models, there are explanations for this reduction. These include overfitting on the exogenous variables, as adding a lot of extra parameters to a model might very well hurt the out-of-sample forecasting performance. Also, there could be an overall lack of predictive power of the exogenous variables with the target. However, these possible explanations are not investigated further, as this is not the objective of this paper.

4.1.2 Country-Specific MSPE

This section discusses the performance of the proposed models further, by evaluating the MSPE per country. Figure 6 presents the MSPE per country for all discussed models, except ARMAX and ARX. It presents the results for three countries: Turkey, France, and Spain. These countries have been chosen because the results for them illustrate how different countries have very different model performances. The ARMAX and ARX models are excluded based on their relatively high MSPE. Including these models would lead to a difficult-to-interpret bar plot, as the ARMAX (or ARX) results would dominate the plot, leading to bars seemingly close to zero for all other models. However, a table of the full results, with relative MSPEs for all models for all countries, is presented in Appendix D.

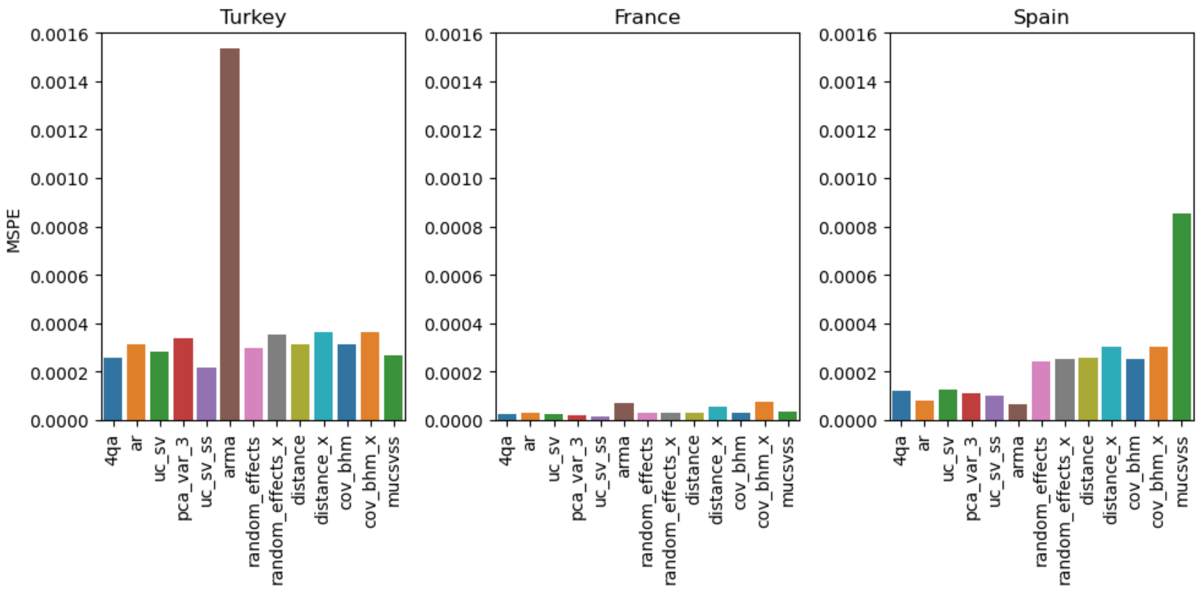


Figure 6: MSPE of the different models for Turkey, France and Spain, one-quarter ahead, 2000-2023

Looking at Figure 6, the plots show a large difference in the performance of the models between countries. The models perform remarkably well for France, and substantially worse for Turkey, with Spain's model performance being somewhere in between. This result is in line with expectations, as the Turkish economy has suffered high inflation at both the beginning and end of the out-of-sample period. With higher inflation comes higher variance in inflation, which is a reason why models would have more difficulty in predicting inflation in such countries. Furthermore, we see that for Turkey the ARMA model is by far the worst performing model of the models displayed here. This model has issues with the high volatility in inflation in Turkey in our sample. For France, there is much less variation in the model performance, indicating that there are countries for which all models perform relatively well. However, the best performing model is clearly the UC-SV-SS model (see also Table 9 in Appendix D). For Spain, Figure 6 clearly shows that the M-UC-SV-SS model performs substantially worse than all other models, and the ARMA model is actually among the best performing models. This is also the case for other countries, such as Italy or the United States. This shows that the M-UC-SV-SS model has trouble with forecasting inflation for countries where other models perform well. This again indicates that having a prior based on geospatial distance does not work in this scenario.

4.2 Coefficient Comparison

Next, this section discusses the coefficients of four different models. When comparing the coefficients of an AR model, RE-BHM, DB-BHM, and COV-BHM, it is important to note that the coefficients corresponding to the constant, as well as the first and second lag, can be compared across the models. This is possible, as these models are all linear models, where the DB-BHM, RE-BHM, and COV-BHM impose restricting priors on the values the coefficients can take. Therefore, we expect the coefficients of the AR model to differ more among the countries than the corresponding coefficients of the RE-BHM, DB-BHM, and COV-BHM. Figure 7 shows histograms of the common coefficients of the models, together with a kernel density estimate of the distribution. The figures clearly show that for the AR model, the coefficients of the constant and the second lag have a larger dispersion than the corresponding coefficients of the RE-BHM, DB-BHM, and COV-BHM. This indicates that these coefficients thus differ more among countries for the AR model, than for the RE-BHM, DB-BHM, and COV-BHM. However, this result is less clear for the first lag, which seems to have the same magnitude of dispersion for all four models under consideration.

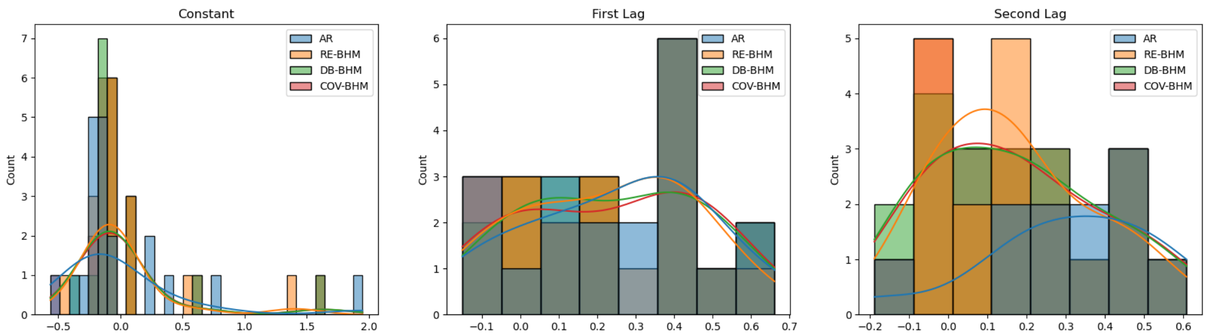
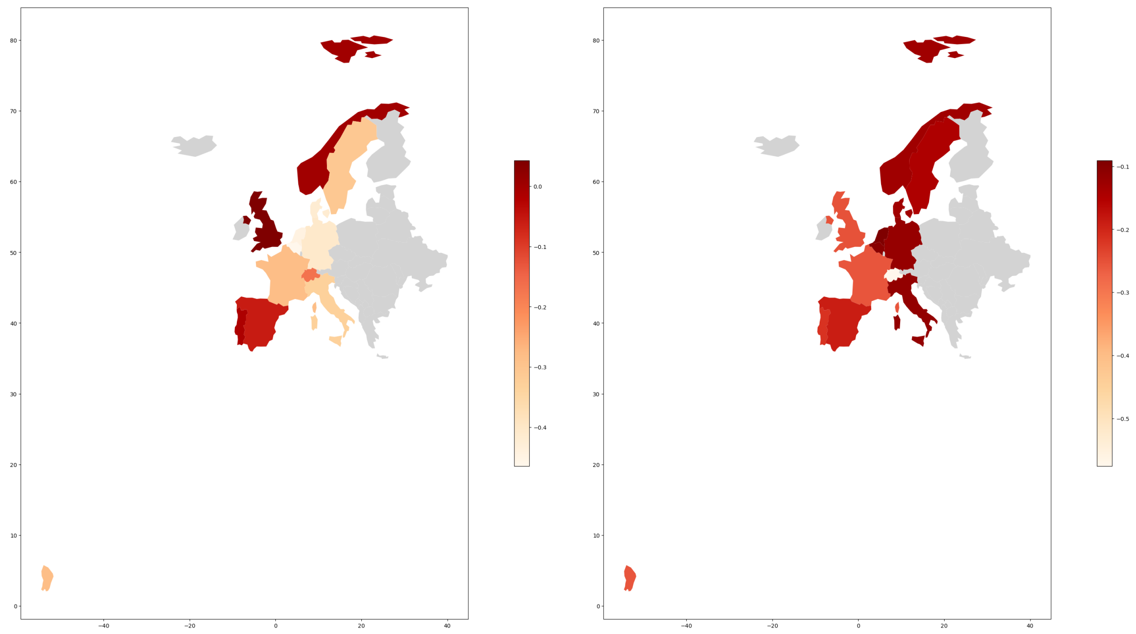


Figure 7: Coefficients of the AR model, RE-BHM, DB-BHM, and COV-BHM, 1977-2000

Next, we look at the coefficients on a country level of both an AR model and DB-BHM. Note that, as explained before, the DB-BHM is a type of linear AR model, with restricting priors on the coefficients. We can therefore compare the coefficients across the models, and expect the coefficients of the DB-BHM to be closer across countries, than the coefficients of the AR model. Furthermore, note that besides economic ties, cultural ties can be expected to play a role in inflation levels as well. Cultural differences might lead to differing inflation levels if these cultural differences lead to differing tolerance levels for inflation. Some countries such as Germany have taken a harsh stance on inflation, because of hyperinflation in the 1920s that is still remembered. Other countries, such as Spain and Italy, have historically taken on a more lenient stance, and are thus 'culturally' more tolerant to higher inflation levels.

Figure 8 shows the intercepts of the AR models, as well as the DB-BHM, for the European countries in our sample. The figure shows that there is a difference between the coefficients of these models.

The coefficients of the AR model show a lot of dissimilarity across countries. This shows that this model is quite flexible in the way it allows every country individually to have optimal coefficients, unconstrained by any priors. What is especially remarkable, is the divergence between the coefficients of Norway and Sweden. These countries are expected to have similar inflation levels, due to both their close economic ties, as well as their cultural similarities. Therefore, the economic shocks to the countries' economies would be expected to be similar, as well as the political and monetary reaction to these shocks. However, fitting the best AR



(a) AR Model (b) Distance Based Model

Figure 8: Intercepts of the models per country, 1977-2000

model for these countries clearly gives different values for the intercept. This indicates further that even though countries might be culturally and economically close, inflation might still differ substantially from time to time. This further supports the previous conclusion that it is difficult to use distances between countries in Bayesian hierarchical models to forecast inflation.

Looking further at Figure 8, it also shows the intercepts of the DB-BHM across countries. Because of the prior introduced in this model, the coefficients of the European countries are very similar. This can be seen from the similar colours in the graph for the Germanic and Nordic countries, as well as Italy. Also, we see similar colours for France, The United Kingdom, Portugal, and to a lesser extent, Spain. Maybe some readers might have expected Italy to share more characteristics with countries such as France. However, the large German minority in the northern regions of the country, and its close proximity to the Germanic countries, do make the country similar to Germany as well. Therefore, these results are not unexpected.

What might be more unexpected is the close similarity between the UK and France. If we look at the geopolitical proximity of countries in the last few decades, France and the UK have been strong allies, on both a military and economic level. Figure 8 shows that the intercepts are indeed almost equal. Looking back at the AR model intercepts, the figure shows that the intercepts of these countries differ quite a bit. This demonstrates that the DB-BHM is too restrictive when estimating the coefficients of the model, which could contribute to its relatively poor performance, compared to a UC-SV or AO-RW model.

4.3 Density Forecasts

Figure 9 shows the different density and point forecasts of the models for the United States for two different points in time: Q4 2022, and Q1 2006. These two periods are chosen because they correspond to a high and low inflationary period respectively. This allows to show how the forecasts compare to each other when having different values for inflation. For the sake of clarity, only the densities of the UC-SV, UC-SV-SS, DB-BHM, RE-BHM, and COV-BHM are displayed. We have chosen these models as they are the best-performing models for which we can obtain the densities. For the same reason, we have chosen to only display the point forecasts of the AO-RW, PCA-VAR and AR models, as well as the actual value of inflation (note that in Q4 2022, the AR and PCA-VAR forecast are almost the same and thus hard to distinguish). The graph clearly demonstrates that there is quite some difference between the density forecasts of the different models. This further illustrates the difference in performance of the models. However, the graph also shows that the UC-SV and UC-SV-SS models have quite similar forecasts, as do the RE-BHM, DB-BHM, and COV-BHM. This is to be expected, as the UC-SV-SS model is a generalisation of the UC-SV model, and the RE-BHM, DB-BHM, and COV-BHM impose somewhat similar priors on the coefficients. Note, however, that even though in these instances the DB-BHM, RE-BHM, and COV-BHM predict almost equal densities, this is not always the case in the sample. As can be inferred from the results in Table 1, the three models do not always produce densities and medians that are similar.

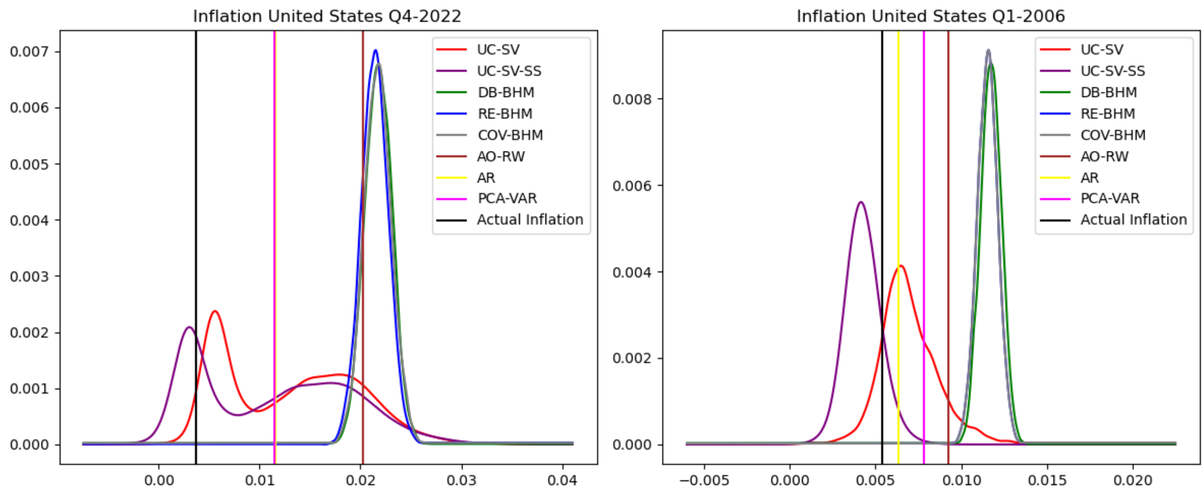


Figure 9: Density and point forecasts

We furthermore see that in these cases, the UC-SV(-SS) model performs better than the other density-producing models, which is in line with the conclusions from Table 1. The graphs also show that the AR model outperforms the AO-RW model for these two quarters. However, as is concluded from Table 1, even though this is the case in the quarters presented here, overall, the AR model underperforms the AO-RW model.

Moreover, we see that in periods with relatively stable inflation (Q1 2006) the forecasts produced are relatively similar, which can be seen from the closely grouped densities and point forecasts. When looking at a high-inflationary period (Q4 2022), the figure shows that the densities and point forecasts deviate significantly more (notice that the x-axes of the figures differ in values). This is expected, as inflation is typically less stable in high inflationary periods. Therefore, depending on the weights the models put on previous observations, the forecasts differ more among the models.

4.4 Robustness Checks

Next, as a robustness check, we investigate the performance of the models when using core CPI rather than headline CPI. The difference between these two inflation measures is that core CPI does not include food and energy prices when calculating inflation. This could lead to a potentially worse performance of the UC-SV-SS model since especially these prices are known to be seasonal. The data on core CPI comes from the same dataset as headline CPI. For this inflation measure, data is available for both inflation and the usual covariates between Q1 1977 and Q1 2023 for twelve countries. These countries are Canada, Denmark, France, Germany, Italy, Japan, Portugal, Sweden, Switzerland, The Netherlands, The United Kingdom, and The United States. Note that this is a subset of the twenty countries we have for headline CPI. So, we can not directly compare the results of the evaluation metrics in this section and in Section 4.1. However, we can still investigate the predictive power of the models when using core CPI.

Table 2: Average MSPE (normalized with UC-SV = 1), MAE (normalized with UC-SV = 1), R^2, and Mincer-Zarnowitz statistics (p-value in between brackets) for core inflation, one-quarter ahead, out-of-sample, 2000-2023

	MSPE	MAE	R^2	MZ-intercept	MZ-slope	MZ-R^2	MZ-Wald-Stat
UC-SV-SS	0.4878	0.6588	0.5148	0.0005	0.9602	0.5193	4.64(0.01)
ARMA	0.8079	0.8915	0.2153	0.0013	0.6536	0.2993	51.12(0.00)
AO-RW	0.8266	0.9186	0.1737	0.0008	0.8315	0.1826	5.46(0.00)
PCA-VAR	0.8642	0.9624	0.1386	0.0022	0.6960	0.2582	81.42(0.00)
AR	0.8785	0.9440	0.1467	0.0010	0.6588	0.2052	31.37(0.00)
UC-SV	1.0000	1.0000	0.0053	0.0020	0.5256	0.0368	16.05(0.00)
RE-BHM	1.6213	1.2310	-0.6095	0.0059	-0.4065	0.0520	339.43(0.00)
RE-BHM-X	1.6951	1.2588	-0.6827	0.0046	-0.2301	0.0244	352.58(0.00)
COV-BHM	1.7047	1.2684	-0.6961	0.0058	-0.3910	0.0569	388.40(0.00)
DB-BHM	1.7121	1.2715	-0.6996	0.0058	-0.3930	0.0578	391.10(0.00)
ARX	1.8876	1.3878	-0.8334	0.0031	0.1936	0.0504	396.48(0.00)
DB-BHM-X	2.0430	1.3929	-1.0281	0.0045	-0.1983	0.0285	529.07(0.00)
COV-BHM-X	2.0760	1.4083	-1.0656	0.0045	-0.1944	0.0283	547.73(0.00)
M-UC-SV-SS	2.8173	1.7702	-1.8023	0.0038	0.0308	0.0018	887.46(0.00)
ARMAX	53.6276	4.5978	-51.0891	0.0037	0.0039	0.0008	21780.78(0.00)

Table 2 presents the same evaluation metrics as Table 1, but now using core CPI as a measure for inflation. The table shows that the UC-SV-SS model remains to be the best model when using core CPI. The out-of-sample performance of this model is considerably better than other models based on the high R^2 and the relatively low Mincer-Zarnowitz Wald test statistic. This indicates that seasonal patterns remain in inflation data, even when excluding food and energy prices. This supports the claim of Lis and Porqueddu (2018), who show that even when excluding food and energy prices, there is still seasonality to be found in the HICP inflation in the euro area. Although we use core CPI rather than HICP and we do not only consider the euro area, we can draw a similar conclusion about seasonal effects in inflation measures that exclude food and energy prices.

For all models, the p-value of the Wald test is below the 5% significance level, which indicates that there is some bias and/or inefficiency in the forecasts of the models.

In general, the performance of the models seems to be somewhat similar to the performance when using headline CPI. First of all, adding covariates always reduces model performance. Second, the order of the models based on MSPE is fairly similar to that in Table 1. The most striking difference is the increase in performance of the ARMA model. However, this can be explained by the absence of Turkey in this data set. As we concluded, based on Figure 6, the ARMA model performs very badly for Turkey. This could indicate that the bad relative performance of the ARMA model in Table 1 is driven by this outlier. Thus, removing Turkey from the dataset should increase the performance of the ARMA model, which supports the result in Table 2.

Table 3: Average MSPE (normalized with UC-SV = 1), MAE (normalized with UC-SV = 1), R^2, and Mincer-Zarnowitz statistics (p-value in between brackets), out-of-sample, four-quarters ahead, 2000-2023

	MSPE	MAE	R^2	MZ-intercept	MZ-slope	MZ-R^2	MZ-Wald-Stat
COV-BHM	0.9503	0.9908	0.2100	-0.0010	1.0240	0.2130	3.16(0.04)
AR	0.9533	0.9981	0.2450	-0.0010	1.0760	0.2520	6.78(0.00)
AO-RW	0.9550	0.9848	0.2080	0.0020	0.7350	0.2440	40.50(0.00)
DB-BHM	0.9565	0.9945	0.2050	-0.0000	0.9790	0.2080	2.97(0.05)
RE-BHM	0.9672	1.0033	0.1960	-0.0010	1.0550	0.2010	4.96(0.01)
ARMA	0.9707	0.9925	0.2320	0.0010	0.8250	0.2440	11.39(0.00)
RE-BHM-X	0.9899	1.0088	0.1770	0.0020	0.8890	0.1950	17.97(0.00)
UC-SV	1.0000	1.0000	0.1670	0.0020	0.8130	0.1820	16.21(0.00)
DB-BHM-X	1.0864	1.0675	0.0970	0.0030	0.6280	0.1680	68.33(0.00)
COV-BHM-X	1.0886	1.0709	0.0950	0.0030	0.6320	0.1700	72.68(0.00)
PCA-VAR	1.1096	1.0648	0.0430	0.0030	0.6440	0.1370	91.08(0.00)
UC-SV-SS	1.4031	1.2411	-0.1690	0.0050	0.3250	0.0480	199.86(0.00)
M-UC-SV-SS	4.4708	2.1356	-2.7230	0.0060	0.0590	0.0110	2422.88(0.00)
ARX	17.6768	2.4716	-12.9920	0.0060	-0.0320	0.0120	9528.73(0.00)
ARMAX	2228.4067	12.0363	-1762.8430	0.0060	-0.0040	0.0270	1312027.94(0.00)

To check whether the results are robust to different forecast horizons, we also forecast four-quarters ahead (one year) instead of one-quarter ahead. Table 3 shows the performance of all models when forecasting one-year ahead headline CPI. The MSPE and MAE are again relative to the benchmark UC-SV model and the p-values of the Wald test statistic are again given in brackets.

The table shows that in this case, the performance of the models changes substantially. The COV-BHM performs best based on the MSPE, followed by the AR and AO-RW models. These models are followed by the DB-BHM and RE-BHM, which both outperform the UC-SV benchmark in this instance. From comparing the RE-BHM to the DB-BHM and COV-BHM, we conclude that it is possible to benefit from the correlation of inflation rates between countries when forecasting one-year ahead, as the DB-BHM and COV-BHM slightly outperform the RE-BHM. The DB-BHM is the only model for which we cannot reject the null hypothesis of the Mincer-Zarnowitz Wald test at a 5% significance level. This indicates that for all other models, there is evidence that the forecasts are inefficient and/or biased. Nevertheless, from the relatively close MSPEs, we conclude that the Bayesian models we consider do not add predictive power over simpler already existing models such as the AR and AO-RW models. On a relative scale, we do however notice that hierarchical models perform much better when we consider a longer forecast horizon.

What is furthermore noticeable to see, is that the UC-SV-SS model underperforms the UC-SV benchmark when predicting one-year ahead. The increased forecasting power of the UC-SV-SS model due to the use of the seasonality of the data seems to disappear when forecasting four-quarters ahead instead of one. We suspect that this happens because seasonality plays a less prominent role when forecasting year-ahead differences as the seasonal effects are negated when you take the difference between two time periods in identical seasons. Hence, the flexibility of the UC-SV-SS model is not required in this case, and the additional unknown seasonal coefficients solely lead to increased variance during estimation and prediction.

For the M-UC-SV-SS, ARX and ARMAX models the conclusions do not substantially change, as they still significantly underperform the UC-SV benchmark.

5 Conclusion

This paper investigates whether existing inflation forecasts can be improved upon by using seasonality features and imposing similarity between countries' inflation rates via hierarchical model structures. We argue that inflation is a global phenomenon, which is why a prior based on geographical distances between countries is of special interest. This paper investigates this by analyzing inflation between Q1 1977 and Q1 2023 for a set of 20 countries, using headline consumer price index as a measure of inflation. As a benchmark, we use a wide variety of existing time series models, with of special interest the unobserved component stochastic volatility (UC-SV) model introduced by Stock and Watson (2007). We introduce three Bayesian hierarchical models (BHMs), namely the random effects model, which is a BHM with a normal prior, a sample covariance based model, which is a BHM with a prior based on the sample covariance matrix of inflation and a distance based model. The distance based model uses a Matérn kernel to transform geospatial distances between countries into a covariance matrix for the coefficients. We also extend the UC-SV model by adding a seasonality component, which leads to the unobserved component stochastic volatility stochastic seasonality (UC-SV-SS) model. We further extend the UC-SV-SS model to a multivariate setting (M-UC-SV-SS) by making use of the Matérn kernel to incorporate a geospatial prior.

First, we conclude on the effectiveness of including seasonality in inflation forecasting models. Exploratory analysis of quarterly inflation per country shows that seasonality was not only clearly present in all countries of the sample, but also increased in magnitude over time. Using several performance metrics, such as the MSPE, MAE and the Mincer-Zarnowitz regression, we find that the unobserved component stochastic volatility model with stochastic seasonality outperforms all other models for all proposed metrics. This shows the importance of incorporating seasonality when modelling inflation after 2000. Somewhat surprisingly, this result does not disappear when forecasting core inflation, which is already cleared from energy and food prices, which are generally very seasonal. On the contrary, the improvement in forecasting performance from including a stochastic seasonality component in the unobserved component model becomes larger, as measured by the change in relative MSPE, when using core inflation.

However, one can notice that the improvement in forecasting performance stemming from stochastic seasonality disappears when considering a forecasting horizon of one year. In this case, models with stochastic seasonality perform much worse than their counterparts without seasonality features. We thus conclude that stochastic seasonality yields significant outperformance over existing inflation forecasting models for short forecast horizons (one-quarter ahead), where seasonality plays a big role. When forecasting one-year ahead, including stochastic seasonality worsens performance.

Moving forward, we can conclude on whether imposing similarity in inflation patterns across countries can improve the forecasting performance of inflation models. The results for this innovation are mixed. For the main forecasting setting explored in this paper - one-quarter ahead headline inflation between 2000 and 2023 - no BHM outperforms a simple AR model or an Atkeson-Ohanian random walk. Hence, imposing restrictive priors does not improve one-quarter ahead forecasting performance. In contrast, the simpler regularisation imposed on cross-sectional inflation by PCA-VAR performs well, with better out-of-sample results than the UC-SV. These results hold both for headline and core inflation forecasts.

In the same setting, we also find that the distance based model performs worse than the random effects model and that the M-UC-SV-SS model performs worse than the UC-SV-SS model, which suggests that the geospatial prior we incorporate was either too strong or does not help with forecasting.

The interpretation is somewhat flipped when considering one-year ahead headline inflation forecasts. In this case, all three BHMs rank in the top five of all explored models, with the sample covariance based model performing best. In this case, the hierarchical models with priors that restrict cross-country differences also perform marginally better than their unrestricted counterparts.

Overall, the results of the robustness checks show that the relative inflation forecasting performance of the models presented in this paper is sensitive to the forecast horizon. Incorporating seasonality improves one-quarter ahead forecasts, but worsens one-year ahead forecasts. Multivariate BHMs perform poorly in one-quarter ahead settings, but well for one-year ahead predictions. As there is no obvious reason for this sensitivity, it merits to be the subject of future research.

Wrapping up, our research shows that a prior based on geospatial distance adds little to no predictive power to the random effects model. However, this might be caused by the use of a geospatial proxy for distance. Inflation could be more similar between countries that are similar in a political-economical way, not necessarily countries that have a similar geographical location. It would therefore make more sense to have a distance function that proxies the political-economical distance between countries, not necessarily the geospatial distance. However, we do not have access to the data needed to construct such a distance, and if such data were available it could potentially be published with a significant delay. Hence, it could be of interest for further research to find proxies that better measure the political and economic distance between countries to obtain a better distance metric. Furthermore, as the random effects model introduced in this paper seems to have a too restrictive prior, it would be of interest to further research the use of weaker priors for the random effects model to try to improve the forecasting performance.

References

Atkeson, A., & Ohanian, L. E. (2001). Are Phillips Curves Useful for Forecasting Inflation? *Federal Reserve bank of Minneapolis quarterly review*, *25*(1), 2–11.

Bańbura, M., & Rünstler, G. (2011). A look into the factor model black box: Publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting*, *27*(2), 333–346.

Bank for International Settlements. (2024). *Interest Rate Statistics 2024*. `https://data.bis.org/topics/CBPOL/data`. (Accessed: 2024-03-16)

Boughton, J. M., & Branson, W. H. (1988). *Commodity Prices As a Leading Indicator of Inflation* (Tech. Rep.). Cambridge, Mass., USA: National Bureau of Economic Research.

Chen, Y.-c., Turnovsky, S. J., & Zivot, E. (2014). Forecasting inflation using commodity price aggregates. *Journal of Econometrics*, *183*(1), 117–134.

Ciccarelli, M., & Mojon, B. (2010). Global Inflation. *The Review of Economics and Statistics*, *92*(3), 524–535.

Fama, E. F. (1975). Short-Term Interest Rates as Predictors of Inflation. *American Economic Review*, *65*(3), 269–282.

Fisher, I. (2006). *Mathematical Investigations in the Theory of Value and Prices, and Appreciation and Interest.* Cosimo, Inc.

Groen, J. J., Paap, R., & Ravazzolo, F. (2013). Real-Time Inflation Forecasting in a Changing World. *Journal of Business & Economic Statistics*, *31*(1), 29–44.

Ha, J., Kose, M. A., & Ohnsorge, F. (2023). One-stop source: A global database of inflation. *Journal of International Money and Finance*, *137*, 102896.

Harvey, A. C. (1990). Forecasting, Structural Time Series Models and the Kalman Filter. *Cambridge University Press*.

Koulakiotis, A., Lyroudi, K., & Papasyriopoulos, N. (2012). Inflation, GDP and Causality for European Countries. *International Advances in Economic Research*, *18*, 53–62.

Laflèche, T., & Armour, J. (2006). Evaluating Measures of Core Inflation. *Bank of Canada Review*, *2006*(Summer), 19–29.

Lindley, D. V., & Smith, A. F. (1972). Bayes Estimates for the Linear Model. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *34*(1), 1–18.

Lis, E., & Porqueddu, M. (2018). The role of seasonality and outliers in HICP inflation excluding food and energy. *Economic Bulletin Boxes*, *2*.

Mayer, T., & Zignago, S. (2011). *Notes on CEPII's distances measures: The GeoDist database* (Tech. Rep.). Paris, France: CEPII.

Medeiros, M. C., Vasconcelos, G. F., Veiga, Á., & Zilberman, E. (2021). Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods. *Journal of Business & Economic Statistics*, *39*(1), 98–119.

Mincer, J. A., & Zarnowitz, V. (1969). *Economic forecasts and expectations: Analysis of forecasting behavior and performance.* National Bureau of Economic Research, Inc.

Organisation for Economic Co-operation and Development. (2024). *GDP Growth Data 2024*. `https://stats.oecd.org/index.aspx?queryid=350`. (Accessed: 2024-03-16)

Organization, I. L. (2024). *Global Employment Trends 2024.* Retrieved from https://www.ilo.org/shinyapps (Accessed: 2024-03-16)

Stock, J. H., & Watson, M. W. (1999). Forecasting Inflation. *Journal of monetary economics*, *44*(2), 293–335.

Stock, J. H., & Watson, M. W. (2007). Why Has US Inflation Become Harder to Forecast? *Journal of Money, Credit and banking*, *39*, 3–33.

Stock, J. H., & Watson, M. W. (2016). Core Inflation and Trend Inflation. *Review of Economics and Statistics*, *98*(4), 770–784.

A Summary statistics of datasets

Table 4: Summary statistics for quarterly changes in global commodity indices, 1977-2023. Metrics are in decimals (0.01=1% change).

	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	25th Percentile	Median	75th Percentile	Maximum
Crude Oil	0.0261	0.1838	1.1176	8.1897	-0.5852	-0.0563	0.0131	0.1180	1.1419
Natural Gas	0.0258	0.1868	0.9388	3.5839	-0.5855	-0.0629	0.0095	0.0858	0.8832
Agriculture	0.0055	0.0558	0.2057	1.7596	-0.2106	-0.0270	0.0003	0.0388	0.2065
Metals & Minerals	0.0125	0.0993	-0.3984	2.7807	-0.4835	-0.0447	0.0065	0.0673	0.2661
Precious Metals	0.0170	0.0907	1.2193	3.9569	-0.1909	-0.0420	0.0087	0.0596	0.4231

Table 5: Summary statistics for quarterly inflation rates, 1977-2023. Inflation rates in decimals (0.01=1% inflation).

Country	Count	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	25th Percentile	Median	75th Percentile	Maximum
Australia	180	0.0102	0.0090	0.6263	0.7674	-0.0194	0.0041	0.0075	0.0167	0.0409
Belgium	98	0.0058	0.0068	1.6501	4.0867	-0.0077	0.0019	0.0044	0.0082	0.0333
Canada	185	0.0086	0.0085	0.8001	0.7207	-0.0153	0.0029	0.0072	0.0126	0.0359
Denmark	98	0.0050	0.0060	1.5641	4.4919	-0.0037	0.0010	0.0039	0.0086	0.0330
France	86	0.0040	0.0050	0.7998	1.1749	-0.0050	0.0000	0.0031	0.0075	0.0229
Germany	76	0.0051	0.0062	0.9797	2.1844	-0.0061	0.0018	0.0040	0.0081	0.0285
Italy	101	0.0051	0.0064	2.7826	14.0496	-0.0048	0.0020	0.0049	0.0072	0.0431
Japan	149	0.0036	0.0074	1.0617	1.6426	-0.0122	-0.0020	0.0023	0.0072	0.0319
Korea	95	0.0062	0.0057	0.3665	-0.1230	-0.0051	0.0021	0.0057	0.0100	0.0208
Netherlands	95	0.0057	0.0078	1.1354	4.3655	-0.0137	0.0018	0.0048	0.0097	0.0405
New Zealand	143	0.0064	0.0062	1.0424	2.9333	-0.0081	0.0027	0.0055	0.0096	0.0352
Norway	95	0.0058	0.0069	0.0266	2.3294	-0.0167	0.0030	0.0056	0.0094	0.0285
Portugal	101	0.0053	0.0095	0.8916	3.5662	-0.0106	-0.0012	0.0052	0.0113	0.0502
South Africa	77	0.0135	0.0078	0.4031	0.5994	-0.0070	0.0085	0.0126	0.0169	0.0347
Spain	101	0.0057	0.0106	0.0843	-0.2183	-0.0199	-0.0020	0.0058	0.0136	0.0329
Sweden	92	0.0047	0.0075	1.4008	3.9393	-0.0145	0.0010	0.0034	0.0067	0.0327
Switzerland	67	0.0021	0.0059	0.3157	0.0500	-0.0107	-0.0015	0.0010	0.0068	0.0169
Turkey	60	0.0238	0.0153	0.5772	0.5770	-0.0037	0.0125	0.0245	0.0301	0.0661
United Kingdom	87	0.0049	0.0045	0.3109	0.1217	-0.0060	0.0013	0.0047	0.0081	0.0189
United States	185	0.0090	0.0085	0.5352	3.2899	-0.0283	0.0040	0.0079	0.0118	0.0395

Table 6: Summary statistics for quarterly GDP growth rates, 1977-2023. Inflation rates in decimals (0.01=1% GDP growth).

Country	Count	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	25th Percentile	Median	75th Percentile	Maximum
Australia	180	0.0076	0.0103	-1.9961	17.7674	-0.0695	0.0041	0.0075	0.0112	0.0425
Belgium	98	0.0045	0.0179	-0.4768	35.9043	-0.1139	0.0022	0.0045	0.0075	0.1175
Canada	185	0.0060	0.0129	-2.5485	43.9907	-0.1100	0.0020	0.0069	0.0113	0.0908
Denmark	98	0.0038	0.0123	-0.5756	10.5941	-0.0604	-0.0006	0.0047	0.0093	0.0591
France	86	0.0033	0.0251	1.8773	35.5376	-0.1315	0.0003	0.0043	0.0064	0.1754
Germany	76	0.0035	0.0090	-2.2981	12.6145	-0.0468	0.0001	0.0048	0.0081	0.0225
Italy	101	0.0014	0.0205	0.7189	29.8412	-0.1168	-0.0024	0.0022	0.0052	0.1372
Japan	149	0.0054	0.0135	-1.6421	10.9594	-0.0776	-0.0005	0.0061	0.0122	0.0552
Korea	95	0.0094	0.0098	-1.2712	5.2476	-0.0328	0.0050	0.0087	0.0140	0.0299
Netherlands	95	0.0040	0.0132	-2.3855	25.1262	-0.0835	0.0013	0.0049	0.0073	0.0638
New Zealand	143	0.0065	0.0170	1.6005	38.2873	-0.1011	0.0015	0.0070	0.0121	0.1414
Norway	95	0.0041	0.0118	-0.5474	5.8265	-0.0519	-0.0002	0.0044	0.0094	0.0394
Portugal	101	0.0033	0.0236	-0.6457	30.7364	-0.1511	-0.0011	0.0048	0.0087	0.1460
South Africa	77	0.0050	0.0262	-2.2230	33.9814	-0.1689	0.0024	0.0054	0.0102	0.1373
Spain	101	0.0048	0.0258	-1.3453	38.1652	-0.1764	0.0022	0.0066	0.0097	0.1624
Sweden	92	0.0057	0.0147	-1.3349	16.9210	-0.0798	0.0018	0.0065	0.0119	0.0740
Switzerland	67	0.0050	0.0130	-0.6585	21.4442	-0.0658	0.0024	0.0050	0.0077	0.0687
Turkey	60	0.0119	0.0320	0.6873	10.7816	-0.1071	0.0032	0.0149	0.0231	0.1641
United Kingdom	87	0.0038	0.0292	-2.1520	40.5663	-0.2032	0.0030	0.0058	0.0074	0.1678
United States	185	0.0067	0.0110	-1.2399	28.1310	-0.0789	0.0035	0.0074	0.0108	0.0776

Table 7: Summary statistics for quarterly central bank interest rates, 1977-2023. Interest rates in decimals (0.01=1%).

Country	Count	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	25th Percentile	Median	75th Percentile	Maximum
Australia	180	0.0710	0.0516	0.8715	-0.0949	0.0010	0.0350	0.0533	0.1074	0.2225
Belgium	98	0.0158	0.0153	0.5358	-1.0513	0.0000	0.0001	0.0100	0.0290	0.0475
Canada	185	0.0546	0.0453	0.8188	-0.0715	0.0025	0.0125	0.0450	0.0877	0.2041
Denmark	98	0.0140	0.0195	0.4398	-1.2225	-0.0075	-0.0060	0.0086	0.0308	0.0540
France	86	0.0129	0.0137	0.7872	-0.5821	0.0000	0.0000	0.0100	0.0200	0.0475
Germany	76	0.0129	0.0141	0.8443	-0.6045	0.0000	0.0000	0.0100	0.0215	0.0458
Italy	101	0.0168	0.0161	0.5570	-0.9475	0.0000	0.0003	0.0108	0.0300	0.0550
Japan	149	0.0222	0.0243	0.6903	-0.7922	-0.0010	0.0005	0.0083	0.0450	0.0900
Korea	95	0.0285	0.0139	0.0783	-1.2006	0.0050	0.0158	0.0275	0.0404	0.0525
Netherlands	95	0.0152	0.0151	0.5934	-0.9790	0.0000	0.0000	0.0100	0.0271	0.0475
New Zealand	143	0.0577	0.0400	1.0323	0.9102	0.0025	0.0250	0.0542	0.0769	0.1800
Norway	95	0.0250	0.0212	0.9875	-0.3089	0.0000	0.0100	0.0175	0.0354	0.0700
Portugal	101	0.0168	0.0161	0.5489	-0.9874	0.0000	0.0003	0.0108	0.0300	0.0533
South Africa	77	0.0701	0.0245	0.9496	0.2448	0.0350	0.0550	0.0667	0.0750	0.1350
Spain	101	0.0165	0.0156	0.4697	-1.1682	0.0000	0.0003	0.0108	0.0300	0.0475
Sweden	92	0.0143	0.0161	0.4361	-1.1625	-0.0050	0.0000	0.0112	0.0275	0.0458
Switzerland	67	0.0005	0.0097	1.5450	2.5397	-0.0075	-0.0075	0.0013	0.0037	0.0325
Turkey	60	0.1063	0.0528	1.0328	0.0821	0.0450	0.0683	0.0800	0.1477	0.2400
United Kingdom	87	0.0249	0.0225	0.3607	-1.6140	0.0010	0.0050	0.0075	0.0471	0.0733
United States	185	0.0470	0.0415	0.9828	0.8125	0.0013	0.0100	0.0458	0.0681	0.1837

Table 8: Summary statistics for quarterly unemployment rates. 1977-2023, Unemployment rates in decimals (0.01=1% unemployment).

Country	Count	Mean	Standard Deviation	Skewness	Kurtosis	Minimum	25th Percentile	Median	75th Percentile	Maximum
Australia	180	0.0668	0.0180	0.7156	-0.1378	0.0330	0.0540	0.0620	0.0793	0.1170
Belgium	98	0.0740	0.0112	-0.4937	-0.8295	0.0490	0.0653	0.0765	0.0830	0.0930
Canada	185	0.0816	0.0181	0.7579	0.1347	0.0460	0.0690	0.0770	0.0920	0.1390
Denmark	98	0.0557	0.0128	0.4355	-0.6200	0.0330	0.0470	0.0535	0.0640	0.0850
France	86	0.0895	0.0110	0.3440	0.1232	0.0680	0.0813	0.0890	0.0960	0.1210
Germany	76	0.0613	0.0255	0.5631	-0.8765	0.0300	0.0380	0.0530	0.0795	0.1150
Italy	101	0.0954	0.0194	0.0148	-0.9256	0.0560	0.0800	0.0940	0.1110	0.1360
Japan	149	0.0304	0.0092	0.8922	-0.3378	0.0190	0.0240	0.0270	0.0380	0.0540
Korea	95	0.0342	0.0062	1.3139	2.7542	0.0240	0.0300	0.0330	0.0370	0.0600
Netherlands	95	0.0460	0.0141	0.5297	-0.4001	0.0210	0.0360	0.0430	0.0560	0.0810
New Zealand	143	0.0587	0.0191	0.9426	0.5654	0.0320	0.0440	0.0560	0.0675	0.1140
Norway	95	0.0365	0.0064	0.1147	-0.2929	0.0210	0.0320	0.0360	0.0410	0.0510
Portugal	101	0.0852	0.0353	0.8598	-0.1380	0.0380	0.0620	0.0760	0.1070	0.1780
South Africa	77	0.2490	0.0380	0.5009	-0.5928	0.1830	0.2200	0.2420	0.2810	0.3400
Spain	101	0.1581	0.0519	0.4050	-0.8078	0.0790	0.1140	0.1520	0.1960	0.2690
Sweden	92	0.0714	0.0125	0.0042	-0.6554	0.0470	0.0630	0.0720	0.0810	0.0990
Switzerland	67	0.0447	0.0063	-0.9278	1.4647	0.0250	0.0415	0.0460	0.0490	0.0580
Turkey	60	0.1043	0.0187	0.4428	-0.6689	0.0730	0.0888	0.1020	0.1185	0.1470
United Kingdom	87	0.0566	0.0133	0.6313	-0.7060	0.0360	0.0480	0.0520	0.0620	0.0850
United States	185	0.0617	0.0175	0.8132	0.6038	0.0330	0.0500	0.0580	0.0720	0.1290

B Derivation of Posterior Distribution for the Random Effects Model

We will consider a simplified random effects model with a prior mean on the coefficients of 0 to simplify derivations. This leads to the following hierarchical priors:

$$\alpha_i \sim N(0, \tau_\alpha^{-1}), \tag{30}$$
$$\beta_i \sim N(0, \tau_\beta^{-1}), \tag{31}$$
$$\gamma_i \sim N(0, \tau_\gamma^{-1}), \tag{32}$$
$$\epsilon_{i,t} \sim N(0, \tau_\epsilon^{-1}), \tag{33}$$
$$\tau_\alpha, \tau_\beta, \tau_\gamma, \tau_\epsilon \sim \Gamma(1, 1). \tag{34}$$

The likelihood function is given by:

$$p(y_{t,i}|\alpha_i, \beta_i, \gamma_i, \tau_\epsilon) = \prod_{t=1}^{T}\prod_{i=1}^{N} \mathcal{N}(y_{t,i}|\alpha_i + \beta_i x_{t-1,i} + \gamma_i w_{t-1}, \tau_\epsilon^{-1}), \tag{35}$$

where $\mathcal{N}(\cdot|\mu, \sigma^2)$ denotes the normal distribution with mean μ and variance σ^2.

In Equations (36)-(38) we specify that all coefficients and the error term follow a normal distribution with zero mean and constant variance. In Equations (39)-(42) we specify that the inverse of each of these variance terms follows a Gamma distribution.

This posterior distribution of this specific BHM can be derived analytically. The prior distributions are:

$$p(\alpha_i|\tau_\alpha) = \mathcal{N}(\alpha_i|0, \tau_\alpha^{-1}), \tag{36}$$
$$p(\beta_i|\tau_\beta) = \mathcal{N}(\beta_i|0, \tau_\beta^{-1}), \tag{37}$$
$$p(\gamma_i|\tau_\gamma) = \mathcal{N}(\gamma_i|0, \tau_\gamma^{-1}), \tag{38}$$
$$p(\tau_\alpha) = \Gamma(\tau_\alpha|1, 1), \tag{39}$$
$$p(\tau_\beta) = \Gamma(\tau_\beta|1, 1), \tag{40}$$
$$p(\tau_\gamma) = \Gamma(\tau_\gamma|1, 1), \tag{41}$$
$$p(\tau_\epsilon) = \Gamma(\tau_\epsilon|1, 1). \tag{42}$$

Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_N)$, and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N)$ denote the vectors of country-specific coefficients. The posterior distribution can be written as:

$$
\begin{aligned}
&p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\beta, \tau_\gamma, \tau_\epsilon | \mathbf{y}, \mathbf{X}, \mathbf{W}) && (43)\\
&\quad \propto \prod_{t=1}^{T} \prod_{i=1}^{N} \mathcal{N}(y_{t,i} | \alpha_i + \beta_i x_{t-1,i} + \gamma_i w_{t-1}, \tau_\epsilon^{-1}) && (44)\\
&\quad \times \prod_{i=1}^{N} \mathcal{N}(\alpha_i | 0, \tau_\alpha^{-1}) \times \prod_{i=1}^{N} \mathcal{N}(\beta_i | 0, \tau_\beta^{-1}) && (45)\\
&\quad \times \prod_{i=1}^{N} \mathcal{N}(\gamma_i | 0, \tau_\gamma^{-1}) \times \Gamma(\tau_\alpha | 1, 1) && (46)\\
&\quad \times \Gamma(\tau_\beta | 1, 1) \times \Gamma(\tau_\gamma | 1, 1) \times \Gamma(\tau_\epsilon | 1, 1), && (47)
\end{aligned}
$$

where $\mathbf{y} = (y_{1,1}, \ldots, y_{T,N})$, $\mathbf{X} = (x_{0,1}, \ldots, x_{T-1,N})$, and $\mathbf{W} = (w_0, \ldots, w_{T-1})$ denote the observed data.

Due to the conjugacy of the normal likelihood with the normal priors on the coefficients and the conjugacy of the normal likelihood with the gamma priors on the precision parameters, the posterior distributions of the parameters have closed-form expressions. To derive the posterior distributions for the parameters, we use the following properties of conjugate priors:

1. If the likelihood is normally distributed and the prior on the mean is normal, the posterior of the mean is also normally distributed.
2. If the likelihood is normally distributed and the prior on the precision (inverse variance) is gamma-distributed, the posterior of the precision is also gamma-distributed.

For each α_i:

$$
\begin{aligned}
p(\alpha_i | \mathbf{y}, \mathbf{X}, \mathbf{W}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \tau_\alpha, \tau_\epsilon) &\propto \prod_{t=1}^{T} \mathcal{N}(y_{t,i} | \alpha_i + \beta_i x_{t-1,i} + \gamma_i w_{t-1}, \tau_\epsilon^{-1}) \times \mathcal{N}(\alpha_i | 0, \tau_\alpha^{-1})\\
&\propto \exp\left[-\frac{\tau_\epsilon}{2} \sum_{t=1}^{T} (y_{t,i} - \alpha_i - \beta_i X_{t-1,i} - \gamma_i w_t) - \frac{\tau_\alpha}{2} \alpha_i^2 \right]
\end{aligned}
$$

To simplify subsequent operations, we let $z_{t,i} := y_{t,i} - \beta_i X_{t-1,i} - \gamma_i w_{t-1}$. It follows that

$$
\begin{aligned}
&\exp\left[-\frac{\tau_\epsilon}{2}\sum_{t=1}^{T}(y_{t,i}-\alpha_i-\beta_i X_{t-1,i}-\gamma_i w_t)-\frac{\tau_\alpha}{2}\alpha_i^2\right]\\
&=\exp\left[-\frac{\tau_\epsilon}{2}\sum_{t=1}^{T}z_{t,i}^2-\frac{\tau_\epsilon}{2}\left(-2\alpha_i\sum_{t=1}^{T}z_{t,i}+\left(T+\frac{\tau_\alpha}{\tau_\epsilon}\right)\alpha_i^2\right)\right]\\
&\propto\exp\left[-\frac{1}{2}\left((T\tau_\epsilon+\tau_\alpha)\left(\alpha_i^2-\frac{2\tau_\epsilon\sum_{t=1}^{T}z_{t,i}}{T\tau_\epsilon+\tau_\alpha}\alpha_i\right)\right)\right]\\
&=\exp\left[-\frac{1}{2}(T\tau_\epsilon+\tau_\alpha)\left(\left(\alpha_i-\frac{\tau_\epsilon\sum_{t=1}^{T}z_{t,i}}{T\tau_\epsilon+\tau_\alpha}\right)^2-\left(\frac{\tau_\epsilon\sum_{t=1}^{T}z_{t,i}}{T\tau_\epsilon+\tau_\alpha}\right)^2+\frac{\tau_\epsilon\sum_{t=1}^{T}z_{t,i}^2}{T\tau_\epsilon+\tau_\alpha}\right)\right]\\
&\propto\mathcal{N}(\alpha_i|\mu_{\alpha_i},\sigma^2_{\alpha_i})
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_{\alpha_i}&=\frac{\tau_\epsilon\sum_{t=1}^{T}(y_{t,i}-\beta_i x_{t-1,i}-\gamma_i w_{t-1})}{T\tau_\epsilon+\tau_\alpha},\\
\sigma^2_{\alpha_i}&=\frac{1}{\tau_\alpha+T\tau_\epsilon}.
\end{aligned}
$$

For the following derivations, we skip the algebraic details for brevity. For each β_i:

$$
\begin{aligned}
p(\beta_i|\mathbf{y},\mathbf{X},\mathbf{W},\boldsymbol{\alpha},\boldsymbol{\gamma},\tau_\beta,\tau_\epsilon)&\propto\prod_{t=1}^{T}\mathcal{N}(y_{t,i}|\alpha_i+\beta_i x_{t-1,i}+\gamma_i w_{t-1},\tau_\epsilon^{-1})\times\mathcal{N}(\beta_i|0,\tau_\beta^{-1})\\
&\propto\mathcal{N}(\beta_i|\mu_{\beta_i},\sigma^2_{\beta_i}),
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_{\beta_i}&=\frac{\tau_\epsilon\sum_{t=1}^{T}x_{t-1,i}(y_{t,i}-\alpha_i-\gamma_i w_{t-1})}{\tau_\beta+\tau_\epsilon\sum_{t=1}^{T}x_{t-1,i}^2},\\
\sigma^2_{\beta_i}&=\frac{1}{\tau_\beta+\tau_\epsilon\sum_{t=1}^{T}x_{t-1,i}^2}.
\end{aligned}
$$

Similarly, for each γ_i:

$$
\begin{aligned}
p(\gamma_i|\mathbf{y},\mathbf{X},\mathbf{W},\boldsymbol{\alpha},\boldsymbol{\beta},\tau_\gamma,\tau_\epsilon)&\propto\prod_{t=1}^{T}\mathcal{N}(y_{t,i}|\alpha_i+\beta_i x_{t-1,i}+\gamma_i w_{t-1},\tau_\epsilon^{-1})\times\mathcal{N}(\gamma_i|0,\tau_\gamma^{-1})\\
&\propto\mathcal{N}(\gamma_i|\mu_{\gamma_i},\sigma^2_{\gamma_i})
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_{\gamma_i}&=\frac{\tau_\epsilon\sum_{t=1}^{T}w_{t-1}(y_{t,i}-\alpha_i-\beta_i x_{t-1,i})}{\tau_\gamma+\tau_\epsilon\sum_{t=1}^{T}w_{t-1}^2},\\
\sigma^2_{\gamma_i}&=\frac{1}{\tau_\gamma+\tau_\epsilon\sum_{t=1}^{T}w_{t-1}^2}.
\end{aligned}
$$

For τ_α:

$$p(\tau_\alpha|\boldsymbol{\alpha}) \propto \prod_{i=1}^{N} \mathcal{N}(\alpha_i|0, \tau_\alpha^{-1}) \times \Gamma(\tau_\alpha|1,1)$$
$$\propto \Gamma(\tau_\alpha|a_\alpha, b_\alpha),$$

where

$$a_\alpha = 1 + \frac{N}{2},$$
$$b_\alpha = 1 + \frac{1}{2}\sum_{i=1}^{N} \alpha_i^2.$$

Similarly, for τ_β:

$$p(\tau_\beta|\boldsymbol{\beta}) \propto \prod_{i=1}^{N} \mathcal{N}(\beta_i|0, \tau_\beta^{-1}) \times \Gamma(\tau_\beta|1,1)$$
$$\propto \Gamma(\tau_\beta|a_\beta, b_\beta),$$

where

$$a_\beta = 1 + \frac{N}{2},$$
$$b_\beta = 1 + \frac{1}{2}\sum_{i=1}^{N} \beta_i^2.$$

For τ_γ:

$$p(\tau_\gamma|\boldsymbol{\gamma}) \propto \prod_{i=1}^{N} \mathcal{N}(\gamma_i|0, \tau_\gamma^{-1}) \times \Gamma(\tau_\gamma|1,1)$$
$$\propto \Gamma(\tau_\gamma|a_\gamma, b_\gamma),$$

where

$$a_\gamma = 1 + \frac{N}{2},$$
$$b_\gamma = 1 + \frac{1}{2}\sum_{i=1}^{N} \gamma_i^2.$$

Finally, for τ_ϵ:

$$p(\tau_\epsilon|\mathbf{y}, \mathbf{X}, \mathbf{W}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}) \propto \prod_{t=1}^{T}\prod_{i=1}^{N} \mathcal{N}(y_{t,i}|\alpha_i + \beta_i x_{t-1,i} + \gamma_i w_{t-1}, \tau_\epsilon^{-1}) \times \Gamma(\tau_\epsilon|1,1)$$
$$\propto \Gamma(\tau_\epsilon|a_\epsilon, b_\epsilon),$$

where

$$a_\epsilon = 1 + \frac{NT}{2},$$
$$b_\epsilon = 1 + \frac{1}{2}\sum_{i=1}^{N} \tau_i^2$$

This concludes the derivation of the posterior distribution of the random effects BHM. In general, for more complicated BHMs, it might not be possible to derive the analytical posterior distribution.

C PCA VAR Scree Plot

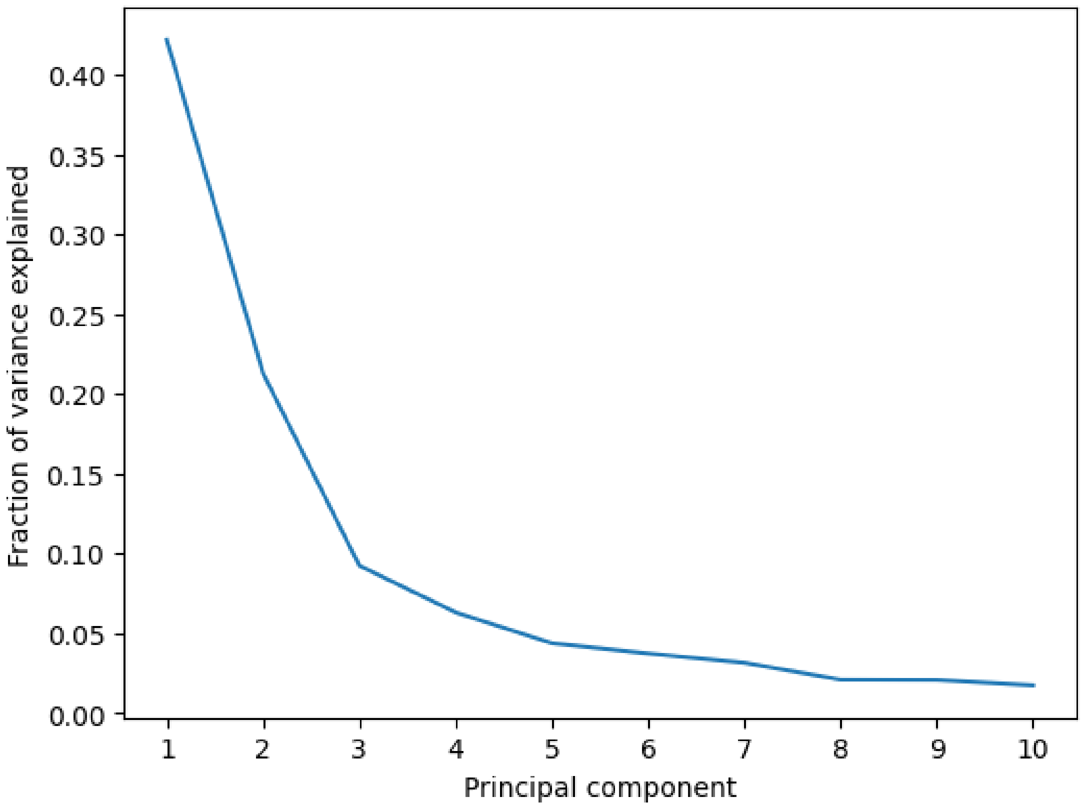


Figure 10: Scree plot of the PCA VAR model

From Figure 10, it is clear that three principal components are a sensible choice for capturing a large portion of the variance of the inflation across all countries.

D MSPE Per Country

Table 9: Country-specific relative MSPE per model, out-of-sample, 2000-2023

Country	AO-RW	AR	UC-SV	PCA-VAR	UC-SV-SS	ARX	ARMA	ARMAX	RE-BHM	RE-BHM-X	DB-BHM	DB-BHM-X	COV-BHM	COV-BHM-X	M-UC-SV-SS
Australia	0.7953	0.8641	1.0000	0.9803	0.8077	0.9319	0.7780	99.5329	0.9772	0.8976	0.9861	0.9541	0.9863	0.9558	0.9656
Belgium	0.8928	0.9861	1.0000	0.9979	0.9962	1.3160	1.4407	40.9111	0.8617	0.8553	0.9031	1.0918	0.8385	1.0123	0.9142
Canada	1.0497	1.2877	1.0000	1.2145	0.7543	1.3565	1.0445	5.8909	1.8572	1.6110	1.8837	1.6382	1.8813	1.6320	4.7086
Denmark	0.8977	1.2572	1.0000	1.0216	0.6768	3.3726	1.3201	1444.2704	1.0848	1.2005	1.0827	1.5153	1.0981	1.5542	1.1920
France	0.9046	1.1164	1.0000	0.8132	0.5761	3.0531	2.8288	13.0225	1.1017	1.2428	1.0923	2.2646	1.0766	3.0388	1.4570
Germany	0.7216	1.5646	1.0000	1.3878	0.8364	2.5342	1.4449	6.3418	1.0969	1.1204	1.0808	1.2056	1.1170	1.3955	1.3519
Italy	0.7472	1.0417	1.0000	1.2962	1.0876	1.4450	1.0802	2.5533	0.8979	0.9659	0.8877	1.0914	0.8796	1.0509	9.6954
Japan	0.9979	1.1708	1.0000	1.9415	1.1689	1.6599	1.0712	4.0744	1.4530	1.1857	1.4317	1.3023	1.4344	1.3123	1.1818
Korea	0.9301	1.0825	1.0000	1.4709	0.8200	1.8039	2.4550	154.1959	0.9991	0.9531	1.0191	0.9538	1.0208	1.0152	1.1778
Netherlands	0.9302	1.2201	1.0000	1.1497	0.8437	2.1502	1.0661	7.2977	0.9882	1.0008	0.9845	1.1185	0.9901	1.2155	1.2183
New Zealand	1.0289	1.0059	1.0000	1.9499	0.9720	1.3950	0.9711	4.9500	1.1010	1.2228	1.1110	1.3438	1.1085	1.3405	31.7271
Norway	1.0042	0.8562	1.0000	1.1077	1.0099	48.0645	1.6499	55.4528	1.0473	1.0760	0.9994	1.5040	1.0149	1.3958	2.7815
Portugal	0.9129	0.8943	1.0000	0.6363	0.4044	1.1570	0.5928	12.9065	1.3130	1.4222	1.4004	1.9957	1.3844	1.9957	1.1224
South Africa	1.4398	0.9132	1.0000	2.7243	0.8023	7.7164	1.7183	1826.2013	1.1777	1.2056	1.2956	2.3210	1.2506	1.6878	1.4149
Spain	0.9470	0.6415	1.0000	0.8660	0.9165	0.9540	0.4949	1.2249	1.9175	1.9957	2.0476	2.4044	1.9963	2.3834	6.7687
Sweden	0.8101	1.0547	1.0000	1.0782	0.6346	5.1183	0.8360	152.0086	1.3544	1.5346	1.5228	2.1545	1.4067	2.2374	3.6724
Switzerland	0.6734	1.0563	1.0000	1.0329	0.8254	48.7268	1.0571	242.3268	1.1885	0.9733	0.8323	1.6353	1.2844	1.7526	0.9166
Türkiye	0.8915	1.0874	1.0000	1.1678	0.7704	55.6929	5.3531	1443.3916	1.0292	1.2128	1.0918	1.2663	1.0786	1.2530	0.9359
United Kingdom	0.8716	1.0461	1.0000	0.7237	0.6323	6.9872	132.2926	103.4021	1.0458	1.2833	1.1598	2.1590	1.4306	2.8182	1.2392
United States	1.0082	1.0830	1.0000	1.0223	0.7483	1.1715	0.8194	2.9207	1.6852	1.4621	1.7127	1.4509	1.7096	1.4420	22.2352

E M-UC-SV-SS Model Performance for Subset of Countries

Table 10: Average MSPE (normalized with UC-SV = 1), MAE (normalized with UC-SV = 1), R^2, and Mincer-Zarnowitz statistics, out-of-sample, 2000-2023, for the United States, The Netherlands, Germany and Belgium

	MSPE	MAE	R^2	MZ-intercept	MZ-slope	MZ-R^2	MZ-Wald-Stat
UC-SV-SS	0.8505	0.8748	0.1270	0.0020	0.6650	0.1710	9.95(0.00)
UC-SV	1.0000	1.0000	-0.0270	0.0030	0.4280	0.0300	10.89(0.00)
M-UC-SV-SS	1.2642	1.1936	-0.2980	0.0050	0.0930	0.0030	55.87(0.00)

F UC-SV-SS Sensitivity to Θ

Table 11: Average MSPE (normalized with UC-SV = 1), MAE (normalized with UC-SV = 1), R^2, and Mincer-Zarnowitz statistics, out-of-sample, 2000-2023, varying values of Θ

	MSPE	MAE	R^2	MZ-intercept	MZ-slope	MZ-R^2	MZ-Wald-Stat
UC-SV-SS ($\Theta = 0.02$)	0.7784	0.8408	0.3540	0.0010	0.8490	0.3670	16.54(0.00)
UC-SV-SS ($\Theta = 0.0002$)	0.8005	0.8564	0.3470	0.0010	0.8430	0.3600	18.12(0.00)
UC-SV-SS ($\Theta = 0.002$)	0.8009	0.8504	0.3360	0.0010	0.8440	0.3490	16.81(0.00)
UC-SV	1.0000	1.0000	0.1710	0.0020	0.7840	0.1870	16.11(0.00)