



EUROPEAN CENTRAL BANK

EUROSYSTEM

## Working Paper Series

Marta Bańbura, Danilo Leiva-León,  
Jan-Oliver Menz

### Do inflation expectations improve model-based inflation forecasts?

Revised June 2023

No 2604 / October 2021

## **Abstract**

Those of professional forecasters do. For a wide range of time series models for the euro area and its member states we find a higher average forecast accuracy of models that incorporate information on inflation expectations from the ECB's SPF and Consensus Economics compared to their counterparts that do not. The gains in forecast accuracy from incorporating inflation expectations are typically not large but statistically significant in some periods. Both short- and long-term expectations provide useful information. The professional forecasters expectations seem to help to correct the upward forecast bias in the low inflation period and to make the model forecasts more robust, in particular in the environment of high volatility. By contrast, incorporating expectations derived from financial market prices or those of firms and households does not lead to systematic improvements in forecast performance. The analysis is undertaken for headline inflation and inflation excluding energy and food and both point and density forecast are evaluated using real-time data vintages over 2001-2022.

**JEL Classification:** C53, E31, E37

**Keywords:** Forecasting, inflation, inflation expectations, Phillips curve, Bayesian VAR

## Non-technical summary

The ECB strategy review has concluded that price stability is best maintained by aiming at a symmetric two percent inflation over the medium term as target. Within this context, inflation expectations play a key role in the conduct of monetary policy through their influence on how actual inflation might deviate from the target. Due to this intrinsic relationship between inflation and inflation expectations a natural question that emerges is to what extent the latter can help to obtain better forecasts of the former. This paper addresses this question within the context of econometric models. The followed approach to do so consists of comparing the performance associated with two versions of econometric models used to forecast inflation. The first version of models includes, among other macroeconomic variables, information on inflation expectations. Instead, the second version of models does not include data on inflation expectations. In this way, the value added of the incorporation of data on inflation expectations into econometric models used to forecast inflation can be quantified. This type of evaluation is carried out across different (i) state-of-the-art econometric models, (ii) measures of inflation, (iii) measures of inflation expectations, and (iv) geographic regions within the euro area.

A comprehensive out-of-sample evaluation based on euro area real-time data over 2001-2022 suggests that the incorporation of inflation expectations of professional forecasters (those collected by the ECB in its Survey of Professional Forecasters and by Consensus Economics) into econometric models does help to increase the accuracy of the latter when forecasting inflation. These forecasting gains are typically not large but significant in some periods for some models. Considering that inflation is very difficult to forecast with precision, any systematic improvements are useful for policy makers. Both short- and long-term expectations provide useful information. In general, these results also hold (i) when performing similar evaluations at the country level and (ii) for the recent post-COVID-19 sample. Expectations seem to help to correct the upward forecast bias in the low inflation period and to make the model forecasts more robust, in particular in the environment of high volatility. Last but not least, no forecasting gains are, in general, obtained when using inflation expectations of firms and households or derived from financial market prices.

In addition, this paper compares the predictive accuracy associated with model-based forecasts of inflation to that of inflation expectations directly used as forecasts. The results point to the “supremacy” of inflation expectations when forecasting euro area headline inflation in the

pre-pandemic sample, in that their predictions can be considered as a benchmark hard to beat by sophisticated econometric models. This is not always the case for the countries and in the more recent sample.

Overall, this paper shows that policy makers can benefit from incorporating information on inflation expectations from professional forecasts into econometric models devised to forecast inflation as such expectations appear to contain relevant information beyond what is already captured by other predictors of inflation. More generally, the results suggest that policy makers should pay attention to developments in those measures of expectations.

# 1 Introduction

Inflation expectations are usually closely monitored at central banks as they are believed to be an important determinant of current inflation.<sup>1</sup> In particular, it has been argued, that anchored inflation expectations help to stabilise inflation through agents reacting less strongly to economic shocks (Bernanke, 2010). Following its recent strategy review, the ECB has defined a new inflation target – symmetric two percent inflation over the medium term – arguing that it is “expected to contribute to a more solid anchoring of longer-term inflation expectations” which, in turn, “is essential for maintaining price stability”.<sup>2</sup>

Macroeconomic models often link current inflation to inflation expectations. One prominent example of such a relationship is the New Keynesian Phillips curve, which is a key ingredient of many structural and semi-structural models implemented at central banks and other institutions.<sup>3</sup> Inflation expectations also feature prominently in explanations put forward in order to interpret the puzzling behaviour of inflation in the aftermath of the global financial crisis. For example, Coibion and Gorodnichenko (2015) and Friedrich (2016) claim that it was the explicit behaviour of households’ inflation expectations which gave rise to the surprising inflation development after the global financial crisis.

A natural question thus is whether inflation expectations should be taken into account when forecasting inflation *out of sample* and if so in which manner. Inflation expectations are often not explicitly included in (reduced form) models routinely used to forecast inflation. One reason for this could be unavailable or only imperfect proxies of inflation expectations of economic agents in a given economy and lack of consensus as to which measures of expectations are the most relevant. Popular indicators of inflation expectations are professional forecasts as they are available for many economies and often over longer time samples, which is typically needed to evaluate a forecasting model.<sup>4</sup> However, they are often criticised as not representative of expectations in the economy at large. On the other hand, measures of inflation expectations of

---

<sup>1</sup>E.g. Clark and Davig (2008), Nunes (2010), Adam and Padula (2012), Canova and Gambetti (2010) or Fuhrer (2012) show that inflation expectations are a significant factor in explaining inflation in the United States.

<sup>2</sup>See ECB (2021); remarkably, the overview note mentions the term “inflation expectations” 12 times.

<sup>3</sup>Examples include the New Area-Wide Model II (Coenen et al., 2018) and the ECB-BASE (Angelini et al., 2019), the main macro models for the euro area at the ECB.

<sup>4</sup>For example, in the NAWM II, long-term inflation expectations from the ECB’s Survey of Professional Forecasters are used as a proxy for the unobserved perceived inflation objective. In the ECB-BASE, long-term inflation expectations are represented by long-term inflation forecasts from Consensus Economics. Time series models used to forecast inflation also typically rely on professional forecasts as measures of expectations, see e.g. Faust and Wright (2013) for the US or Bańbura and Bobeica (2023) for the euro area.

households and firms or those derived from financial market prices are subject to other pitfalls such as limited availability, measurement issues or short sample (see e.g. ECB, 2006). Another reason could be that observed measures of inflation expectations might not carry any additional information beyond what is already captured by other predictors of inflation. Existing studies usually report gains from incorporating observed measures of inflation expectations into econometric models but they typically focus on a particular measure, model and economy or do not perform out-of-sample forecast evaluations, as discussed in the literature review below. However, the out-of-sample perspective is important as contemporaneous correlations make it often difficult to disentangle *in sample* the “marginal” importance of various inflation determinants.

In this paper, we undertake an extensive evaluation of the usefulness of observed measures of inflation expectations in forecasting inflation out of sample. Contrary to the previous literature, we adopt a very broad take on this issue, considering a wide range of reduced form (time series) models, different measures of inflation expectations, several economies and two inflation indices. In terms of models, we cover main Phillips curve and Bayesian VAR (BVAR) specifications that have been shown to perform well for inflation forecasting in previous work (see the references in Section 3). In order to evaluate the “marginal” gain due to inflation expectations, for each model type we compare the performance of a version that incorporates a measure of expectations to its counterpart that does not. The main results are focused on forecasting euro area inflation, based on both headline HICP and HICP excluding energy and food components (“core HICP”), using the ECB’s Survey of Professional Forecasters (SPF) as the measure of expectations. But we also run analogous exercises for several individual countries of the euro area<sup>5</sup> and also consider Consensus Economics forecasts, measures of expectations of households and firms collected by the European Commission as well as those based on inflation-linked swap rates (where available and feasible). Whenever possible we use real-time data in order to appropriately assess the information content of various indicators. In addition to average point forecast accuracy, density forecasts and changes in forecast performance over time are investigated as well. We also assess the absolute performance of the models compared to the expectations and to popular benchmarks. We study separately the pre-pandemic period (up to 2019) and the more recent volatile sample, affected by the consequences of COVID-19 pandemic and the Russian invasion of Ukraine.

---

<sup>5</sup>These include Germany, France, Italy, Spain, the Netherlands, Belgium, Austria and Finland.

We find that incorporating expectations based on *professional forecasts* into models results in more accurate forecasts in the majority of cases. Both long- and short-term expectations appear to carry useful off-model information. This applies in particular to the euro area and the expectations based on the SPF but holds in general also for the countries we consider and the expectations based on the Consensus Economics forecast. The result applies also to the recent (post-)COVID-19 sample. Thus, inflation expectations embedded in professional forecasts do improve model-based forecasts of inflation.

The gains in forecast accuracy are typically not large, in particular, when forecasting the core component of inflation for some countries<sup>6</sup>. On the other hand, inflation is difficult to forecast (see e.g. [Stock and Watson, 2007](#)) and any systematic improvements are useful. The gains from incorporating inflation expectations into models are occasionally statistically significant, with the relative performance of models with and without expectations changing substantially over time. In particular, in the low inflation period expectations seem to help to correct the upward forecast bias from models assuming a constant mean of inflation. They also appear to make the model forecasts more robust, in particular in the environment of high volatility that characterises the most recent sample. Finally, the improvements from including the expectations appear more erratic at the tails, suggesting that they might be more indicative for the centre of the predictive distribution of inflation.

What regards measures of expectations of firms and households or those based on swaps - model forecasts typically do not benefit from incorporating such information. This is different from what has been found in some other studies (see e.g. [Basselier et al., 2018](#); [Moretti et al., 2019](#); [Álvarez and Correa-López, 2020](#)) and deserves further analysis.<sup>7</sup>

Finally, the horse race of models delivers a clear message that points to the “supremacy” of inflation expectations of professional forecasters when forecasting euro area headline HICP inflation in the pre-pandemic sample, in that their predictions can be considered as a benchmark very hard to beat by sophisticated econometric models, at least in terms of point forecast. This is in line with the findings of [Ang et al. \(2007\)](#) and [Faust and Wright \(2013\)](#) for the US and

---

<sup>6</sup>One reason for worse relative performance of models with expectations for core inflation is that the professional forecast we use refer to headline HICP. Expectations based on the SPF are also available for HICP excluding food and energy, however the short sample available prohibits meaningful evaluations at the moment.

<sup>7</sup>The difference with respect to [Álvarez and Correa-López \(2020\)](#) could be related to the fact that we evaluate the usefulness of the measures in forecasting out of sample. For example, measures of inflation expectations of households in the euro area exhibit a high contemporaneous correlation with actual inflation, which could explain their good performance in conditional in-sample forecasting exercises.

[Grothe and Meyler \(2015\)](#) and [Bańbura et al. \(2021\)](#) for the euro area. By contrast, inflation expectations are not always more accurate than model forecasts in more recent sample or for the countries considered. The latter might explain the more “erratic” improvements when incorporating the expectations in country models.

To conclude, inflation expectations of professional forecasters appear to contain useful information or judgment that should be used to complement the information from other predictors of inflation when producing model-based inflation forecasts. More generally, the results suggest that policy makers should pay attention to developments in those measures of expectations. In that sense, our paper also contributes to the resurrecting debate whether or not inflation expectations matter for inflation ([Rudd, 2021](#)).

The rest of this paper is organised as follows. Section 2 describes existing studies analysing the usefulness of inflation expectations for forecasting inflation. Section 3 provides details on the set of models used to forecast inflation. Section 4 presents the results when forecasting euro area inflation. Section 5 shows the forecasting results when focusing on individual euro area countries. Section 6 focuses on the (post-)pandemic period. Section 7 concludes. A detailed description of the data set and additional results are provided in the appendix.

## 2 Related literature

Existing studies, with few exceptions, report gains from using information from observed measures of inflation expectations in time series or reduced form models applied to forecasting inflation. The models typically belong to a Bayesian (V)AR or a Phillips curve (possibly non-linear or embedded in a bigger model) family. The expectations are mostly those of professional forecasters although some selected studies also consider those of firms and consumers or those based on financial market prices. In terms of how expectations are used, they serve: i) as “boundary” values (nowcasts and long-term “anchors” or trends) ii) as explanatory variables iii) to tilt or constrain the model forecasts and/or iv) to inform the model parameters.

[Faust and Wright \(2013\)](#) compare the forecasting performance of a large set of different models for United States inflation and show that nowcasts and long-term predictions from subjective forecasts (such as from the Blue Chip survey or from the SPF) provide very good “boundary values” for models, in particular that a simple autoregressive “glide path” between the survey



assessment of inflation in the current quarter and the long-term survey forecast value is very hard to beat. [Clark and Doh \(2014\)](#) report good forecasting performance of models in which trend inflation is proxied by long-term SPF forecasts compared to alternative specifications. [Chan et al. \(2018\)](#) show that long-term Blue Chip forecasts help to pin down the inflation trend and therefore improve model fit and forecast accuracy. [Hasenzagl et al. \(2022\)](#) stress the importance of using inflation expectations (consumers' and professionals' one-year-ahead forecasts) to identify trend inflation and the Phillips curve in the US and report significant forecasting gains for both headline and core inflation. [Jarociński and Lenza \(2018\)](#) report that linking the unobserved inflation trend to long-term inflation forecasts from Consensus Economics in a Phillips curve embedded in a dynamic factor model results in improved forecast performance for inflation excluding energy and food in the euro area. [Bańbura and Bobeica \(2023\)](#) for the euro area show good forecasting performance of Phillips curves linking the inflation trend to long-term inflation forecasts from Consensus Economics. Within unobserved component Phillips curve models, [Stevens and Wauters \(2021\)](#) find that imposing a common trend for euro area inflation and its SPF forecasts tends to improve the out-of-sample forecasting performance whereas [Basselier et al. \(2018\)](#) conclude that qualitative business price expectations from European Commission surveys provide useful information for inflation forecasts in both the euro area and in Belgium. [Chan and Song \(2018\)](#) find that financial market prices help to pin down the uncertainty around US inflation trend but not the trend itself.

[Stockhammar and Österholm \(2018\)](#) show that both short-run and long-run survey inflation expectations improve the forecasting performance of Swedish inflation when included in a BVAR. [Moretti et al. \(2019\)](#) apply dynamic model averaging to a large number of Phillips curve models and on the basis of inclusion probabilities conclude that inflation expectations, in particular those based on inflation-linked swap rates, have been the single most important determinant of euro area core inflation since 2001. [Álvarez and Correa-López \(2020\)](#) find that expectations of consumers and firms lead to more accurate conditional inflation forecasts compared to professional forecasts and expectations based on financial market prices. [Kulikov and Reigl \(2019\)](#) also in a conditional forecast framework show that inflation expectations and in particular those based on financial market prices, explain a large part of the dynamics of euro area inflation since 2012.

Krüger et al. (2017) find that tilting the starting point of forecasts from BVARs to SPF nowcasts improves the overall accuracy of such forecasts for the US. Tallman and Zaman (2020) find substantial improvements in inflation forecasts from simple VARs when they are tilted to short- and long-term forecasts from the SPF in the US. Ganics and Odendahl (2021) find gains from using the one- and two-year-ahead expectations from the euro area SPF in BVARs via tilting and soft conditioning. Bańbura et al. (2021) analyse for euro area data how to best combine subjective forecasts from the SPF and model forecasts from several BVARs and recommend tilting the model forecasts only to the first moments of the SPF (thus ignoring the information from the second) prior to performing forecast combination. Galvao et al. (2021) also find improvements in forecast accuracy when tilting model forecasts to the mean of professional forecasts for output growth and inflation in the UK.

Wright (2013) shows gains in forecasting performance from using long-term Blue Chip forecasts as priors for BVAR steady states. Frey and Mokinski (2016) use the US SPF nowcasts to inform the parameters of a VAR and report better forecasting performance compared to a VAR not using such information.

Regarding studies that find less role for inflation expectations in explaining inflation, one example is Forbes et al. (2021) who argue that commodity prices and the exchange rate are more important for inflation in the United Kingdom. Cecchetti et al. (2017) are even more forcefully negative and state that inflation expectations have no effect on inflation once a local mean of inflation is taken into account.

In this paper we do not evaluate the advantages of entropic tilting as this is extensively analysed for a similar set of models by Bańbura et al. (2021). We also do not consider a “glide path” model here as short-term (current quarter) inflation expectations are not available for our “main” measure of inflation expectations for the euro area (the SPF). Finally, we only use the first moment (mean) of the expectations given the findings of Bańbura et al. (2021) and Galvao et al. (2021) (see also Clements, 2014, 2018).

### 3 Empirical framework

The purpose of this section is threefold. First, we describe the wide range of models to forecast inflation used in this paper. In order to answer our main research question, for each specification

we construct two versions: one that includes information on inflation expectations, and another version that does not incorporate such information. As the first robustness check and in order to cover specifications often used in the literature, for each model we employ two alternative specifications: one that only includes information on inflation and another one that contains information on inflation along with other macroeconomic variables. Second, we describe the data employed to estimate the models. Third, we provide information about the design of the real-time forecasting exercises and the evaluation metrics.

### 3.1 Models

Let  $\pi_t = 400 \times \ln\left(\frac{P_t}{P_{t-1}}\right)$  denote the annualised quarter-on-quarter inflation rate, where  $P_t$  is the appropriate price index, expressed at the quarterly frequency. Further, let  $\pi_t^A = \frac{1}{4} \sum_{i=0}^3 \pi_{t-i} = 100 \times \ln\left(\frac{P_t}{P_{t-4}}\right)$  denote the annual inflation rate and  $\pi_t^{Exp}$  the expectation of  $\pi_{t+h}^A$  given the information up to  $t$  (we drop the reference to the horizon of the expectations to simplify the notation).

The models employed to provide forecasts of  $\pi_t$  are listed in Table 1, and are detailed as follows:

#### 1. Autoregressive Distributed Lag (ADL) models with time-varying trend inflation

Let  $\hat{\pi}_t = \pi_t - \bar{\pi}_t$  denote the inflation gap, where  $\bar{\pi}_t$  is the inflation trend. The first model is specified as follows:

$$\hat{\pi}_t = \alpha \hat{\pi}_{t-1} + \beta y_t + \nu_t, \quad \nu_t \sim N(0, \sigma^2), \quad (1)$$

where  $\alpha$  and  $\beta$  denote the autoregressive and slope coefficients, respectively, and  $y_t$  is the output gap.

- In the version *not incorporating inflation expectations* we explore two variants in defining  $\bar{\pi}_t$ . First, the inflation trend is assumed to be constant and given by the sample mean ( $\bar{\pi}_t \equiv \mu_\pi$ ), denoted by ‘M’ in Table 1. Second, trend inflation is defined by the exponentially-weighted moving average (EWMA) of past inflation ( $\bar{\pi}_t = \phi \sum_{j=0}^{\infty} (1 - \phi)^j \pi_{t-j}$ ) with a “smoothing” parameter  $\phi$ , denoted by ‘E’.<sup>8</sup>

---

<sup>8</sup>In the forecasting exercises the parameter  $\phi$  is set equal to 0.05 as in Bańbura and Bobeica (2023).

Table 1: Overview of modelling approaches

Not incorporating inflation expectations	Incorporating inflation expectations
1. <i>ADL models with time-varying trend inflation</i>	
1a. Model includes only inflation rate	
1b. Model includes inflation rate and output gap (Phillips curve)	
Trend is constant, captured by the mean ('M') or trend is EWMA of past inflation ('E')	Trend is captured by long-term inflation expectations
2a. <i>ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility</i>	
2a. Model includes only inflation rate	
2b. Model includes inflation rate and output gap (Phillips curve)	
Trend is a random walk	Trend is a random walk linked to long-term inflation expectations via a measurement equation
3. <i>Bayesian VARs with democratic priors and stochastic volatility</i>	
3a. Model includes only inflation rate	
3b. Model includes inflation rate, real GDP growth and short-term interest rate	
The priors on the unconditional mean are loose	The mean of the prior on the unconditional mean is given by long-term expectations We use standard ('S') and tight ('T') priors
4. <i>Bayesian VARs with time-varying trends and stochastic volatility</i>	
4a. Model includes only inflation rate	
4b. Model includes inflation rate, real GDP growth and short-term interest rate	
Trend is a random walk	Trend is a random walk linked to long-term expectations via a linear measurement equation
5. <i>Phillips curves with constant coefficients</i>	
Backward looking Phillips curve	Hybrid Phillips curve, including one-year-ahead inflation expectations
6. <i>Bayesian VARs with Minnesota priors and stochastic volatility</i>	
6a. Model includes only inflation rate	
6b. Model includes inflation rate, real GDP growth and short-term interest rate	
	Long- ('L') or short-term ('S') inflation expectations are included as endogenous variables

- In the version *incorporating inflation expectations* the inflation trend is given by long-term inflation expectations ( $\bar{\pi}_t = \pi_t^{Exp}$ ). For the HICP excluding energy and food the trend is adjusted by the difference of historical means of the expectations and of the target variable ( $\bar{\pi}_t = \pi_t^{Exp} - (\mu_{Exp} - \mu_{\pi})$ ) and corrects for the fact that the expectations concern headline inflation and that inflation excluding energy and food has been systematically lower over the sample considered (bias correction).<sup>9</sup>

<sup>9</sup>The bias corrected version of the specification results in higher forecast accuracy than the uncorrected version. The means are computed in real time by only using the data available at the respective point of the evaluation sample.

We consider a specification that only includes information on inflation, where  $\beta = 0$ , (referred to as 1a. in Table 1) and a specification that incorporates information on inflation and the output gap (1b.). Note that the latter specification can be thought of as a backward looking Phillips curve for the inflation gap.

Equation (1) is estimated and the forecasts are simulated using Bayesian techniques. The prior for the coefficients and the variance of the residuals is normal-inverse-gamma with Minnesota-type settings. The inflation trends are assumed constant over the forecast horizon and are added back to the forecasts of the inflation gaps to obtain inflation forecasts. Such models have been previously used for forecasting inflation in e.g. [Faust and Wright \(2013\)](#) or [Bańbura and Bobeica \(2023\)](#).

## 2. ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility

The second model, proposed by [Chan et al. \(2018\)](#), represents a generalisation of the first model where both, the coefficients and the residual variance, are allowed to exhibit changes over time:

$$(\pi_t - \bar{\pi}_t) = \alpha_t(\pi_{t-1} - \bar{\pi}_{t-1}) + \beta_t y_t + \nu_t, \quad \nu_t \sim N(0, \sigma_{\nu,t}^2), \quad (2)$$

$$\bar{\pi}_t = \bar{\pi}_{t-1} + e_t, \quad e_t \sim N(0, \sigma_{e,t}^2). \quad (3)$$

The coefficients and log volatility of the residuals are assumed to follow random walks. Also, the inflation trend follows a random walk, as specified in Equation (3).

- In the version *not incorporating inflation expectations* no further equations are included.
- In the version *incorporating inflation expectations* the inflation trend is also linked to the long-term inflation expectations via a measurement equation with time-varying coefficients:

$$\pi_t^{Exp} = a_t + b_t \bar{\pi}_t + u_t, \quad u_t \sim N(0, \sigma_{u,t}^2). \quad (4)$$

Note that in the latter version, inflation expectations are allowed to be a biased measure of the inflation trend since the intercept and slope coefficients in Equation (4) are not restricted to be  $a_t = 0$  and  $b_t = 1$ , respectively.

Similarly to the first model, we consider two alternative specifications, one without information on output gap, where  $\beta_t = 0$  (2a. in Table 1), and another specification that includes data on output gap (2b.).

The estimation is carried out in a Bayesian setting following Chan et al. (2018).<sup>10</sup> Previous work by Chan et al. (2018) and Bańbura and Bobeica (2023) has reported good forecasting performance of this model for US and euro area inflation, respectively.

### 3. Bayesian VARs with democratic priors and stochastic volatility

This model consists of a vector autoregression where the priors are chosen to line up model’s long-term forecasts with long-term (inflation) expectations (see Wright, 2013). In doing so, the VAR is specified for the variables in deviation from their unconditional mean,  $\mu$ , sometimes referred to as the “steady state”:

$$y_t - \mu = \sum_{i=1}^p B_i (y_{t-i} - \mu) + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \quad (5)$$

where the log volatilities of the shocks follow random walks (as in Clark, 2011, see also below).

- In the version *not incorporating inflation expectations* the priors on  $\mu$  are loose.
- In the version *incorporating inflation expectations* at each point of the evaluation sample the mean of the prior on  $\mu$  is set to the long-term inflation expectation from the latest available survey at that point in time. We consider the standard setting for the variance of the prior (denoted by ‘S’ in Table 1) as well as very tight priors (denoted by ‘T’). For the case of HICP excluding energy and food the prior is adjusted for the difference in historical averages, similarly as in Model 1.

We consider a univariate specification of the model with *democratic* priors that only includes data on inflation,  $y_t = \pi_t$ , (3a. in the Table 1) and a multivariate specification where  $y_t$  contains data on real GDP growth, inflation and the short-term interest rate (3b.). In the latter case, the prior for the short-term interest rate is non-informative, as expectations data of sufficient length is not available. For GDP growth we use the corresponding long-term expectations. Three-variable VARs including a measure of real

---

<sup>10</sup>We use the codes provided by Joshua Chan on his website.

activity, of inflation and a short-term interest rate have often been used to analyse and forecast inflation (see e.g. [Cogley and Sargent, 2002](#); [Cogley et al., 2010](#)).<sup>11</sup>

The settings of the standard priors for  $\mu$  and the estimation follows [Villani \(2009\)](#) and [Clark \(2011\)](#).<sup>12</sup> Also, we assume Minnesota-type priors for the autoregressive coefficients,  $B_i$ , see below.

#### 4. Bayesian VARs with time-varying trends and stochastic volatility

The VAR model is specified for the variables in deviation from their “local” mean, which is allowed to vary over time as a random walk:

$$y_t - \mu_t = \sum_{i=1}^p B_i (y_{t-i} - \mu_{t-i}) + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \quad (6)$$

$$\mu_t = \mu_{t-1} + \eta_t, \quad \eta_t \sim N(0, V_t). \quad (7)$$

- In the version *not incorporating inflation expectations* no further equations are included.
- In the version *incorporating inflation expectations* the local mean is linked to the long-term expectations for GDP growth and inflation via a measurement equation:

$$y_t^{Exp} = \mu_t + g_t, \quad g_t \sim N(0, G_t). \quad (8)$$

Similarly as in Model 3, we consider a univariate specification of the model that includes only data on inflation,  $y_t = \pi_t$  (4a. in the Table 1), and a multivariate specification with  $y_t$  containing data on real GDP growth, inflation and the short-term interest rate (4b.).

The log volatilities of the shocks are assumed to follow random walks (the matrices  $V_t$  and  $G_t$  are diagonal, see below for the details on  $H_t$ ). Also, the priors for the autoregressive coefficients,  $B_i$ , are Minnesota-type. The settings of the priors and estimation follows [Bańbura and van Vlodrop \(2018\)](#), who document good forecasting performance of this model compared to other VAR specifications. Similar models were proposed by [Garnier et al. \(2015\)](#), [Crump et al. \(2016\)](#), [Mertens \(2016\)](#) and [Del Negro et al. \(2017\)](#).

---

<sup>11</sup>Unemployment rate rather than GDP growth is often used as a measure of real activity. We use this variable as a robustness check.

<sup>12</sup>More precisely, the loose, standard and tight priors correspond to the prior variance for  $\mu$  of 1000, 0.05 and 0.005, respectively.

## 5. *Phillips curves with constant coefficients*

We also use a similar version of Model 1, where instead of letting long-term inflation expectations influence the inflation trend, we incorporate short-term inflation expectations as an additional regressor in the forecasting equation. Precisely, we consider the following version of the Phillips curve:

$$\pi_t = c + \alpha\pi_{t-1} + \beta y_t + \gamma\pi_t^{Exp} + \nu_t, \quad \nu_t \sim N(0, \sigma^2), \quad (9)$$

where, in this case,  $\pi_t^{Exp}$  denotes short-term (one-year-ahead) inflation expectations.

- In the version *not incorporating inflation expectations* the coefficient  $\gamma$  in Equation (9) is set to zero.
- In the version *incorporating inflation expectations* no modifications to Equation (9) are made.

The estimation approach is the same as for Model 1 in that Bayesian methods are employed.

It should be pointed out that this formulation, further augmented by a supply shock proxy, has been previously used in different studies to understand the drivers of inflation, see [IMF \(2013\)](#), [Ciccarelli and Osbat \(2017\)](#), [Bobeica and Sokol \(2019\)](#) or [Moretti et al. \(2019\)](#).

Given recent interest in analysing risks to (future) inflation by means of quantile regressions (see e.g. [López-Salido and Loria, 2020](#); [Banerjee et al., 2020](#); [Korobilis et al., 2021](#)), for robustness, we also compare performance of this model to a Bayesian quantile regression counterpart, where  $\nu_t$  has the asymmetric Laplace distribution, see [Yu and Moyeed \(2001\)](#) and [Kozumi and Kobayashi \(2011\)](#).<sup>13</sup>

## 6. *Bayesian VARs with “Minnesota” priors and stochastic volatility*

We also include in our set of competing models standard BVARs, which are typically used in macroeconomic applications:

$$y_t = c + \sum_{i=1}^p B_i y_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \quad (10)$$

---

<sup>13</sup>We thank Matteo Mogliani and the members of the ESCB Expert Group on Macro at Risk for sharing their codes for this model.



where the intercept and autoregressive coefficients are assumed to remain constant, while the log volatilities of the shocks vary over time following random walks.

- In the version *not incorporating inflation expectations* no further variables are included.
- In the version *incorporating inflation expectations* data on either short- (denoted by ‘S’) or long-term (denoted by ‘L’) inflation expectations are included to the vector  $y_t$ .

We consider a specification of the model that includes only inflation,  $y_t = \pi_t$ , (6a. the Table 1) and a specification where  $y_t$  contains real GDP growth, inflation and the short-term interest rate (6b.). The settings of the priors and estimation follows [Bańbura and van Vlodrop \(2018\)](#).

In a recent work, [Stockhammar and Österholm \(2018\)](#) find that inclusion of inflation expectations in BVARs tends to improve forecast precision for Swedish inflation.

## 7. Benchmarks

Lastly, we also employ a couple of widely used *benchmark* models to forecast inflation. The first *benchmark* is the unobserved components stochastic volatility model (UCSV) of [Stock and Watson \(2007\)](#):

$$\begin{aligned}\pi_t &= \tau_t + \varepsilon_t, \\ \tau_t &= \tau_{t-1} + \eta_t,\end{aligned}$$

where  $\tau_t$  is the permanent component of inflation, or the trend, and  $\varepsilon_t$  and  $\eta_t$  are characterised by stochastic volatility.<sup>14</sup> The forecast from this model is given by the estimate of the trend:  $\pi_{t+h|t} = \tau_{t+h|t}$ , and is obtained by Bayesian simulation.

The second *benchmark* is the random walk (RW) model of [Atkeson and Ohanian \(2001\)](#):

$$\pi_{t+h|t} = \pi_t^A,$$

where the forecast is set to the latest observed annual inflation rate.

<sup>14</sup>More precisely we adopt the non-centered parameterisation of the UCSV model where  $\varepsilon_t = \exp(h_0 + \omega_h h_t)\tilde{\varepsilon}_t$ ,  $h_t = h_{t-1} + u_t$  and  $\tilde{\varepsilon}_t$  and  $u_t$  are  $N(0,1)$ . Analogous assumptions are taken for  $\eta_t$ , see [Chan \(2018\)](#).

For all the BVAR models we use independent normal priors for the coefficients  $B_i$ . The prior means are equal to 0. Following the “Minnesota” convention, the coefficients for more distant lags are “shrunk” more (have tighter priors around 0). The prior variances are also adjusted for relative differences in predictability. The overall degree of shrinkage is set to the standard value of 0.2. The draws of the coefficients  $B_i$  are obtained equation by equation as suggested by [Carriero et al. \(2019\)](#) with the correction in [Carriero et al. \(2021\)](#).<sup>15</sup> The prior for the intercept  $c$  (where applicable) is non-informative.<sup>16</sup> Similar convention is applied for models 1 and 5. The time-varying variances in the BVAR models are parameterised as  $H_t = A^{-1}\Lambda_t(A^{-1})'$ , where  $\Lambda_t = \text{diag}(\sigma_{\epsilon,1,t}^2, \dots, \sigma_{\epsilon,N,t}^2)$ ,  $\log(\sigma_{\epsilon,i,t}^2)$  follow independent random walks and  $A$  is a lower diagonal matrix with ones on the diagonal.

Note that in models 1-4 the long-term expectations inform the evolution of the low-frequency movements (trends) of the variables. Instead, in models 5-6 information on either short- or long-term expectations are used as additional explanatory variables.

## 3.2 Data

The variety of models described in Section 3.1 uses data on headline inflation, core inflation (defined as headline inflation excluding food and energy components), short- and long-term inflation expectations, real GDP, real GDP growth (short- and long-term) expectations, and the short-term interest rate. We focus the following description on the euro area. The construction of data sets for the euro area countries follows similar steps, however data availability is often more limited and non-public data sources are used in certain cases.

To simulate the environment faced by policy makers and forecasters in practice and to appropriately assess the information content of various indicators, the exercises rely on real-time data. The cut-off dates are those for the ECB’s Survey of Professional Forecasters (SPF) over 2001Q1-2021Q3. The data for expectations are unrevised. For the remaining series, we mainly rely on the ECB’s real-time data base (RTDB)<sup>17</sup>. We use seasonally (and working day) adjusted data on HICP and GDP. As the data for seasonally adjusted headline HICP are not available in the RTDB, we use the real-time vintages stored in the ECB’s Statistical Data Warehouse as

<sup>15</sup>We thank these Authors for sharing their code for the corrected algorithm.

<sup>16</sup>See e.g. [Kadiyala and Karlsson \(1997\)](#), [Bańbura et al. \(2010\)](#) and [Carriero et al. \(2019\)](#) for more details on this type of models.

<sup>17</sup>See [Giannone et al. \(2012\)](#) and [RTDB in ECB’s Statistical Data Warehouse](#)

of 2006 and for earlier vintages we seasonally adjust the data obtained from the RTDB using X11. Further, as data for core inflation are not available in the RTDB, we use ECB’s Statistical Data Warehouse (SDW) as of 2006 and we construct pseudo real-time data for earlier vintages. If a “full” quarter of data is not available for a monthly series we take an average of available months. For the alternative measures of inflation expectations we assume the same cut-off dates. The output gap is obtained by applying (in real time) the Christiano-Fitzgerald filter ([Christiano and Fitzgerald, 2003](#)) to log real GDP, where we keep the cycles shorter than 15 years.<sup>18</sup>

The data on GDP and the short-term interest rate have been backdated to 1970 using the Area Wide Model (AWM) data base ([Fagan et al., 2005](#)). The data for the SPF expectations have been backdated using Consensus Economics forecasts and go back to 1990. For the latter, the forecasts for the euro area prior to 2003 are obtained by aggregating the available forecasts for the countries. Table A in the appendix provides the details.

Figure 1 plots the measures of inflation and inflation expectations for the euro area considered in this paper. We can note different properties of different measures of expectations. In particular, long-term expectations from the SPF and Consensus Economics (right and left upper panel, respectively) decline from little above 3% in 1990 to somewhat below 2% in 1999 and vary little thereafter, largely reflecting low frequency movements of inflation. By contrast, shorter-term inflation expectations (from the SPF, Consensus Economics and household and firm expectations collected in a survey of the European Commission (in the lower left panel)) follow inflation movements more closely, but the correlation appears more contemporaneous. The expectations based on five-year, five-year forward inflation linked swap rates (lower right panel) lie somewhere in between - they are more volatile than the long-term expectations from the surveys and less volatile than the shorter-term expectations. None of the measures appears to indicate the recent surge in inflation well in advance.

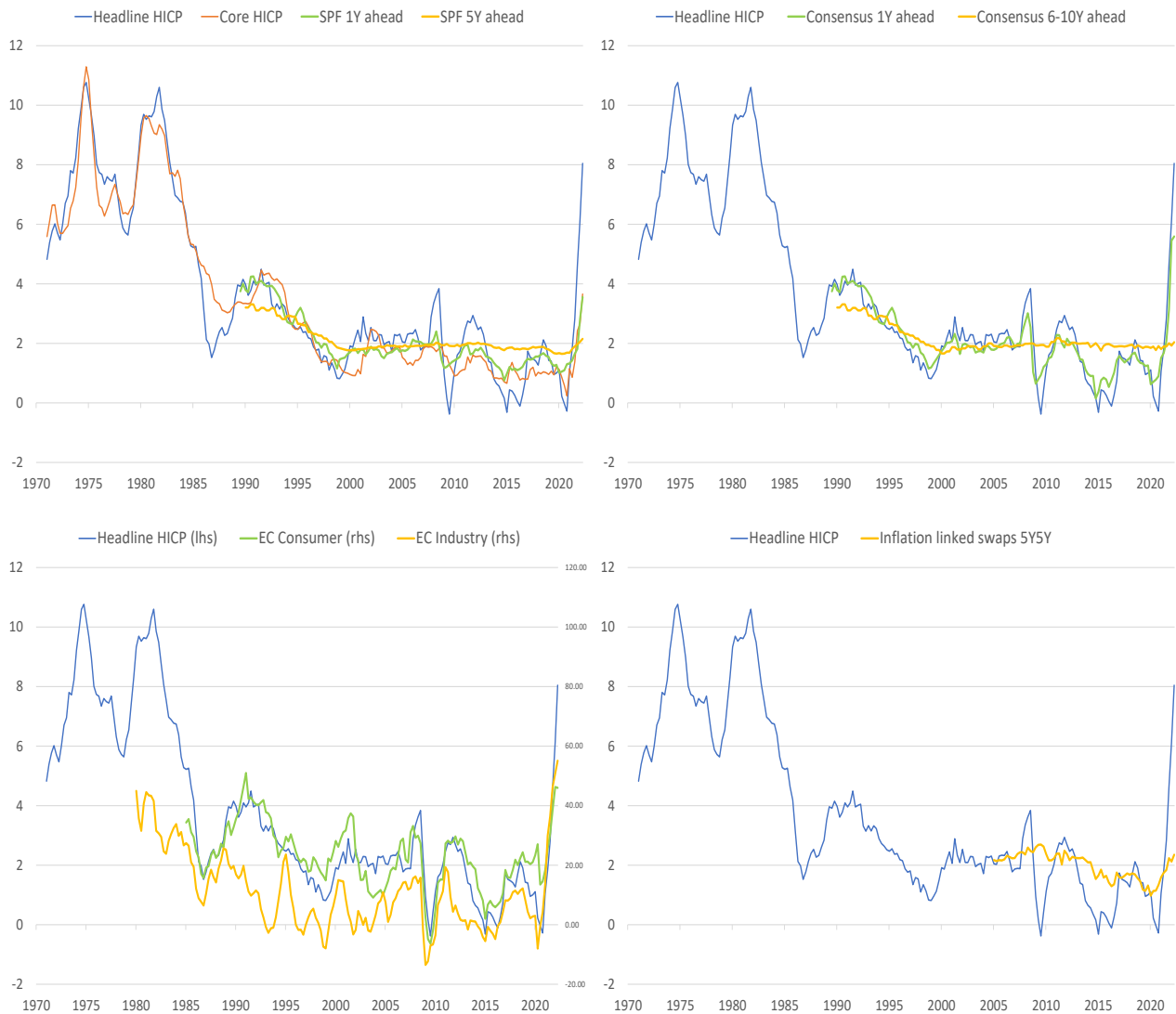
### 3.3 Real-time forecasting design

We consider two alternative target variables to forecast: the annual inflation rate based on headline HICP and the annual inflation rate based on HICP excluding energy and food components. As explained above the models are estimated with data at the quarterly frequency, employing

---

<sup>18</sup>This measure of economic slack has performed well compared to several alternatives in an extensive forecast evaluation of Phillips curve models undertaken by [Bańbura and Bobeica \(2023\)](#).

Figure 1: Euro area inflation and inflation expectations



Note: EC denotes the expectations from the European Commission surveys. For headline and core inflation annual rates of change are plotted.

the annualised quarter-on-quarter inflation rates. Consequently, the forecasts for the target variable are obtained by taking an average of the appropriate quarter-on-quarter inflation rate forecasts:  $\pi_{t+h|t}^A = \frac{1}{4} \sum_{i=0}^3 \pi_{t+h-i|t}$ .

For each of the real-time vintages we produce forecasts from the models described in Section 3.1. The target forecast period matches that of the respective one-year-ahead and two-year ahead inflation expectations in the SPF. Forecasts are obtained by simulation from the posterior

distributions of the parameters (including the volatilities) and the residuals. The point forecasts are taken as the median of the predictive distribution. As we evaluate the forecasts with real-time data, we have to deal with the “ragged edge” of the vintages. We simulate the parameters based on a “balanced” data set and we take the ragged edge into account when simulating the forecasts.<sup>19</sup> More in detail, the forecasts  $h$ -steps-ahead are obtained in an *iterative* fashion. For models 1, 2 and 5 the explanatory variables are first forecast with an AR(4) process.<sup>20</sup> Then we iteratively obtain forecasts for  $\pi_{t+i}$ ,  $i = 1, \dots, h$ . For models 3, 4 and 6 we cast the VARs in a state space representation and we generate the forecasts “conditional” on the ragged edge using the simulation smoother of [Durbin and Koopman \(2002\)](#) (see e.g. [Bańbura et al., 2015](#)). In models 1 and 2 it is assumed that the long-term expectations remain constant over the forecast horizon. The estimation sample starts in 1990 and is recursive, that is, extended at each subsequent point of the evaluation sample.

Our main evaluation criterion is the Root Mean Squared Forecast Error (RMSFE). However, we also evaluate the density forecasts by means of the Continuous Ranked Probability Score (CRPS)<sup>21</sup> and quantile scores, and investigate how relative forecast performance changes over time and whether the differences are statistically significant by means of the fluctuation test of [Giacomini and Rossi \(2010\)](#).

Given the available real-time vintages our baseline evaluation period is 2001Q4-2019Q4 for the one-year-ahead horizon and 2002Q4-2019Q4 for the two-year ahead horizon. We also provide a separate forecast evaluation for the one-year-ahead horizon of the period 2020Q4-2022Q2, which is a particular episode characterized by high volatility, caused by the consequences of COVID-19 pandemic and the Russian invasion of Ukraine.

## 4 Forecasting euro area inflation

The purpose of this section is twofold. First, we assess the extent to which accounting for information on inflation expectations in econometric models would help to increase the accuracy

---

<sup>19</sup>“Ragged edge” means that, in a given vintage, the last observation does not fall in the same period for all the variables. For example, we might have GDP only until Q3 but inflation already for Q4. In the “balanced” version we discard the quarters at the end of the sample for which not all the variables are available.

<sup>20</sup>The estimation of these autoregressive models is carried out using standard Bayesian methods with priors as described in Section 3.1.

<sup>21</sup>This scoring rule is less sensitive to extreme outcomes compared to the log predictive score.

of inflation forecasts. Second, we are also interested in evaluating the accuracy of inflation expectations when used directly as forecasts and compared to that of econometric models. To that end, we perform a horse race forecasting exercise that involves all the models described in Section 3.1, and where the target variable is the annual inflation rate.

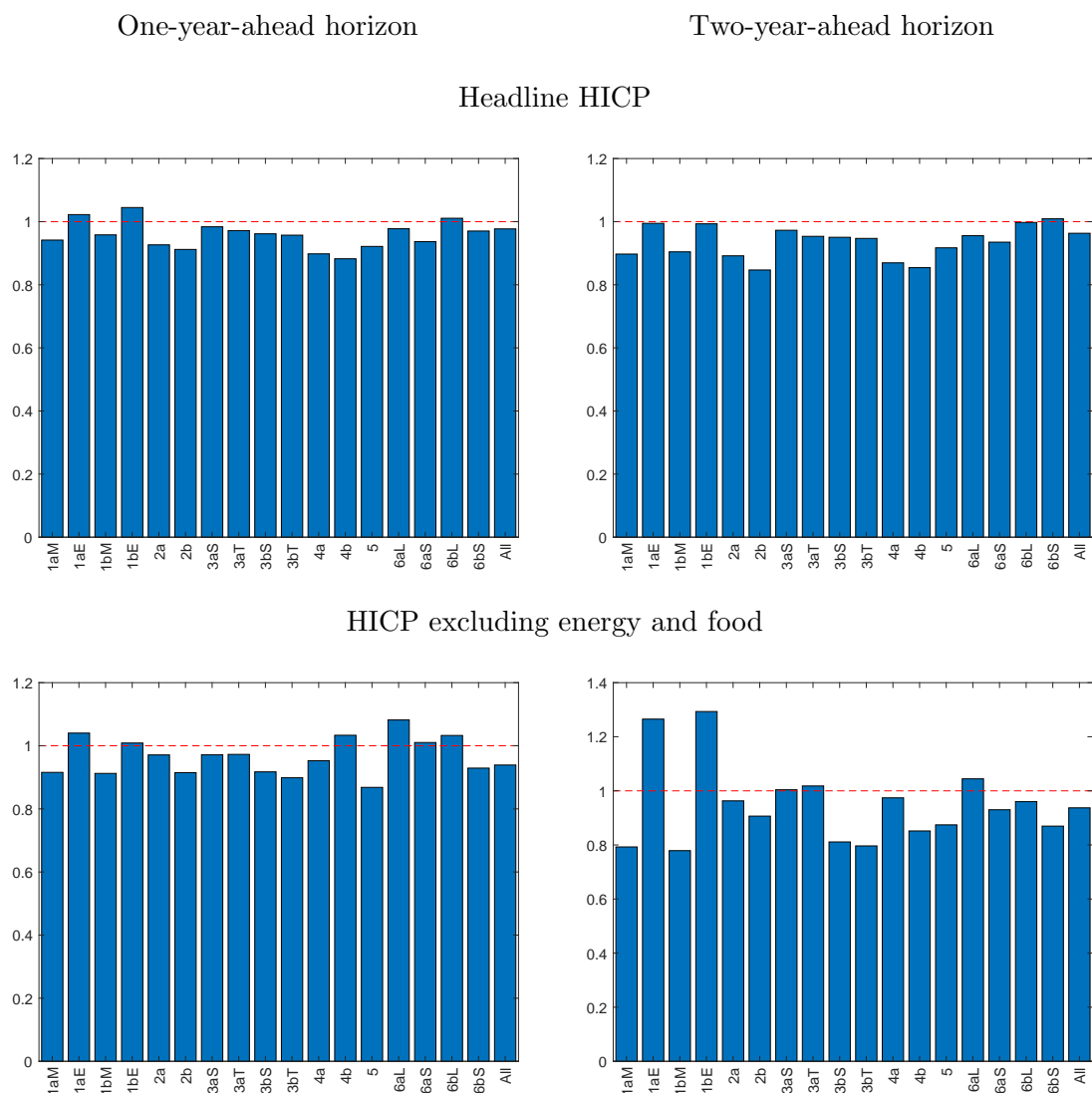
#### 4.1 How helpful are inflation expectations for model-based forecasts?

We begin by evaluating the *relative* predictive ability of all the models under consideration when forecasting euro area inflation based on HICP (*headline* inflation) and on HICP excluding energy and food (*core* inflation). For each model class 1 to 6, as described in Section 3.1, we divide the RMSFE of the version incorporating expectations from the SPF by the RMSFE of the version not incorporating such information, and report this ratio. We also compare the RMSFE of the median forecast of all models incorporating expectations to the RMSFE of the median forecast of all the models not incorporating them. Accordingly, a value of the ratio lower than one indicates that the SPF expectations help to improve forecast accuracy when such information is included into the corresponding model.

Figure 2 presents the *relative* ratios, this information is reported for both the one- and two-year-ahead horizons. For the case of *headline* inflation, relative forecasting accuracy increases for almost all model versions. The gains are modest - up to 10% depending on the model and horizon. The improvements are the largest for the models where the expectations are used to pin down the inflation trend relative to a model where the trend would simply follow a random walk (models 2 and 4). Interestingly, in the ADL model the EWMA appears to capture the trend inflation at least as well as and in many cases better than the expectations (even more so for HICP excluding energy and food for the two-year-ahead horizon). For the case of *core* inflation, incorporating expectations helps for most of the models although whether there is an improvement and its size varies more strongly across model versions and forecast horizons (gains up to 20%). Larger gains are attained for the longer forecast horizon of two years for both variables.

Similar messages overall emerge when evaluating the accuracy of density forecasts provided by the models, which is measured by the CRPS and the quantile scores at 10% and 90% and shown in Figures A1 and A2 in the appendix. At the same time the improvements from including the expectations appear more erratic at the tails, suggesting that they might be more indicative for

Figure 2: Incorporating information from SPF expectations into models, relative RMSFE



Note: The figure shows the RMSFE of the model version incorporating the SPF expectations divided by the RMSFE of the version not incorporating such information. The RMSFE is computed over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

the centre of the predictive distribution of inflation.

Finally, when all models with or without surveys are pooled (model class denoted by ‘All’ in Figure 2) improvements of the former type are rather small. This might suggest that incorporating inflation expectations makes the individual model forecasts more “robust”, a feature

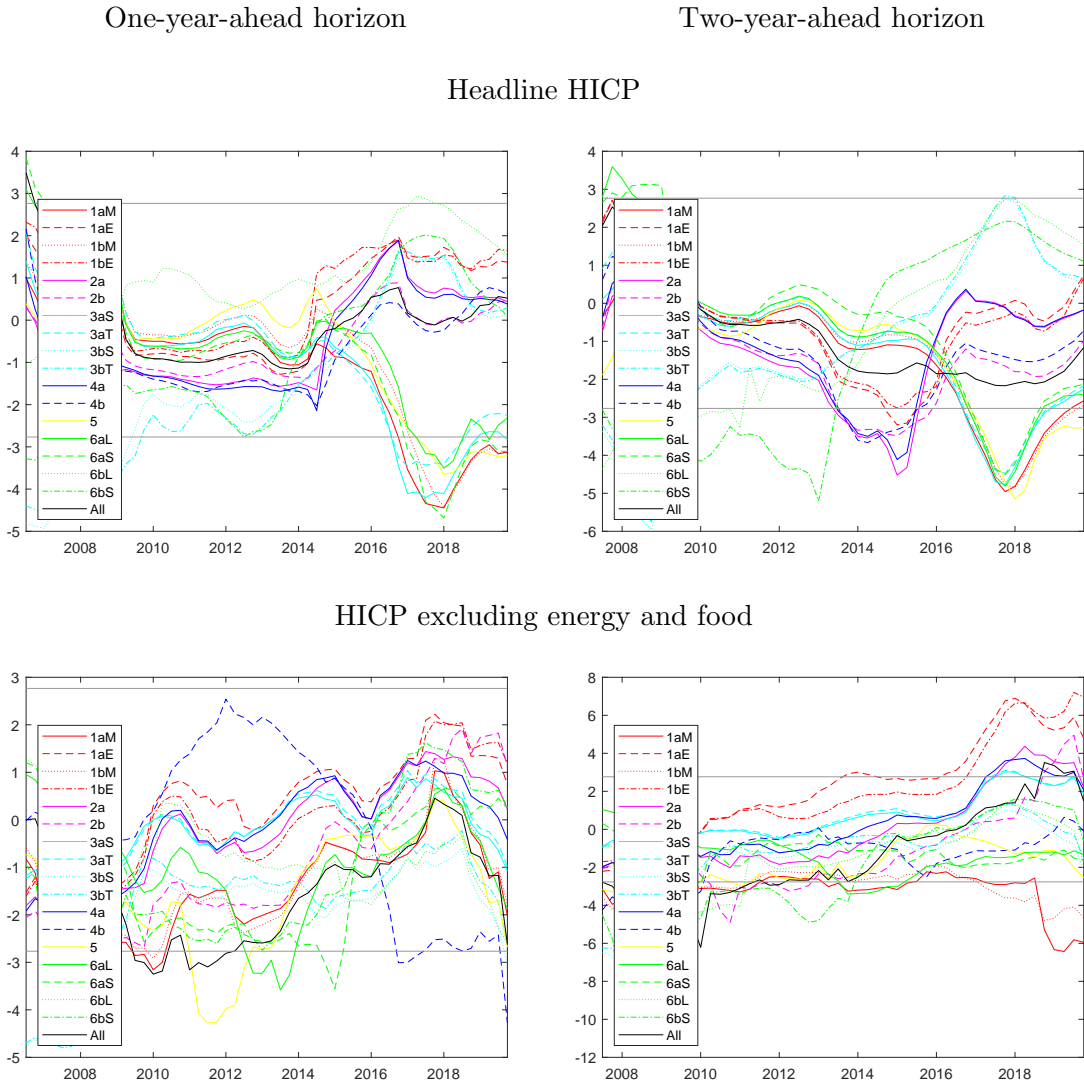
that can be also achieved by pooling. This leads us to our first main result, which is that the incorporation of inflation expectations into the models helps to increase forecasting accuracy, although such help is not large.

To assess in more detail the significance of the differences in forecasting performance and how it evolves over time, we compute the [Giacomini and Rossi \(2010\)](#) fluctuation test statistics – based on a rolling window of 20 quarters – and the associated critical values for headline HICP and for HICP excluding energy and food, respectively. The null hypothesis of equal forecasting performance is rejected when the test statistic is outside the interval given by the critical values. The values of the test statistics below the interval mean that the model that incorporates expectations was performing significantly better than the model that does not (and vice versa for test statistics values above the interval).

Figure 3 shows that the incorporation of inflation expectations into models tends to occasionally provide predictive gains that are statistically significant, and that the periods of such better performance differ across model classes. In particular, for the case of *headline* inflation, the relative predictive gains of some models when including information on inflation expectations have become statistically significant in recent years, this is the case for both one- and two-year ahead forecast horizons (although, note that such predictive gains are more frequent for the longer horizon). It is worth noting that this is the period of low inflation and these models incorporate an assumption of constant inflation mean (models 1a, 3, 5 and 6). In other words, in low inflation period inflation expectations seem to help to correct the upward bias of inflation forecasts based on historical average of inflation. In contrast, for models that explicitly allow for a time-varying mean of inflation (models 1b, 2 and 4) including the expectations leads to a deterioration in relative performance in the recent period, most likely reflecting the upward bias of expectations themselves (see Figure 1). For these models expectations result in better relative performance in earlier years. These observations are in line with the results of [Bańbura and Bobeica \(2023\)](#) for HICP inflation excluding energy. In the case of *core* inflation the patterns are less clear, nevertheless including information on expectations leads to statistically significant improvements in forecast accuracy for some models in some periods. This constitutes our second main result, which suggests that the relative performance of models with and without expectations changes over time and the gains from incorporating inflation expectations are statistically significant in some periods for some models.



Figure 3: Incorporating information from SPF expectations into models, test of relative forecast performance over time



Note: The figure shows the [Giacomini and Rossi \(2010\)](#) fluctuation test statistics for a rolling window of 20 quarters. Grey lines show the critical values for the 90 % confidence interval. The null of equal forecasting performance is rejected when the test statistic is outside the interval. The values of test statistics below the interval mean that a model that incorporates SPF expectations was performing significantly better than a model that does not (and vice versa for test statistics values above the interval). The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

## 4.2 How accurate are model-based forecasts versus inflation expectations as forecasts?

In order to shed some light on this question, we evaluate the *absolute* predictive ability of all the competing models and compare it against the performance of the benchmark models, as described in Section 3.1, and of inflation expectations used directly as forecasts. Figure 4 shows the associated RMSFE. The results for headline HICP show that whereas most models produce more accurate forecasts than the benchmarks (the UCSV and the RW), in terms of the RMSFE, none of them is better than the forecasts produced by the SPF.<sup>22</sup> Hence, our third main result points to a “supremacy” of inflation expectations when forecasting euro area *headline* inflation, at least in terms of point forecast, in that their predictions can be considered as a benchmark hard to beat by sophisticated econometric models. As for the benchmarks, the UCSV is better than the random walk for headline inflation.<sup>23</sup>

In the case of *core* inflation, the SPF forecasts are available only since recently, making comparisons of accuracy difficult. Hence, we proceed to compare the performance of the models only with respect to that of the UCSV and RW benchmarks. Most of the models under consideration produce better forecasts of *core* inflation than both benchmarks, although the gains tend to be relatively small in some cases.<sup>24</sup>

Figure A3 in the appendix shows the RMSFE of the best 10 models for each inflation measure and each forecast horizon: this information is also shown for the case of combination of the models. The best models in terms of forecast accuracy vary with the inflation measure and the forecast horizon. However, they typically correspond to the versions of models incorporating the information from inflation expectations. Also, note that pooling the forecasts from all the models also seems to offer a good hedge against model uncertainty, especially for HICP excluding energy and food, where the performance of the pooled forecast is comparable to that of the best

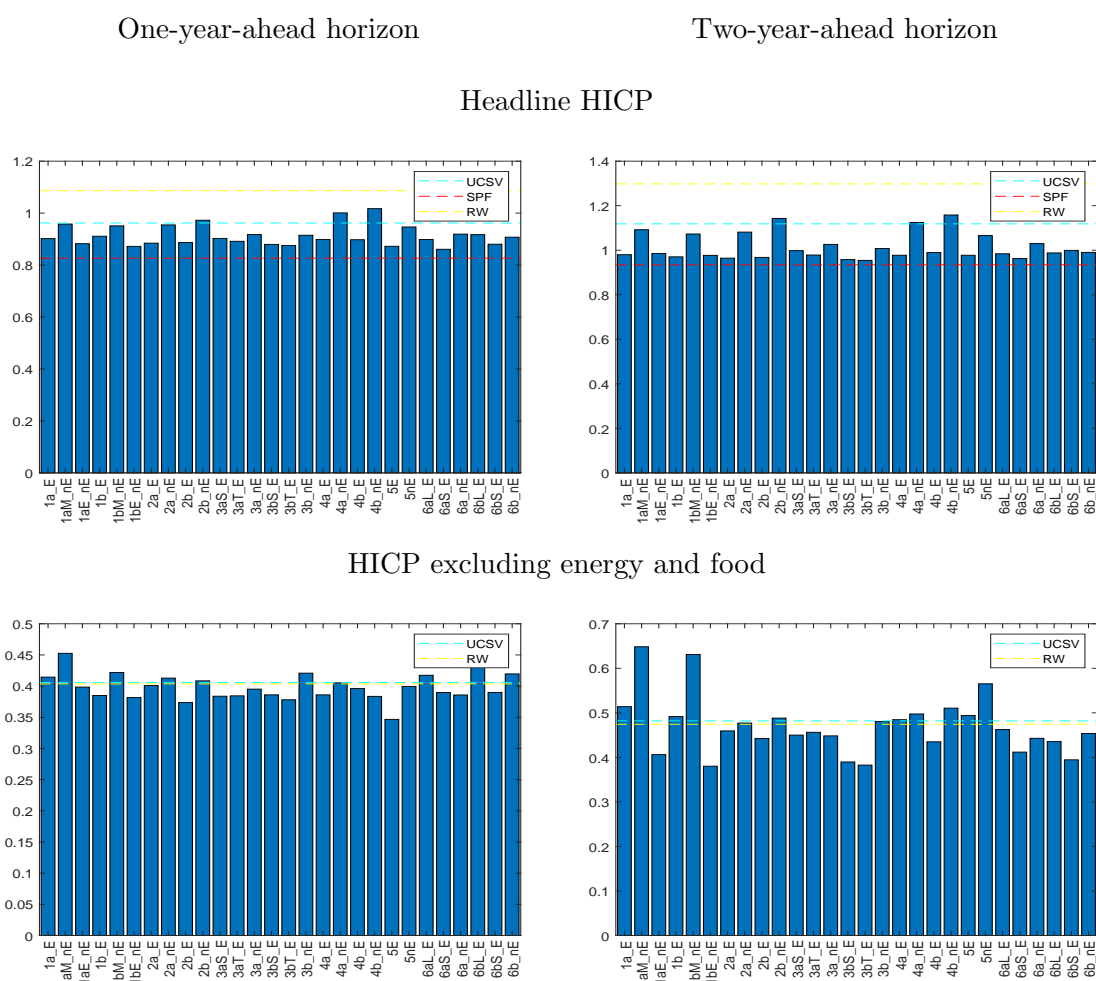
---

<sup>22</sup>Based on the fluctuation test of [Giacomini and Rossi \(2010\)](#) we can conclude that while some models perform significantly better than the UCSV benchmark around the mid-2000s, no model is able to provide significantly better forecasts than the SPF for almost the entire sample period. On the contrary, the performance of the SPF forecasts has been improving relative to most of the models over the sample considered and they tend to perform significantly better than many models during recent years. Compared to the UCSV benchmark the models tend to also produce more accurate *density* forecasts, although, the forecasting gains are more sizeable for the longer forecast horizon. The results are available upon request.

<sup>23</sup>Similar results have been reported by [Stock and Watson \(2007\)](#) and [Bańbura and Bobeica \(2023\)](#).

<sup>24</sup>The same message can be also obtained when evaluating the significance of model’s forecasts improvements, with respect to the UCSV benchmark, based on the [Giacomini and Rossi \(2010\)](#) fluctuation test statistics and when comparing the models’ predictive ability based on density forecast with the CRPS.

Figure 4: All models, RMSFE



Note: The figure shows the (absolute) RMSFE of all the models. The RMSFE is computed over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. 'E' refers to a model version that incorporates the SPF expectations whereas 'nE' denotes a model version that does not. See Section 3.1 and Table 1 for the detailed description of the models.

model ex-post.

The results are robust to changes in model specifications. Particularly, we replaced real GDP by data on unemployment rate and we also added the crude oil price (in euro) to the specifications for headline HICP. Our main results remain unchanged in that (i) inflation expectations help to reduce the RMSFE, although the forecast gains are rather small, and (ii) all models are beaten by SPF inflation expectations when used directly as forecasts.

We also evaluated a restricted version of model 2 where the expectations are assumed to be an unbiased measure of inflation trend.<sup>25</sup> Precisely, in Equation (4) we fix the coefficients to  $a_t = 0$  and  $b_t = 1$ . For headline HICP the forecasts are only slightly less accurate suggesting that long-term SPF expectations have essentially been an unbiased measure of the trend. For core HICP the restrictions lead to sizably worse forecast performance, which is only partly alleviated by correcting the mean of the expectations as discussed above. This indicates that expectations for headline HICP might provide more limited information for core HICP.

Finally we also compare the performance of Model 5 to a specification where the same equation is estimated as a Bayesian quantile regression. The results in Figure A4 suggest that the quantile regression does not offer obvious gains in accuracy, whether for point forecasts or for predictive distributions.

### 4.3 Alternative measures of inflation expectations

We assess the ability of inflation expectations other than the SPF in helping models to increase their forecasting performance. These other measures of inflation expectations are derived from alternative surveys or from financial market prices. In particular, we evaluate inflation expectations delivered by (i) Consensus Economics, (ii) European Commission – both from industry and consumer sides – and (iii) inflation-linked swap rates. For each alternative measure of expectations we repeat the real-time forecasting exercises described in Section 3.3 and estimate the models described in Section 3.1, to assess their relative and absolute forecasting performances. Model sets vary across expectation measures, depending on the available forecast horizons and the length of historical data, see below. The cut-off dates of the real-time vintages are the same as for the exercises reported above.<sup>26</sup>

The results for the alternative measures are presented in Figure 5. For each measure circles and triangles correspond to one- and two-year horizon, respectively. We begin by assessing the extent to which the incorporation of Consensus Economics inflation expectations into models helps to increase their forecast accuracy (red markers). For the case of headline HICP, the figure shows

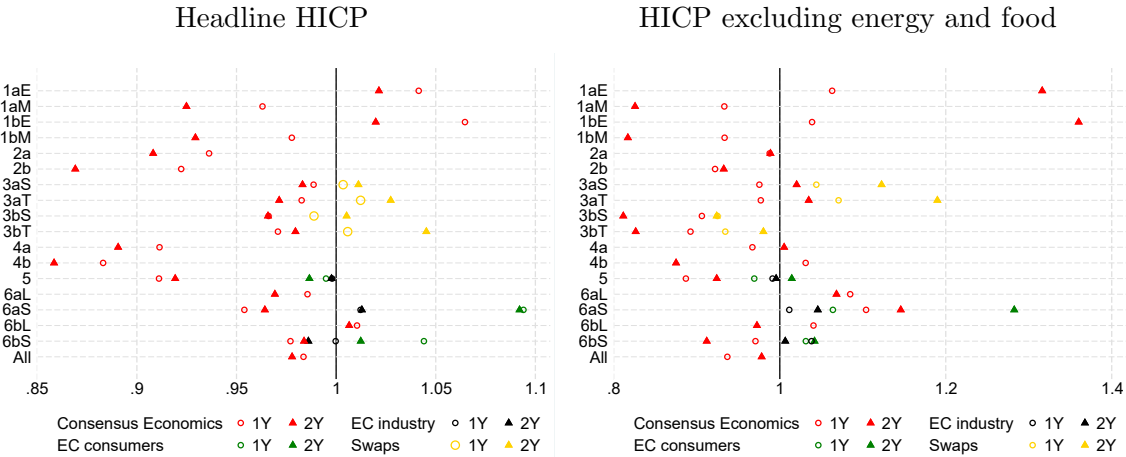
---

<sup>25</sup>We thank an anonymous referee for the suggestion.

<sup>26</sup>Consensus Economics forecast release dates are reasonably close to those of the SPF and we believe that retaining the SPF cut-off dates does not affect the results in a significant manner. For the expectations collected by the European Commission we use the real-time data available in the ECB's SDW as of 2006 and pseudo real-time data before. For the inflation-linked swaps, which are available daily, we take the data available at each SPF cut-off date.

that the incorporation of Consensus Economics inflation expectations helps models to increase their forecast accuracy both for the one- and two-year-ahead horizon, although the gains are not large. Also, note that such gains are slightly smaller when using Consensus Economics than with SPF expectations. For the case of HICP excluding food and energy components, similar messages to those obtained with the SPF emerge in that there are forecast gains associated to the inclusion of Consensus Economics inflation expectations into the models, which are larger and relatively more stable for the longer forecast horizon. In general, the results obtained with Consensus Economics expectations are closely aligned with the ones obtained when using expectations from the SPF.<sup>27</sup>

Figure 5: Incorporating alternative measures of expectations, relative RMSFE



Note: The figure shows the RMSFE of the model version incorporating expectations divided by the RMSFE of the version not incorporating such information. The RMSFE is computed over 2001Q4-2019Q4 for the one-year horizon and over 2002Q4-2019Q4 for the two-year horizon, with the exception of inflation linked swaps for which the respective evaluation samples are 2006Q1-2019Q4 and 2007Q1-2019Q4. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

The European Commission (EC) also provides inflation expectations based on surveys from both industry and consumer sides. The forecast horizons are three and 12 months, for the industry and the consumer survey, respectively. As long-term horizon expectations from these surveys

<sup>27</sup>Similarly as in the case of the SPF, forecasting gains obtained when incorporating Consensus Economics expectations are occasionally statistically significant and change substantially over time, depending on the model and horizon. Further, the best models in terms of forecast accuracy outperform UCSV and RW benchmarks, but none of these models is able to provide more accurate forecasts than the SPF inflation expectations. The results are available upon request.

are not available<sup>28</sup> only models 5 and 6 can be evaluated. It turns out that the incorporation of those expectations into the models does not seem to help improving the forecasts of the latter. The relative RMSFE of the models indicate that there are no gains from incorporating EC industry expectations (black markers), a result that is also validated by the corresponding fluctuation tests (the results are available upon request). Similar results hold when using EC inflation expectations based on consumer survey (green markers). In this case, the incorporation of expectations is even somewhat detrimental to the model-based forecasts. Thus these surveys appear to contain contemporaneous rather than forward looking information on inflation.

Inflation-linked swap rates can be also interpreted as inflation expectations derived from financial market prices. Although these data are available at high frequency, a disadvantage is that they cover a relatively short time span for the euro area. Due to this limitation, we perform the forecast evaluation only with the BVAR models with democratic priors (model 3). In particular, we take the mean of the prior equal to the five-year, five-year forward expected inflation based on the corresponding swap rates (yellow markers).<sup>29</sup> The evaluation period is set to 2006-2019 for one-year-ahead forecasts and to 2007-2019 for two-year-ahead forecasts. Overall, the inclusion of market-based inflation expectations into the models does not provide a significant help in terms of forecast accuracy. The relative RMSFEs are close to one (mostly above one) and the improvements are never statistically significant.<sup>30</sup> It should be noted that the financial market data we use most likely contain other elements apart from inflation expectations, notably risk premia. Evaluation of the usefulness of these data after it has been “corrected” *in real time* for such elements is left for future research.

## 5 Forecasting inflation of individual euro area countries

In this section, we provide a more granular perspective and focus on assessing the extent to which inflation expectations help to improve model-based forecasts of inflation in individual euro area countries. The selected countries are Germany<sup>31</sup>, France, Italy, Spain, the Netherlands,

---

<sup>28</sup>Recently, the ECB and some national central banks have established consumer surveys that also contain long-term inflation expectations making it possible to analyse them in the future.

<sup>29</sup>In the version where also GDP growth is included, we use the SPF expectations for that variable.

<sup>30</sup>For comparison, the relative RMSFEs of models with the SPF are also close to one over this shorter period, but always below one (with the exception of model 3a for core inflation at two-year horizon), occasionally significant and the models including the SPF are always more accurate than their counterparts based on inflation-linked swaps. The results are available upon request.

<sup>31</sup>For Germany, we have used real-time data of the national CPI instead of the HICP due to data limitations.

Belgium, Austria and Finland. As measure of inflation expectations, we take inflation forecasts from Consensus Economics since the SPF is available only at the euro area level and not for the countries. We have transformed the fixed-event (calendar year) Consensus forecasts into fixed-horizon (one- and two-year-ahead) forecasts by computing weighted averages. Also, due to more limited availability of real-time data for some countries, we start forecast evaluations in 2005. Hence, forecast errors are computed over 2005Q4-2019Q4 for the one-year-ahead horizon and over 2006Q4-2019Q4 for the two-year-ahead horizon. In Figure 6, we again plot the RMSFE of models including expectations relative to their counterparts without expectations but on top of that, we also split the results along the country dimension.<sup>32</sup>

---

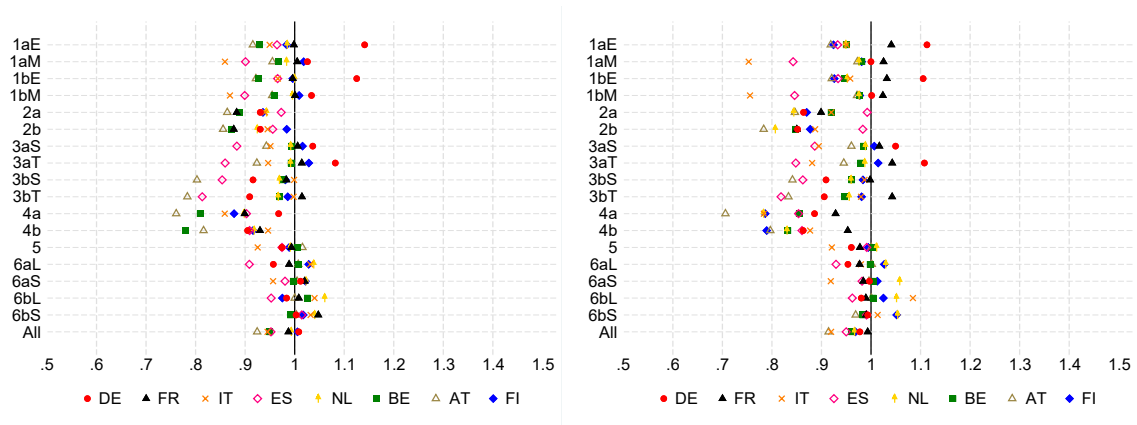
<sup>32</sup>The complete set of results are available upon request.

Figure 6: Country-specific results for headline HICP, relative RMSFE

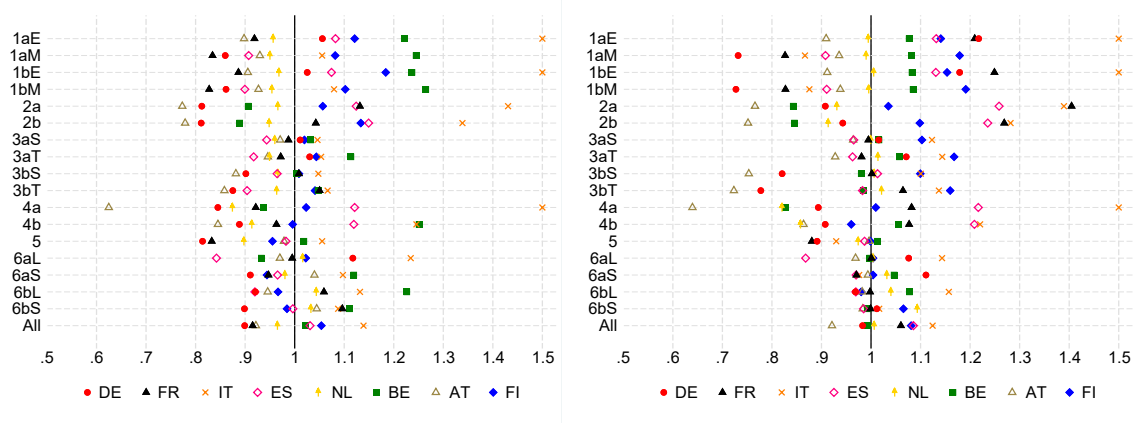
One-year-ahead horizon

Two-year-ahead horizon

Headline HICP



HICP excluding energy and food



Note: The figure shows the RMSFE of the model version incorporating expectations divided by the RMSFE of the version not incorporating such information. The RMSFE is computed over 2005Q4-2019Q4 for the one-year horizon and over 2006Q4-2019Q4 for the two-year horizon for each country. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.



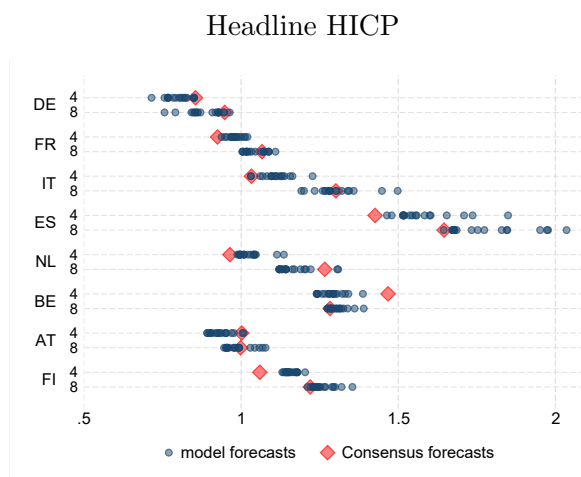
Overall, our analysis supports the main finding obtained for the euro area aggregate that inflation expectations lead to improvements of model-based inflation forecasts albeit the size of the improvement tends to be rather modest. Regarding the inflation measure, the evidence suggests that headline inflation can be predicted more accurately with the help of expectations than core inflation. The forecast horizon does not matter that much, but we find some evidence that expectations help more in the medium-run for headline inflation and more in the short-run for core inflation, whereas the forecasting gains are similar in size. Also in line with the results for the euro area, forecasting improvements are the largest for models 2 and 4, in which expectations are used to inform the inflation trend compared to versions where the trend is proxied with a random walk. Taking a closer look at the countries, expectations lead to better headline inflation forecast in more than half of the models under consideration, except for the one-year-ahead forecasts in Finland. As regards core inflation, adding expectations again does not help much in Finland, in addition to Italy, Belgium and for the one-year-ahead forecasts in France and Spain. Overall, the largest forecasting gains from including expectations can be obtained in Austria.

Next, similar to the euro area, the forecasting performance varies significantly over time, in particular for headline inflation (see Figures A5 and A6 in the appendix). From 2005 to 2009, adding expectations leads to better forecasts in almost all models and countries. From 2010 to 2014, gains from expectations became smaller, but tended to increase again since 2015. Moreover, the size of the forecasting gains in the sub-samples can be fairly large reaching almost 50%. Finally, comparing the model-forecasts including expectations to the Consensus forecasts directly in Figure 7, we find that the models yield more accurate predictions in more than half of the countries and horizons. This is in contrast to our earlier finding for the euro area.

## 6 Forecast evaluation since the COVID-19 pandemic

Macroeconomic instabilities observed since early 2020, triggered by the COVID-19 pandemic and the Russian invasion of Ukraine, induced a deterioration in the performance of models typically used to forecast inflation (see e.g. [Bobeica and Hartwig, 2023](#)). In this section, we provide a closer look at the role of inflation expectations when included in time series models to forecast

Figure 7: Country-specific results, model forecasts compared to Consensus Economics



Note: The figure shows the RMSFE of the model forecasts including expectations compared to the forecasts from Consensus Economics. The results are distinguished across countries and forecasts horizons of 4 quarters and 8 quarters ahead. The fixed-event forecasts from Consensus Economics are transformed into fixed horizon forecasts by computing weighted averages. The RMSFE is computed over 2005Q4-2019Q4 for the one-year horizon and over 2006Q4-2019Q4 for the two-year horizon.

inflation one year ahead over the period 2020Q4-2022Q2.<sup>33</sup> Considering the results presented in previous sections, we restrict this exercise to the use of inflation expectations from the Survey of Professional Forecasters.

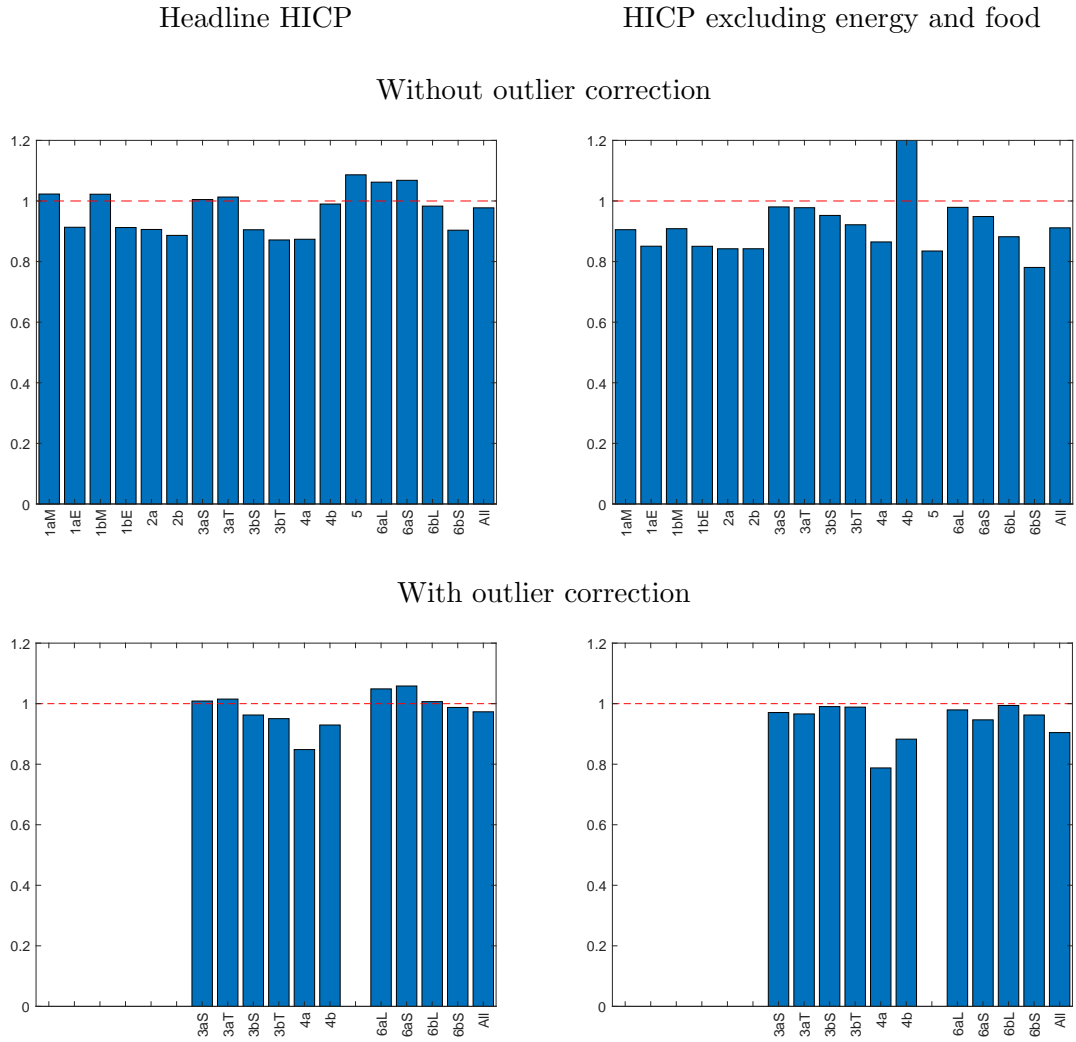
Figure 8 reports the *relative* RMSFE associated with both headline HICP and HICP excluding energy and food, whereas Figure A7 in the appendix shows the *absolute* RMSFE. As is evident in the upper panel of Figure A7, some of the models are more strongly affected by the COVID-19 observations in terms of forecast accuracy loss. This particularly affects Bayesian VARs that include GDP growth. For this reason we also consider BVAR versions with “outlier correction” as proposed by Stock and Watson (2016) and Carriero et al. (2022).<sup>34</sup> The results for these versions are reported in lower panels of Figures 8 and A7.

The first main result is that the inclusion of inflation expectations, on average, has also helped

<sup>33</sup>In this exercise we want that both the forecast *origin* and the *target* period belong to the (post-)COVID-19 sample. Given that for the two-year horizon the resulting evaluation sample would be very short, we do not consider it in this section.

<sup>34</sup>More precisely, the variance of the shocks is rescaled by a time-varying (shock-specific) factor (see also Primiceri and Lenza, 2022, for a related approach). We thank Joshua Chan for sharing his codes to implement the “correction”. Other strategies to minimize the influence of very large fluctuations in the data include adapting the priors (Cascaldi-Garcia, 2022), modelling the innovations as t-distributed (Antolin-Diaz et al., 2021; Karlsson et al., 2021; Bobeica and Hartwig, 2023), “de-COVIDing” (Ng, 2021) or treating the COVID observations as missing (Schorfheide and Song, 2021).

Figure 8: Incorporating information from SPF expectations into models, relative RMSFE: 2020-2022



Note: The figure shows the RMSFE of the model version incorporating the SPF expectations divided by the RMSFE of the version not incorporating such information. The RMSFE is computed over 2020Q4-2022Q2 for one-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

time series models to improve their forecast performance during the period 2020Q4-2022Q2. This is despite the fact that many models appear to perform better than the SPF over this period, which is the second main result. Outlier correction helps to improve the accuracy of the most affected models, on the one hand, while reducing the advantage of adding expectation data on the other. This suggests that the inclusion of expectations into a given model in some way

also works as an outlier correction, helping to “stabilize” the forecasts, in the presence of large swings in the data. This is also related to an earlier observation on the advantages of inclusion of expectations versus forecast combination.

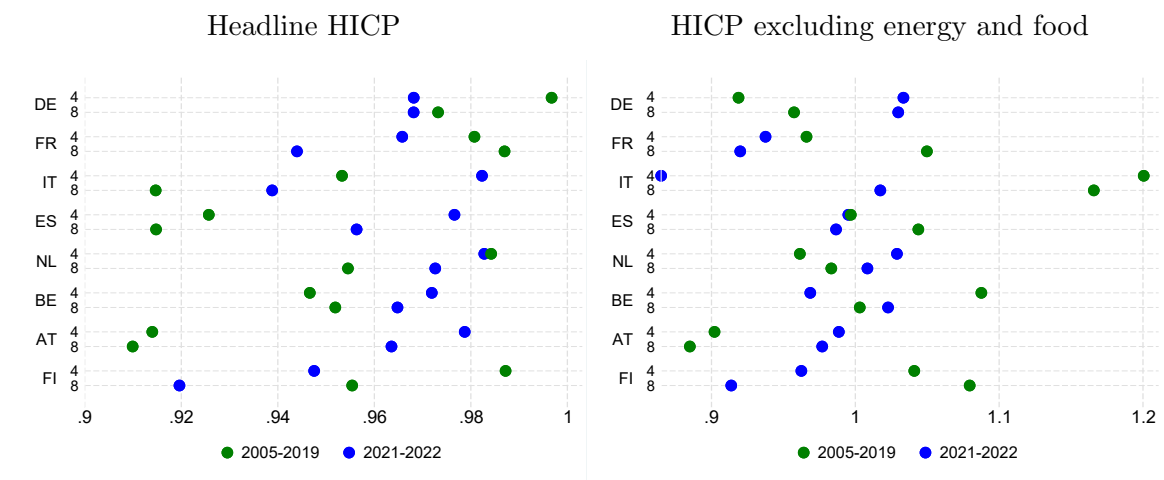
When looking at the disaggregated perspective across euro area countries, the results point to a similar conclusion as the one obtained at the aggregate level, although, there is a considerable heterogeneity across countries. For models applied to country-specific data we employ the measures of inflation expectations that are collected by Consensus Economics. To facilitate interpretation, Figure 9 plots the relative RMSFE, between the model versions that do and do not contain information on inflation expectations, averaged across the different models listed in Table 1. The average relative RMSFE are computed for (i) different countries, (ii) two forecast horizons: four- and eight-quarter ahead, and (iii) two sample periods: 2005-2019 and 2021-2022. The average relative RMSFE are computed for both headline HICP and HICP excluding food and energy. For the case of headline inflation, the results indicate that the inclusion of expectations from Consensus Economics into models used to forecast country-specific inflation also contributed to sharpen their accuracy during the period 2021-2022, this is the case for all countries and forecast horizons under consideration. Instead, for country-specific inflation based on HICP excluding energy and food, the inclusion of Consensus Economics inflation expectations into models does not always lead to forecast improvements during the period 2021-2022, which is consistent with the same type of forecast evaluation associated with the period 2005-2019.

## 7 Conclusions

This paper evaluates the extent to which the incorporation of inflation expectations into econometric models helps to improve inflation forecasts. In order to quantify the value added of information on inflation expectations within this context, we compare the predictive accuracy associated with two variants of univariate and multivariate time series models. The first variant includes information on inflation expectations, while the second variant does not include such information. This type of comparison is carried out in a real-time environment and from a comprehensive perspective which covers different types of models, measures of inflation and inflation expectations, and levels of geographic aggregation.

The main results suggest that inflation expectations provided by the Survey of Professional

Figure 9: Country-specific results, relative RMSFE averaged over models



Note: The figure shows ratios of RMSFE averaged over models. For both headline HICP and HICP excluding energy and food, the ratios for each model are constructed as the RMSFE of the model version incorporating expectations divided by the RMSFE of the version not incorporating such information, and are computed over two different samples: 2005-2019 and 2021-2022. The vertical axis makes reference to both countries and forecast horizons (four- and eight-quarter ahead). See Section 3.1 and Table 1 for the detailed description of the models used in the calculation of the averages.

Forecasters or Consensus Economics forecasts do improve model-based forecasts of inflation. Such improvements are modest but statistically significant in some periods. This finding applies both for the euro area economy as well as for several euro area countries. By contrast, the forecasting performance of models does not improve when using inflation expectations of firms and households collected by the European Commission or based on financial market prices. In case of the former the expectations appear to contain contemporaneous rather than forward looking information. For the latter, usefulness of such data when corrected for risk premia and also when a longer time series is available is a question left for future research.

We also compare the predictive performance of model-based forecasts of inflation with that of inflation expectations used as forecasts. The results point to the “supremacy” of SPF inflation expectations when forecasting euro area headline inflation in the pre-COVID-19 sample, in that their predictions turn to be a benchmark very hard to beat by sophisticated econometric models. By contrast, the relative performance of the SPF deteriorates in the most recent sample.

Overall, the results presented in this paper illustrate that policy makers can benefit from incorporating information on inflation expectations from professional forecasts into the econometric models used to forecast inflation as such expectations appear to contain relevant information beyond what is already captured by other predictors of inflation.

That being said, the evaluation period in this paper is relatively short, while there is evidence for the US that the usefulness of expectations for signalling inflation developments might be changing over time/across regimes (see e.g. [Mertens, 2016](#); [Mertens and Nason, 2020](#)). Analysis of such variation, also including the current high inflation regime is an interesting avenue for future research.

## References

- Adam, K. and Padula, M. (2012). Inflation Dynamics and Subjective Expectations in the United States. *Journal of Business & Economic Statistics*, 30(3):432–453.
- Álvarez, L. J. and Correa-López, M. (2020). Inflation Expectations in Euro Area Phillips Curves. *Economics Letters*, 195(C).
- Ang, A., Bekaert, G., and Wei, M. (2007). Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better? *Journal of Monetary Economics*, 54(4):1163–1212.
- Angelini, E., Bokan, N., Christoffel, K., Ciccarelli, M., and Zimic, S. (2019). Introducing ECB-BASE: The Blueprint of the New ECB Semi-Structural Model for the Euro Area. Working Paper Series 2315, European Central Bank.
- Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2021). Advances in Nowcasting Economic Activity: Secular Trends, Large Shocks and New Data. CEPR Discussion Papers 15926, C.E.P.R. Discussion Papers.
- Atkeson, A. and Ohanian, L. E. (2001). Are Phillips Curves Useful for Forecasting Inflation. *Federal Reserve Bank of Minneapolis Quarterly Review*, 25(1):2–11.
- Bañbura, M. and Bobeica, E. (2023). Does the Phillips Curve Help to Forecast Euro Area Inflation? *International Journal of Forecasting*, 39(1):364–390.
- Bañbura, M., Brenna, F., Paredes, J., and Ravazzolo, F. (2021). Combining Bayesian VARs with Survey Density Forecasts. Does it Pay Off? Working Paper Series 2543, European Central Bank.
- Bañbura, M., Giannone, D., and Lenza, M. (2015). Conditional Forecasts and Scenario Analysis with Vector Autoregressions for Large Cross-Sections. *International Journal of Forecasting*, 31(3):739–756.
- Bañbura, M., Giannone, D., and Reichlin, L. (2010). Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics*, 25(1):71–92.
- Bañbura, M. and van Vlodrop, A. (2018). Forecasting with Bayesian Vector Autoregressions with Time Variation in the Mean. Tinbergen Institute Discussion Papers 18-025/IV, Tinbergen Institute.

- Banerjee, R. N., Contreras, J., Mehrotra, A., and Zampolli, F. (2020). Inflation at Risk in Advanced and Emerging Economies. BIS Working Papers 883, Bank for International Settlements.
- Basselier, R., de Antonio Liedo, D., Jonckheere, J., and Langenus, G. (2018). Can Inflation Expectations in Business or Consumer Surveys Improve Inflation Forecasts? Working Paper Research 348, National Bank of Belgium.
- Bernanke, B. (2010). The Economic Outlook and Monetary Policy. *Speech, Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming*, August 27.
- Bobeica, E. and Hartwig, B. (2023). The COVID-19 Shock and Challenges for Inflation Modelling. *International Journal of Forecasting*, 39(1):519–539.
- Bobeica, E. and Sokol, A. (2019). Drivers of Underlying Inflation in the Euro Area Over Time: a Phillips Curve Perspective. Economic Bulletin Article 4/2019, European Central Bank.
- Canova, F. and Gambetti, L. (2010). Do Expectations Matter? The Great Moderation Revisited. *American Economic Journal: Macroeconomics*, 2:183–205.
- Carriero, A., Chan, J., Clark, T. E., and Marcellino, M. (2021). Corrigendum to “Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors” [J. Econometrics 212 (1) (2019) 137–154]. *Journal of Econometrics*.
- Carriero, A., Clark, T. E., and Marcellino, M. (2019). Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors. *Journal of Econometrics*, 212(1):137 – 154. Big Data in Dynamic Predictive Econometric Modeling.
- Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2022). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics*, pages 1–38.
- Cascaldi-Garcia, D. (2022). Pandemic Priors. International Finance Discussion Papers 1352, Federal Reserve Board of Governors.
- Cecchetti, S. G., Feroli, M., Hooper, P., Kashyap, A. K., and Schoenholtz, K. (2017). Deflating Inflation Expectations: The Implications of Inflation’s Simple Dynamics. CEPR Discussion Paper DP11925.



- Chan, J. C., Clark, T. E., and Koop, G. (2018). A New Model of Inflation, Trend Inflation, and Long-Run Inflation Expectations. *Journal of Money, Credit and Banking*, 50(1):5–53.
- Chan, J. C. and Song, Y. (2018). Measuring Inflation Expectations Uncertainty Using High-Frequency Data. *Journal of Money, Credit and Banking*, 50(6):1139–1166.
- Chan, J. C. C. (2018). Specification Tests for Time-Varying Parameter Models with Stochastic Volatility. *Econometric Reviews*, 37(8):807–823.
- Christiano, L. J. and Fitzgerald, T. J. (2003). The Bandpass Filter. *International Economic Review*, 44(2):435–465.
- Ciccarelli, M. and Osbat, C. (2017). Low Inflation in the Euro Area: Causes and Consequences. Occasional Paper Series 181, European Central Bank.
- Clark, T. and Davig, T. (2008). An Empirical Assessment of the Relationships Among Inflation and Short- and Long-Term Expectations. Working Paper 08-05, Federal Reserve Bank of Kansas City.
- Clark, T. E. (2011). Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility. *Journal of Business & Economic Statistics*, 29(3):327–341.
- Clark, T. E. and Doh, T. (2014). Evaluating Alternative Models of Trend Inflation. *International Journal of Forecasting*, 30(3):426–448.
- Clements, M. P. (2014). Forecast Uncertainty—Ex Ante and Ex Post: U.S. Inflation and Output Growth. *Journal of Business & Economic Statistics*, 32(2):206–216.
- Clements, M. P. (2018). Are Macroeconomic Density Forecasts Informative? *International Journal of Forecasting*, 34(2):181–198.
- Coenen, G., Karadi, P., Schmidt, S., and Warne, A. (2018). The New Area-Wide Model II: An Extended Version of the ECB’s Micro-Founded Model for Forecasting and Policy Analysis with a Financial Sector. Working Paper Series 2200, European Central Bank.
- Cogley, T., Primiceri, G. E., and Sargent, T. J. (2010). Inflation-Gap Persistence in the US. *American Economic Journal: Macroeconomics*, 2(1):43–69.

- Cogley, T. and Sargent, T. (2002). Evolving Post-World War II U.S. Inflation Dynamics. In *NBER Macroeconomics Annual 2001, Volume 16*, pages 331–388. National Bureau of Economic Research, Inc.
- Coibion, O. and Gorodnichenko, Y. (2015). Is the Phillips Curve Alive and Well after All? Inflation Expectations and the Missing Disinflation. *American Economic Journal: Macroeconomics*, 7(1):197–232.
- Crump, R. K., Eusepi, S., and Moench, E. (2016). The Term Structure of Expectations and Bond Yields. Staff Reports 775, Federal Reserve Bank of New York.
- Del Negro, M., Giannone, D., Giannoni, M. P., and Tambalotti, A. (2017). Safety, Liquidity, and the Natural Rate of Interest. *Brookings Papers on Economic Activity*, 48(1 (Spring)):235–316.
- Durbin, J. and Koopman, S. (2002). A Simple and Efficient Simulation Smoother for State Space Time Series Analysis. *Biometrika*, 89(3):603–615.
- ECB (2006). Measures of Inflation Expectations in the Euro Area. Monthly Bulletin July, European Central Bank.
- ECB (2021). An Overview of the ECB’s Monetary Policy Strategy. [https://www.ecb.europa.eu/home/search/review/html/ecb.strategyreview\\_monpol\\_strategy\\_overview.en.html](https://www.ecb.europa.eu/home/search/review/html/ecb.strategyreview_monpol_strategy_overview.en.html).
- Fagan, G., Hendry, J., and Mestre, R. (2005). An Area-Wide Model for the Euro Area. *Economic Modelling*, 22(1):39–59.
- Faust, J. and Wright, J. H. (2013). Forecasting Inflation. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, pages 2–56. North Holland.
- Forbes, K., Kirkham, L., and Theodoridis, K. (2021). A Trendy Approach to UK Inflation Dynamics. *The Manchester School*, 89(S1):23–75.
- Frey, C. and Mokinski, F. (2016). Forecasting with Bayesian Vector Autoregressions Estimated Using Professional Forecasts. *Journal of Applied Econometrics*, 31(6):1083–1099.
- Friedrich, C. (2016). Global Inflation Dynamics in the Post-Crisis Period: What Explains the Twin Puzzle? *Economics Letters*, 142:31–34.

- Fuhrer, J. (2012). The Role of Expectations in Inflation Dynamics. *International Journal of Central Banking*, 8:137–165.
- Galvao, A. B., Garratt, A., and Mitchell, J. (2021). Does Judgment Improve Macroeconomic Density Forecasts? *International Journal of Forecasting*, 37(3):1247–1260.
- Ganics, G. and Odendahl, F. (2021). Bayesian VAR Forecasts, Survey Information and Structural Change in the Euro Area. *International Journal of Forecasting*, 37(2):971–999.
- Garnier, C., Mertens, E., and Nelson, E. (2015). Trend Inflation in Advanced Economies. *International Journal of Central Banking*, 11(4):65–136.
- Giacomini, R. and Rossi, B. (2010). Forecast Comparisons in Unstable Environments. *Journal of Applied Econometrics*, 25(4):595–620.
- Giannone, D., Henry, J., Lalik, M., and Modugno, M. (2012). An Area-Wide Real-Time Database for the Euro Area. *The Review of Economics and Statistics*, 94(4):1000–1013.
- Grothe, M. and Meyler, A. (2015). Inflation Forecasts: Are Market-Based and Survey-Based Measures Informative? Working Paper Series 1865, European Central Bank.
- Hasenzagl, T., Pellegrino, F., Reichlin, L., and Ricco, G. (2022). A Model of the Fed’s View on Inflation. *The Review of Economics and Statistics*, 104(4):686–704.
- IMF (2013). Chapter 3, The Dog That Didn’t Bark: Has Inflation been Muzzled or Was It Just Sleeping? In *World Economic Outlook, April*, pages 1–17. International Monetary Fund, Washington.
- Jarociński, M. and Lenza, M. (2018). An Inflation-Predicting Measure of the Output Gap in the Euro Area. *Journal of Money, Credit and Banking*, 50(6):1189–1224.
- Kadiyala, K. R. and Karlsson, S. (1997). Numerical Methods for Estimation and Inference in Bayesian VAR-models. *Journal of Applied Econometrics*, pages 99–132.
- Karlsson, S., Mazur, S., and Nguyen, H. (2021). Vector Autoregression Models with Skewness and Heavy Tails. Working Papers 2021:8, Örebro University, School of Business.
- Korobilis, D., Landau, B., Musso, A., and Phella, A. (2021). The Time-Varying Evolution of Inflation Risks. Working Paper Series 2600, European Central Bank.

- Kozumi, H. and Kobayashi, G. (2011). Gibbs Sampling Methods for Bayesian Quantile Regression. *Journal of Statistical Computation and Simulation*, 81(11):1565–1578.
- Krüger, F., Clark, T. E., and Ravazzolo, F. (2017). Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts. *Journal of Business & Economic Statistics*, 35(3):470–485.
- Kulikov, D. and Reigl, N. (2019). Inflation Expectations in Phillips Curves Models for the Euro Area. Bank of Estonia Working Papers wp2019-8, Bank of Estonia.
- López-Salido, J. D. and Loria, F. (2020). Inflation at Risk. Finance and Economics Discussion Series 2020-013, Board of Governors of the Federal Reserve System (U.S.).
- Mertens, E. (2016). Measuring the Level and Uncertainty of Trend Inflation. *The Review of Economics and Statistics*, 98(5):950–967.
- Mertens, E. and Nason, J. M. (2020). Inflation and Professional Forecast Dynamics: An Evaluation of Stickiness, Persistence, and Volatility. *Quantitative Economics*, 11(4):1485–1520.
- Moretti, L., Onorante, L., and Zakipour Saber, S. (2019). Phillips Curves in the Euro Area. Working Paper Series 2295, European Central Bank.
- Ng, S. (2021). Modeling Macroeconomic Variations After COVID-19. *National Bureau of Economic Research*.
- Nunes, R. (2010). Inflation Dynamics: The Role of Expectations. *Journal of Money, Credit and Banking*, 42:1161–1172.
- Primiceri, G. and Lenza, M. (2022). How to Estimate a Vector Autoregression After March 2020. *Journal of Applied Econometrics*, 37(4):688–699.
- Rudd, J. B. (2021). Why Do We Think That Inflation Expectations Matter for Inflation? (And Should We?). Finance and Economics Discussion Series 2021-062, Board of Governors of the Federal Reserve System (U.S.).
- Schorfheide, F. and Song, D. (2021). Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic. NBER Working Papers 29535, National Bureau of Economic Research, Inc.

- Stevens, A. and Wauters, J. (2021). Is Euro Area Lowflation Here to Stay? Insights from a Time-Varying Parameter Model with Survey Data. *Journal of Applied Econometrics*, 36(5):566–586.
- Stock, J. H. and Watson, M. M. (2007). Why has U.S. Inflation Become Harder to Forecast? *Journal of Money, Credit, and Banking*, 39(7):3–34.
- Stock, J. H. and Watson, M. W. (2016). Core Inflation and Trend Inflation. *The Review of Economics and Statistics*, 98(4):770–784.
- Stockhammar, P. and Österholm, P. (2018). Do Inflation Expectations Granger Cause Inflation? *Economia Politica: Journal of Analytical and Institutional Economics*, 35(2):403–431.
- Tallman, E. W. and Zaman, S. (2020). Combining Survey Long-Run Forecasts and Nowcasts with BVAR Forecasts Using Relative Entropy. *International Journal of Forecasting*, 36(2):373–398.
- Villani, M. (2009). Steady-State Priors for Vector Autoregressions. *Journal of Applied Econometrics*, 24(4):630–650.
- Wright, J. H. (2013). Evaluating Real-Time VAR Forecasts with an Informative Democratic Prior. *Journal of Applied Econometrics*, 28(5):762–776.
- Yu, K. and Moyeed, R. A. (2001). Bayesian Quantile Regression. *Statistics & Probability Letters*, 54(4):437–447.

## Appendix A Description of the data set

Variable	Source	Description
<b>Consumer Prices</b>		
Headline inflation	<a href="#">RTDB</a> , <a href="#">SDW</a>	Harmonised index of consumer prices (HICP), seasonally & calendar adjusted. EA: real-time data; Countries: real-time data if available, pseudo real-time data if not. Missing seasonal adjustment added by using X11. Start of series: 1970 (EA, DE), 1978 (NL), 1980 (FR), 1981 (IT), 1987 (BE, FI), 1995 (ES), 1999 (AT). DE uses CPI instead of HICP.
Core inflation	<a href="#">SDW</a>	HICP excluding energy & food, seasonally & calendar adjusted. EA: real-time data as of 2006; pseudo real-time data before; Countries: real-time data if available, pseudo real-time data if not. Missing seasonal adjustment added by using X11. Start of series: 1970 (EA, DE), 1980 (FR), 1981 (IT), 1985 (NL), 1987 (BE, FI), 1995 (ES), 1999 (AT). DE uses CPI instead of HICP.
<b>Inflation Expectations</b>		
Long-run SPF	<a href="#">SPF</a>	Five-year-ahead inflation expectations for euro area headline inflation. Start of series 1999, backdated to 1990 using long-run CE.
Short-run SPF	<a href="#">SPF</a>	One-year-ahead inflation expectations for euro area headline inflation. Start of series 1999, backdated to 1990 using short-run CE.
Long-run CE	<a href="#">CE</a>	Average 6-10-year-ahead expectations for headline inflation (CPI, HICP). Start of series: 1990 (DE, FR, IT), 1995 (BE, ES, FI, NL), 1999 (AT). Euro area series starts in 2003, backdated to 1990 using average forecast from available countries.
Short-run CE	<a href="#">CE</a>	One-year ahead inflation expectations for headline inflation. Derived by weighting current and next calendar year expectations. Start of series: 1990 (DE, FR, IT), 1995 (BE, ES, FI, NL), 1999 (AT). Euro area series starts in 2003, backdated to 1990 using average forecast from available countries.
Households	<a href="#">ECBCS</a>	One-year ahead (qualitative) inflation expectations for euro area headline inflation. Start of series: 1985.
Firms	<a href="#">ECBCS</a>	One-quarter ahead (qualitative) expectations for firm's selling prices. Start of series: 1985.
Financial markets	<a href="#">Refinitiv</a>	Five-year-five-year inflation expectations for euro area headline inflation derived from inflation-linked swap rates. Start of series: 2005.

---

### Macro Variables

---

Real GDP	<a href="#">RTDB</a> , <a href="#">AWM</a>	Chain-linked volume, seasonally & calendar adjusted. EA: real-time data; Countries: real-time data if available, pseudo real-time if not. Missing seasonal adjustment added by using X11. Start of series: 1970 (DE), 1978 (NL), 1980 (FR), 1981 (IT), 1987 (BE, FI), 1995 (ES), 1999 (AT). Euro area series starts in 1999, backdated to 1970 using AWM.
Output gap	Own calculation	Christiano-Fitzgerald filter applied to log real GDP in real-time, cycles shorter than 15 years.
Unemployment	<a href="#">RTDB</a> , <a href="#">AWM</a>	Unemployment rate of the euro area, seasonally adjusted. Real-time data. Start of series 1990, backdated to 1970 using AWM.
Long-run SPF	<a href="#">SPF</a>	Five-year-ahead expectations for euro area real GDP. Start of series 1999, backdated to 1990 using CE.
Long-run CE	<a href="#">CE</a>	Average 6-10-year-ahead expectations for real GDP growth. Start of series: 1990 (DE, FR, IT), 1995 (BE, ES, FI, NL), 1999 (AT). Euro area series starts in 2003, backdated to 1990 using average forecast from available countries.
Interest rate	<a href="#">RTDB</a> , <a href="#">AWM</a>	Three month nominal interest rate. EA: real-time data; Countries: real-time data if available, pseudo real-time if not. Start of series 1970 (DE), 1978 (NL), 1981 (FR, IT), 1987 (BE, FI), 1995 (ES), 1999 (AT). Euro area series starts in 1999, backdated to 1970 using AWM.
Oil price	<a href="#">RTDB</a> , <a href="#">AWM</a>	Brent crude oil price expressed in euro. Real-time data. Start of series 1985, backdated to 1970 using AWM.

---

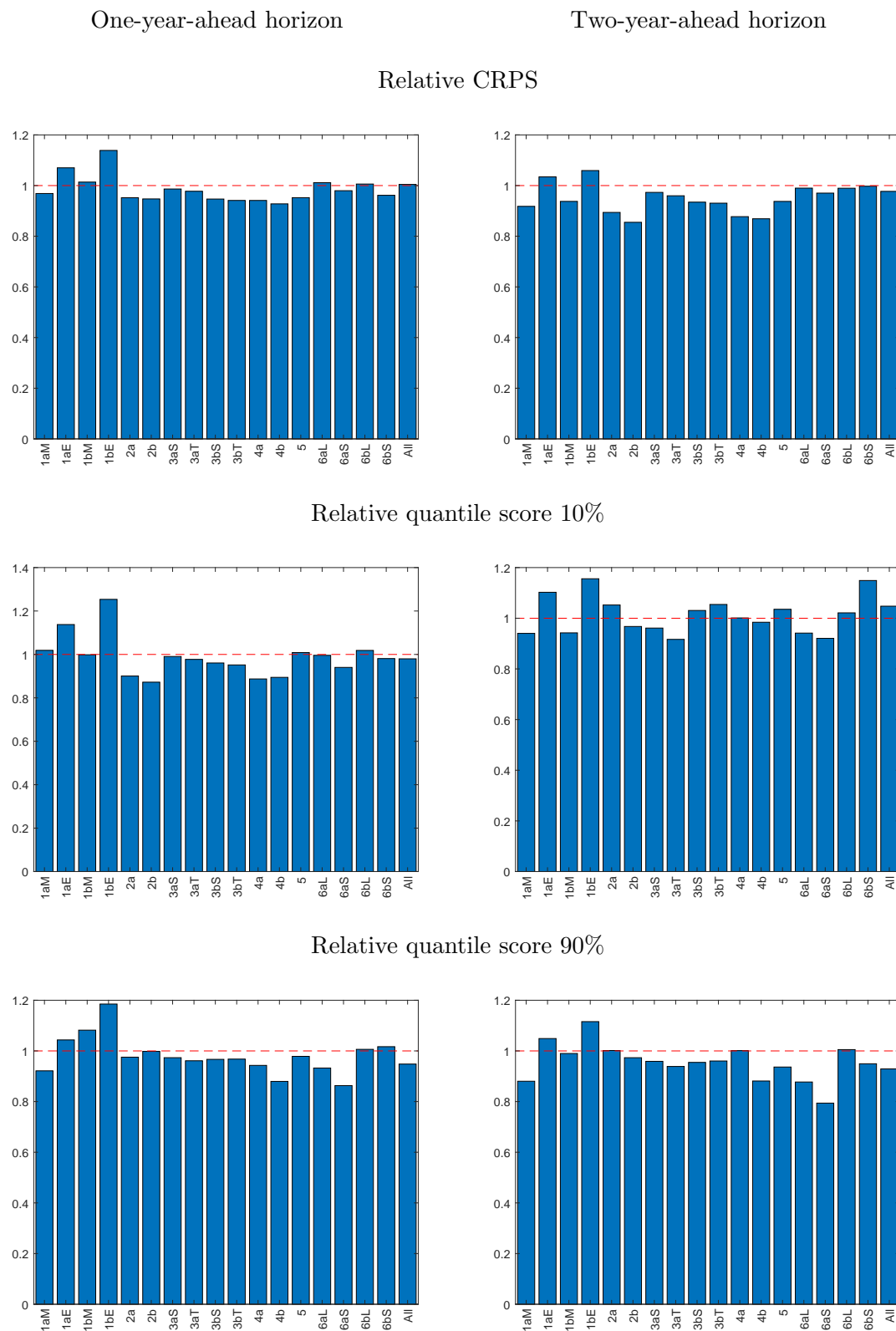
Note: *RTDB*: ECB's Real-time data base, *SDW*: ECB's Statistical Data Warehouse, *SPF*: ECB's Survey of Professional Forecasters, *CE*: Consensus Economics, *ECBCS*: European Commission's Business and Consumer Surveys, *AWM*: Area Wide Model Data Base, *AT*: Austria, *BE*: Belgium, *DE*: Germany, *EA*: Euro area, *ES*: Spain, *FI*: Finland, *FR*: France, *IT*: Italy, *NL*: Netherlands.

For some countries, the sources indicated above were supplemented by (non-public) data available at respective central bank.

## Appendix B Additional results

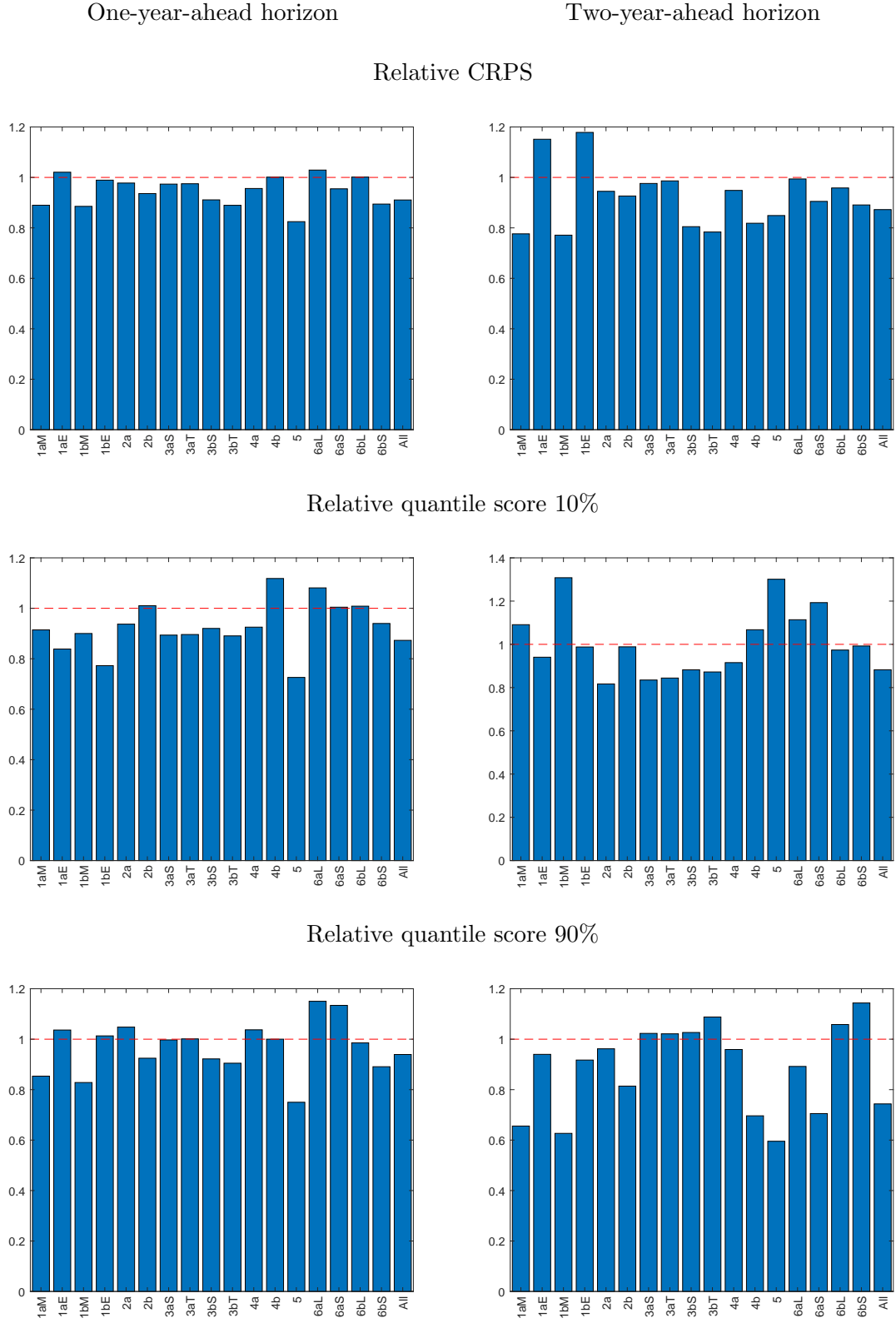


Figure A1: Incorporating information from SPF expectations into models, Headline HICP



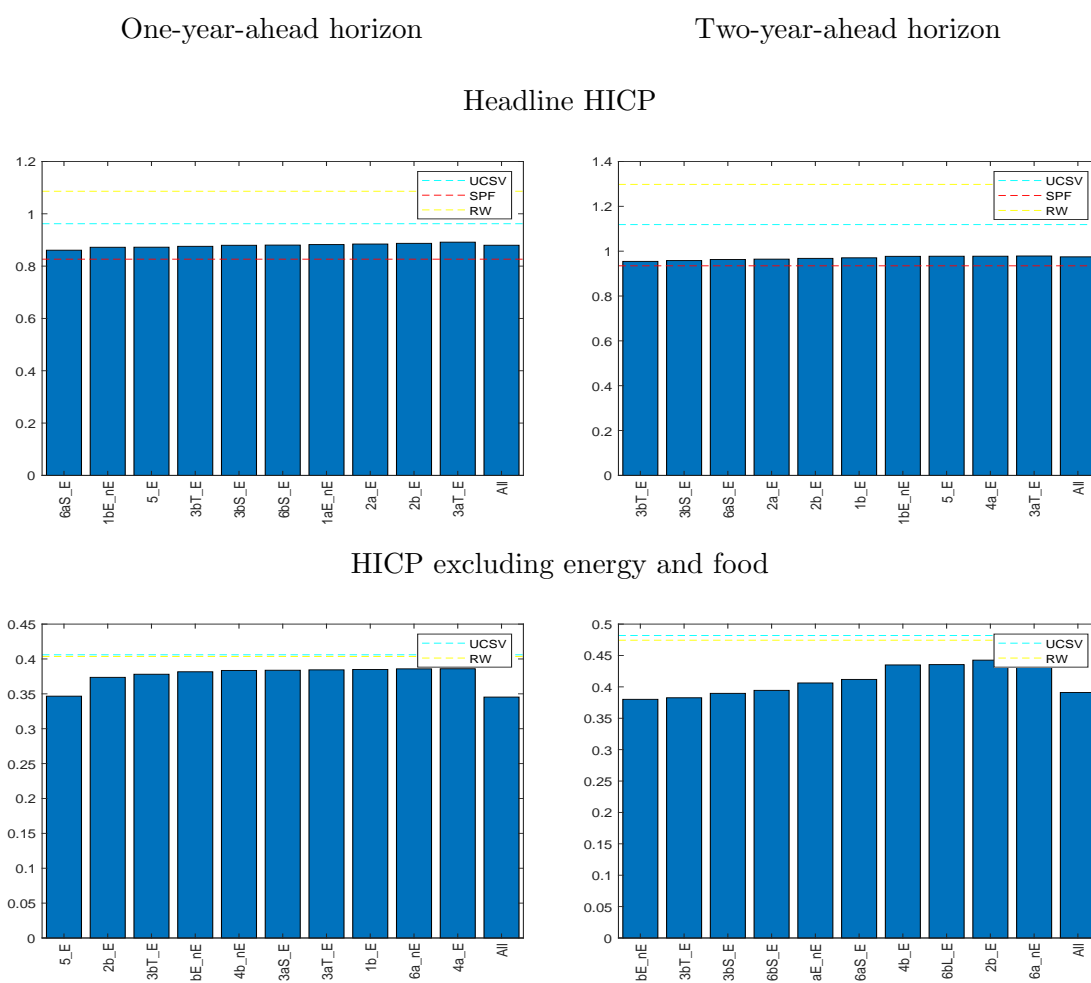
Note: The figure shows the CRPS (quantile score) of the model version incorporating expectations divided by the CRPS (quantile score) of the version not incorporating such information. The CRPS and quantile scores are computed over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

Figure A2: Incorporating information from SPF expectations into models, HICP excluding energy and food



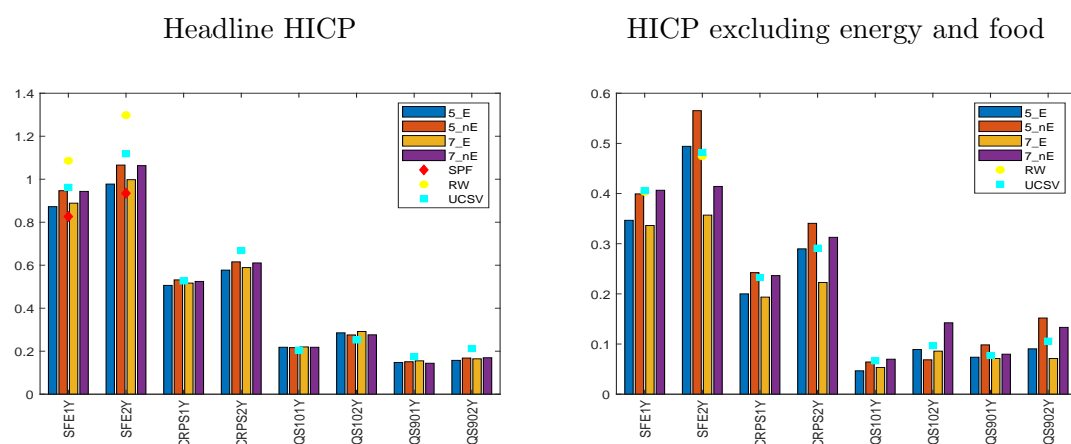
Note: The figure shows the CRPS (quantile score) of the model version incorporating expectations divided by the CRPS (quantile score) of the version not incorporating such information. The CRPS and quantile scores are computed over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. ‘a’ and ‘b’ refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

Figure A3: Best performing models, RMSFE



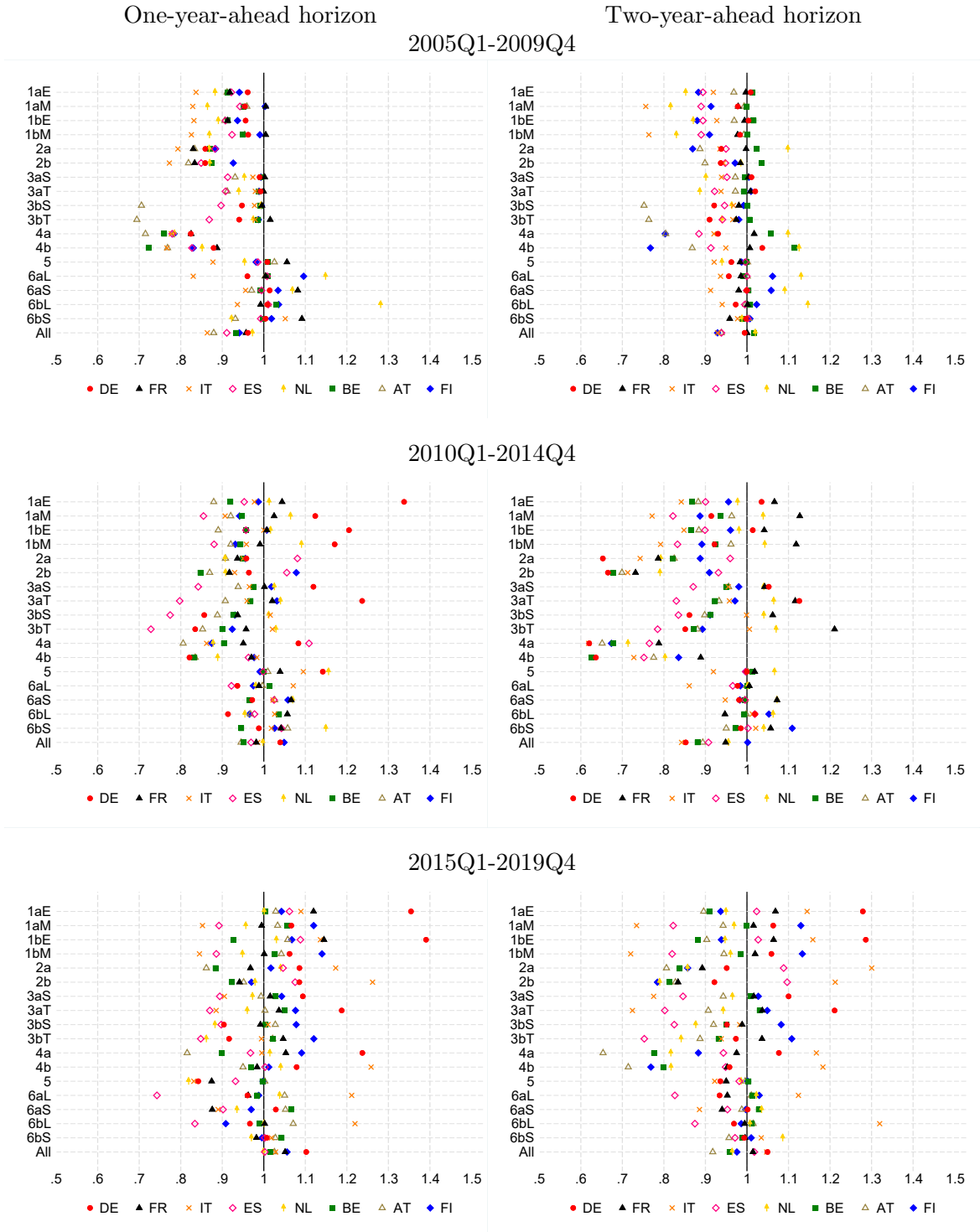
Note: The figure shows the RMSFE of the best 10 models and of the combination of all the models. The RMSFE is computed over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. 'E' refers to a model version that incorporates the SPF expectations whereas 'nE' denotes a model version that does not. See Section 3.1 and Table 1 for the detailed description of the models.

Figure A4: Comparison of Model 5 to a Bayesian quantile regression



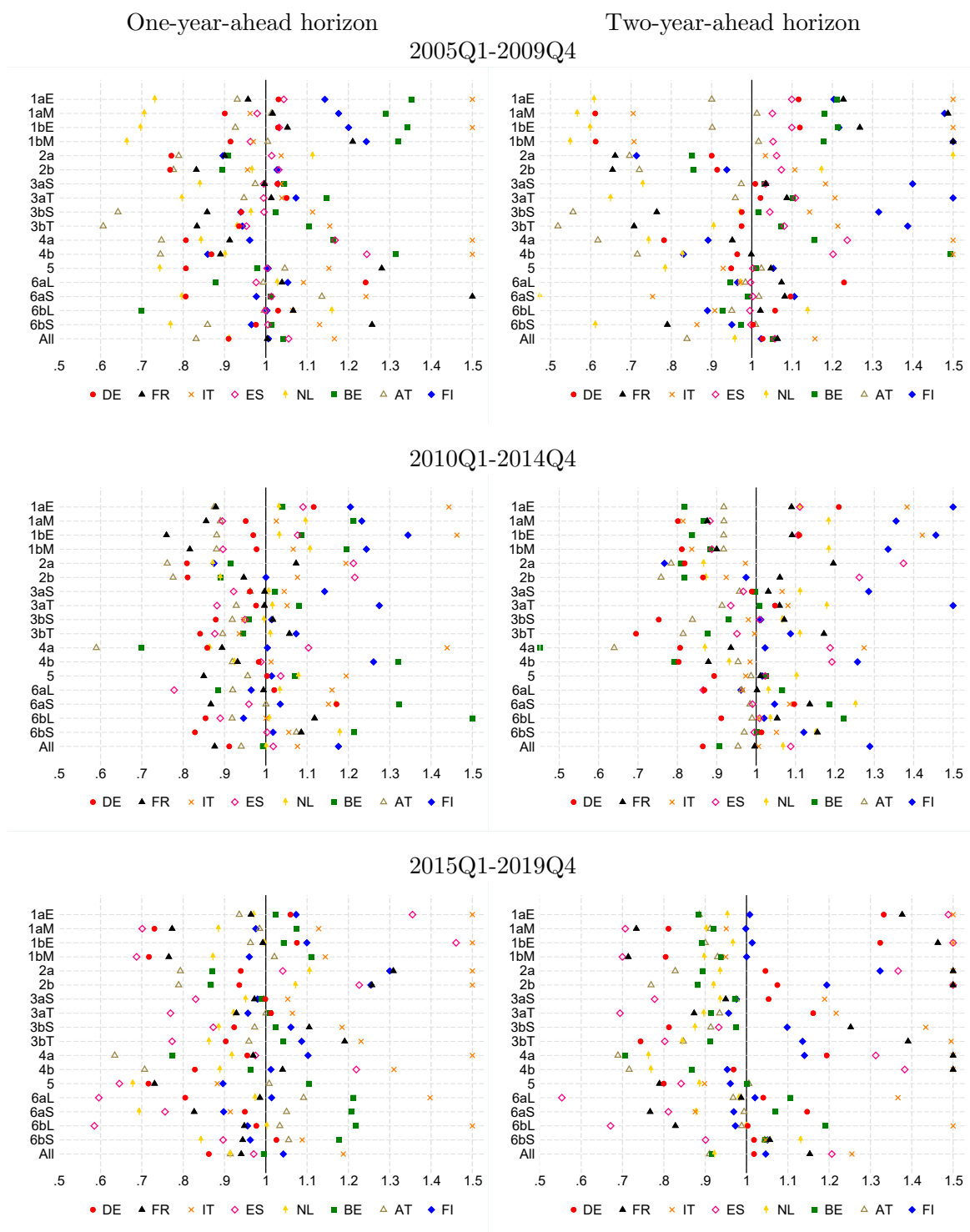
Note: The figure shows different (absolute) accuracy measures over 2001Q4-2019Q4 for one-year-ahead horizon and over 2002Q4-2019Q4 for two-year-ahead horizon. The numbers denote the model classes: 5: Phillips curves with constant coefficients, 7: as in 5 but in a Bayesian quantile regression setup. 'E' refers to a model version that incorporates the SPF expectations whereas 'nE' denotes a model version that does not. x-axis labels indicate the accuracy measures: RMSFE (SFE), CRPS and quantile scores (QS) at 10% and 90% at one-year (1Y) and two-year (2Y) horizon.

Figure A5: Headline HICP, incorporating information from SPF expectations into models, relative RMSFE



Note: The figure shows the RMSFE of the model version incorporating expectations divided by the RMSFE of the version not incorporating expectations for each country. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. Values above 1.5 are truncated for sake of comparability.

Figure A6: HICP excluding energy and food, incorporating information from SPF expectations into models, relative RMSFE



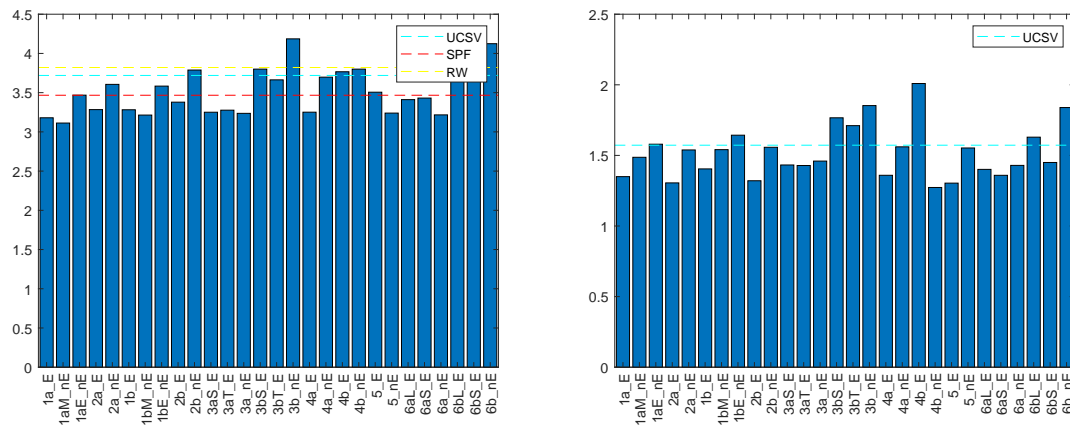
Note: The figure shows the RMSFE of the model version incorporating expectations divided by the RMSFE of the version not incorporating expectations for each country. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. Values above 1.5 are truncated for sake of comparability.

Figure A7: All models, RMSFE: 2020-2022

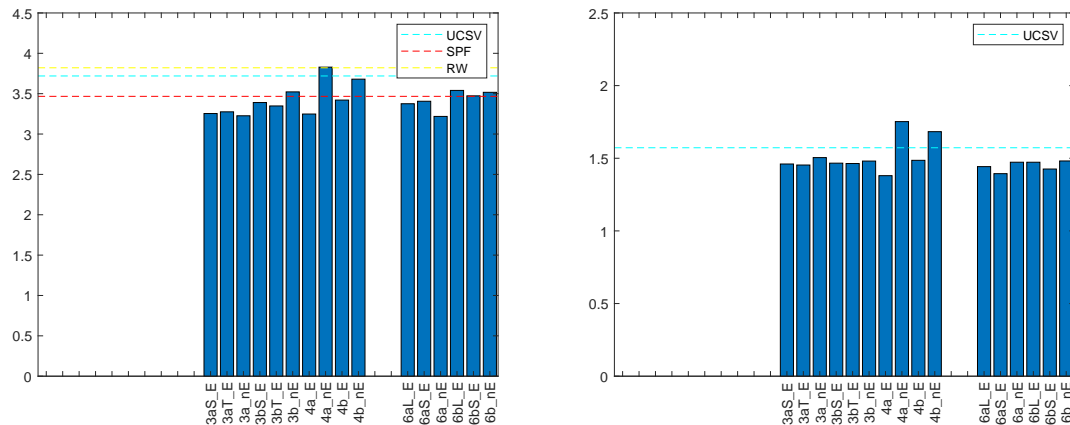
Headline HICP

HICP excluding energy and food

Without outlier correction



With outlier correction



Note: The figure shows the (absolute) RMSFE of all the models. The RMSFE is computed over 2020Q4-2022Q2 for one-year-ahead horizon. The numbers denote the model classes: 1: ADL models with time-varying trend inflation, 2: ADL models with time-varying trend inflation, time-varying coefficients and stochastic volatility, 3: Bayesian VARs with democratic priors, 4: Bayesian VARs with time-varying trends, 5: Phillips curves with constant coefficients, 6: Bayesian VARs with Minnesota priors. 'a' and 'b' refer to univariate and multivariate models, respectively. See Section 3.1 and Table 1 for the detailed description of the models.

## Acknowledgements

We would like to thank Friedrich Fritzer, Jana Jonckheere, Harri Ponka and Ide Kearney for providing results for Austria, Belgium, Finland and the Netherlands, respectively; Aidan Meyler for sharing data on inflation expectations; and Fulvia Marotta for research assistance. We also thank Elena Bobeica, Ide Kearney, Elmar Mertens, anonymous referees, the members of the ESCB Expert Group on Inflation Expectations, of the ESCB Working Group on Forecasting, of the ECB Forecasting and Policy Modelling Division and the participants of seminar at the University of Melbourne, of Workshop on New Directions for Inflation Forecasting, of Bundesbank Workshop on Revisiting and Improving Prediction Tools for Central Banks, of Workshop on Survey Data and Probabilistic Expectations at the University of Heidelberg, of 8th annual conference of the IAAE, and of SUERF-ECB-Bank of Finland-Bank of Italy-OeNB workshop on Challenges and Recent Advances in Modelling and Forecasting Inflation for their comments. The views expressed in this paper are of the authors only and do not necessarily reflect those of the European Central Bank or the Deutsche Bundesbank.

## Marta Bańbura

European Central Bank, Frankfurt am Main, Germany; email: [marta.banbura@ecb.europa.eu](mailto:marta.banbura@ecb.europa.eu)

## Danilo Leiva-León

European Central Bank, Frankfurt am Main, Germany; email: [danilo.leiva-leon@ecb.europa.eu](mailto:danilo.leiva-leon@ecb.europa.eu)

## Jan-Oliver Menz

Deutsche Bundesbank, Frankfurt am Main, Germany; email: [jan-oliver.menz@bundesbank.de](mailto:jan-oliver.menz@bundesbank.de)

## © European Central Bank, 2023

Postal address 60640 Frankfurt am Main, Germany

Telephone +49 69 1344 0

Website [www.ecb.europa.eu](http://www.ecb.europa.eu)

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from [www.ecb.europa.eu](http://www.ecb.europa.eu), from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

PDF

ISBN 978-92-899-4857-9

ISSN 1725-2806

doi:10.2866/162786

QB-AR-21-095-EN-N