EUROPEAN CENTRAL BANK
EUROSYSTEM

Working Paper Series

Carlos Montes-Galdón, Joan Paredes, Elias Wolf

Conditional density forecasting: a tempered importance sampling approach

No 2754 / December 2022

Abstract

This paper proposes a new and robust methodology to obtain conditional density forecasts, based on information not contained in an initial econometric model. The methodology allows to condition on expected marginal densities for a selection of variables in the model, rather than just on future paths as it is usually done in the conditional forecasting literature. The proposed algorithm, which is based on tempered importance sampling, adapts the model-based density forecasts to target distributions the researcher has access to. As an example, this paper shows how to implement the algorithm by conditioning the forecasting densities of a BVAR and a DSGE model on information about the marginal densities of future oil prices. The results show that increased asymmetric upside risks to oil prices result in upside risks to inflation as well as higher core-inflation over the considered forecasting horizon. Finally, a real-time forecasting exercise yields that introducing market-based information on the oil price improves inflation and GDP forecasts during crises times such as the COVID pandemic.

JEL CLASSIFICATION SYSTEM: C11, C53, E31, E37

KEYWORDS: Forecasting, inflation-at-risk, Bayesian Analysis, Importance Sampling

Non-technical summary

Policymakers use forecasting densities obtained from macroeconometric models to analyze the evolution of forecast uncertainty. It is of the utmost importance for them to understand risks to their forecasts and how risks arising in one economic variable may translate to other variables. An example based on financial market data could be the significant increase of upside risks to future energy prices stemming from the Russian invasion of Ukraine. While there exists a number of methods in the literature to condition model-based forecasting densities on external information about the future paths of some variables included in the model, less research considers conditioning those forecasting densities not only on the mean but also on higher-order information about their marginal distributions. But limiting the information set to an expected path of the model variables ignores the evolution of risks around these forecasts.

To tackle this issue, this paper proposes a new methodology to condition model-based forecasting densities on off-model information about some of its marginal distributions. Our methodology can be understood as an extension of the conditional forecasting literature, and we will refer to it as conditional density forecasting. While the motivation for our method is similar to the idea of entropic tilting developed in Robertson, Tallman, and Whiteman (2005) our methodology is more flexible with regards to the information that can be introduced to the distributions. Our algorithm is based on a tempered importance sampling procedure which makes it more robust with regards to the support of the densities, and avoids well-known problems that arise when using entropic tilting in practice. Additionally, we provide several extensions that make our method applicable for a wide range of forecasting models and their respective densities.

We illustrate our algorithm by conditioning the forecasting densities of a BVAR and a DSGE model on information about the marginal densities of oil price futures that can be obtained from their options. Since these option-implied forecasting distributions exhibit large positive skewness and increased volatility over the full forecasting horizon, we use a multivariate skew T distribution to appropriately capture these features. We document the transmission of upside risks from oil price futures to inflation risks in the euro area after the onset of the Russian invasion of Ukraine. Furthermore, median forecasts of core inflation remain persistently higher over the forecasting horizon compared to the symmetric model-based forecast densities. In a real-time forecasting exercise, we evaluate the improvements of our conditional density forecasts and find substantial increases in the forecasting performance during the onset of the Covid pandemic.

1 Introduction

For policymakers, it is of the utmost importance to understand risks to their forecasts and how changes in the risks arising in one economic variable can translate to other variables. This has been clear for example in 2022 as a consequence of the Russian invasion of Ukraine, when it has been a challenging task for policymakers and researchers to gauge the macroeconomic impact of rising risks in energy prices.

Policymakers and researchers use macroeconometric models to capture the transmission of changes in one to the rest of variables in the model, as well as to produce forecasts and understand their balance of risks. When off-model information on the expected evolution of selected model variables is available, the literature provides several methods on how to condition model-based multivariate forecasts on the future paths of those variables. These methods usually assume that those paths correspond to a central moment of those forecasts (see for example Waggoner and Zha (1999) or Bańbura, Giannone, and Lenza (2015)). This could also be understood as a scenario analysis or a conditional forecast. For instance, given an expected path of energy prices, researchers can obtain a model consistent path for inflation or real GDP.

However, limiting the information set to a central future path ignores the evolution of risks around these forecasts. So far, less research has considered conditioning forecasting densities not only on the mean but also on second or even higher moments of some of their marginal distributions (one notable approach is the entropic tilting methodology developed by Robertson, Tallman, and Whiteman (2005)). For example, policymakers could have views not only about the central tendency of the evolution of a variable such as energy prices, but also about other moments such as their variance, which accounts for uncertainty, and their skewness, which allows to consider asymmetric upside and downside risks. Therefore, this paper proposes a robust methodology that allows researchers to condition a model based multivariate forecasting density on information about the marginal densities of some selected variables in the model, rather than just on their first moment. Specifically, researchers might want to inform the unconditional forecasting density from a macroeconometric model with market-based expectations about the marginal densities of certain economic variables. Therefore, our methodology can be understood as an extension of the conditional forecasting literature, and thus we will refer to it as conditional density forecasting. As an illustration, one can use the implied probability densities obtained from option prices of futures on a specific model variable. Alternatively, information about the probability densities could be based on other econometric models (Giacomini and Ragusa (2014)), surveys or expert knowledge/judgement. Additionally, the information set could be based on assumptions made by researchers or policymakers to conduct risk analysis under certain scenarios.

Thus we first propose an algorithm that uses tempered importance sampling to re-weight the model-based forecasting densities to the off-model specified target marginal densities. While the motivation for our method is similar to the idea of entropic tilting developed in Robertson, Tallman, and Whiteman (2005) and Krüger, Clark, and Ravazzolo (2017), our methodology is more flexible with regards to the information that can be introduced to the distributions via our algorithm. Due to the tempering steps of our algorithm, it is also more robust with regards to the support of the densities, and avoids well-known problems that arise when using entropic tilting in practice.

Second, we demonstrate our methodology by conditioning the forecasting distribution obtained from a BVAR model and a DSGE model on asymmetric forecasting densities of the future oil price. Information about these densities at different horizons is derived from option prices of future contracts on oil and implies asymmetric forecasting densities over the full sample period. Similar to the work of Adrian, Boyarchenko, and Giannone (2019), we model these densities using the multivariate skew-T distribution of Azzalini and Capitanio (2003), which we fit to the option-implied moments. In a first exercise, we document the transmission of upside risks from future oil prices to inflation risks in the euro area after the onset of the Russian war in Ukraine. Since the invasion, option-implied densities exhibit high volatility and large positive skewness at all available forecasting horizons. We find that conditioning forecast densities on market-based upside risks to oil prices would imply upside risks to inflation and core inflation well above the forecast implied from the BVAR model for all periods. Furthermore, the median forecast of core inflation would also remain elevated over the forecast period.
In a second exercise, we investigate if historically introducing information from oil-options would have improved the accuracy of our model-based density forecasts for GDP and inflation in a BVAR model. While the results indicate no substantial gains in moderate times, we find substantial increases in the forecasting performance during the onset of the Covid pandemic.

The the paper is structured as follows: section 2 introduces the necessary information on importance sampling methods and describes our proposed algorithm. Section 3 provides an example of our methodology based on forecasting conditional distributions of inflation, given external information on oil prices. Section 4 proposes several extensions that make our method applicable for a wide range of forecasting models and their respective densities. Section 5 concludes and gives an outlook on further research.

2 Methodology

This section describes first our methodology and presents the technical details on importance sampling methods that are the work-horse of our algorithm. As introduced in section 1 the aim of our method is to condition the forecasting distribution of a model on off-model information on the marginal densities of a subset of variables. Formally, this amounts to adapting observations generated from a model distribution Q_θ with density $q_\theta(y_i)$ and parameter vector $\theta \in \Theta$ to another parametric distribution P_η with density $p_\eta(y_i)$ and $\eta \in H$ that satisfies the assumptions made by the researcher and that come from outside the original macroeconometric model. This off-model information about P_η includes, but is not limited to, moments or parameters of $p_\eta(y_i)$. Furthermore, Q_θ and P_η are not required to belong to same family of distributions such that the parameter spaces Θ and H can have different support and dimensions.[1] Most importantly, the external information may also be restricted to only a subset of variables included in $q_\theta(y_i)$, so that information is only available for some of the marginal densities of $p_\eta(y_i)$.

Our methodology is closely related to the entropic tilting methodology of Robertson, Tallman, and Whiteman (2005). Entropic tilting uses an optimization procedure to reweight a distribution so that it satisfies some conditions or moments. Yet, the use of entropic tilting has some drawbacks. First, while in theory entropic tilting is a powerful non-parametric tool to introduce information into a model based density forecast, its performance crucially hinges on the support of the original distribution Q_θ. Even if in theory this distribution might be unbounded, in practice researchers work with a finite set of draws from the distribution to implement a change in the distribution using entropic tilting. If the original draws of the distribution do not have enough support for the final density $p_\theta(y_i)$, implying a big change in the distributions, the methodology will yield unfavourable results and the algorithm might even fail to find a solution to the entropic tilting optimization procedure. Second, the methodology also struggles to find a solution as more conditions are introduced in the optimization problem. A similar point is raised in Krüger, Clark, and Ravazzolo (2017) who show that entropic tilting with higher dimensional distributions results in poor approximations of the final density since the weights of the reweighting step become very unevenly distributed.

[1] For example if Q_θ is a normal distribution with $\theta = (\mu, \sigma)$ and P_η is a T distribution with η denoting the degrees of freedom, then $\Theta = \mathbb{R} \times \mathbb{R}_{>0}$ and $H = \mathbb{N}$ the set of natural numbers.

Our methodology thus aims to provide a robust and flexible alternative to entropic tilting that can be applied in various circumstances. To overcome the aforementioned problems our methodology is based on tempered importance sampling methods which we adapt to our needs. With our methodology, we are able to move the draws from the original distribution slowly to the final distribution that incorporates researcher's additional information or judgement, even if the original draws do not cover the support of the final distribution.

2.1 Importance Sampling

The cornerstone of our methodology is importance sampling as introduced by Kloek and Dijk (1978). Since then, importance sampling has been applied in various scientific fields and its theoretical properties are well understood.[2] Importance sampling is helpful when a researcher wants to evaluate the properties of a distribution, but has only access to draws from other distribution that might be similar or not to the final one. In our case, we make use of importance sampling as follows. Suppose that the researcher wanted to introduce external information into a model-based density forecast. First, given a set of i.i.d. draws $\{y_i\}_{i=1}^N$ from the model-based forecasting density $y_i \sim q_\theta(y)$, the researcher could re-weight those draws so that the final forecasting density, $p_\eta(y_i)$, satisfies the information that they aim to introduce. The importance weights are calculated as the ratio

$$w_i = \frac{p_\eta(y_i)}{q_\theta(y_i)}$$

which are normalized to sum to 1 using $W_i = \frac{w_i}{\sum_i^N w_i}$. The tuples $\{y_i, W_i\}_{i=1}^N$ provide a particle approximation of $p_\eta(y_i)$ given by

$$\hat{p}_\eta(dy_i) = \sum_i^N W_i \delta_{y_i}(dy_i). \tag{1}$$

Resampling the draws using a multinomial distribution with support points y_i and weights W_i yields

$$\tilde{y}_i \sim \mathcal{MN}(y_i|W_i)$$

[2]For example, a thorough treatment of importance sampling can be found in Doucet, Freitas, and Gordon (2001).

and adapts the draws to the density $p_\eta(y_i)$.[3]

In typical applications of importance sampling, $q_\theta(y_i)$ serves as a proposal density while $p_\eta(y_i)$ is the target density. Furthermore, given that samples from $q_\theta(y)$ are drawn i.i.d, the strong Law of Large Numbers ensures that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i}^{N} h(\tilde{y}_i) \xrightarrow{a.s.} E_P[h(y)] = \int h(y)dP_\eta(y) \tag{2}$$

This insures that moments of interest such as the mean as well as the quantiles and P-values of the target distribution can be computed from the adapted draws $\{\tilde{y}_i\}_{i=1}^N$.[4]

If both density functions are reasonably close with most of their probability mass concentrated in the same regions of space, importance sampling can be applied to easily adapt draws from $q_\theta(y)$ to $p_\eta(y)$ and is commonly used to obtain samples from a distribution that is hard to sample from. Similarly to entropic tilting, importance sampling becomes infeasible if the Kullback-Leibler Divergence between the two distribution is large.

It is also well known that the quality of the importance sampler depends on the variance of the weights W_i. A high variance of the weights implies that only few draws are resampled, leading to a large approximation error a phenomenon dubbed weight degeneracy. Therefore, the right choice of the proposal density is crucial for a good importance sampling approximation. In many applications, researchers try to choose the proposal distribution such that the approximation error is small. However, in our case, we assume that the proposal density

[3] More formally, importance sampling constitutes a change of the measure of a random variable from one measure Q_θ to another measure P_η. Two different measures Q and P are related by

$$E_P[Y] = E_Q[\Lambda Y] \quad \text{with} \quad \Lambda = \frac{dP}{dQ}$$

(see Theorem 10.6 in Klebaner (2012)). The ratio Λ is called the Radon-Nikodym derivative. Given the two measures are two absolutely continuous probability distributions Q_θ and P_η, the Radon-Nikodym derivative is given by the ratio of the respective density functions

$$\Lambda = \frac{p_\eta(y)}{q_\theta(y)}$$

which is equal to the unnormalized weights $w(y)$ that are calculated in the correction step of importance sampling. Rewriting the integral in 2 as

$$E_{P_\eta}[y] = \int y w(y) dQ_\theta(y)$$

shows that importance sampling exploits the relationship above with a finite sample approximation.

[4] Additionally, a Central Limit Theorem can be established that enables statistical inference on the quantities computed from the draws

is predetermined and it is a model-based (forecasting) density $q_\theta(y_i)$. Yet the target density $p_\eta(y_i)$ might be far apart with a high Kullback-Leibler divergence. For example, external information that will improve or alter the forecast density of a model might often imply a very different mean, variance or skewness of $p_\eta(y_i)$ compared to $q_\theta(y_i)$. Additionally, this problem becomes even more severe if the dimension of the model implied density $q_\theta(y_i)$ is large such that the probability mass is concentrated in a small region of the high dimensional space (see for example the exhibition in Betancourt (2017)).

Moreover, in many applications researchers do not know a closed form solution to evaluate the model-based density $q_\theta(y_i)$. This happens for example when trying to evaluate the (marginal) forecasting density of a BVAR model accounting for parameter uncertainty. Last but not least, external information might only be available for transformations of the model variables of interest which naturally implies a change in the respective marginal distributions of the variables that needs to be accounted for. For example, market-based options are available for the level of oil prices, but in macroeconomic models usually variables appear in log-levels or growth rates.

The aforementioned problems render standard importance sampling methods infeasible for most applications we target with our methodology. To overcome these problems, our algorithm resorts to tempered importance sampling methods that are discussed in the next section.

2.2 Tempered Importance Sampling

Tempering importance sampling has its roots in the annealed importance sampling methodology from Neal (2001) and was introduced to the DSGE modelling literature in Herbst and Schorfheide (2014) and Herbst and Schorfheide (2019). To remedy the aforementioned problem of uneven weights when both the proposal and target density are far away from each other, the idea of tempered importance sampling is to adapt the draws via a sequence of bridge densities that assign more equal weights to the proposed draws and eventually converges to the true target. As shown in Herbst and Schorfheide (2019), an easy way to define such a sequence of bridge distributions is to use an inflated variance that is sequentially reduced to its actual level. More formally, let $p(y_i|\mu_\eta, \Sigma_\eta)$ be a target density with first and second moments that might depend on a set of model parameters η.[5] Tempered Importance Sampling uses a sequence of N_ϕ bridge distributions

$$p_n(y_i|\mu_\eta, \Sigma_\eta/\phi_n) \quad \text{with} \quad 0 < \phi_1 < ... < \phi_{N_\phi} = 1 \tag{3}$$

that converge towards the target distribution $p(y_i|\mu_\eta, \Sigma_\eta)$ for $\phi_n \to 1$. Starting from a low value of ϕ_1, $p_1(y_i|\mu_\eta, \Sigma_\eta/\phi_1)$ assigns weights to the proposed draws $\{y_i\}_{t=1}^N$ that are more evenly distributed. It is important to note, that there is freedom in defining the tempering method of the function.[6] After proposing a first set of N particles from $q(y_i|\mu_\theta, \Sigma_\theta)$, a combination of importance sampling and MCMC-methods move the proposed particles towards the target distribution via the bridge distributions by cycling through the following three steps until $\phi_{N_\phi} = 1$:

1. **Correction:** Compute new importance weights

$$W_{i,n} \propto \frac{p_n(y_{i,n-1}|\mu_\eta, \Sigma_\eta/\phi_n)}{p_{n-1}(y_{i,n-1}|\mu_\eta, \Sigma/\phi_{n-1})}$$

2. **Selection:** Resample the draws

$$\tilde{y}_{i,n} \sim \mathcal{MN}(y_{i,n-1}|W_{i,n})$$

3. **Mutation:** Propagate the resampled particles $\{\tilde{y}_i\}_{i=1}^N$ using M steps of an MH-Algorithm with a transition Kernel

$$y_{i,n} \sim K_n(y_n|\tilde{y}_{i,n})$$

that has the stationary distribution $p_n(y_i|\mu_\eta, \Sigma_\eta/\phi_n)$

In each iteration, the correction and selection steps adapt the draws $\{y_{i,n-1}\}_{i=1}^N$ of the previous iteration to the n^{th} brigde distribution using an importance sampling step (compare Section

[5]Note that the density can also depend on additional parameters controlling the shape or Kurtosis which are dropped in the following exposition for notational convenience.

[6]In general, other tempering methods are possible. Yet, it has to be satisfied that

$$p_n(y_i|\mu_\eta, \Sigma_\eta) = p(y_i|\mu_\eta, \Sigma_\eta) \quad \text{for} \quad \phi_{N_\phi} = 1$$

and

$$Var[W_{i,n}] \to 0 \quad \text{for} \quad \phi_n \to 0.$$

We propose another definition in section 4

2.1). Once the particles are adapted to $p_n(y_i|\mu_\eta, \Sigma_\eta/\phi_n)$, the mutation step moves the resampled particles to a region with a higher probability density using a Metropolis Hastings sampler.

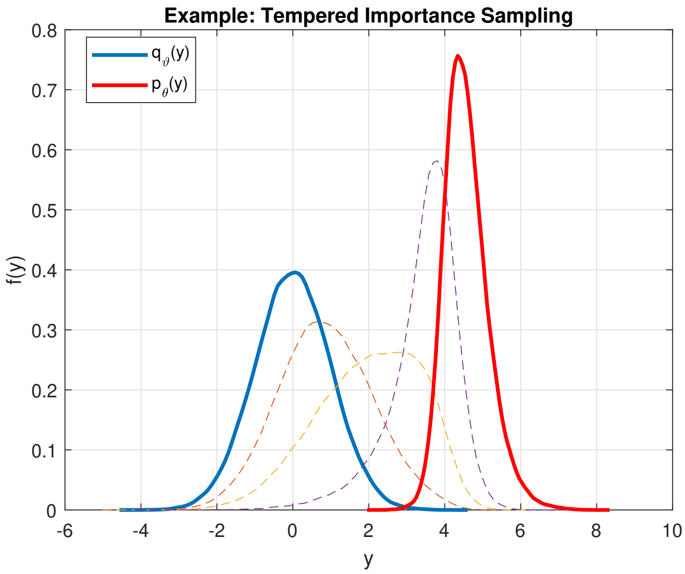

Figure 1: Proposal, Target and Bridge Distributions

Figure 1 provides an illustration of this idea in a simple univariate example. The graph shows Kernel-density estimates of the draws at different stages of the tempered importance sampling procedure. The draws are proposed from a standard normal distribution (blue solid line) and adapted to a skew-T distribution with mean 5.16 and positive skewness (red solid line). Clearly, the probability masses of the two densities are concentrated in different regions of the real line which renders a standard importance sampler infeasible.[7] However, the intermediate bridge distributions (dashed lines) are close to each other by construction and thus provide a suitable proposal density for the next tempering step. This prevents the importance weights in each iteration from deteriorating. Additionally, the Metropolis Hastings step insures that the proposed draws are concentrated in regions with high probability and sequentially mutates the draws towards the target distribution. Since $p_n(y_i|\mu_\eta, \Sigma_\eta/\phi_n)$ eventually converges to the target distribution, the particles are gradually adapted to the final distribution. Convergence results and Central Limit Theory similar to simple importance sampling are based on the work of Chopin (2004) and established in Herbst and Schorfheide (2014).

It is important to note that without the mutation step, the importance sampling procedure

[7]The Kullback-Leibler Divergence of the two densities is 107.22

would face the same limitations as entropic tilting or plain vanilla importance sampling. If there is no empirical support for the draws from the proposal density, the weights of the brige distributions will deteriorate as there are no draws in the region of space where the target density has most probability mass. However, including the mutation step moves the particles and ensures that there are draws in the region of the target densities.

Our algorithm is based on a tempered importance sampler to adapt draws from the model based distribution Q_θ to the target distribution P_η. The ability to mutate the proposed draws from the model density $q_\theta(y_i)$ towards the target density makes our methodology more robust then the entropic tilting methodology of Robertson, Tallman, and Whiteman (2005). The next section introduces our algorithm.

2.3 Our proposed algorithm

Let y_i be a vector with dimension $L \times 1$ drawn from $Q_\theta(y)$ and let $y_i^e \in y_i$ be a vector with a subset of elements in y_i with dimension $L^e\times$ such that $L^e < L$. Our algorithm proceeds in to steps: First elements of the vector y_i^e are adapted to the target density $p(y_i|\mu_\eta, \Sigma_\eta)$ that satisfies the imposed restrictions. Second, the corresponding values of the vector y_i^{-e} with length $L^{-e} = L - L^e$ that holds the remaining elements of y_i are recovered conditional on the final values of y_i^e. To overcome the problems of weight degeneracy, initial draws are proposed based on the density $q(y_i|\mu_\theta, \Sigma_\theta)$ and subsequently adapted to $p(y_i|\mu_\eta, \Sigma_\eta)$ in a sequence of tempering iterations. This gives the following algorithm:

1. Draw $y_{i,1} \sim q(y_i|\mu_\theta, \Sigma_\theta)$ for $i = 1, .., N$
2. Select the subset $y_{i,1}^e \in y_{i,1}$ for which there exists external information on the transformation $h(y_{i,1}^e)$.
   (a) Obtain initial importance weights $W_{i,1} \propto p_1(h(y_{i,1}^e)|\mu_\eta, \Sigma_\eta/\phi_1)$
   (b) Resample $y_{i,1}^e \sim \mathcal{MN}(y_{i,1}^e|W_{i,1})$
3. For $n = 2 : N_{\phi_N}$
   (a) **Correction**: Obtain weights $W_{i,n} \propto \frac{p_n(h(y_{i,n-1}^e)|\mu_\eta, \Sigma_\eta/\phi_n)}{p_{n-1}(h(y_{i,n-1}^e)|\mu_\eta, \Sigma_\eta/\phi_{n-1})}$

(b) **Selection:** Resample $y^e_{i,n} \sim \mathcal{MN}(y^e_{i,n-1}|W_{i,n})$

(c) **Mutation:** For $j = 1 : H$

i. Draw $\hat{y}^e_{i,n} \sim q(y^e_i|y^e_{i,n}, \mu_\theta, c_n\Sigma_\theta)$

ii. Compute

$$\alpha = \frac{p_n(h(\hat{y}^e_{i,n})|\mu_\eta, \Sigma_\eta/\phi_n)}{p_n(h(y^e_{i,n})|\mu_\eta, \Sigma_\eta/\phi_n)} \times \left|\frac{\det(\mathcal{J}_{h^{-1}}(y^e_{i,n}))}{\det(\mathcal{J}_{h^{-1}}(\hat{y}^e_{i,n}))}\right| \tag{4}$$

where $\mathcal{J}_{h^{-1}}(y)$ denotes the Jacobian of the inverse transformation of $h(y)$.

iii. Draw $u \sim U(0, 1)$.

Iff $u < \alpha$:

$$\text{Set} \quad y^e_{i,n} = \hat{y}^e_{i,n}$$

4. Draw the other variables y^{-e}_i from conditional density

$$y^{-e}_i \sim q(y^{-e}_i|y^e_{i,N_\phi}, \mu_{\theta,-e|e}, \Sigma_{\theta,-e|e})$$

Compared to previous applications of sequential importance sampling applied in the existing literature, our algorithm has two important extensions. First, since the draws are proposed based on $q(\tilde{y}_i|\mu_\theta, \Sigma_\theta)$ but evaluated given the transformation $h(y)$, we need to adjust the acceptance ratio of the Metropolis Hastings sampler to target the correct posterior density. Using a change of variables argument yields that the proposal density for the acceptance ratio α is given by

$$q(y|\mu_\theta, \Sigma_\theta) \times |\det(\mathcal{J}_{h^{-1}}(y)|. \tag{5}$$

This leads to the Jacobian correction in An example of why we need this transformation, which we will use later in our application, is related to oil prices. Economists have access to option-based oil price forecasting densities. However, in many applications, the price of oil does not enter in our models as a level variable. Researchers usually take transformations such as a natural logarithm, a growth rate or other transformations to stationarize the level of the oil price.

Plugging in the latter expression for the proposal density leads to the acceptance ratio given in equation 4. Of course, this implies that our algorithm requires the transformation $h(y)$ to be bijective and differentiable. However, this holds for many transformations of interest such as the ones mentioned before.

Second, since we only seek to include external information for a subset y_i^e of the elements of y_i the other variables in the model y_i^{-e} need to be recovered from the mutated particles. We achieve this by using the conditional distribution $q(y_i^{-e}|\mu_{\theta,-e|e}, \Sigma_{\theta,-e|e}, y^e_{i,N_\phi})$ to mutate the other elements of y_i conditional on the final particles y^e_{i,N_ϕ}. Note that this is akin to the principle of a Gibbs sampling step. Hence, for our methodology it is necessary that one can draw from the conditional density of $q(\tilde{y}_i|\mu_\theta, c_n\Sigma_\theta)$.

Finally, the algorithm can be used with a predetermined tempering schedule such that the sequence $\{\phi_i\}_{i=0}^{N_\phi}$ is deterministic as in Neal (2001) or Herbst and Schorfheide (2014) but also with adaptive implementations as for example in Herbst and Schorfheide (2019). The same holds for the scaling factor c_n of the covariance matrix Σ_θ of the proposal density q that can be adapted to target a specific acceptance rate of the MH-step in every new iteration.[8]

2.4 Relationship to entropic tilting

Given the aim of our new methodology as alternative to entropic tilting, this section illustrates how the tempered importance sampling approach is related to the entropic tilting methodology of Robertson, Tallman, and Whiteman (2005). The idea of entropic tilting is to reweight draws from a model-based distribution $F(y)$ to adapt them to a target distribution $F'(y)$. $F'(y)$ is found by minimizing the following optimization problem

$$D(F|F') = \int f'(y)\frac{f'(y)}{f(y)}dy \quad \text{s.t.} \quad \int f'(y)g(y)dy = \bar{g} \quad \text{and} \quad \int f'(y)dy = 1 \tag{6}$$

[8] For example, following Herbst and Schorfheide (2019), c_n can be updated recursively by the following formula

$$c_n = c_{n-1}f(\hat{A}_{n-1}) \quad \text{with} \quad f(x) = a + b\frac{e^{c(x-\bar{s})}}{1+e^{c(x-\bar{A})}}$$

where $\hat{A}_{n-1}$ is an estimate of the acceptance rate of the MH step in the previous iteration and $0 < \bar{A} < 1$ is the targeted acceptance ratio set be the researcher. The constants a, b and c control the speed of the adjustment and should satisfy $a + 0.5b = 1$ and $c > 0$.

Hence, the distribution F_θ is the closest density that satisfies a number of constraints $\bar{g}$ imposed by the researcher. As shown in the original paper the solution to this problem is given as

$$f'(y) = f(y) \exp\left(\gamma' g(y)\right) \tag{7}$$

where γ is the vector of Lagrange multipliers associated with the constraints $\bar{g}$ imposed under F'. It's value is given by

$$\gamma = \arg\left[\int f(y) \exp\left(\gamma' g(y)\right)(g(y) - \bar{g})\right] \tag{8}$$

Based on equation (8), it is straight forward to show that this condition is minimized if γ is set such that

$$\exp\left(\gamma' g(y)\right) = \frac{f'(y)}{f(y)} \tag{9}$$

This implies that draws from F_ϑ are resampled using the importance weights $W(y)$. Given that the minimization procedure in equation 8 finds the optimal solution, the target density of entropic tilting and importance sampling coincide.

While entropic tilting imposes constraints on the moments of the target density without an assumption about the parametric form of the target density, our methodology imposes these restrictions via the choice of a specific parametric density. In case only few restrictions are placed on the target density, the resulting density of entropic tilting might differ from the density that is imposed in our proposed method. However, as the number of restrictions on the moments of the target density increases, the space of density functions that satisfy these restrictions shrinks and the density of entropic tilting will be equal to the imposed density of our importance sampling method.

As outlined before, while our methodology requires an additional assumption, it allows to move the particles in the mutation step of the tempered importance sampling procedure. This enables a more robust and versatile method to include external information into the forecast densities or perform scenario analysis based on counterfactuals.

3 Application: The transmission of oil-price risks to inflation

To showcase how our algorithm functions, we use our methodology to condition the density forecast of a small euro area BVAR model and a large DSGE model on option-implied densities of oil price futures. The idea is to introduce oil price forecasting densities based on market data, which might not be symmetric, into the different models and to explore how that additional information tilts the forecasting distribution of variables such as inflation and real GDP.

Our application also contributes to the macro-at-risk literature that started with the seminal paper of Adrian, Boyarchenko, and Giannone (2019) and has spawned an active field of research to model and evaluate asymmetric risks of macroeconomic variables such as GDP or inflation. Recent contributions from Wolf (2022) or Delle Monache, De Polis, and Petrella (2021) propose different univariate modelling approaches to model time-varying asymmetries of the forecasting densities of macroeconomic variables using a skewed distribution of the shocks whose moments are time-varying based on auto-regressive components and exogenous variables. Furthermore, the work of Montes-Galdón and Ortega (2022) extends this approach to Bayesian VAR models exploiting a specific representation of the skewed Normal distribution by Azzalini (2013) for the structural shocks of the model. While the aforementioned papers seek to capture the evolution of risks based on some latent state variables, the application in this paper aims to directly introduce off-model information about the full[9] marginal distribution of some variables in the models and analyse the effect on other marginals based on correlations captured by the model. That is, we explore how possibly asymmetric risks in one variable translate to risks in other variables.

Based on the results of Breeden and Litzenberger (1978), it is possible to infer probabilities about the value of an underlying asset at the date of expiry from derivative prices observed in the market. The resulting probabilities can then be used to construct the full probability density of the underlying variable at the expiration date (see for example Vincent-Humphreys and Puigvert Gutiérrez (2010)). Option-implied probability densities are derived for different objects such as exchange rates, interest rates or oil prices, and are regularly published by the European Central Bank, the Bank of England or the Federal Reserve.

[9] In the sense of using information beyond the second moment

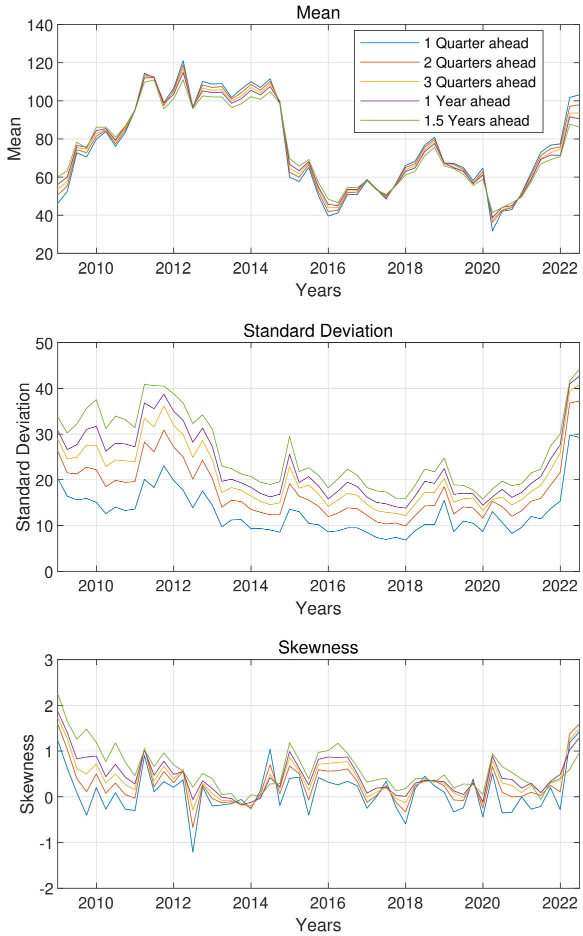

Figure 2: Option-implied moments of future Oil-Price densities

Figure 2 shows the first three moments of the option implied probability density functions of oil prices using quarterly data from 2008 to 2022 for different forecast horizons obtained from the ECB's Statistical Data Warehouse. Most notably, the probability density of the future oil price exhibits large fluctuations in the evolution of skewness for all horizons over the full sample. Remarkably, in 2022, with the beginning of the Russian invasion of Ukraine, market-based data shows that agents expected significant asymmetric upside risks to the price of oil. We first need to find a distribution that can explain those fluctuations and which will be the target distribution $P_\eta(y_i^e)$ in our methodology as defined in the previous section. Following the recent literature on skewed densities spawned by Adrian, Boyarchenko, and Giannone (2019) we model the marginal forecasting densities of the price of oil at any time period and at different horizons as a multivariate Skew-T distribution as introduced by Azzalini and Capitanio (2003). We outline further details of the density function of the multivariate Skew-T distribution and our fitting procedure in the next section.

3.1 Properties of the Multivariate Skew-T Distribution

A random vector $y \in \mathbb{R}^p$, follows a a multivariate Skew-T

$$y \sim \mathcal{MST}(y|\xi, \Omega, \lambda, \nu)$$

where $\xi \in \mathbb{R}^p$ determines the location, Ω is a $p \times p$ Covariance matrix, $\lambda \in \mathbb{R}^p$ is the shape parameter and $\nu \in \mathbb{N}$ gives the degrees of freedom. Based on the construction of skewed densities established in Azzalini (2013), the corresponding density function is given by

$$\tau(y|\xi, \Omega, \lambda, \nu) = 2t_p(y|\xi, \Omega, \nu) T_1 \left\{ \lambda' z \times \left(\frac{\nu + p}{\nu + Q(z)} \right)^{1/2} \middle| \nu + p \right\} \tag{10}$$

with $z = \omega(y - \xi)$, $Q(z) = z'\bar{\Omega}^{-1}z$, correlation matrix $\bar{\Omega} = \omega^{-1}\Omega\omega^{-1}$ and

$$t_p(y|\xi, \Omega, \nu) = \frac{\Gamma((\nu + p)/2)}{|\Omega|^{1/2}(\nu\pi)^{p/2}\Gamma(\nu/2)} \left(1 + \frac{Q(z)}{\nu}\right)^{-(\nu+p)/2} \tag{11}$$

As shown in Proposition 3 in Arellano-Valle and Genton (2010) the multivariate skew-T distribution is closed under marginalization. For a partition $y = (y_1, y_2)$, with dimensions p_1 and p_2 and parameters (ξ, Ω, λ), the marginal distributions of y_i with $i = 1, 2$ are given by

$$y_i \sim \mathcal{MST}_i(\xi_i, \Omega_{ii}, \lambda_{i(j)}, \nu) \tag{12}$$

with

$$\lambda_{i(j)} = \frac{\lambda_i + \bar{\Omega}_{ii}^{-1}\bar{\Omega}_{ij}\lambda_j}{\sqrt{1 + \lambda_j'\tilde{\Omega}_{ii|j}\lambda_j}} \quad \text{and} \quad \tilde{\Omega}_{ii|j} = \bar{\Omega}_{jj} - \bar{\Omega}_{ji}\bar{\Omega}_{ii}^{-1}\bar{\Omega}_{ij} \tag{13}$$

Expression (13) shows that the shape parameter of the marginal distribution is a weighted sum of all elements of the vector of individual shape parameters λ, with weights depending on the correlation between y_i and y_j. Thus, $\lambda_i = 0$ does generally not imply that the marginal distribution of y_i is symmetric.[10] Figure 3 illustrates this in a simple two-dimensional example with $\xi_i = 0$ and $\omega_i = 1$ for $i = \{1, 2\}$ and a correlation coefficient of $\rho = 0.8$. The shape parameters are given as $\lambda_1 = -2$ and $\lambda_2 = 0$. Based on equation 13, the values for the shape parameters of the marginal distributions are then given by

$$\lambda_{i(j)} = \frac{\lambda_i + \rho\lambda_j}{\sqrt{1 + \lambda_j^2(1 - \rho^2)}} \tag{14}$$

which yields $\lambda_{1(2)} = -2$ and $\lambda_{2(1)} = -1.024$. Thus, the positive correlation between y_1 and y_2 introduces negative skewness in both marginal distributions. From expression (14) it is also clear that a negative correlation of y_1 and y_2 has an offsetting effect on the marginal shape parameter such that the marginal distributions are skewed in opposite directions. We use exactly

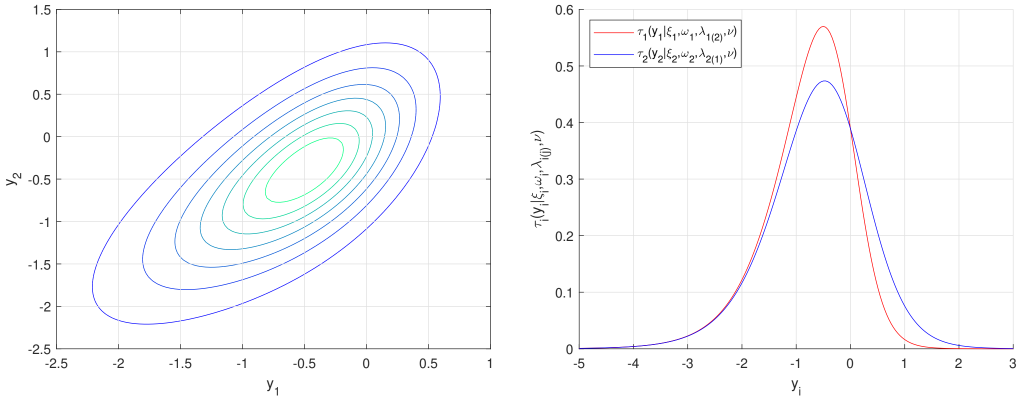

Figure 3: Bivariate Skew-T Density with Marginal Densities

this property of the multivariate skew-T distribution to analyse the effects of tail-risks in the marginal forecast densities of oil-prices on other macroeconomic variables.

We use our algorithm to introduce information about the shape of the distribution of the oil price and analyze how this will affect risks to other macroeconomics variables such as GDP or measures of inflation. Depending on the correlations between the variables, tilting the distribution of the oil prices based on the option-implied moments will result in asymmetric density forecasts for other variables than the oil price.

[10] As shown in Arellano-Valle and Genton (2010) a necessary and sufficient condition for $\lambda_{i(j)} = 0$ is that $\lambda_i = -\bar{\Omega}_{ii}^{-1}\bar{\Omega}_{ji}\lambda_j$.

3.2 Fitting a skew-T distribution to oil price forecasts

To fit the multivariate Skew-T distribution to the option implied moments, we obtain external information on the mean μ_i^{oil}, standard deviation σ_i^{oil} and skewness γ_i^{oil} of the marginal forecast densities from derivatives on the Price of Brent Crude Oil for forecast horizons from $i = 1$ up to $i = 6$ quarters.[11] From Proposition 3 in Arellano-Valle and Genton (2010), it follows that the marginal forecast densities of the oil price for each quarter are univariate skew-T distributions. Thus, we can use the results from Azzalini and Capitanio (2003) to obtain the parameters of the marginal skew T-densities. We match the option implied moments to the theoretical moments of the skew-T distribution by solving the following equations

$$\gamma_i^{oil} = \kappa_{i(j)} \left[\frac{\nu\left(3-\delta_{i(j)}^2\right)}{\nu-3} - \frac{3\nu}{\nu-2} + 2\kappa_{i(j)}^2 \right] \left[\frac{\nu}{\nu-2} - \kappa_{i(j)}^2 \right] \tag{15}$$

$$\kappa_{i(j)} = \frac{\sqrt{v}\Gamma\left(\frac{1}{2}(\nu-1)\right)}{\sqrt{\pi}\Gamma\left(\frac{1}{2}\nu\right)}\delta_{i(j)} \tag{16}$$

$$\lambda_{i(j)} = \frac{\delta_{i(j)}}{\sqrt{1-\delta_{i(j)}^2}} \tag{17}$$

$$\sigma_i^{oil} = \omega_i \sqrt{\left[\frac{\nu}{\nu-2} - \kappa_{i(j)}^2 \right]} \tag{18}$$

$$\mu_i^{oil} = \xi_i + \omega_i \kappa_{i(j)} \tag{19}$$

with respect to the parameters $\xi_i, \omega_i, \lambda_{i(j)}$ for $i = 1, ..., 6$. To provide a maximum amount of flexibility for the PDF we set $\nu = 5$.[12] After obtaining the shape parameters of each marginal density $\lambda_{i(j)}$ we obtain the shape parameters of the joint distribution in a second step using expression 13. We obtain an estimate of the correlation matrix based on the draws from our model. Subsequently, we jointly solve for each λ_i which is given as

$$\lambda_i = \lambda_{i(-i)} \sqrt{1 + \lambda_{-i}' \tilde{\Omega}_{ii|-i} \lambda_{-i}} - \bar{\Omega}_{ii}^{-1} \bar{\Omega}_{-ii}' \lambda_{-i} \tag{20}$$

[11] Since options are only traded for horizons 1,...,4 and 6 quarters ahead, we interpolate the values for the 5 quarter using a cubic spline.

[12] Based on the results in Azzalini and Capitanio (2003) page 17, this is the smallest value of ν for which the first 4 moments of the multivariate Skew-T are defined.

Hence, while we allow for changes in the mean, standard deviation and skewness of the target distribution we assume that the model-based correlation between the different forecasting horizons does not change with regards to the target distributions. Once we have obtained the fitted parameter vectors $\hat{\xi}$, $\hat{\Omega} = \hat{\omega}^{-1}\bar{\Omega}\hat{\omega}^{-1}$ and $\hat{\lambda}$ as solutions to equations (15) - (20), we can specify the target density of our algorithm as

$$p_\eta(y_i^{oil}) = \tau(y^{oil}|\hat{\xi}, \hat{\Omega}, \hat{\lambda}, \nu) \tag{21}$$

3.3 Introducing market-based densities information in a BVAR model

Once we have the target density obtained using the methodology in the previous section for the price of oil, we need to incorporate it into model-based forecasts to infer how other variables are impacted. In our example, we obtain $q_\theta(y)$ from the forecasting density of a reduced form BVAR model

$$y_t = \zeta + A_1 y_{t-1} + ... + A_s y_{t-s} + u_t \quad \text{with} \quad u_t \sim \mathcal{N}(0, \Sigma_u) \tag{22}$$

As the endogenous variables of the model, we include the log of the price of oil, the log of real GDP, the log of prices including and excluding energy as well as the log of the US/Dollar exchange rate, log of employment and the long and short term interest rates.

We estimate the BVAR model using Bayesian methods, under a Minnesota prior. To find the optimal hyperparameters of the prior, we use the hierarchical approach of Giannone, Lenza, and Primiceri (2015) that is based on maximizing the marginal data density of the BVAR model with respect to those hyperparameters. Since we estimate the model using Bayesian methods we obtain sets with I posterior draws for both the intercept ζ_i and the slope coefficients $A_{j,i}$, $j = 1, .., 5$ as well as the elements of Σ_i. Additionally, we use the novel methodology of Lenza and Primiceri (2020) to deal with the Covid period in the first quarters of 2020. We sample with replacement from the posterior draws to generate the model consistent forecasts up to h periods, $y_i = [y'_{i,t}, y'_{i,t+1}, ..., y'_{i,t+h}]'$. Thus, we also incorporate parameter uncertainty from the posterior densities into our risk analysis. The density of the proposal distribution for the i^{th}

draw y_i is then given by the multivariate forecasting distribution of the model and takes the form

$$q_\theta(y_i) = \varphi(y|\mu_i, \Sigma_i) \tag{23}$$

where $\varphi(...)$ denotes the density function of the multivariate normal distribution with mean μ_i and Variance Covariance Matrix Σ_i. More details on how to obtain this forecasting density are provided in the Appendix.

3.4 Adjusting the forecast densities

Once we have obtained the parameters for the multivariate Skew-T distribution, we use the algorithm described in Section 2.3 to adjust the marginal forecast densities of oil prices from the BVAR to the the option implied forecast densities. Based on 25000 draws for μ and Σ from the BVAR we generate 50000 model consistent forecasts y_i. For our algorithm, we use the the adaptive tempering schedule of Herbst and Schorfheide (2019) to obtain the optimal values for N_ϕ and ϕ_n. In each iteration, we optimize ϕ_n such that the inefficiency ratio is equal to a target ratio $r^* > 1$

$$\phi_n = \text{argmin} \; \frac{1}{M} \sum_i^M \left[\frac{W_{i,n}}{\frac{1}{M} \sum_{i=1}^M W_{i,n}} \right]^2 - r^* \tag{24}$$

with $W_{i,n}$ given as in Section 2.3 Step 3 (a). Setting r^* closer to 1 results in a better approximation of $p_\eta(y^{oil})$ but also increases the number of tempering steps. To obtain a precise approximation of the target density we set $r^* = 1.01$. For the mutation of the particles in Step 3 (c) we set $H = 10$. Given this set-up, we first use our method to investigate the effect of the strong increase in oil prices on inflation due to the begin of the war in Ukraine in the first quarter of 2022. Subsequently, we evaluate the gains of introducing external information from options into forecasting densities in a real-time forecasting exercise.

3.5 Results from the BVAR

As can be seen from figure 2, option-implied moments of the forecasting distribution of the oil price have sharply increased over the first two quarters in 2022, with option implied skewness peaking in the last quarter. We estimate the BVAR model using data up to the first quarter of

2022 and introduce the information of the option-implied densities at the $4th$ of March. The option-implied moments for all forecasting horizons h at that point are displayed in Table 1. Since there are no options with an expiry date of 15 months traded the missing values for the 5 month-ahead forecasting densities are interpolated using a cubic spline.

h	**Mean**	**SD**	**Skewness**
1	110.2	38.88	1.8
2	103.16	40.64	1.56
3	98.92	40.59	1.28
4	95.3	41.1	1.14
5	92.13	41.88	1.09
6	89.75	41.99	1.01

Table 1: Option-implied moments

While the mean of the distribution is monotonically decreasing by approximately 20 euros per barrel over the forecasting horizon, uncertainty (standard deviation) is increasing from 38.88 to 41.99 at the same time. Additionally, the distribution is significantly skewed to the right with values larger than 1 for all horizons indicating increased upside risks to the price of oil. Yet, the skewness is also decreasing over time, which indicates that risks become more symmetric over time. Based on the properties of the Skew-T distribution described in 3.1, we use our algorithm to introduce the information about the forecasting densities of oil to our model to investigate if and how the forecasting distributions of inflation are affected. By introducing information about the full density, we can gauge the effects on the point forecasts as well as the effect that market-based oil upside risks have on other model variables. The skewness implied by the densities of the oil price will affect the distributions of other model variables depending on the correlations implied by the BVAR. Given the debate about the pass-through of high energy prices to inflation, we are particularly interested in the effect on both, inflation with and without energy prices. Figure 4 shows the fitted marginal skew T-densities that we obtain from the values in Table 1 together with the histogram of the final particles $\{y^{Oil}_{i,N_\phi}\}_{i=1}^{N}$ with $N = 50000$. Based on our targeted inefficiency ratio $r^* = 1.01$, the algorithm required a number of $N_\phi = 40$ tempering steps to move the original particles form the proposal distribution to the target distribution. The draws are very well adapted to the target distribution that is implied by the options and provide a precise approximation. The positive skewness and increasing volatility is clearly visible from the theoretical distribution as well as from the adapted

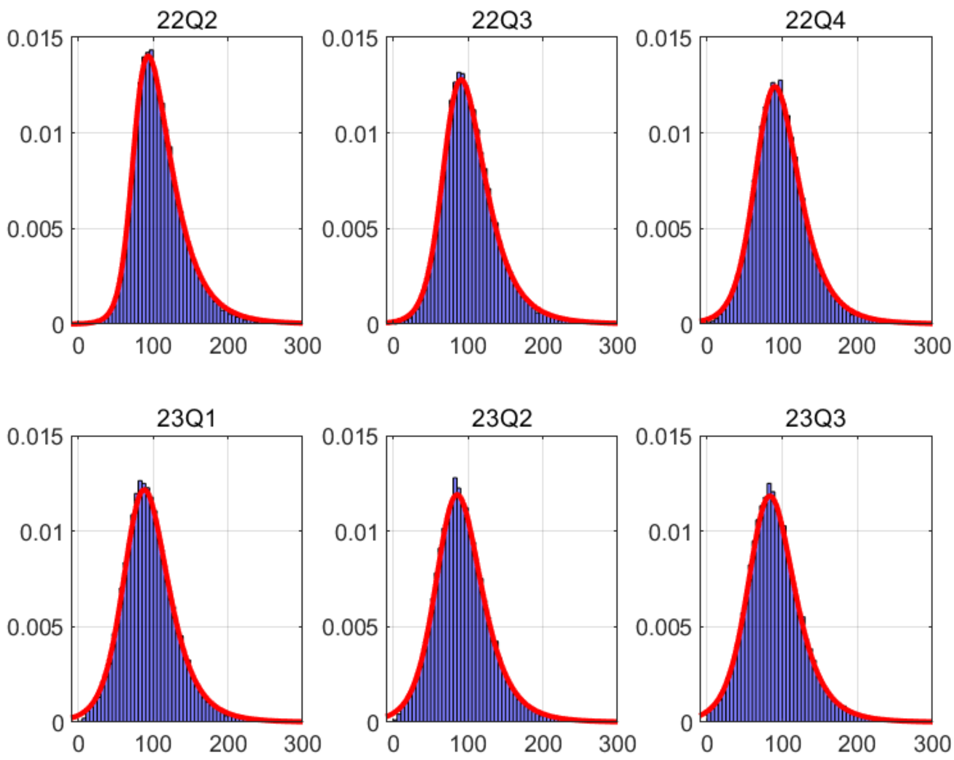

Figure 4: Option-implied probability density functions

draws.

Figure 5 shows the resulting densities for inflation and core inflation. The shaded areas show the 16, 25, 75 and 84 percent quantiles of resulting forecasting distribution of the annualized inflation rate together with the median given by the solid black line. Additionally, the dotted red lines show the 16 and 84 percent quantiles of the original distribution. Figure 8 in the Appendix also shows the histograms of the annualized growth rates of the oil-price, inflation and core inflation for all forecasting horizons. In both cases, introducing the information of the options results in an upward shift of the full distributions. In case of inflation, the new median nearly coincides with the original 84 percent quantile of the original distribution in the first two periods. The forecasting density of core inflation is similar to the original model in the first period but subsequently deviates from the original model with significantly higher values over the rest of the forecasting horizon. Additionally, the positive skewness in the distribution of the oil prices results in upside risks to inflation as can be seen from the quantiles and the histograms in Figure 8. Hence, while headline inflation (i.e. including energy prices) reacts contempora-

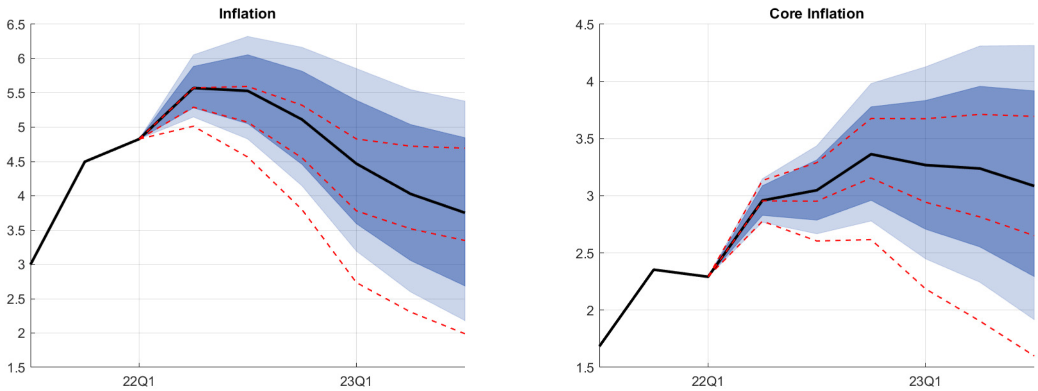

Figure 5: Option-implied forecasting densities

neously, the effects on core inflation appear one quarter later, reflecting second round effects arising from the upside risks in the distribution of the price of oil. Finally, the median forecast for core inflation remains persistently elevated over the forecasting horizon compared to the original forecast from the BVAR.

3.6 Forecasting GDP and Inflation

To evaluate the effect of conditioning on information about the market-based forecasting distribution of our BVAR model, we look at the probabilistic forecasting performance in a real time forecasting exercise to forecast GDP, inflation and core inflation. We estimate the same BVAR as in section 3.3 using data vintages starting in the last quarter of 2013 up until the third quarter of 2021. With the onset of the Covid pandemic we again use the method of Lenza and Primiceri (2020). Subsequently, we use our algorithm to impose the option-implied distribution at the end of the quarter to the forecasting density of the oil-price. Since our methodology seeks to incorporate information about the full distribution, we use the continuous ranked probability score (CRPS) as the metric to evaluate the density forecasts. The CRPS generalizes the Mean Squared Error to take into account the the full forecasting distribution. It can be formalised as,

$$CRPS(F,x) = \int_{-\infty}^{\infty} (F(y) - \mathbf{1}(y-x))^2 \, dy \tag{25}$$

where x is the realized value, F is the cumulative distribution function implied by the density forecast of the model and $\mathbf{1}(...)$ denotes the heavyside function. Figure 6 shows the ratios of the

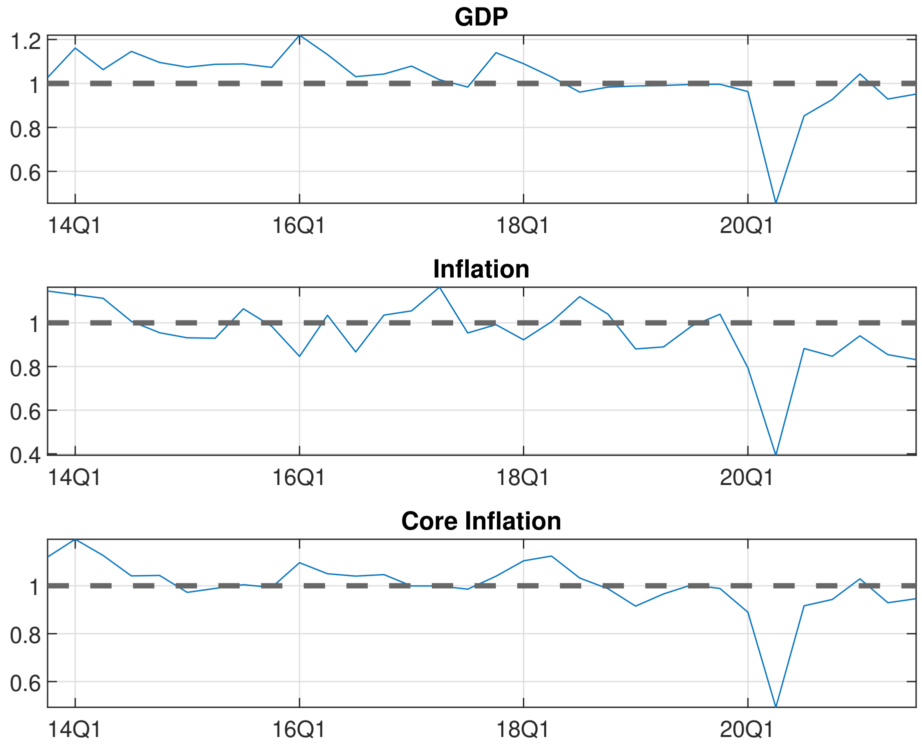

Figure 6: Continuously Ranked Probability Scores

mean CRPS for the symmetric density forecasts under Q_θ and the skew-T forecasts under P_η

$$R_t = \frac{\frac{1}{P}\sum_{i=1}^{P} CRPS(Q_\theta, x_{t+i}^k)}{\frac{1}{P}\sum_{i=1}^{P} CRPS(P_\eta, x_{t+i}^k)} \tag{26}$$

for $k =$ GDP, inflation, core inflation. The tables with values for each period are included in the Appendix. The results indicate that while including additional information on the distribution from the options does not increase predictive accuracy in moderate periods with stable economic conditions, it strongly increases the probabilistic forecasts accuracy in times of economic turmoil during the onset of the Covid pandemic in the first and second quarter of 2020. This is in line with other findings on skewed density forecasts such as Adrian, Boyarchenko, and Giannone (2019) who show that conditional forecast densities of macroeconomic variables are symmetric in normal times but become skewed in times of crisis.

3.7 Tilting the forecast densities from a DSGE model

Finally, we show that our methodology can also be applied in large models, such as a DSGE. Note that the reduced form solution of a DSGE model can be written as,

$$x_t = J + Qx_{t-1} + G\varepsilon_t \tag{27}$$

where x_t is a vector of endogenous and exogenous state variables in the model, and ε_t is a vector of i.i.d. structural shocks. Yet, note that in general the number of structural shocks in a DSGE model is smaller than the vector of state variables. Usually, for the estimation of a DSGE, there is a subset of variables that are observed, y_t, so that,

$$y_t = Hx_t \tag{28}$$

where H is a selection matrix. With these two equations in hand, we show in the Appendix how to construct the proposal density $q_\theta(y)$ as in the case for the BVAR.

We then repeat the analysis in section 3.5 using the ECB's New Area Wide Model II (NAWM II) from Coenen et al. (2018). The NAWM II is an estimated dynamic, stochastic, general equilibrium (DSGE) model of the euro area as a whole. The model incorporates a rich financial sector that allows for (i) accounting for a genuine role of financial frictions in the propagation of economic shocks as well as macroeconomic policies and for the presence of shocks originating in the financial sector itself, (ii) capturing the prominent role of bank lending rates and the gradual interest-rate pass-through in the transmission of monetary policy in the euro area, and (iii) providing a structural framework useable for assessing the macroeconomic impact of the ECB's large-scale asset purchases conducted in recent years. For the exercise in this section, we slightly modify the model to account for a faster pass-through of oil prices to inflation. In the original version of the model, there is one foreign intermediate-good firm that sells its goods in the domestic euro area market. The marginal costs of this firm is a weighted average of oil prices and foreign prices. Then, the firm is subject to staggered price contracts à la Calvo when setting the final domestic price, which introduces a sluggish price adjustment.

For this analysis, we separate the problem into two firms. One that sets domestic prices for imported oil, and the second one that only takes care of foreign prices. The firm that sets oil prices has a smaller Calvo parameter, reflecting a faster pass-through of oil prices into final import prices, and thus, into the private consumption deflator and HICP in the model.

Figure 7 shows the results of introducing market based data on oil prices in the forecasting distribution of the NAWM II. In the figure, the blak lines show the model-based forecast, while the blue shaded areas represent the tilted distributions so that the distribution of the price of oil matches the market-based measures. In the model, the price of oil behaves as a supply side shock. Thus, once we incorporate information from the markets that assumes that the distribution of the price of oil is skewed to the upside, the transmission channel in the model indicates significant downside risks to the real economy, which in the figure is represented by annual GDP growth, and upside risks to inflation, represented by inflation in the private consumption deflator. The figure also shows that the final distributions are skewed, and the asymmetries are inherited from the skewness in the market-based options.

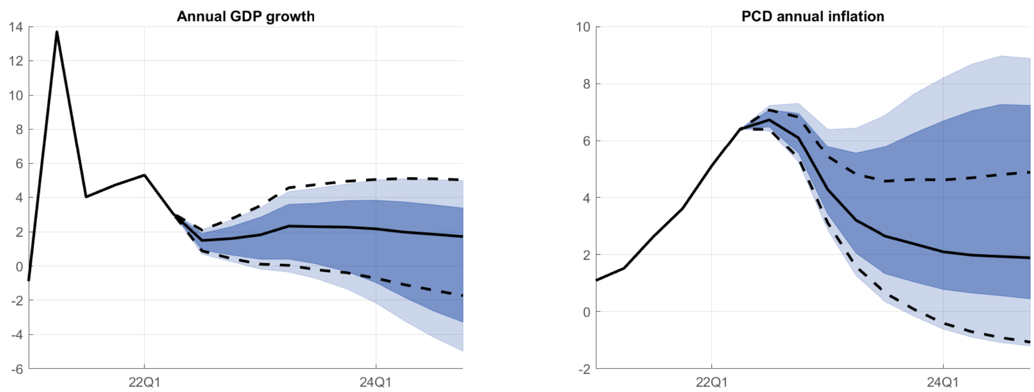

Figure 7: Option-implied forecasting densities in the NAWM II

4 Extensions

This section proposes several extensions of our core algorithm that make our methodology applicable in case some of the assumptions for the algorithm in section 2 are not satisfied. As described in section 2.3, our methodology requires knowledge of the conditional distribution of the proposal distribution to recover the values of the variables y^{-e}. Second, we use a tempering method that requires a target density with a parameter to control the scale of the distribution. In this section we propose two remedies in case the application at hand does not meet these requirements.

First, if the conditional distribution is not available, but the researcher has access to draws from any arbitrary distribution, it is possible to approximate the proposal density $q(x)$ with a Gaussian mixture density of the form

$$q(x) = \sum_{k=1}^{K} \pi_k \varphi(x|\mu_k, \Sigma_k). \tag{29}$$

with weights $0 < \pi_k < 1$ that satisfy $\sum_{k=1}^{K} \pi_k = 1$. The Gaussian mixture density then has the conditional density

$$q(x_1|x_2) = \sum_{k=1}^{K} \left[\frac{\pi_k \varphi(x_2|\mu_{k,2}, \Sigma_{k,22})}{\sum_{l=1}^{L} \pi_l \varphi(x_2|\mu_{l,2}, \Sigma_{l,22})} \right] \varphi(x_1|x_2, \mu_{k,1|2}, \Sigma_{k,1|2}) \tag{30}$$

that can be used to sample from in step 4.

With regards to the second point, we propose another way to define the bridge distributions as given in Neal (2001)

$$p_n(y_i) = p_\eta(y_i)^{\phi_n} q_\theta(y_i)^{(1-\phi_n)} \tag{31}$$

In this specification, the bridge distributions are given by the geometric average of the model-implied density and the target density. Gradually increasing ϕ from 0 towards 1 will sequentially adapt the particles proposed by the model to the final distribution.[13] Expression 31 provides an attractive alternative if the scale of the distribution is not captured by a specific parameter. It is also possible to implement an adaptive tempering schedule instead of working with a predetermined sequence for $\{\phi_n\}_{n=1}^{N_\phi}$. With these remedies in hand, we consider our methodology applicable to a wide variety of problems.

[13] In our application, we experimented with both specifications and found that in both cases, our algorithm is able to adapt draws to the target distributions even if the Kullback-Leibler divergence is high.

5 Conclusion

In this paper, we develop a methodology that can be used to condition probabilistic forecasts of a model on off-model information about the marginal distributions of some of the model variables. More technically, the algorithm uses the tempered importance sampling method of Neal (2001) and Herbst and Schorfheide (2014) to adapt draws from a model-based distribution to a target distribution that satisfies the external information one intends to condition the forecast on. Our algorithm allows applications where the proposed draws are far away from the target density in a Kullback-Leibler sense as well as conditioning on information on transformations of the model variables. This makes our method superior to the entropic tilting methodology of Robertson, Tallman, and Whiteman (2005) whose method is similar in spirit but less robust and less flexible.

We illustrate our algorithm by introducing off-model information about the distribution of future oil prices into the forecasting densities of a BVAR. The information is obtained from option prices and indicates significant amounts of skewness at all available forecasting horizons. Using the results of Azzalini and Capitanio (2003) we model the option-implied marginal forecasting densities as skew T and apply our methodology to investigate the transmission of upside risks to future oil prices on future inflation and core inflation in the first quarter of 2022. Due to the war in Ukraine, option-implied forecasting distributions of oil prices exhibit large positive skewness and increased volatility over the full forecasting horizon. We find that adapting the forecasting distributions of the BVAR to the option-implied densities results in upside risks to inflation and core inflation. Furthermore, median forecasts of core inflation remain persistently higher over the forecasting horizon compared to the symmetric forecast densities of the BVAR. We also investigate the forecasting accuracy of the density forecasts in real time. We focus on the probabilistic forecasts of GDP, inflation and core inflation over the period of 2013 Q4 up to 2021 Q4. Based on the CRPS, our results indicate that introducing information about the marginal distribution of oil prices improves forecasts for GDP and inflation measures during the Covid pandemic compared to symmetric forecasting distributions. This is in line with results of Adrian, Boyarchenko, and Giannone (2019) who find skewness on conditional forecasting densities in times of economic turmoil.

Our methodology as well as our application is widely applicable and provides several extensions for further research such as introducing information from traded derivatives with other underlyings such as interest rates or exchange rates. Additionally, the model implied forecasting densities are not limited to time-series models such as BVARs but can also be applied to forecasting distributions of DSGE models or semi-structural models.

References

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone (2019). "Vulnerable growth". *American Economic Review* 109(4), pp. 1263–89.

Arellano-Valle, R. and Marc Genton (Dec. 2010). "Multivariate extended skew-t distributions and related families". *Metron - International Journal of Statistics* LXVIII, pp. 201–234.

Azzalini, Adelchi (2013). *The Skew-Normal and Related Families*. Institute of Mathematical Statistics Monographs. Cambridge University Press.

Azzalini, Adelchi and Antonella Capitanio (2003). "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution". *Journal of the Royal Statistical Society Series B* 65(2), pp. 367–389.

Bańbura, Marta, Domenico Giannone, and Michele Lenza (2015). "Conditional forecasts and scenario analysis with vector autoregressions for large cross-sections". *International Journal of Forecasting* 31(3), pp. 739–756.

Betancourt, Michael (2017). *A Conceptual Introduction to Hamiltonian Monte Carlo*. Tech. rep. ArXiv e-prints.

Breeden, Douglas T and Robert H Litzenberger (1978). "Prices of State-contingent Claims Implicit in Option Prices". *The Journal of Business* 51(4), pp. 621–51.

Chopin, Nicolas (2004). "Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference". *The Annals of Statistics* 32(6), pp. 2385–2411.

Coenen, Günter et al. (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". *ECB Working Paper No 2200*.

Delle Monache, Davide, Andrea De Polis, and Ivan Petrella (2021). "Modeling and forecasting macroeconomic downside risk". *Bank of Italy Temi di Discussione (Working Paper) No* 1324.

Doucet, Arnaud, Nando de Freitas, and Neil J. Gordon, eds. (2001). *Sequential Monte Carlo Methods in Practice*. Statistics for Engineering and Information Science. Springer.

Giacomini, Raffaella and Giuseppe Ragusa (2014). "Theory-coherent forecasting". *Journal of Econometrics* 182(1), pp. 145–155.

Giannone, Domenico, Michele Lenza, and Giorgio Primiceri (2015). "Prior Selection for Vector Autoregressions". *The Review of Economics and Statistics* 97(2), pp. 436–451.

Herbst, Edward and Frank Schorfheide (2014). "Sequential Monte Carlo Sampling for DSGE Models". *Journal of Applied Econometrics* 29(7), pp. 1073–1098.

Herbst, Edward and Frank Schorfheide (2019). "Tempered particle filtering". *Journal of Econometrics* 210(1), pp. 26–44.

Klebaner, Fima C. (2012). *Introduction to stochastic calculus with applications*. 3rd ed. Imperial College Press London.

Kloek, T. and H. K. van Dijk (1978). "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo". *Econometrica* 46(1), pp. 1–19.

Krüger, Fabian, Todd E. Clark, and Francesco Ravazzolo (2017). "Using Entropic Tilting to Combine BVAR Forecasts With External Nowcasts". *Journal of Business & Economic Statistics* 35(3), pp. 470–485.

Lenza, Michele and Giorgio Primiceri (Sept. 2020). *How to Estimate a VAR after March 2020*. Working Paper 27771. National Bureau of Economic Research.

Montes-Galdón, Carlos and Eva Ortega (Mar. 2022). *Skewed SVARs: Tracking the Structural Sources of Macroeconomic Tail Risks*. Working Paper 2208. Banco de Espana.

Neal, Radford (2001). "Annealed importance sampling". *Statistics and Computing* 11, pp. 125–139.

Robertson, John C., Ellis W. Tallman, and Charles H. Whiteman (2005). "Forecasting Using Relative Entropy". *Journal of Money, Credit and Banking* 37(3), pp. 383–401.

Vincent-Humphreys, Rupert de and Josep Maria Puigvert Gutiérrez (2010). *A quantitative mirror on the Euribor market using implied probability density functions*. ECB Working Paper 1281.

Waggoner, Daniel and Tao Zha (1999). "Conditional Forecasts In Dynamic Multivariate Models". *The Review of Economics and Statistics* 81(4), pp. 639–651.

Wolf, Elias (Feb. 2022). *Estimating growth at risk with skewed stochastic volatility models*. Working Paper 2022/2. Freie Universität Berlin.

A Appendix

A.1 Deriving the Proposal Density in a VAR model

Rewriting the VAR-Model as a VAR(1) gives

$$y_t = c_i + \Phi_i y_{t-1} + G_i \varepsilon_t \tag{32}$$

where Φ_i is the companion matrix of the i^{th} posterior draw for the slope coefficients and c_i the corresponding vector of intercepts. G_i is a lower-triangular matrix such that $G_i G_i' = \Sigma_{u,i}$ and $\varepsilon_t \sim \mathcal{N}(\varepsilon|0, I)$. Iterating the equation forward in time gives for the h step ahead forecast

$$y_{T+h} = \sum_{j=1}^{h} \Phi_i^{j-1} c_i + \Phi_i^h y_T + \sum_{j=1}^{h} \Phi_i^{h-j} G_i \varepsilon_{T+j} \tag{33}$$

Stacking the realizations over the full forecasting horizon in a vector y_i yields

$$\begin{bmatrix} y_{T+1} \\ y_{T+2} \\ \vdots \\ y_{T+h} \end{bmatrix} = \begin{bmatrix} \tilde{c}_{i,T+1} \\ \tilde{c}_{i,T+2} \\ \vdots \\ \tilde{c}_{i,T+h} \end{bmatrix} + \begin{bmatrix} G_i & 0 & 0 & 0 \\ \Phi_i G_i & G_i & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ \Phi_i^{h-1} G_i & \Phi_i^{h-2} G_i & \cdots & G_i \end{bmatrix} \begin{bmatrix} \varepsilon_{T+1} \\ \varepsilon_{T+2} \\ \vdots \\ \varepsilon_{T+h} \end{bmatrix}$$

with $\tilde{c}_{i,T+h} = \sum_{j=1}^{h} \Phi_i^{j-1} c_i + \Phi_i^h y_T$. Redefining the terms results in the simple expression

$$y_i = \mu_i + \mathcal{G}_i \varepsilon \tag{34}$$

where $\mu = [\tilde{c}_{i,T+1}, ..., \tilde{c}_{i,T+h}]'$. The matrix $\mathcal{G}_i$ is a lower-triangular matrix that captures the correlations of the reduced form errors $u_t = G\varepsilon_t$ between the model variables as well as between the different time periods. These correlations depend on the values of the posterior draws for Φ_i and $\Sigma_{u,i}$. Given the distributional assumption about ε it follows that

$$y_i \sim \mathcal{N}(y|\mu_i, \Sigma_i) \tag{35}$$

with $\Sigma_i = \mathcal{G}_i \mathcal{G}_i'$. This gives the proposal density in equation 23.

A.2 Deriving the Proposal Density in a DSGE model

A similar reasoning as in the VAR model can be applied in the case of a DSGE model. The reduced form representation of a linearised DSGE model takes the form,

$$x_t = J + Qx_{t-1} + G\varepsilon_t \tag{36}$$

where x_t is a $N \times 1$ vector of endogenous and exogenous variables in the model, and ε_t is a $P \times 1$ vector of i.i.d. structural shocks. Yet, note that in general $N > P$. This implies that $\Sigma = GG'$ is a reduced rank matrix and thus not invertible. This implies that we cannot proceed exactly as in the case of the VAR model, and as the proposal density in equation 35 cannot be evaluated. However, as in the case of the estimation of a DSGE model, we can focus on a subset of variables in x_t which are assumed to be observed. If H is a matrix that selects some variables (or a combination of them), then we can write,

$$y_t = Hx_t = HJ + HQx_{t-1} + HG\varepsilon_t = \tilde{J} + \tilde{Q}x_{t-1} + \tilde{Q}\varepsilon_t \tag{37}$$

where we assume that the dimension of y_t is $P \times 1$. That is, the number of observed variables is equal to the number of fundamental shocks in the model. Then, given an initial condition for all the variables in the model, x_T, which could be recovered running a Kalman filter for example, we can proceed similarly as in the case of the VAR. Iterating forward, we can get,

$$y_{T+h} = \sum_{j=1}^{h} \tilde{Q}^{j-1}\tilde{J} + \tilde{Q}^{h}x_T + \sum_{j=1}^{h} \tilde{Q}^{h-j}\tilde{G}\varepsilon_{T+j} \tag{38}$$

And with the latter expression in hand, it is straightforward to compute an expresion for the mean and the covariance matrix in 35 using the same matrix representation as in the VAR case.

A.3 Additional Plots

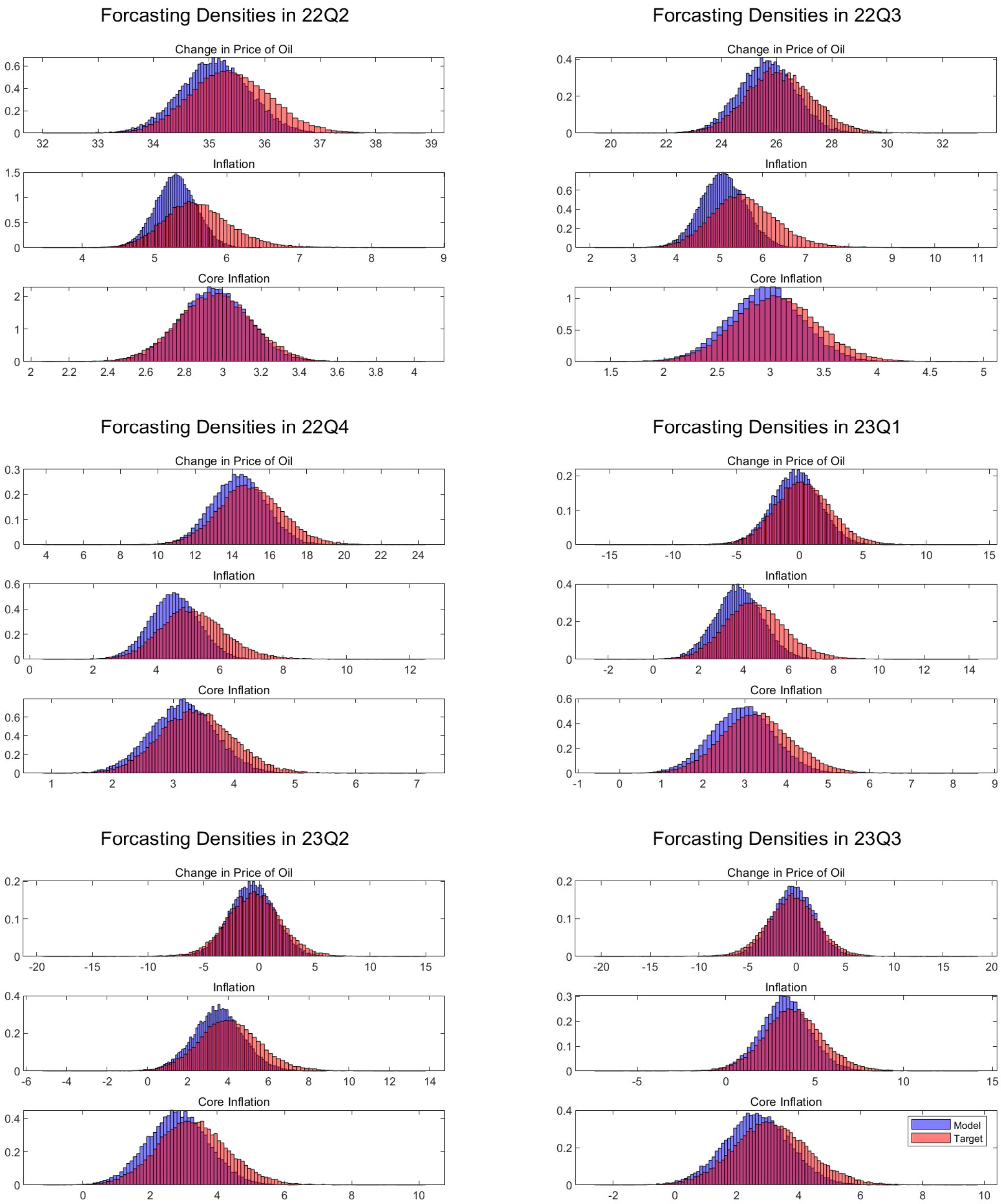

Figure 8: Forecasting densities of annual growth rates

T	h=1	h=2	h=3	h=4	h=5	h=6
14Q1	1.233	1.032	**0.979**	**0.851**	1.253	**0.917**
14Q2	1.258	1.065	**0.993**	1.601	1.045	1.017
14Q3	1.019	1.066	1.129	1.102	1.044	1.017
14Q4	1.023	1.089	**0.985**	1.252	1.265	1.219
15Q1	1.227	1.032	**0.948**	1.121	1.179	1.048
15Q2	1.005	1.039	**0.991**	1.258	1.013	1.144
15Q3	**0.952**	**0.936**	1.234	1.086	1.206	1.118
15Q4	1.053	1.066	1.015	1.124	1.148	1.101
16Q1	1.159	**0.979**	1.093	1.039	1.069	1.079
16Q2	1.308	1.159	1.089	1.287	1.274	1.185
16Q3	1.125	1.032	1.053	1.190	1.231	1.102
16Q4	**0.794**	**0.874**	**0.954**	1.291	1.173	1.046
17Q1	**0.859**	**0.939**	1.033	1.231	**0.946**	1.125
17Q2	1.064	1.124	1.125	**0.962**	1.184	**0.982**
17Q3	1.019	1.068	**0.901**	1.159	**0.888**	1.171
17Q4	**0.880**	1.004	1.108	**0.676**	1.358	1.210
18Q1	1.133	1.067	**0.885**	1.504	1.387	**0.962**
18Q2	1.134	0.982	1.207	1.573	**0.863**	**0.942**
18Q3	1.091	1.034	1.153	1.041	**0.984**	**0.893**
18Q4	**0.851**	**0.970**	1.034	1.033	**0.838**	**0.968**
19Q1	0.924	1.101	1.030	**0.848**	**0.963**	**0.991**
19Q2	1.135	1.042	**0.922**	**0.946**	**0.987**	1.004
19Q3	1.028	**0.919**	**0.968**	**0.987**	1.007	**0.865**
19Q4	1.041	**0.985**	**0.994**	1.005	**0.904**	**0.927**
20Q1	**0.997**	**0.996**	1.004	**0.790**	**0.847**	1.014
20Q2	1.001	**0.974**	**0.637**	**0.861**	**0.846**	**0.878**
20Q3	**0.835**	**0.439**	**0.271**	**0.296**	**0.214**	**0.146**
20Q4	**0.729**	**0.927**	**0.748**	1.122	**0.967**	1.057
21Q1	**0.696**	1.014	1.012	1.089	1.091	1.006
21Q2	**0.742**	1.086	1.181	1.112	1.177	1.084
21Q3	**0.661**	1.199	1.067	1.117	1.242	1.080
21Q4	**0.566**	1.099	1.156	1.085	1.298	1.218

Table 2: Ratios of CRPS for GDP

T	h=1	h=2	h=3	h=4	h=5	h=6
14Q1	1.652	1.269	1.048	1.071	**0.888**	1.322
14Q2	2.008	1.310	**0.908**	**0.999**	1.220	**0.776**
14Q3	2.923	1.201	**0.974**	**0.966**	**0.874**	**0.8**
14Q4	1.463	1.077	1.163	**0.937**	**0.829**	**0.79**
15Q1	**0.935**	1.298	**0.866**	**0.893**	**0.761**	1.395
15Q2	**0.692**	1.014	**0.919**	**0.928**	1.125	1.01
15Q3	**0.982**	**0.913**	**0.871**	**0.970**	**0.899**	1.152
15Q4	1.454	**0.919**	1.171	1.201	1.174	1.075
16Q1	**0.794**	1.282	1.211	1.250	1.178	**0.936**
16Q2	**0.470**	**0.997**	**0.856**	**0.788**	1.178	1.175
16Q3	1.551	**0.813**	**0.939**	1.035	1.163	1.151
16Q4	**0.466**	**0.736**	1.142	1.332	**0.980**	1.032
17Q1	**0.589**	1.150	1.063	1.169	1.151	1.166
17Q2	**0.967**	**0.938**	1.121	1.163	1.215	1.247
17Q3	1.354	1.186	1.123	1.105	1.217	1.065
17Q4	**0.887**	**0.941**	1.079	1.121	1.056	**0.831**
18Q1	**0.946**	**0.862**	1.148	1.083	**0.850**	1.309
18Q2	**0.578**	1.015	**0.983**	**0.950**	1.234	**0.855**
18Q3	1.550	1.191	**0.922**	1.163	**0.926**	**0.875**
18Q4	2.096	1.087	1.113	1.116	1.007	**0.881**
19Q1	1.238	**0.960**	1.168	1.260	1.017	**0.875**
19Q2	**0.529**	1.250	1.163	1.030	**0.883**	**0.862**
19Q3	1.168	1.132	**0.851**	**0.877**	**0.838**	**0.783**
19Q4	1.264	**0.910**	**0.910**	1.004	**0.885**	1.046
20Q1	1.313	**0.978**	**0.912**	**0.960**	1.016	1.253
20Q2	**0.726**	1.319	**0.904**	**0.836**	**0.665**	**0.777**
20Q3	**0.188**	**0.237**	**0.518**	**0.321**	**0.395**	**0.433**
20Q4	1.72	**0.897**	**0.688**	**0.829**	**0.861**	**0.883**
21Q1	**0.666**	**0.667**	**0.899**	**0.887**	**0.915**	**0.844**
21Q2	1.01	**0.815**	1	**0.93**	**0.988**	1.135
21Q3	**0.635**	**0.872**	**0.917**	**0.827**	**0.941**	1.168
21Q4	**0.591**	**0.848**	1.019	1.213	1.041	**0.897**

Table 3: Ratios of CRPS for Inflation

T	h=1	h=2	h=3	h=4	h=5	h=6
14Q1	**0.928**	**0.93**	1.061	1.217	1.26	1.333
14Q2	**0.996**	1.003	1.063	1.215	1.345	1.314
14Q3	**0.926**	**0.984**	1.009	1.277	1.306	1.426
14Q4	1.033	**0.941**	**0.994**	**0.987**	1.301	1.16
15Q1	1.022	1.096	**0.988**	1.145	1.077	1.029
15Q2	1.017	1.067	1.034	**0.938**	**0.871**	1.041
15Q3	1.001	1.013	**0.94**	**0.869**	1.055	1.096
15Q4	1.025	1.044	**0.947**	**0.941**	1.104	1.065
16Q1	1.003	**0.885**	**0.986**	1.054	1.055	1.033
16Q2	1.068	1.207	1.052	**0.969**	1.146	1.162
16Q3	1.033	1.073	1.062	1.104	1.159	**0.923**
16Q4	**0.925**	1.004	**0.96**	1.125	0.933	1.175
17Q1	1.003	**0.91**	1.003	**0.949**	1.128	1.133
17Q2	**0.952**	**0.951**	**0.914**	1.105	1.107	1.081
17Q3	**0.991**	**0.987**	1.051	1.079	1.058	**0.928**
17Q4	1.028	1.062	1.052	1.105	**0.817**	**0.898**
18Q1	1	**0.983**	1.049	**0.834**	**0.922**	1.175
18Q2	1.079	1.072	**0.964**	**0.862**	1.205	1.258
18Q3	1.011	1.055	**0.962**	1.148	1.239	1.192
18Q4	1.063	1.148	**0.972**	1.167	1.165	**0.845**
19Q1	1.054	**0.944**	1.025	1.054	**0.876**	1.166
19Q2	**0.937**	**0.997**	**0.991**	**0.873**	1.037	**0.847**
19Q3	1.02	1.051	**0.926**	1.167	**0.895**	**0.903**
19Q4	**0.984**	1.011	1.077	**0.96**	**0.925**	1.021
20Q1	**0.986**	1.013	**0.933**	**0.89**	1.041	1.008
20Q2	**0.992**	**0.924**	**0.71**	**0.868**	**0.705**	**0.897**
20Q3	**0.829**	0.633	**0.563**	**0.384**	**0.357**	**0.248**
20Q4	**0.632**	**0.983**	**0.797**	**0.989**	**0.92**	**0.965**
21Q1	**0.837**	1.006	**0.927**	**0.986**	1.033	1.062
21Q2	1.249	**0.922**	1.057	1.013	1.123	1.055
21Q3	**0.72**	**0.893**	**0.943**	**0.98**	1.086	1.157
21Q4	**0.593**	**0.871**	1.094	**0.966**	1.203	1.253

Table 4: Ratios of CRPS for Core Inflation

Acknowledgements

This paper has benefited from valuable comments of Frank Schorfheide, Matteo Ciccarelli, Marta Bańbura, Catalina Martínez Hernández and Giorgio Primiceri.
This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Carlos Montes-Galdón
European Central Bank, Frankfurt am Main, Germany; email: Carlos.Montes-Galdon@ecb.europa.eu

Joan Paredes
European Central Bank, Frankfurt am Main, Germany; email: joan.paredes@ecb.europa.eu

Elias Wolf
European Central Bank, Frankfurt am Main, Germany; Freie Universität Berlin, Berlin, Germany; email: e.wolf@fu-berlin.de

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-5402-0 ISSN 1725-2806 doi:10.2866/781711 QB-AR-22-119-EN-N