NBS Working paper

02/2026

A robust approach to tilting: parametric relative entropy

Carlos Montes-Galdón, Joan Paredes, Elias Wolf

NATIONAL BANK OF SLOVAKIA
EUROSYSTEM

www.nbs.sk

Publisher

research@nbs.sk

Contact
National Bank of Slovakia
Imricha Karvaša 1
813 25 Bratislava

Electronic version
https://nbs.sk/en/publications/research-papers-working-and-occasional-papers-wp-op/

ISSN 2585-9269 (electronic version)

A robust approach to tilting: parametric relative entropy

Carlos Montes-Galdón[a], Joan Paredes[a,b], and Elias Wolf[c]

[a]*European Central Bank*

[b]*National Bank of Slovakia*

[c]*European Stability Mechanism Bank and Universität Bonn*

February 2026

Abstract

We introduce a novel methodology, ”parametric tilting,” for incorporating external information into econometric model-based density forecasts. Unlike traditional entropic tilting, which can generate unrealistic or unstable distributions under certain conditions, parametric tilting ensures more reliable and numerically stable results. Our approach leverages the flexibility of the skew-T distribution, which captures key moments of macroeconomic time series, and minimizes the Kullback-Leibler divergence between the target and model-based distributions. This method overcomes limitations of entropic tilting, such as multimodal or degenerate distributions, providing a robust alternative for policymakers and researchers aiming to integrate external views into probabilistic forecasting frameworks.

JEL CLASSIFICATION SYSTEM: C14, C53, E52

KEYWORDS: Forecasting, Kullback-Leibler Information Criterion, Entropic Tilting

Non-Technical Summary

In this paper, we present a new and robust approach to integrate external information into model-based predictions. Economists and policymakers often need to adjust their forecasts based on new data, such as higher-frequency observations, expert surveys, or economic theory. One standard method for doing this is called ”entropic tilting”, introduced into the econometric literature by Robertson, Tallman, and Whiteman (2005). However, in real-world situations, this approach can fail to produce accurate or reasonable results, especially when the new information differs greatly from the model's predictions.

Our proposed method, which we call ”parametric tilting,” offers a solution to these problems. Our proposed method leverages closed-form solutions for various moments of interest that are available for the skew-T distribution of Azzalini (2013). This distribution is highly flexible and can provide a better fit for complex data patterns often seen in economic variables. The parameters of the skew-T distribution are determined by minimizing the Kullback-Leibler divergence between the target and the original model-based distribution, subject to moment-constraints that are specified by the researcher. The resulting optimization problem allows the forecast to adjust more smoothly to the external information, generating predictions that are more stable and easier to interpret.

We show through examples that parametric tilting succeeds in cases where the traditional method struggles, making it a useful tool for economists and policymakers who need reliable forecasts to guide decisions in uncertain environments.

1 Introduction

Accurate economic forecasting is critical for policymakers and researchers, especially in times of uncertainty when decisions must be based on the most reliable available information. A common challenge in econometric forecasting is the need to adjust model-based predictions to account for external information, such as high-frequency data, expert opinions, or evolving economic conditions. This adjustment process is crucial for producing forecasts that accurately reflect the complexities of real-world situations.

One widely used method to incorporate this external information is entropic tilting, introduced by Robertson, Tallman, and Whiteman (2005). Entropic tilting modifies the probability distribution generated by an econometric model to align it with external constraints, typically by minimizing the Kullback-Leibler divergence between the original and adjusted distributions.The original application embedded theoretical constraints within the density forecasts of a Vector Autoregression (VAR) model. Subsequent studies expanded on this methodology. For instance, Giacomini and Ragusa (2014) applied entropic tilting to adjust a model-based forecasting distribution, incorporating Euler conditions as constraints in their consumption forecasts. Beyond constraints based on economic theory, entropic tilting has been instrumental in integrating additional external information into model-based density forecasts. An example is Krüger, Clark, and Ravazzolo (2017), who incorporates nowcasts into medium-term BVAR-based density forecasts. Similarly, Banbura, Brenna, Paredes, and Ravazzolo (2021) used the Survey of Professional Forecasters (SPF) data to adjust the model-based forecasts. This technique is also particularly valuable for policymakers, as it facilitates the combination of model-based density forecasts with expert judgment, thus enhancing the understanding of risk balances when merging forecasts of different nature.

Although this method has proven conceptually simple and effective in some cases, it frequently encounters significant challenges in practice. Specifically, entropic tilting can produce multi-modal or degenerate distributions, particularly when the external information significantly differs (measured by the Kullback-Leibler divergence criteria) from the model's predictions, or if the latter is non-Gaussian. These issues can lead to unrealistic or unstable forecasts, limiting the practical applicability of the method. An illustration of such impractical distributions is observable in Banbura, Brenna, Paredes, and Ravazzolo (2021), particularly during the integration of external information in the COVID-19 period, a task that proved challenging for the conventional entropic tilting method.

To address these limitations, we propose a new approach, called parametric tilting, which enhances the flexibility and robustness of traditional tilting by employing the skew-T distribution of Azzalini (2013) as a general distribution able to capture key features of real-world economic data, such as asymmetry and heavy tails. By minimizing the Kullback-Leibler divergence between the original model and the skew-T distribution (including new moment conditions), our method ensures that the final forecast distribution accurately reflects the external information while maintaining stability and coherence.

The advantages of parametric tilting are twofold: it provides a more reliable adjustment process that avoids the pitfalls of entropic tilting, and it ensures that the resulting forecast distribution remains well-behaved, even in cases where the original data is non-Gaussian or the external information deviates substantially from the model. Through various examples, we demonstrate that our approach consistently outperforms traditional entropic tilting, offering a robust tool for policymakers and researchers who require reliable forecasts for decision-making.

Our paper is structured as follows. Section 2 reviews the foundations of entropic tilting. Section 3 introduces our parametric tilting framework. Section 4 provides an example that compares our methodology with entropic tilting. Section 5 concludes and provides an outlook on further research.

2 Entropic tilting

Incorporating conditioning information into forecast distributions is a common challenge in economic and financial modeling. Policymakers and researchers frequently face the need to adjust model-based density forecasts based on external data or expert judgment, such as nowcasts, survey data, or theoretical constraints. Entropic tilting has emerged as a popular method to address this challenge by allowing the re-weighting of forecast distributions to align with new information while remaining as close as possible to the original distribution. Introduced by Robertson, Tallman, and Whiteman (2005), entropic tilting applies Kullback-Leibler divergence as a criterion for minimizing the difference between the adjusted and original distributions. This technique ensures that the updated forecast incorporates the additional information without deviating unnecessarily from the model's inherent structure, offering a balance between model coherence and external inputs.

Consider a density forecast for a set of n different variables. For each variable that we would like to tilt, we start with the $n \times k$ matrix that contains the model-based forecasting density - that is, k draws for each variable y_i, each initially associated with some weight $\pi_i, = 1, ..., k$ that reflects the distribution π. In the following exposition we refer to the discrete probabilities of the draws using indexes i while teh variable without index refers to the implied probability distributions. The basic idea of tilting is to find a new set of weights, say π_i^*, such that the reweighted distribution π^* satisfies a user-specified set of moment conditions, while remaining "close" to the original π. To measure the closeness of both probability distributions, entropic tilting uses the Kullback-Leibler (KL) divergence criterion defined as

$$K(\pi_i^*, \pi_i) = \sum_{i=1}^{k} \pi_i^* \log \Lambda_i \quad \text{where} \quad \Lambda_i \equiv \frac{\pi_i^*}{\pi_i}. \tag{1}$$

In measure theory, the factor Λ_i is called the Radon-Nikodym derivative while econometricians might be more familiar with this expression as the likelihood ratio. Hence, the new weights π_i^* are found by minimizing $K(\pi_i^*, \pi_i)$, subject to the following constraints:

$$\pi_i^* \geq 0\,, \tag{2}$$

$$\sum_{i=1}^{k} \pi_i^* = 1\,, \tag{3}$$

$$\sum_{i=1}^{k} \pi_i^* g(y_i) = \bar{g}. \tag{4}$$

The first two constraints are trivial and imply that the new weights should be positive and should sum to 1, to ensure that the final weights also imply a valid density function.

The third constraint imposes one or a set of moment restrictions, i.e. the weighted average of a function of the draws from the forecasting distribution should be equal to a desirable value, different from the corresponding sample value. For example, tilting the original density to a new **mean** $\bar{g} := \bar{\mu}$ implies $g(y_i) = y_i$. To match the **median** d, the corresponding moment function for the restrictions in the optimization problem is given by $g(y_i) = I(y_i < d) - 0.5$ where $I(y_i < d)$ is an indicator function which takes value 1 when $I(y_i < d)$ is satisfied. To match higher moments such as the **variance** or **skewness**, the respective moment functions are defined as $g(y_i) = (y_i - \bar{\mu})^2$ and $g(y_i) = ((y_i - \bar{\mu})/s)^3$ where s is the empirical standard deviation of the original forecasting distribution. Depending on the number of assumptions a researcher is willing to make or the richness of information available, entropic tilting allows to include several moment conditions by including constrains represented as empirical averages of the form of equation (4) using different specifications of $g(y)$.

To find the solution to the constrained optimization problem the Lagrangian can be rewritten in terms of expectations under the original distribution π

$$\min_{\Lambda} \mathcal{L} = E_{\pi}[\Lambda_i \log(\Lambda_i)] - \kappa' E_{\pi}[\Lambda_i(g(y_i) - \bar{g})] - \rho E_{\pi}[\Lambda_i - 1] \tag{5}$$

where

$$E_{\pi}[y] = \sum_{i=1}^{k} \pi_i y_i.$$

Equation (5) gives the First Order Conditions:

$$\frac{\partial \mathcal{L}(\Lambda_i)}{\partial \Lambda_i} = 1 + \log(\Lambda_i) - \gamma'(g(y_i) - \bar{g}) - \mu = 0 \tag{6}$$

which implies that the Radon Nikodym derivative is proportional to

$$\Lambda_i \propto \exp(\kappa' g(y_i)). \tag{7}$$

Therefore, to compute Λ_i it is crucial to first solve for the optimal multipliers κ, which is given as the solution to the set of moment conditions

$$E_{\pi}[\exp(\kappa' g(y_i))(g(y_i) - \bar{g})] = 0. \tag{8}$$

In practice κ is found numerically as:

$$\hat{\kappa} = \text{argmin} \sum_{i=1}^{k} \pi_i \exp(\kappa'[g(y_i) - \bar{g}]) \tag{9}$$

Eventually, combining equation (7) and the constraint of equation (3) yields that draws from the original distribution π are resampled using an importance sampling step with corresponding weights

$$\Lambda_i = \frac{\exp\left(\kappa' g(y_i)\right)}{E_\pi[\exp\left(\kappa' g(y_i)\right)]}. \tag{10}$$

to change the distribution of the draws to the tilted density π^*. Given an empirical set of draws of the model based distribution, this implies that the probabilities of the tilted distribution can be calculated using

$$\pi_i^* = \frac{\pi_i \exp(\kappa' g(y_i))}{\sum_{i=1}^{k} \pi_i \exp(\kappa' g(y_i))}. \tag{11}$$

This reveals the two difficulties that can arise in practice linked to the support of the proposal and target densities π and π^*: first of all, given the definition of the Radon-Nikodym derivative Λ_i, π^* has to assign non-zero probabilities to the same set as π. This means that an event that is impossible under π is also impossible under π^* such that the ratio is well defined for all y_i.[1] If the constraints require a high probability mass in a region where π has low probabilities and therefore no sampled draws, the Radon-Nikodym derivative Λ_i does not exist and the optimization for the lagrange multipliers κ will not be successful or yield unsatisfying results.

Second, the quality of the importance sampling step also depends on the support of the proposal and target densities π and π^*. Hence, if the imposed moments of π^* are too far away from the original density, the resampling step results in a degenerate distribution π^*, even if a solution for Λ_i exists. To show this more formally, consider the Effective Sample Size (ESS)

$$ESS = \frac{1}{E_\pi[\Lambda^2]} \tag{12}$$

a common measure to gauge the quality of an importance sampler. The ESS quantifies how efficiently expectations under π^* can be estimated using samples drawn from π with values between $0 < ESS < 1$. A value close to 1 indicates that the

[1] This condition is formally known as as absolute continuity

quality of the importance sampling approximation of π^* is high while a value close to 0 indicates a deteriorating approximation. Starting from the inverse and rearranging terms yields

$$\frac{1}{ESS} = \frac{\sum_{i=1}^{k} \pi_i [\exp(\kappa' g(y_i))]^2}{\left[\sum_{i=1}^{k} \pi_i (\exp(\kappa' g(y_i))\right]^2} \tag{13}$$

$$= 1 + \frac{\sum_{i=1}^{k} \pi_i [\exp(\kappa' g(y_i))]^2 - \left[\sum_{i=1}^{k} (\pi_i \exp(\kappa' g(y_i))\right]^2}{\left[\sum_{i=1}^{k} (\pi_i \exp(\kappa' g(y_i))\right]^2} \tag{14}$$

$$= 1 + \sum_{i=1}^{k} \pi_i \left[\frac{\exp(\kappa' g(y_i)) - \left[\sum_{i=1}^{k} (\pi_i \exp(\kappa' g(y_i))\right]}{\left[\sum_{i=1}^{k} (\pi_i \exp(\kappa' g(y_i))\right]^2} \right]^2 \tag{15}$$

$$= 1 + \sum_{i=1}^{k} \pi_i \left[\Lambda_i - 1\right]^2 \tag{16}$$

Therefore the Inefficiency Ratio of the resampling step in entropic tilting is given by

$$ESS = \frac{1}{1 + \sum_{i=1}^{k} \pi_i \left[\Lambda_i - 1\right]^2} \tag{17}$$

From expression (17), it follows that when the weights of the two distributions, π and π^*, are similar, the terms $[\Lambda_i - 1]^2$ remain small, and the inefficiency ratio approaches one. In contrast, substantial differences between the weights π_i and π_i^* increase the dispersion of the importance weights, thereby enlarging the sum over $[\Lambda_i - 1]^2$ and reducing the effective sample size (ESS). Consequently, when the tilted distribution π^* implied by the moment conditions in (8) departs substantially from the baseline distribution π of the model, the importance sampling approximation deteriorates. This degradation reflects the growing dissimilarity between the two distributions, which manifests in a smaller ESS and, hence, lower efficiency of the reweighted sample. While on the one hand, failure of entropic tilting could simply be interpreted as an indication of unrealistic or misspecified scenarios. On the other hand, it is of great value for policy makers and researchers to retain the ability to also incorporate new and atypical information in extreme circumstances e.g. during crises times. Naturally, we consider the method proposed in this paper as a suggestion to target the later case. In the next subsection, we illustrate some of these caveats.

2.1 Entropic tilting: A caveat

A well-known result in the literature is that, if the original distribution is Gaussian, the problem of tilting it towards a different value of the mode (mean/median) and/or the variance has a closed-form solution, such that the final distribution is also Gaussian with an analytical solution for the mode and the variance.

Figure 1 further illustrates this an example of a Gaussian distribution. The original distribution in blue has a mean (mode/median) of 2.4, and we would like to tilt it to a value of 1.7. Entropic tilting works as expected and changes the location of the original distribution to the tilted distribution in red.

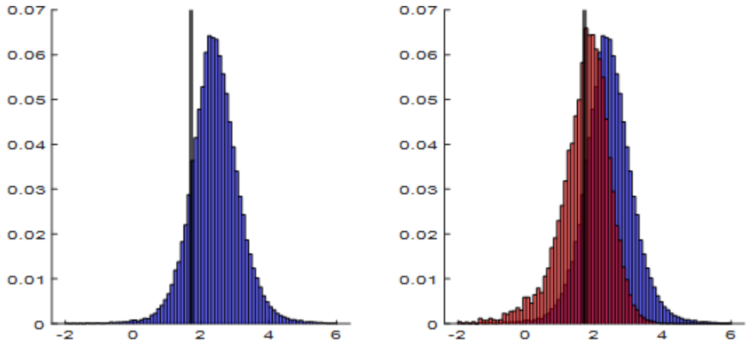

Figure 1: Tilting a Gaussian distribution with entropic tilting

But, when the original distribution is non-Gaussian, closed-form solutions may not be available. The non-parametric approach could still be able to find a solution to the mathematical problem described above, but it is not ensured that the final distribution is for example uni-modal, [2] or that the problem finds a well-balanced set of new weights, avoiding degeneracy. This problem is exacerbated if the target mean is very far from the support of the original empirical distribution. In that case, the latter does not have enough support for the new target values and the re-weighting faces numerical issues resulting in a degenerate distribution ϕ^*.

In Figure 2 we can see an example of a distribution with non-Gaussian features, as it incorporates fat tails and is not fully symmetric. In this case, entropic tilting produces a bimodal distribution when imposing a mean of 1.7 compared to the original of 2.4, as the final mean is far away from the original one. This solution, given the small deviation from Gaussianity in the original distribution, may be unreasonable for many problems in economics.

[2] A non uni-modal distribution would not be by itself a problem if well-grounded in theory, see Shiller (2000) and his work on market bubbles for example, but here the non uni-modality would be just of a technical nature.

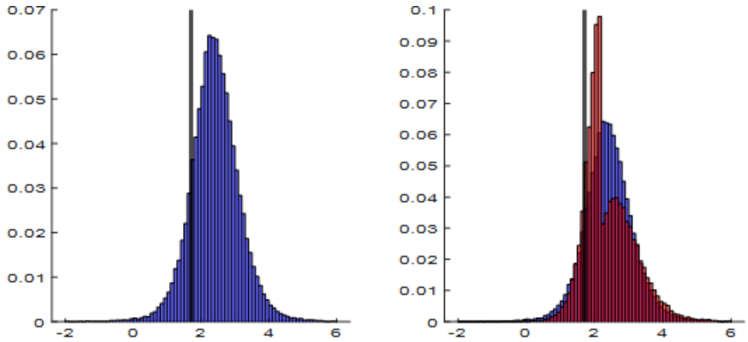

Figure 2: Tilting a non-Gaussian distribution with entropic tilting

3 Towards a robust approach. Parametric tilting

In addressing the challenges inherent in the non-parametric approach previously outlined, our proposed strategy involves restructuring the optimization problem to yield a final density with a parametric form. The central issue with the non-parametric method is its somewhat opaque nature; it functions as a 'black box' that merely re-weights the original density in an attempt to achieve the imposed moment conditions. This process does not account for the final density's shape. As demonstrated earlier, it is entirely possible for this method to meet a specified mean value while inadvertently producing a possibly unreasonable distribution and instability during resampling.

To circumvent this issue, we slightly reframe the optimization problem. Diverging from the previous approach based on re-weighing the empirical draws with new weights directly determined from the optimization in equation (5), we consider the KL divergence in terms of parametric densities to impose the constraints:

$$KL(P||Q) = \sum_{i=1}^{k} P(y_i) \log \left(\frac{P(y_i)}{Q(y_i)} \right) \tag{18}$$

here y_i corresponds to the value of each draw in the model-based forecasting density, $P(y_i)$ is the pdf of the original density, and $Q(y_i)$ is the parametrically tilted parametric density that we will use to impose the constraints. Opting for a parametric density, as opposed to directly adjusting the weights, affords researchers greater control over the final density's shape and facilitates the assurance of unimodality, if so desired. We will delve into the specifics of our chosen target density and discuss approaches for scenarios where the original density is unknown.

3.1 A skew-t distribution

To provide a flexible and theoretically well-behaved density, we propose using a skew-t distribution Azzalini (2013) as the target distribution in the KL divergence.[3] The skew-t distribution is a flexible, parametric density that allows us to have fat tails as well as asymmetries that can be controlled by the four parameters defining the distribution. Therefore, it is possible to find solutions to the optimization problem that reflect the asymmetric or leptokurtic behavior of the target densities and can thereby accommodate a variety of assumptions.Additionally, with the recent introduction of the Macro at Risk literature by Adrian, Boyarchenko, and Giannone (2019), the skew-t distribution has gained wide popularity among researchers to capture non-normal features of macroeconomic variables (see for example López-Salido and Loria (2024) or De Santis and Van der Veken (2020)).

The skew-t distribution can be parameterised as follows. First, we write that a variable Y has a skew-t (ST) distribution,

$$Y \sim ST(\xi, \omega, \alpha, \nu) \tag{19}$$

where ξ is a location parameter, ω is the scale, α is the slant parameter that determines the skewness of the distribution, and ν is the degrees of freedom.

Second, the pdf of Y can be written in terms of a standardized skew-t distribution. Define $y = \omega^{-1}(y - \xi)$. Then the density function at y is $\omega^{-1}t(y;\alpha,\nu)$, where

$$t(y;\alpha,\nu) = 2t(y;\nu)T(\alpha y;\nu) \tag{20}$$

where $t(y;\nu)$ is the pdf of a standard t-distribution with degrees of freedom ν and $T(.;\nu)$ denotes the cumulative distribution function of a t-distribution. Note that if the degrees of freedom are large enough $(\nu \to \infty)$ and $\alpha = 0$, the distribution collapses to a Gaussian distribution.

For our purposes, it is important to note that the ST distribution has a closed form solution for different moments such as the mean, the variance, the skewness, and for the mode (see Azzalini (2013)). These four moments are non-linear functions of the four parameters that determine the shape of the distribution,

[3] In principle, many other distributions could be used, for example, a skew normal distribution.

$(\xi, \omega, \alpha, \nu)$. Define first,

$$b_\nu = \frac{\sqrt{\nu\Gamma(0.5(\nu-1))}}{\sqrt{\nu\Gamma(0.5\nu)}} \tag{21}$$

$$\delta = \frac{\alpha}{\sqrt{1+\alpha^2}} \tag{22}$$

$$\sigma_z = \sqrt{\frac{\nu}{\nu-2} - (b_\nu\delta)^2} \tag{23}$$

where $\Gamma(...)$ represents the Gamma function. Then, we can define the moments of the distribution as,

$$\mu = E(y)\zeta + \omega b_\nu \delta, \quad \nu > 1 \tag{24}$$

$$\sigma^2 = \text{var}(y) = \omega^2\sigma_z^2, \quad \nu > 2 \tag{25}$$

$$\gamma = \text{skew}(y) = \frac{b_z\delta}{\sigma_z^{3/2}}\left[\frac{\nu((3-\delta^2)}{(\nu-3} - \frac{3\nu}{\nu-2} + 2(b_\nu\delta)^2\right], \quad \nu > 3 \tag{26}$$

Furthermore, the mode of the skew-t distribution is unique and given by

$$m = \text{mode}(y) = \epsilon + y_0\omega \tag{27}$$

with

$$y_0 = \text{argmin}\left(y\sqrt{\nu+1}T(\omega(y);\nu+1) - \frac{t(\omega(y);\nu+1)\nu\alpha}{\sqrt{\nu+1}}\right) \tag{28}$$

and w(y) defined as

$$\omega(y) = \alpha y\sqrt{\frac{\nu+1}{\nu+y^2}} \tag{29}$$

3.2 Tilting towards a skew-t distribution

With all the elements in hand, now we can define our parametric tilting problem as follows,

$$\min_{\xi,\omega,\alpha,\nu} \sum_{i=1}^{k} P(y_i)\log\left(\frac{P(y_i)}{ST(y_i)}\right) \tag{30}$$

Subject to one or more restrictions on the moments of the distribution,

$$\mu = \bar{\mu} \tag{31}$$

$$\sigma = \bar{\sigma} \tag{32}$$

$$\gamma = \bar{\gamma} \tag{33}$$

$$m = \bar{m} \tag{34}$$

That is, rather than stating the problem in terms of finding the final weights, we minimize the KL divergence over the parameters that determine the shape of the ST distribution while satisfying the moment constraints we want to impose. Several comments are in order. First, we define the KL divergence in terms of discrete densities. This is because the original distribution is generally a set of empirical draws from a model or a combination of models. Thus, we need to evaluate the KL-divergence for each of the draws. Second, we will not always know the shape of the original distribution, $P(y_i)$, and thus, it would not be easy to evaluate it.

If the original distribution $P(y_i)$ is known, we can just discretize the support of the two density functions $P(Y)$ and $Q(Y)$ around the values $y_i \in Y$ from the original model. With those values in hand, we can easily evaluate the KL divergence in the minimization problem to impose the constraint.

However, if the original distribution is unknown, we can find a discrete approximation of $P(y_i)$ using a histogram with appropriate scaling and bandwidth selection to accurately cover the range of the original draws.[4] Evaluating the pdf of the skew-t distribution over the grid obtained from the bin edges of this histogram then yields the corresponding values for the discrete approximation of the target $ST(y_i)$. Note that it is important to align the widths of the histogram approximation for both densities to ensure consistency in the optimization routine as well as in the construction of the resampling weights in the final steps.

We eventually find the solution to the problem numerically using standard solvers.[5] Note however that we need to keep the degrees of freedom of the distribution, ν, as a discrete value. Therefore, we proceed as follows. We construct first

[4]To obtain an approximation of a probability density that integrates to 1 from a standard histogram that counts the frequency of observations, each bin needs to be re-scaled by the bin value $b_i = \frac{c_i}{kw_i}$, where k is the number of observations, c_i is the number of elements in and w_i is the width of the i^{th} bin.

[5]Due to the closed form solution of the ST distribution, we can easily solve the optimization problem using standard solvers available in common programming languages such as MATLAB, Python or R.

a grid for the degrees of freedom. For each value of the grid, we obtain optimal values for the location $\hat{\xi}$, scale $\hat{\omega}$ and slant parameters $\hat{\alpha}$ that minimize the KL divergence, subject to the moment constraints that we would like to impose. Then, we select the minimum value out of the grid and the associated parameters of the optimized skew-t distribution.

In the last step of our problem, we need to finally re-sample the original density forecast, according to the new weights based on the importance ratio of the model based density and the fitted ST distribution.In case the original distribution needs to be represented as a histogram, the importance weights are calculated for each bin resulting in so called histogram reweighting.[6]

4 Application: from theory to practice

4.1 Theory: Exploring the Limits

As an initial example, we consider the same problem as in the previous section. First, we consider a parametric tilting in which we impose only the mode of the final distribution. Compared to the non-parametric tilting approach, we should emphasize that in this case, given that the skew-t distribution has a closed form solution for the mode, we do not need to make any other assumption about the mean or the median of the distribution.

First, Figure 4 shows both the original and the tilted distribution that come out from our parametric minimization problem. Given that the final mode is far away from the original model-based one, we see that the problem endogenously yields a skewed distribution which would signal upside risks. Moreover, the final distribution is now well behaved and unimodal.

We can also consider imposing additional moment restrictions. For example, on top of the mode, policy makers might want to impose that the mean of the target distribution is the same as the model-based one (Figure 3). In this sense, policy makers would have in mind that the most likely scenario is still their judgemental forecast, but that the expectation is the model-based one, showing asymmetric risks.

[6] While histogram reweighting has a long tradition in physics, similar techniques have also recently been proposed in economic applications such as portfolio allocation or analyzing demographic parity by Chakraborty, Bhattacharya, and Pati (2024).

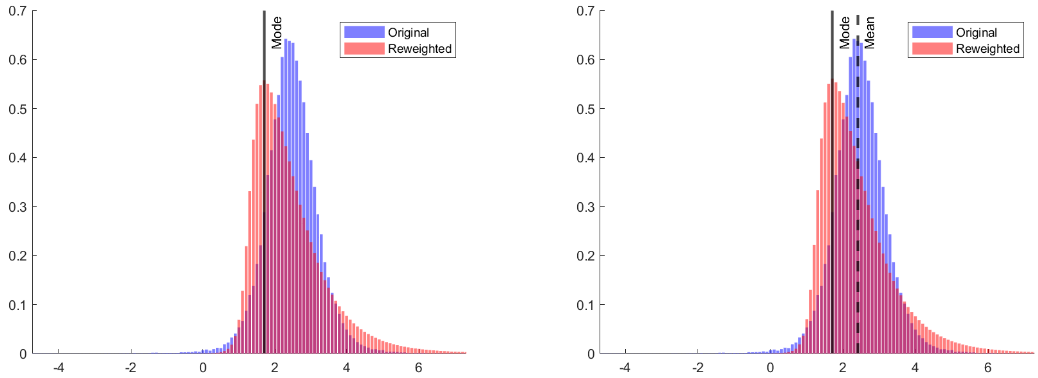

Figure 3: Tilting the mode of a non-Gaussian distribution with parametric tilting

Another possibility is that policymakers might want to keep the original model-based uncertainty (Figure 4). In this case, it is therefore possible to restrict both the mode and the variance of the target distribution.

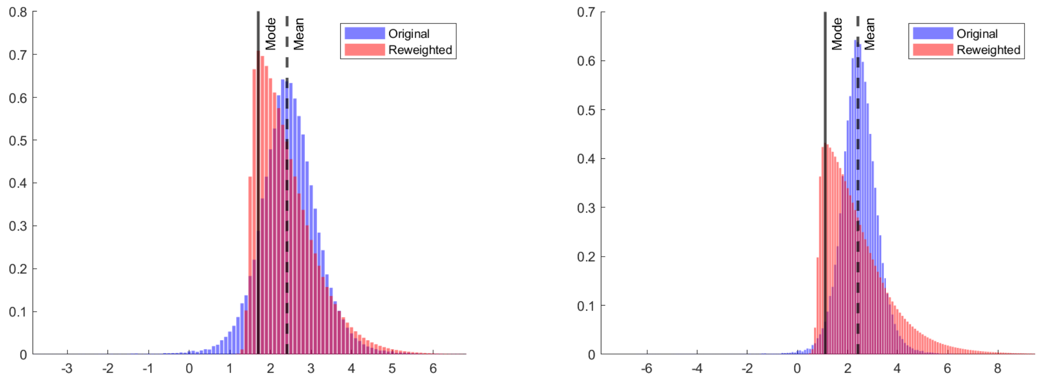

Figure 4: Pushing the limits: Tilting mode, mean and variance of a non-Gaussian distribution with parametric tilting

In this example, tilting also towards the mean does not change the shape of the final distribution in a meaningful way. However, keeping the original variance reduces the weight of extreme values in the right tails of the distribution. At the end, this would be a choice of the forecaster, which could be also a result of a previous evaluation of the forecast performance under different alternatives.

Finally, we consider a more extreme case, to show that our parametric tilting approach can handle somewhat more difficult problems than the non-parametric approach. In the original problem, we tilted the mode of the distribution from 2.4 to 1.7. We now reduce the target mode to 1.1, that is, further away to the left from the original one. Figure 4 shows that even in this case, our robust approach can find the solution to the problem, although the final distribution is significantly more skewed than in the previous case.

4.2 Practice: Tilting to the SPF during Covid

We now illustrate how our proposed method would have performed in a real-time setting. In Figure 5, we replicate Figure 5 from the updated version of Banbura, Brenna, Paredes, and Ravazzolo (2021), where the authors assess how the traditional tilting method would have operated during the onset of the COVID-19 pandemic.[7] In our figure, we display the one-year-ahead real GDP growth density forecast obtained from three approaches that are able to incorporate external information: the original optimal pool combination, the entropic tilting method and the parametric tilting method proposed in this paper.

We focus on the forecast for the fourth quarter of 2020, produced during the first quarter of 2020. At that time, participants in the Survey of Professional Forecasters (SPF) had already begun to observe the economic impact of COVID-19, whereas the models included in the optimal pool had not yet incorporated any data reflecting these developments. Starting with the forecast for 2020Q4, the uncertainty assessment of SPF respondents adjusts to the new circumstances, and the realised outcomes (around minus 5 percent GDP growth) lie well within the support of the SPF predictive distribution at the time.[8] By contrast, the forecast from the Bayesian VARs combination (optimal pool depicted by the blue line in all panels) for 2020Q4 fails to reflect the unfolding crisis, as the underlying data still contained limited information about the pandemic.

SPF participants reported a mean forecast of 2.8 percent negative growth, with an associated variance of around 15.3. As discussed in Banbura, Brenna, Paredes, and Ravazzolo (2021), there are at least three ways to incorporate this information using entropic tilting. The first approach is to include the SPF as an additional model within the combination (black line in top panel). The second approach is to tilt the predictive distribution towards the SPF mean only (black line bottom left panel). The third approach is to tilt the distribution towards both the SPF mean and variance (black line in bottom right panel).

An important limitation of the tilting procedure becomes evident when inspecting the tilted predictive distributions for 2020Q4 in the two bottom panels. When the target moments used for tilting are substantially distant from those im-

[7] We are very thankful to the authors for providing to us the replication codes.

[8] See top-left panel in Figure 5 of Banbura, Brenna, Paredes, and Ravazzolo (2021).

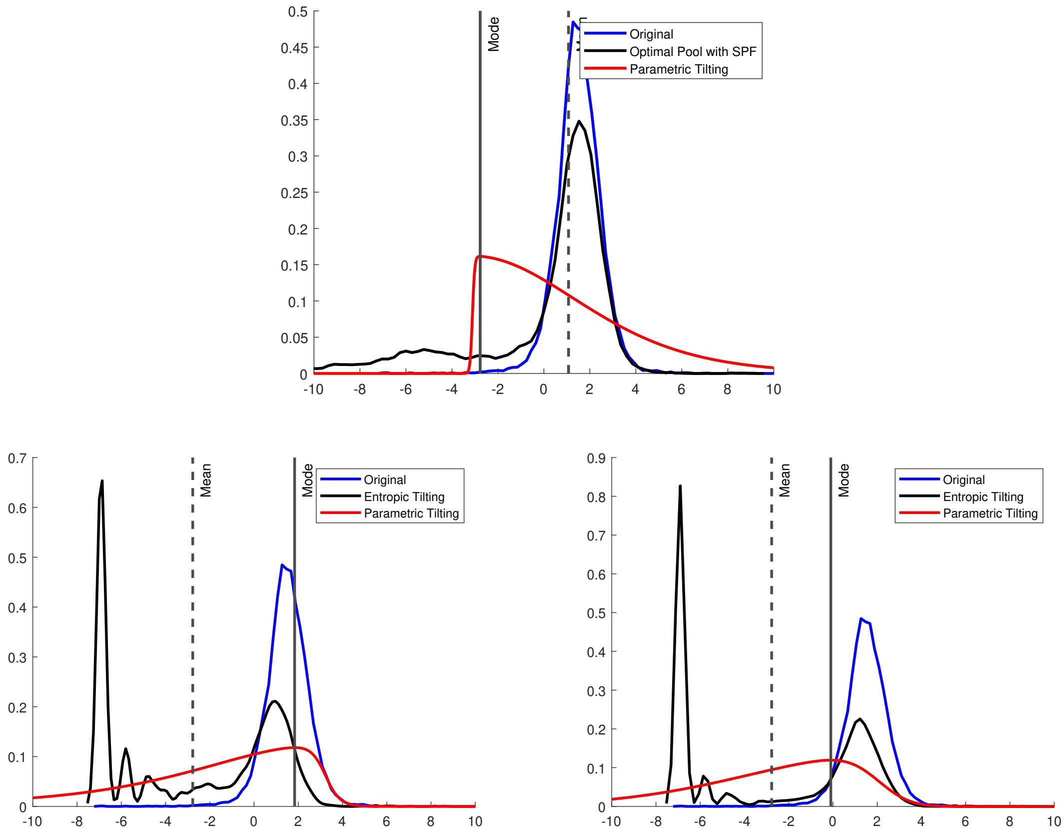

Figure 5: One-year-ahead real GDP growth density forecasts. Parametric vs entropic tilting with restrictions on the mode, mean and variance.

plied by the original distribution, resulting in a lack of overlapping support, the tilted distribution can degenerate or exhibit undesirable properties, such as bimodality.

In contrast, our proposed method (red lines in Figure 5), across its different specifications, succeeds in combining the external information while preserving a smooth parametric form. In the top panel, we impose the SPF mean as the mode of the new distribution. This adjustment implies that the most likely scenario corresponds to a negative GDP growth rate, accompanied by positive skewness, as the density of the original optimal pool still exhibited positive growth. Conversely, when the SPF mean is imposed as the mean of the parametric distribution, the mode remains closer to that of the original distribution but the left tail becomes considerably fatter, indicating heightened downside risks. Finally, imposing both the mean and variance of the SPF shifts the mode of the resulting distribution further toward zero, while preserving the pronounced left fat tail.

Overall, these three specifications illustrate how the parametric tilting method behaves depending on how the practitioner incorporates the external information.

5 Conclusion

In this paper, we introduced a new methodology called parametric tilting to address the limitations found in traditional entropic tilting methods, particularly their struggles with producing reasonable distributions in practical applications. Parametric tilting offers a more robust and flexible approach to integrating external information into model-based density forecasts, providing solutions that are more reliable and numerically stable.

Parametric tilting not only ensures that any distribution can be centered around a specified forecast baseline but also guarantees unimodality, which is a common requirement in economic and policy applications. It retains the capacity to faithfully reflect the balance of risks inherent in the original distribution while offering the flexibility to incorporate various moments (mean, variance, skewness) of the original distribution into the final one. This flexibility makes it highly adaptable to the diverse requirements faced by policymakers and researchers in scenarios such as risk assessments, nowcasting, and scenario analysis.

As demonstrated in our application, the parametric tilting method successfully generates well-behaved distributions even when external constraints are imposed far from the original model distribution. This includes extreme cases where the original distribution may exhibit non-Gaussian features such as fat tails and asymmetry. By leveraging the skew-t distribution of Azzalini (2013), our approach can capture complex behaviors often observed in macroeconomic time series, ensuring that the final forecast remains coherent with both the external information and the model's internal structure.

In contrast to non-parametric methods, which often function as "black boxes" that re-weigh samples without guaranteeing the shape or behavior of the final distribution, parametric tilting provides transparency and control over the final outcome. By framing the optimization in terms of Kullback-Leibler divergence and selecting a parametric target distribution, researchers can better manage the resulting forecast's properties, such as skewness and tail behavior, which are critical for realistic scenario analysis and policy decision-making.

Given its versatility and robustness, parametric tilting should become a valuable tool in the forecaster's toolkit. It opens avenues for future research to explore further enhancements, such as incorporating more sophisticated external information sources or extending the methodology to multi-dimensional frameworks. By doing so, it can further contribute to the development of more accurate and reliable forecasting models in both theoretical and practical settings.

References

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

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

Banbura, Marta, Federica Brenna, Joan Paredes, and Francesco Ravazzolo (2021). "Combining Bayesian VARs with survey density forecasts: does it pay off?" *ECB Working Paper* 2543.

Chakraborty, Abhisek, Anirban Bhattacharya, and Debdeep Pati (2024). *Constrained Reweighting of Distributions: an Optimal Transport Approach*. arXiv: 2310.12447 [stat.ML]. URL: https://arxiv.org/abs/2310.12447.

De Santis, Roberto A and Wouter Van der Veken (2020). "Forecasting macroeconomic risk in real time: Great and Covid-19 Recessions". *ECB Working Paper*(2436).

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

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.

López-Salido, David and Francesca Loria (2024). "Inflation at risk". *Journal of Monetary Economics* 145, p. 103570.

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.

Shiller, Robert J. (2000). *Irrational Exuberance*. 1st. Princeton University Press: Princeton, NJ.