BANCO DE PORTUGAL

Economic Research Department

UNCERTAINTY AND RISK ANALYSIS OF MACROECONOMIC FORECASTS: FAN CHARTS REVISITED

Álvaro Novo
Maximiano Pinheiro

WP 19-03 December 2003

Please address correspondence to Maximiano Pinheiro, Economic Research Department, Banco de Portugal, Av. Almirante Reis nº 71, 1150-165 Lisboa, Portugal, Tel.#351-213128394; Fax#351-213107803; email: mpinheiro@bportugal.pt, or Álvaro António Novo, Economic Research Department, Banco de Portugal, Av. Almirante Reis nº 71, 1150-165 Lisboa, Portugal, Tel.#351-213128107; Fax#351-213107804; emai: anovo@bportugal.pt

Uncertainty and Risk Analysis of Macroeconomic Forecasts: Fan Charts Revisited*

Álvaro A. Novo† Maximiano Pinheiro‡

Version: December 5, 2003

Abstract

Since 1996 the Bank of England (BoE) has been publishing estimates of probability distributions of the future outcomes of inflation and output growth. These density forecasts, known as "fan charts", became very popular with other central banks (e.g. Riksbank) as a tool to quantify uncertainties and risks of conditional point forecasts. The BoE's procedure is mainly a methodology to determine the distribution of a linear combination of independent random variables. In this article, we propose an alternative methodology that addresses two issues with the BoE procedure that may affect the estimation of the densities. The first issue relates to a statistical shortcut taken by the BoE that implicitly considers that the mode of the linear combination of random variables is the (same) linear combination of the modes of those variables. The second issue deals with the assumption of independence, which may be restrictive. An illustration of the new methodology is presented and its results compared with the BoE approach.

Keywords: Density forecasting; Fan charts; Balance of risks; Uncertainty.
JEL classification: C50, C53

*We would like to thank Ana Cristina Leal, Carlos Coimbra, Carlos Robalo Marques, Claes Berg, Fátima Teodoro, Francisco Dias, José Ferreira Machado, Meghan Quinn, Paulo Esteves, Pedro Duarte Neves, Pedro Próspero, Ricardo Felix and Susana Botas, as well as the members of the Monetary Policy Committee and the Working Group on Forecasting of the Eurosystem, for very helpful comments. The usual disclaimers apply.

†Banco de Portugal, Research Department and ISEGI-Universidade Nova, Av. Almirante Reis, 71, 1150-012 Lisbon, Portugal, anovo@bportugal.pt.

‡Banco de Portugal, Research Department and ISEG-Technical Univ. Lisbon. Av. Almirante Reis, 71, 1150-012 Lisbon, Portugal, mpinheiro@bportugal.pt.

1 Introduction

Although forecasts are inherently uncertain, quite often the associated uncertainty (dispersion of the distribution) and risks (degree of asymmetry) are not duly acknowledged and quantified by forecasters. Many institutions choose to publish only point forecasts, while others attempt to explicitly consider uncertainty by publishing forecasting ranges, computed either on an ad hoc basis or by taking into account past forecasting errors. Instead of a quantified and integrated assessment of uncertainty and risks of the forecasts, most institutions only present a qualitative assessment. Since typically institutions do not rely on a single econometric model to produce their forecasts, and furthermore these include (to a larger or a lesser extent) subjective judgments about future economic developments, the quantification of uncertainty and risks by institutional forecasters is not straightforward.

An important contribution towards an explicit quantification of uncertainties and risks, while allowing for subjective judgmental elements, originated from the Bank of England in early 1996 and it was followed soon after by the adoption of similar approaches by other central banks, e.g., the Riksbank (the Swedish central bank). The Bank of England and the Riksbank have since published in their quarterly Inflation Reports estimates of the probability distributions of the future outcomes of inflation and output growth. These "density forecasts" are represented graphically as a set of prediction intervals with different probability coverages. The resulting chart has become known as a "fan chart".

Although the statistical methods used by the Bank of England and the Riksbank to produce their fan charts have some differences, the differences are relatively minor when compared with the similarities. Descriptions of these methods are provided in Britton, Fisher and Whitley (1998) for the Bank of England and in Blix and Sellin (1998) for the Riksbank.[1] Hereafter,

[1] In spite of very similar statistical procedures, the institutional frameworks in which they operate are somewhat different. For a description of these institutional frameworks,

for the sake of expositional simplicity, we will refer to the statistical method as the "Bank of England (BoE) approach".

The BoE approach can be described as a method of aggregation of probability distributions of the "input variables", which include (i) the errors of conditioning variables (such as external demand, oil price, exchange and interest rates, fiscal developments) and (ii) the "pure" errors in the "variables to forecast", i.e., the forecasting errors that remain after explicitly correcting for the errors in the conditioning variables. To this purpose, the BoE approach assumes that the forecasting errors can be expressed as a linear combination of these input variables. In addition, it assumes that the marginal probability distributions of the input variables are two-piece normal (*tpn*). The *tpn* distribution, discussed in detail in John (1982),[2] is defined by only three parameters – mode μ, "left- standard deviation", σ_1, and "right-standard deviation", σ_2. The *tpn* is formed by taking two appropriately scaled halves of normal distributions with parameters (μ, σ_1^2) and (μ, σ_2^2), collapsing to the normal if $\sigma_1 = \sigma_2$. In order to apply the BoE aggregation procedure, sample and/or judgmental information has to be provided to determine the three parameters of each input variable distribution.

In this article, we start by raising two statistical issues about the procedure underlying the BoE approach. The first issue relates to a statistical shortcut considered by the BoE approach when aggregating the distributions of the input variables. The BoE implicitly considers that the mode of the linear combination of input variables can be obtained as the (same) linear combination of the modes of these variables. We will illustrate that this is a potentially poor approximation to the mode of the forecasting errors when the distributions of some inputs are skewed.

The second issue relates to the fact that the BoE procedure assumes the independence of all the input variables. In this respect, it is worth mention-

which is beyond our concern, see e.g. Bean (2001) and Blix and Sellin (1999), respectively for the UK and Sweden.

[2] See also Johnson, Kotz and Balakrishnan (1994).

ing that the construction of the BoE fan charts is derived from the historical forecast errors of the variables concerned in a sample of previous forecasting exercises. The base levels of sample variances are typically adjusted to reflect judgments on the level of uncertainty over the forecasting horizon. As to the correlations between the forecast errors, these are usually ignored, although in some cases judgmental adjustments to the results of the statistical procedure are considered to proxy the effects of correlation between errors.[3] The Riksbank variant described by Blix and Sellin (1998) explicitly considers a non-diagonal covariance matrix, allowing for linear correlations between the conditioning variables. In their analysis, however, these correlations do not affect the skewness of the distributions of the forecasted variables. In a more recent paper, Blix and Sellin (2000) consider a bivariate case which takes into account the effect of correlation on skewness, but this interesting approach is not easily generalized to higher dimensions. In our view, the assumption of independence of all the input variables, in spite of being very convenient for simplifying the aggregation procedure, is restrictive. According to our experience, for a typical conditioning variable, such as the oil price, one can expect serial correlations of the errors at different horizons. In addition, one can expect contemporanous correlations between the errors of different conditioning variables. Hence, in our opinion relaxing the independence assumption, at least partially, is a step forward towards a more realistic uncertainty and risk assessment of the forecasted variable.

Our methodology attempts at relaxing the above simplifying hypothesis of the BoE approach, while keeping the spirit of the underlying procedure.[4]

---

[3]The BoE forecasting team argues that no significant sample correlations are found between the errors of the conditioning variables in their past forecasts series and that stochastic simulations also tend to show little impact on the results if these samples correlations were considered.

[4]In a recent paper, Cogley, Morozov and Sargent (2003) construct BoE's fan charts based on the forecast densities generated from BVAR with drifting coefficients and stochastic volatilities. These densities are modified in a second stage by incorporating judgmental information, using the relative entropy method of Robertson, Tallman and Whiteman (2002). This method does not suffer from the aforementioned BoE approach short-

However, the task is not straightforward. For instance, one natural alternative to assuming the independence of all input variables is to assume a non-diagonal linear correlation matrix between these variables. But it is important to notice that assuming a non-diagonal correlation matrix is not a fully specified alternative to the assumption of independence. The independence also implies imposing restrictions on all the cross-moments of order higher than two. Therefore, by simply specifying the marginal distributions and the correlation matrix, in general, we have multiple joint distributions of the input variables compatible with those assumptions.

This identification problem can only be solved by introducing further *a priori* restrictions. It does not, however, seem reasonable to require the forecaster to specify cross-moments of an order higher than two, on which typically he/she has little knowledge. Instead, we propose to look for solutions which correspond to a distribution of the forecast error belonging to the same family as the one assumed for the marginal distributions of the input variables. The problem with this strategy is that we lose the guarantee of existence of a convenient solution in the asymmetric case. When such solution does not exist, we suggest alternative routes. First, we broaden the search by defining a larger family of distributions. If still no such solution exists, we propose either to revise the assumptions of the exercise or to use an "approximation of last resort", which consists in picking the distribution of the predefined family that is closer (in some multivariate sense) to the set of multiple solutions.

The "*tpn* environment" suggested by the BoE, in spite of its interpretational convenience, does not facilitate the algebra required to deal with the problem. For this reason, we opted to postulate that the marginal distributions of the input variables belong to another three parameters distribution family, which we name 'skewed generalized normal' (*sgn*). As with the *tpn*,

comings. On the other hand, it only introduces judgments in the resulting forecasting error distribution. Hence, it departs from the original BoE spirit, which introduces risks in the input variables distributions.

the *sgn* collapses to the normal distribution in the particular case of absence of skewness. The *sgn* may viewed at the core of a larger family of distributions, including the "convoluted *sgn*" distributions, directly derived from the *sgn*.

The article proceeds as follows. Section 2 reviews the BoE approach, pointing out the shortcomings that motivate the development of an alternative methodology. This methodology, based on the skewed generalized normal distribution, is developed in Section 3. Section 4 applies the new method to an illustrative euro-area inflation forecasting exercise. Finally, concluding remarks are presented.

2 The Bank of England Approach

2.1 Linear combinations

The methodology suggested by the BoE and herein further developed deals with finding the probability distribution of a linear combination of the input variables. We start by motivating the linear approximation in an uniequational setting. A multiequational approach is also suggested. The framework that follows reflects our own perspective on the issue and not necessarily the BoE's.

In the uniequational setting, let us denote the variable to forecast by y_{t+H}, where t and H stand for the last observed period and the forecasting horizon, respectively. Let us also denote by x_{t+h} $(h = 1, \ldots, H)$ the vectors $(K \times 1)$ containing the paths of conditioning variables from period $t+1$ to period $t+H$. The central or baseline scenario for y_{t+H}, conditional on a given path for the conditioning variables $\{x^0_{t+1}, x^0_{t+2}, \ldots, x^0_{t+H}\}$ will be represented by $\hat{y}(x^0_{t+H}, \ldots, x^0_{t+1})$. Typically, this central forecast is not the direct outcome of a single econometric model, or even a combination of forecasts produced by a suite of econometric models.[5] Nonetheless, these econometric models

[5]Bean (2001, p. 438) makes this point very clear in the UK case, and the thrust

can provide the forecasters with a local linear approximation of the central forecast response to small changes in the paths of the conditioning variables of the type:

$$\widetilde{y}(x^1_{t+H}, \ldots \quad , x^1_{t+1}) = \hat{y}(x^0_{t+H}, \ldots, x^0_{t+1}) + \gamma'_0(x^1_{t+H} - x^0_{t+H}) + \\ + \gamma'_1(x^1_{t+H-1} - x^0_{t+H-1}) + \cdots + \gamma'_{H-1}(x^1_{t+1} - x^0_{t+1}) \qquad (1)$$

where $\{x^1_{t+1}, x^1_{t+2}, \ldots, x^1_{t+H}\}$ is an alternative path for the conditioning variables and γ_τ ($\tau = 0, \ldots, H-1$) are $K \times 1$ vectors of interim multipliers. The latter give the estimated responses of y_{t+H} to changes in the various elements of $x_{t+H-\tau}$. After observing both y_{t+H} and x_{t+h} ($h = 1, \ldots, H$), we can define the "pure" forecast error ϵ_{t+H} as:

$$\epsilon_{t+H} = y_{t+H} - \widetilde{y}(x_{t+H}, x_{t+H-1}, \ldots, x_{t+1}) \qquad (2)$$

Note that ϵ_{t+H} is not the conventional forecasting error resulting from the central conditional forecast $\hat{y}(x^0_{t+H}, \ldots, x^0_{t+1})$. Instead, ϵ_{t+H} is computed using the adjusted forecast (given by (1)) that takes into consideration the observed deviations in the conditioning variables from the assumed baseline paths $\{x^0_{t+1}, x^0_{t+2}, \ldots, x^0_{t+H}\}$. The error ϵ_{t+H} reflects the shocks that affected the economy during the forecasting horizon, the estimation errors of the multipliers γ, the errors resulting from taking a linear approximation of a (possibly) non-linear model, as well as any other misspecifications of the model (e.g. omitted variables or structural breaks). From (1) and (2), we can write the (conventional) overall forecast error as:

$$e_{t+H} \quad = \quad y_{t+H} - \hat{y}(x^0_{t+H}, ..., x^0_{t+1}) = \gamma'_0(x_{t+H} - x^0_{t+H}) +$$

of his words applies to most institutions publishing forecasts: "The suite of econometric models is an essential tool, but the quarterly [central] projections are not simply the result of running either the MM [the main Bank of England macroeconometric model], or the suite, mechanically. All economic models are highly imperfect reflections of the complex reality that is the UK economy and at best they represent an aid to thinking about the forces affecting economic activity and inflation. The MPC [Monetary Policy Committee] is acutely aware of these limitations. Moreover, a considerable amount of judgment is required to generate the projections. In making those judgments, the MPC draws on a range of additional sources of information about economic developments."

$$+\gamma_1'(x_{t+H-1} - x^0_{t+H-1}) + \cdots + \gamma_{H-1}'(x_{t+1} - x^0_{t+1}) + \epsilon_{t+H} \quad (3)$$

The extension to a multiequational approach is straightforward. Let y_{t+H} be a $P \times 1$ vector of endogenous variables instead of a scalar variable. We can rewrite (1) as:

$$\begin{aligned}\widetilde{y}_{t+H}(x^1_{t+H}, \ldots, x^1_{t+1}) = \widehat{y}_{t+H}(x^0_{t+H}, \ldots, x^0_{t+1}) + \Gamma_0(x^1_{t+H} - x^0_{t+H}) + \\ +\Gamma_1(x^1_{t+H-1} - x^0_{t+H-1}) + \cdots + \Gamma_{H-1}(x^1_{t+1} - x^0_{t+1}) \quad (4)\end{aligned}$$

where Γ_τ are $P \times K$ matrices of coefficients of the final form of the linear model taken as a local approximation to the "forecast generating process". The expression for the pure forecasting error in the multiequational case is identical to (2), with ϵ_{t+H}, y_{t+H} and $\widetilde{y}(x_{t+H}, x_{t+H-1}, \ldots, x_{t+1})$ being $P \times 1$ vectors instead of scalars. The multiequational version of the $P \times 1$ vector of conventional forecast errors e_{t+H} can be written as:

$$\begin{aligned} e_{t+H} &= \Gamma_0(x_{t+H} - x^0_{t+H}) + \Gamma_1(x_{t+H-1} - x^0_{t+H-1}) + \cdots \\ &\quad \cdots + \Gamma_{H-1}(x_{t+1} - x^0_{t+1}) + \epsilon_{t+H} \quad (5)\end{aligned}$$

The vector of pure forecasting errors ϵ_{t+H} is a vector of "reduced form" errors, but one can envisage its decomposition in structural contributions, each associated with an endogenous variable. Formally, we can write:

$$\epsilon_{t+H} = \Psi_0\nu_{t+H} + \Psi_1\nu_{t+H-1} + \cdots + \Psi_{H-1}\nu_{t+1} \quad (6)$$

where ν_{t+h} stands for the vector of residuals in the structural equations of the dynamic simultaneous equations model that is implicitly taken as the local approximation to the central forecast generating process.[6] As regards the $P \times P$ matrices Ψ_τ ($\tau = 0, \ldots, H-1$), their elements are the responses, after

[6]Let us represent this local linear model by:

$$A_0 y_t = A_1 y_{t-1} + \cdots + B_0 x_t + B_1 x_{t-1} + \cdots + \nu_t$$

where A_τ and B_τ ($\tau = 0, 1, 2, \ldots$) are respectively $P \times P$ and $P \times K$ matrices, with A_0 constrained to be non-singular. Then $\Gamma_0 = A_0^{-1}B_0$, $\Psi_0 = A_0^{-1}$, $\Gamma_\tau = A_0^{-1}B_\tau + A_0^{-1}A_1\Gamma_{\tau-1}$ and $\Psi_\tau = A_0^{-1}A_1\Psi_{\tau-1}$ for any $\tau \geq 1$.

τ periods, of the endogenous variables to shocks in the structural residuals ν. For example, the (i, j)-element of Ψ_0 is the contemporaneous response of the i-th endogenous variable included in y to a shock in the residual of the structural equation normalized in the j-th element of y.

Merging the two latter expressions, we obtain:

$$\begin{aligned} e_{t+H} &= \Gamma_0(x_{t+H} - x^0_{t+H}) + \\ &+\Gamma_1(x_{t+H-1} - x^0_{t+H-1}) + \cdots + \Gamma_{H-1}(x_{t+1} - x^0_{t+1}) + \\ &+\Psi_0\nu_{t+H} + \Psi_1\nu_{t+H-1} + \cdots + \Psi_{H-1}\nu_{t+1} \end{aligned} \tag{7}$$

Note that the conventional forecasting errors e and the deviations $x - x^0$ are observable for past forecasting exercises. Therefore, conditional on a set of estimates of the interim multipliers Γ_τ and Ψ_τ ($\tau = 0, \ldots, H-1$), estimates of ν are easily computed from (7) for these past exercises.[7]

Equations (3) and (7), respectively for the uniequational and multiequational frameworks, present the forecasting errors e_{t+H} as linear combinations of the relevant inputs, which are the"errors" in the conditioning variables $\{(x_{t+1} - x^0_{t+1}), \ldots, (x_{t+H} - x^0_{t+H})\}$ and the pure forecasting errors. These linear equations form the basis for estimating the marginal probability distributions of the forecasting errors $e_{p,t+H}$ ($p = 1, \ldots, P$), from which inter alia confidence bands and fan charts may be constructed. For this purpose, the joint probability distribution of the inputs is required. The BoE approach assumes that these input variables are independent and that their marginal probability distributions are two-piece normal.

[7] Indeed, we have:
$\nu_{t+1} = \Psi_0^{-1}\epsilon_{t+1} = \Psi_0^{-1}\left[e_{t+1} - \Gamma_0\left(x_{t+1} - x^0_{t+1}\right)\right]$
$\nu_{t+2} = \Psi_0^{-1}\left[\epsilon_{t+2} - \Psi_1\nu_{t+1}\right] = \cdots$
$\nu_{t+3} = \Psi_0^{-1}\left[\epsilon_{t+3} - \Psi_1\nu_{t+2} - \Psi_2\nu_{t+1}\right] = \cdots$

2.2 BoE aggregation of the *tpn*

A scalar random variable z has a two-piece normal distribution with parameters $(\mu, \sigma_1, \sigma_2)$ if it has probability density function (pdf):

$$f(z) = \begin{cases} A_1\ \phi(z|\ \mu, \sigma_1^2), & z \leq \mu \\ A_2\ \phi(z|\ \mu, \sigma_2^2), & z \geq \mu, \end{cases} \tag{8}$$

where $A_j = 2\sigma_j/(\sigma_1 + \sigma_2)$ and $\phi(z|\ \mu, \sigma_j^2)$ represents the normal pdf with mean μ and variance σ_j^2, for $j = 1, 2$. The mode of z is μ and John (1982) showed that the mean, variance and third central moment are, respectively:

$$E(z) = \mu + \sqrt{\frac{2}{\pi}}(\sigma_2 - \sigma_1), \tag{9}$$

$$V(z) = (1 - \frac{2}{\pi})(\sigma_2 - \sigma_1)^2 + \sigma_1\sigma_2, \tag{10}$$

$$T(z) = \sqrt{\frac{2}{\pi}}(\sigma_2 - \sigma_1)\left[(\frac{4}{\pi} - 1)(\sigma_2 - \sigma_1)^2 + \sigma_1\sigma_2\right]. \tag{11}$$

If $\sigma_1 = \sigma_2$, then the *tpn* distribution collapses to the (symmetric) normal distribution, but otherwise the density is skewed (to the right if $\sigma_1 < \sigma_2$ and to the left if $\sigma_1 > \sigma_2$).

To keep the notation as simple as possible, let e be the overall forecasting error (e.g. of inflation or output) at a given horizon and let $z = (z_1, \ldots, z_n, \ldots, z_N)'$ represent the vector of N associated input variables. Based on the arguments presented in the previous subsection, we will write e as a linear combination of z:

$$e = \alpha' z \tag{12}$$

where α is an $N \times 1$ vector of coefficients.

By construction, the modes of the input variables (errors in the conditioning variables and pure forecasting errors) are null, i.e., $\mu_n = 0$ $(n = 1, \ldots, N)$. In order to estimate the remaining parameters characterizing the *tpn* of each input variable z_n, two adittional pieces of information are needed. For that purpose, it is very convenient to work in terms of (i) the variance, $V(z_n)$, and (ii) the mode quantile, $P[z_n \leq M(z_n)]$, with the mode set at zero $M(z_n) = 0$. The choice of the mode quantile is aimed at facilitating the interpretation and discussion of the distributional assumptions between the forecaster, on the one hand, and the the public, on the other. With the mode quantile it is easier to convey information on skewness than, say, with the third central moment. For example, placing a value of 40% in the mode quantile immediately suggests (and quantifies) that the risks to the forecast are on the upside. Thus, if for each input variable the forecaster collects the mode quantile and the variance, it can solve the following system:

$$\begin{cases} P[z_n \leq 0] = \sigma_{1n}/(\sigma_{2n} - \sigma_{1n}) \\ V(z_n) = (1 - \frac{2}{\pi})(\sigma_{2n} - \sigma_{1n})^2 + \sigma_{1n}\sigma_{2n} \end{cases} \tag{13}$$

to obtain estimates of σ_{1n} and σ_{2n} With this information, one may compute the mean of each input variable:

$$E(z_n) = \sqrt{\frac{2}{\pi}}(\sigma_{2n} - \sigma_{1n}), \tag{14}$$

which are then linearly aggregated to generate the mean of e. By assumption, in the BoE approach the mode of e, $M(e)$, is also set to zero, which means that the modal point forecast – the baseline – is kept unchanged. Due to the independence assumption, the variance of e is obtained by simply summing the weighted variances of the input variables. Approximating the distribution of e by a *tpn*, its parameters σ_1 and σ_2 are obtained by solving the following system:

$$\begin{cases} \sum_n \alpha_n E(z_n) = \sqrt{\frac{2}{\pi}}(\sigma_2 - \sigma_1) \\ \sum_n \alpha_n^2 V(z_n) = (1 - \frac{2}{\pi})(\sigma_2 - \sigma_1)^2 + \sigma_1\sigma_2. \end{cases} \tag{15}$$

The process is repeated for each relevant forecasting period and summarized in a fan chart format such as Figure 1.[8] This graph depicts the central

[8]More precisely, the distributions are obtained for the fourth quarter of each year in the forecasting horizon and then interpolated between these periods to draw the fan chart.

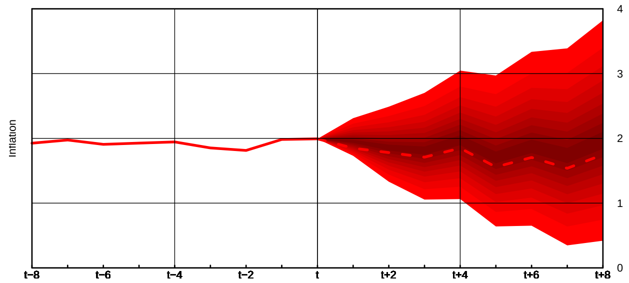

Figure 1: An illustrative fanchart

forecast with confidence bands fanning out around it, each covering an additional 10% of the distribution. If the pdf is asymmetric, the end points of the 90% confidence band do not correspond to the 5th and 95th quantiles. Instead, they are such that they have the same density value and cover 90% of the distribution.

2.3 Wallis (1999) criticisms

Wallis (1999) criticized some aspects of the BoE approach. On the one hand, he argues that the choice of the mode as the central scenario has implicit an "all-or-nothing" loss function, which might be too restrictive. On the other hand, Wallis criticized the method used by the BoE to compute the confidence intervals around the mode, showing that it may distort the perception of the distribution's skewness. For example, if the risks are 'on the upside', the upper limit of the 90% confidence band in the BoE fan chart falls below the 95th quantile. Both criticisms are of a presentational nature, as they do not refer to the validity of the statistical bulk of the procedure. Indeed, conventional confidence intervals can be adopted without requiring any fundamental change to the BoE procedure. In terms of the choice of the central scenario, it is very convenient for the forecasters to use the mode (the most likely value of the variable) and the mode quantile (i.e. the probability of having an outcome not higher than the mode) when communicating with the public. Again, if the forecasters choose to work with mean central forecasts instead of mode central forecasts, for computational and internal purposes, the adaptation of the BoE approach is straightforward. The discussion within the forecasting institution and the presentation to the public can still be based on mode central forecasts.

2.4 The issue of linear aggregation of the modes

A more serious statistical limitation of the BoE approach relates to the derivation of the modes of the forecasting error distributions. From the expression of the *tpn* pdf, it is obvious that in general a linear combination of variables with *tpn* marginal distributions is not *tpn* distributed, even under independence. However, if we are only interested in the three first moments of the distribution, one can argue that the *tpn* assumption is reasonable. The problem is that the statistical procedure suggested by the BoE to deal with the approximation is questionable. By assuming that the mode of e is set to zero (as well as the modes of all input variables), the BoE approach is implicitly taking that the mode of the linear combination of inputs is the linear combination, with the same coefficients, of the modes of those inputs. But clearly this is not the case under asymmetry, as illustrated in Figure 2.

Two identically distributed and independently generated[9] *tpn* variables z_1 and z_2 with $(\mu = 0,\ \sigma_1 = 0.3,\ \sigma_2 = 0.8)$ were randomly drawn (20,000 draws) and then the kernel density of $e \equiv z_1 + z_2$ was computed, resulting in a mode estimate of 0.434, clearly not the linear combination of the input modes.

Under the assumption of independence of the input variables, a better approximation of the mode of the distribution could be based on the (correct) aggregation of the three first moments of the distributions. In general, for $e =$

[9] In sample correlation: 0.0007.

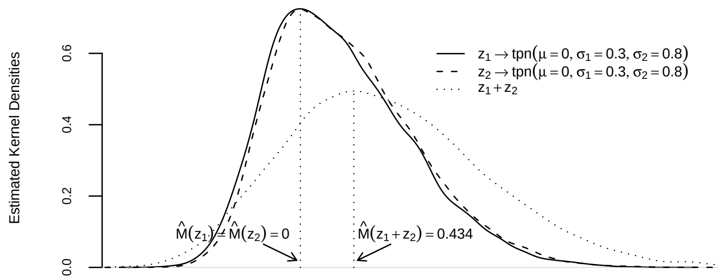

Figure 2: $M(z_1 + z_2)$ vs. $M(z_1) + M(z_2)$

$\alpha' z$, where α and z continue to be $N \times 1$ vectors, respectively of coefficients and independent *tpn* variables, we have:

$$E(e) = \sum_{n=1}^{N} \alpha_n E(z_n) \tag{16}$$

$$V(e) = \sum_{n=1}^{N} \alpha_n^2 V(z_n) \tag{17}$$

$$T(e) = \sum_{n=1}^{N} \alpha_n^3 T(z_n) \tag{18}$$

If we choose to approximate the distribution of e by a *tpn*, we can estimate the three characterizing parameters by the method of moments (MM), first computing the moments of e using (16)-(18) and then estimating the parameters by solving the 3×3 system of equations (9)-(11). In our example above, with $e = z_1 + z_2$, instead of a mode zero as assumed by the BoE, a more correct mode estimate (approximating to the true distribution of e by a *tpn* as suggested) yields 0.414, quite close to the value obtained from the kernel density estimation.

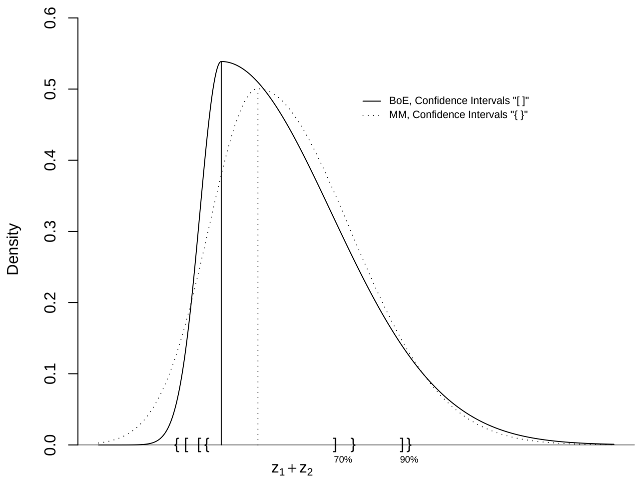

Figure 3: BoE versus MM approaches

Furthermore, if we construct confidence intervals based on the two distributions – BoE and MM – the differences become even more evident. The 70% confidence intervals[10] are $[-0.249, 1.286]$ and $[-0.163, 1.490]$ for the BoE and MM approaches, respectively, while the 90% confidence intervals are $[-0.396, 2.040]$ and $[-0.501, 2.121]$. The differences between the two underlying density functions are further illustrated in Figure 3.

2.5 The issue of independence

A second issue is the assumption of independence of the inputs used in practice to simplify the aggregation problem. In our opinion, this assumption is a limitation in the BoE approach, in particular in the context of dynamic

[10]The confidence intervals are built using the minimum range method described at the end of section 2.2, rather than the percentile method. The choice of 70% in the guassian case would corresponds to approximately a confidence interval one standard deviation to each side of the central forecast. In the *tpn* case that is naturally not the case, but the range of the interval is still approximately two standard deviations.

formulations as (3) and (7). Indeed, one can anticipate that the forecasting errors for a conditioning variable, for example the price of oil or the external demand, are correlated in successive horizons. In addition, one can also anticipate significant contemporaneous correlations between the errors of different conditioning variables.

To illustrate the impact of correlation in the degree of skewness, we follow the suggestion by Blix and Sellin (2000) to derive one of the possible joint distributions of two random variables with given *tpn* marginal distributions and non-zero correlation.[11] They use a result taken from Mardia (1970), who showed that the joint pdf

$$g(z_1, z_2) = \frac{g_1(z_1)g_2(z_2)}{\sqrt{1-\rho^2}} \exp\{-\frac{\rho}{2(1-\rho^2)}[\rho(\check{z}_1^2 + \check{z}_2^2) - 2\check{z}_1\check{z}_2]\}, \quad (19)$$

has as marginal pdf's $g_1(z_1)$ and $g_2(z_2)$. The transformed variables necessary to generate such result, $\check{z}_1$ and $\check{z}_2$, have correlation ρ and are defined as

$$\check{z}_1 \equiv \Phi^{-1}(G_1(z_1)), \quad (20)$$

$$\check{z}_2 \equiv \Phi^{-1}(G_2(z_2)), \quad (21)$$

where $\Phi(\cdot)$ is the cumulative standard normal distribution and $G_1(\cdot)$ and $G_2(\cdot)$ are the cumulative functions of $g_1(\cdot)$ and $g_2(\cdot)$, respectively. With this joint distribution, we were able to generate random samples of variables with marginal distributions *tpn* and correlation $\tilde{\rho}$.[12] We took the marginal *tpn*'s parametrization of the example considered in the previous subsection ($\mu = 0,\ \sigma_1 = 0.3,\ \sigma_2 = 0.8$) to illustrate the effects of correlation on skewness. For each of three different values of ρ (-0.8, 0.0 and 0.8), a bivariate random sample of size 20,000 was generated.[13] Each pair was subsequently

[11] As we noticed in the introduction, there is more than one joint distribution associated with given marginal distributions and correlation.

[12] Lancaster (1957) showed that $|\tilde{\rho}| \leq |\rho|$.

[13] First, using the marginal pdf, a sample was generated for z_1. Second, with this sample and the conditional pdf $g_{1|2}(z_2|z_1) = g(z_1, z_2)/g_1(z_1)$, the rejection method was used to generate the sample for (z_1, z_2). Due to the random nature of the exercise and Lancaster's inequality, the sample correlations were below (in absolute value) the corresponding values of ρ at -0.7498 and 0.7861.

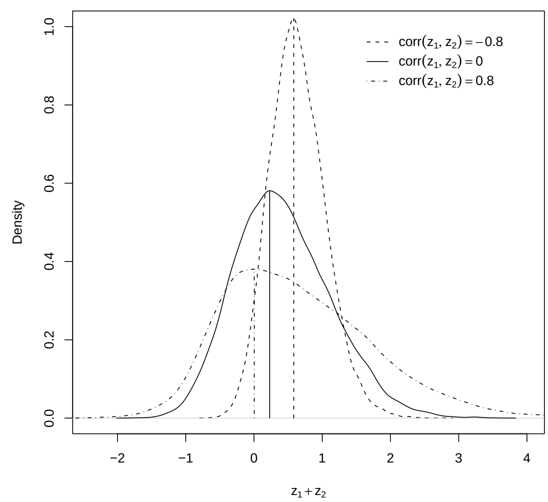

Figure 4: Illustrative sum of *tpn*'s with varying correlations

added, say $e = z_1 + z_2$, and its degree of skewness computed $(T(e)/V(e)^{3/2})$. Figure 4 illustrates that the degree of skewness of e depends on the degree of correlation present in the variables linearly combined. In the zero correlation case the degree of skewness was 0.47. Intuitively, positive correlation increased the degree of skewness (to 0.62), while negative correlation reduced it (to 0.32).

We conclude from this illustration that the degree of skewness of a linear combination of correlated random variables will depend in general on the magnitude of the correlation. Hence, relaxing the independence assumption, at least partially, is a step forward towards more realistic uncertainty and risk assessments of the forecasted variables, due to the likely correlation between the input variables. One natural alternative to assuming the independence of all input variables is to assume a non-diagonal linear correlation matrix between these variables.

Let us continue to denote the overall forecasting error and the $N \times 1$ vector of input variables by e and by z, respectively, with $e = \alpha' z$, as above.

Instead of assuming that z_n and z_q are independent for any $n \neq q$, let us now simply assume that linear correlations between these input variables are available: $R = [r_{nq}]$, where r_{nq} is the linear correlation between z_n and z_q. These correlations may be based on the past observed history of errors, possibly modified by judgements introduced by the forecaster.

As aforementioned, when we assume a non-diagonal correlation matrix, we are not fully specifying an alternative to the assumption of independence. Besides the nullity of the correlation matrix off-diagonal elements, the independence also implies imposing restrictions on all the cross-moments of an order higher than two. Therefore, by simply specifying the marginal distributions $f_n(z_n)$ $(n = 1, \ldots, N)$ and the correlation matrix R, in general we have multiple joint distributions of the input variables $f(z)$ (implying multiple distributions of the forecast error e) compatible with those assumptions.

To keep the problem of non-identification as simple as possible, we suggest that the forecaster should first partition the set of input variables and assume that the input variables belonging to each subset are correlated among themselves but are not correlated with input variables belonging to other subsets. For instance, it may be acceptable to assume *a priori* the independence between two subsets of inputs, the first including the errors in the conditioning variables $\{(x_{t+1} - x^0_{t+1}), \ldots, (x_{t+H} - x^0_{t+H})\}$ and the second including the pure forecasting errors (either in the reduced form ϵ_{t+h} or in the structural form $\nu_{t+1}, \ldots, \nu_{t+H}$). In this case we can rewrite e_{t+H} in (3) and (7) as a sum of two independent variables $e_{1,t+H}$ and $e_{2,t+H}$, with

$$e_{1,t+H} = \Gamma_0(x_{t+H} - x^0_{t+H}) + \Gamma_1(x_{t+H-1} - x^0_{t+H-1}) + \Gamma_{H-1}(x_{t+1} - x^0_{t+1}) \qquad (22)$$

and

$$e_{2,t+H} = \epsilon_{t+H} \qquad (23)$$

in the uniequational framework or

$$e_{2,t+H} = \Psi_0 \nu_{t+H} + \Psi_1 \nu_{t+H-1} + ... + \Psi_{H-1} \nu_{t+1} \qquad (24)$$

in the multiequational framework.

After defining the partition of input variables, we propose to look for solutions which correspond to a distribution of the forecasted variable belonging to the same family as the one assumed for the marginal distributions of the input variables. The "*tpn* environment" of the BoE approach is not the most convenient for this purpose, in particular because the derivatives of order higher than one of the *tpn* pdf are not continuous at the mode of the distribution, which complicates the algebra. In the next section we will present an alternative to the *tpn* distribution, which we denote the "skewed generalized normal" (*sgn*).

3 An Alternative Approach

3.1 The skewed generalized normal distribution

Let z be a (scalar) random variable which results from the linear combination of two independent random variables

$$z = \theta_1 + \theta_2 w + \theta_3 s, \quad \theta_1, \theta_3 \in \mathbb{R}, \quad \theta_2 > 0, \tag{25}$$

where $w \sim N(0,1)$ (standard Gaussian) and $s \sim 2^{1/3} \exp[-2^{1/3}(s+2^{-1/3})]$, for $s \geq -2^{-1/3}$. That is, s results from an exponential distribution with defining parameter set to $2^{1/3}$ after a simple location shift ($-2^{-1/3}$), in order to have $E(s) = 0$. By setting the parameter at $2^{1/3}$, we also have $T(s) = 1$, where as previously $T(\cdot)$ stands for the third central moment. The distribution of z will be denoted

$$z \sim S(\theta_1, \theta_2, \theta_3) \tag{26}$$

and $S(\cdot)$ will be referred to as "skewed generalized normal" (*sgn*). The probability density function, which can be obtained with standard change-of-variable statistical techniques, is given by

$$f(z) = \begin{cases} \frac{1}{\sqrt{2\pi}\theta_2}\, e^{-\frac{1}{2}\left(\frac{z-\theta_1}{\theta_2}\right)^2}, & \theta_3 = 0,\ z \in \Re \\ e^{\eta}\, \frac{2^{1/3}}{|\theta_3|} e^{-\frac{2^{1/3}}{\theta_3} z} \Phi_{\nu,\theta_2}\left(\frac{\theta_3}{|\theta_3|}\, z\right), & \theta_3 \neq 0,\ z \in \Re \end{cases} \tag{27}$$

where $\eta = -1 + 2^{1/3}\theta_1/\theta_3 + 2^{-1/3}(\theta_2/\theta_3)^2$, and $\nu = 2^{-1/3}(\theta_3\eta + 2^{-1/3}\theta_2^2/\theta_3)$. The function $\Phi_{\mu,\sigma}(\cdot)$ represents the Gaussian cumulative distribution function with mean μ and standard deviation σ.

Directly from the definition of z, we have $S(\theta_1, \theta_2, 0) = N(\theta_1, \theta_2^2)$. Therefore, as in the case of the *tpn* (with $\sigma_1 = \sigma_2$), the *sgn* collapses to the normal distribution for the particular case of $\theta_3 = 0$. For the general case $z \sim S(\theta_1, \theta_2, \theta_3)$, by making use of the well-established properties of the normal and exponential distributions:

i) $E(z) = \theta_1$, $V(z) = \theta_2^2 + 2^{-2/3}\theta_3^2$ and $T(z) = \theta_3^3$;

ii) the characteristic function of z is

$$C_z(t) = \exp\{\theta_1 it\}.C_w(\theta_2 t).C_s(\theta_3 t) = \frac{2^{1/3}}{2^{1/3} - \theta_3 it}\, \psi(t) \tag{28}$$

where $\psi(t)$ is the characteristic function of the normal distribution $N(\theta_1 - 2^{-1/3}\theta_3, \theta_2^2)$.

Unlike for the *tpn*, there is no closed expression for the mode of the *sgn*, but given the values of its three parameters, standard simple search algorithms prove to be quite efficient in locating the maximum of the pdf.

Notice that $V(z)$ is a function of θ_3, i.e., it is a function of the third central moment. Clearly, one cannot pick the degrees of skewness and uncertainty (variance) arbitrarily. Indeed, as the *sgn* is only defined for $\theta_2 > 0$, from the variance expression the following condition can be derived:

$$V(z) > 2^{-2/3}\theta_3^2. \tag{29}$$

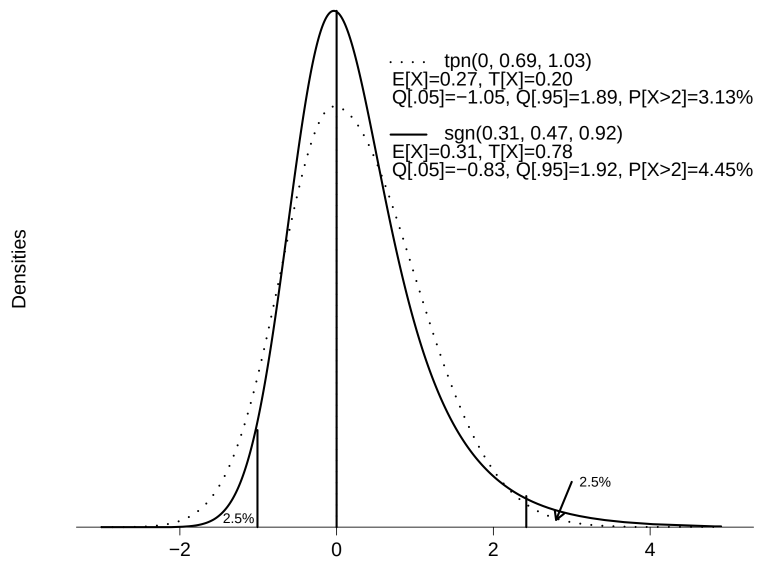

Figure 5: sgn vs. tpn given $M(z) = 0, \ V(z) = 0.75, \ P[z < M(z)] = 0.40$

Condition (29) can be reexpressed in terms of the absolute degree of skewness of z:

$$\left(\frac{\theta_3^2}{V(z)}\right)^{3/2} < 2 \iff \frac{|T(z)|}{V(z)^{3/2}} < 2, \tag{30}$$

that is, the maximum absolute skewness allowed is 2. This restriction is similar in spirit to the condition for the *tpn* described in John (1982), but it is less binding. While for the *tpn* the absolute skewness cannot exceed $(1/2\pi - 1)^{-3/2} - (1/2\pi - 1)^{-1/2} \simeq 0.99527$, for the *sgn* the absolute bound is (slightly more than) twice the *tpn*'s bound.

At this point, it is interesting to compare the *sgn* and the *tpn* distributions. In Figure 5, we carry out the following exercise. Given the mode, 0, variance (standard deviation), 0.75 (0.87), and mode quantile, 0.40, we determine the parameters of the *tpn* and *sgn* and plot the corresponding densities along with some summary statistics. From this exercise, we conclude that the *sgn* has fatter tails (4th moment) than the *tpn*. This information can be complemented with the 5th and 95th quantiles, which are respectively -0.83 and 1.92 for the *sgn* and -1.05 and 1.89 for the *tpn*. Thus, the 90% confidence range for the *tpn* would expand over 2.94 units, while the *sgn* only over 2.75. In this example, the probability that the variable departs from the mode by more than two units, $P(|z| > 2)$, is 4.15% for the *sgn* and 3.13% for the *tpn* case.

3.2 Linear combinations of *sgn* distributions

In general, the *sgn* family, like the *tpn* family, is not closed relative to linear combinations. In particular, if $z = (z_1, ..., z_n, ..., z_N)'$ is a vector of random variables with marginal *sgn* distributions and correlation matrix R, the distribution of $\alpha' z$ is not necessarily *sgn*. More importantly, the *sgn* marginal distributions and the correlation matrix do not uniquely identify a single joint density of z, and several joint distributions will share the same marginal distributions and the same correlation matrix. Hereafter, $z \asymp z^*$ denotes that the two vectors of random variables z and z^* share the same marginal distributions and the same correlation matrix. Due to this identification problem, there is no single solution for the distribution of e. However, we can show that among those possible multiple solutions, if a certain condition is fulfilled, one solution belongs to the *sgn* family and therefore some form of closedness is achieved. The following theorem formally states this result:

Theorem 1 *Let $z = (z_1, \ldots, z_n, \ldots, z_N)'$ be any $N \times 1$ vector of random variables with linear correlation matrix R and marginal distributions*

$$z_n \sim S(\theta_{1n}, \theta_{2n}, \theta_{3n}), \quad (n = 1, ..., N)$$

Also, let $\bar{\Omega} = DRD - 2^{-2/3}\theta_3\theta_3'$, where

$$D = diag\left(\sqrt{\theta_{21}^2 + 2^{-2/3}\theta_{31}^2}, \ldots, \sqrt{\theta_{2n}^2 + 2^{-2/3}\theta_{3n}^2}, \ldots, \sqrt{\theta_{2N}^2 + 2^{-2/3}\theta_{3N}^2}\right)$$

and $\theta_3 = (\theta_{31}, ..., \theta_{3n}, ..., \theta_{3N})'$*. If* $\bar{\Omega}$ *is positive semidefinite and its diagonal elements are all positive, then there are* $N+1$ *independent random variables,* $w = (w_1, \ldots, w_n, \ldots, w_N)'$ *and* s*, with*

$$\begin{aligned} w_n &\sim N(\mu_n, \sigma_{nn}), \qquad n = 1, \ldots, N \\ s &\sim 2^{1/3} \exp[-2^{1/3}(s + 2^{-1/3})], \end{aligned}$$

such that:

$$z \asymp z^* \equiv Bw + \theta_3 s,$$

where B *is the orthogonal matrix of eigenvectors of* $\bar{\Omega}$*,* $\Sigma = diag(\sigma_{11}, \ldots, \sigma_{nn}, \ldots, \sigma_{NN})$ *is the corresponding diagonal matrix of (non-negative) eigenvalues,* $\mu = (\mu_1, \ldots, \mu_n, \ldots, \mu_N)' = B'\theta_1$*, and* $\theta_1 = (\theta_{11}, \ldots, \theta_{1n}, \ldots, \theta_{1N})'$*. The distribution of*

$$\alpha' z^* = (\alpha' B)w + (\alpha'\theta_3)s$$

is sgn *with* $E(\alpha' z^*) = \alpha'\theta_1$*,* $V(a'z^*) = \alpha' DRD\alpha$ *and* $T(a'z^*) = (\alpha'\theta_3)^3$ *(i.e.* $\alpha' z^* \sim S(\alpha'\theta_1, \sqrt{\alpha'\bar{\Omega}\alpha}, \alpha'\theta_3)$*)*

The condition that $\bar{\Omega}$ must be positive semi-definite with positive diagonal elements (hereafter strict PSD condition) may be viewed as some type of multivariate version of condition (29) on the degree of asymmetry. Indeed, note that for $N = 1$ the strict PSD condition collapses to (29).

Unfortunately, in practice the strict PSD condition is very tight, and it is typically violated whenever we assume marked skewness for several input variables and/or strong correlations among those variables. This means that in many relevant cases, there is no *sgn* distribution among those solutions generated from the given marginal *sgn* distributions and correlation matrix. The reason for this is quite simple. According to the previous theorem, if the strict PSD condition is fulfilled, a single elementary source of asymmetry (the scalar random variable s) is enough to generate the asymmetries of the different elements of vector z^*. Indeed, from the definition of z^*, the distribution of a given variable z_n^* is skewed if and only if the corresponding element θ_{3n} of the coefficient vector θ_3 is non-null. As the covariance matrix $V(z^*)$ is given ($V(z^*) = V(z) = DRD$), and the elements of vector θ_3 are uniquely determined by the need to generated the required asymmetries in the marginal distributions, the difference

$$\bar{\Omega} = V(z^*) - 2^{-2/3}\theta_3\theta_3', \tag{31}$$

which should correspond to the covariance matrix of Bw, may not be positive semidefinite.

In such cases, we need to resort to weaker forms of closedness (or "almost closedness") of the *sgn* distributions relative to linear combinations. This implies considering marginal distributions of the input variables that are generated by linear combinations of independent *sgn* variables or, equivalently, by linear combinations of a normal variable and more than one shifted exponential variable, all independent (i.e. we need to replace the scalar s in the the theorem above by a vector of similar independent and identically distributed random variables). Let us name "convoluted *sgn* distribution of order m" the distribution of a linear combination of m independent variables with marginal *sgn* distributions (obviously, these convoluted *sgn* distributions collapse to the basic *sgn* when $m = 1$). Using the latter definition, we can define the "extented *sgn* family or order M" as including all the convoluted *sgn* distributions of order up to M.

If the strict PSD condition is not fulfilled, we suggest broadening the search among the multiple possible solutions for the distribution of e by successively looking for distributions belonging to the extended *sgn* family of orders $M = 2, 3, ..., N$ until a solution is found. Such a search can be implemented in practice by verifying whether the generalized versions of the PSD condition are fulfilled, as stated (for order M) in the following theorem, which is a generalization of Theorem 1 (it collapses to the latter when $M = 1$):

Theorem 2 *Let* $z = (z_1, \ldots, z_n, \ldots, z_N)'$ *be any* $N \times 1$ *vector of random variables with linear correlation matrix* R *and marginal distributions*

$$z_n \sim S(\theta_{1n}, \theta_{2n}, \theta_{3n}), \quad (n = 1, ..., N)$$

Also, let $\bar{\Omega} = DRD - 2^{-2/3}\Theta_3 CC'\Theta_3'$, *where*

$$D = diag\left(\sqrt{\theta_{21}^2 + 2^{-2/3}\theta_{31}^2}, \ldots, \sqrt{\theta_{2n}^2 + 2^{-2/3}\theta_{3n}^2}, \ldots, \sqrt{\theta_{2N}^2 + 2^{-2/3}\theta_{3N}^2}\right),$$

$\Theta_3 = diag(\theta_{31}, \ldots, \theta_{3n}, \ldots, \theta_{3N})$ *and* $C = [c_{nm}]$ *is a* $N \times M$ *matrix, with* M $(\leq N)$ *being the number of non-null diagonal elements of* Θ_3, *such that any element* c_{nm} *is either* 0 *or* 1, *with*

$$\begin{aligned} \sum_{m=1}^{M} c_{nm} &= 1 \;\; \text{if} \;\; \theta_{3n} \neq 0, \;\; 0 \;\; \text{otherwise} \quad (n = 1, \ldots, N) \\ \sum_{n=1}^{N} c_{nm} &\geqslant 1 \quad (m = 1, \ldots, M) \end{aligned}$$

If $\bar{\Omega}$ *is positive semidefinite and its diagonal elements are all positive, then there are* $N + M$ *independent random variables,* $w = (w_1, \ldots, w_n, \ldots, w_N)'$ *and* $s = (s_1, \ldots, s_m, \ldots, s_M)$, *with*

$$\begin{aligned} w_n &\sim N(\mu_n, \sigma_{nn}) \qquad (n = 1, \ldots, N) \\ s_m &\sim 2^{1/3}\exp[-2^{1/3}(s_m + 2^{-1/3})], \qquad (m = 1, \ldots, M) \end{aligned}$$

such that:

$$z \asymp z^{**} = Bw + \Theta_3 Cs,$$

where B *is the orthogonal matrix of eigenvectors of* $\bar{\Omega}$, $\Sigma = diag(\sigma_{11}, \ldots, \sigma_{nn}, \ldots, \sigma_{NN})$ *is the corresponding diagonal matrix of (non-negative) eigenvalues,* $\mu = (\mu_1, \ldots, \mu_n, \ldots, \mu_N)' = B'\theta_1$, *and* $\theta_1 = (\theta_{11}, \ldots, \theta_{1n}, \ldots, \theta_{1N})'$. *The distribution of*

$$\alpha' z^{**} = (\alpha' B)w + (\alpha'\Theta_3 C)s$$

is convoluted sgn of order M with $E(\alpha' z^{**}) = \alpha'\theta_1$, $V(\alpha' z^{**}) = \alpha' DRD\alpha$ and $T(\alpha' z^{**}) = \sum_{m=1}^{M} \delta_m^3$, where $\delta \equiv \alpha'\Theta_3 C$.

Firstly, notice that in Theorem 2 the three first moments of e do not suffice to explicitly determine the $2+M$ parameters of the convoluted *sgn* of order M. Although in principle it is possible to derive the explicit expressions for the probability densities of the convoluted *sgn* of orders 2 and higher, the derivation becomes increasingly cumbersome and the final expressions are quite complicated for $M > 2$. Therefore, whenever we deal with convoluted *sgn*, we suggest approximating those distributions by basic *sgn* distributions, using for that purpose the *sgn* that has the same relevant moments (mean, variance and third central moment) as the "exact" convoluted *sgn*.

As to the generalized PSD condition, for a given M if we can find a matrix C as defined in Theorem 2 such that the corresponding matrix $\bar{\Omega}$ is positive semidefinite with positive diagonal elements, then one of the solutions to our problem takes the form of a convoluted *sgn* of order M, with known three first moments. Intuitively, the role of matrix C is to allocate the different independent elementary sources of asymmetry (the elements of s) to the different variables z_n^{**} for which asymmetry is assumed. The allocation is such that at most one source s_m is allocated to each z_n^{**}, but more than one z_n^{**} may share the same source s_m.

It is also important to remark that the number of possible different choices of matrix C, $\#\{C\}$, which need to be checked until one finds a choice for which the corresponding generalized PSD condition is fulfilled, depends on the number of asymmetric input variables $\overline{N}$ ($\leqslant N$) and it grows exponentially with this number. For $\overline{N} = 1, 2, 3, 4, 5, 6, 7, ...$ we have $\#\{C\} = 1, 2, 5, 15, 42, 202, 877, ...$ For instance, if we have a subset with 3 input variables, 2 of which are asymmetric (the second and the third), the two possible choices of matrix C are the following:

$$\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

If instead we have a subset of 5 input variables ($N = 5$), 3 of which are asymmetric (the first, the third and the fifth) ($\overline{N} = 3$), the 5 possible choices of matrix C are:

$$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

For $\overline{N} \geqslant 4$ the set of choices of C is easily identified using a tree diagram. Figure 5 presents the tree diagram with all the 15 choices of the non-null rows of matrix C for $\overline{N} = 4$. In the diagram, the states A, B, C and D denote the first, second, third and fourth rows of an identity matrix of order 4, respectively. Notice that the number of branches that depart from each knot of the diagram is equal to one plus the number of different states that can be found upstream in a direct line leading to that knot. From each path in the diagram it is straightforward to build the corresponding choice of matrix C. For instance, with a subset of 6 input variables, the second and the fifth being symmetric, the path (A, B, C, A) translates into the matrix (6×3 because the last column was null and therefore may be suppressed in this particular example):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

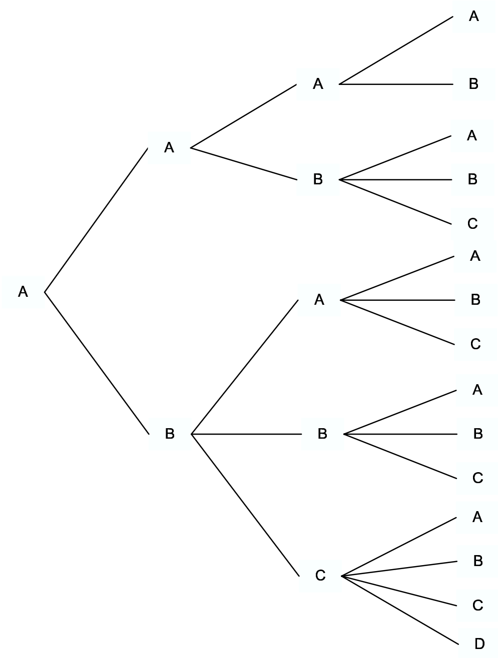

Figure 5 - Tree Diagram for $\bar{N}$= 4

It is worth mentioning that if the input variables refer to several forecasting horizons, the matrix C is subject to some logical restrictions that facilitate the search by organizing it in successive steps, thereby decreasing the number of possible choices. To illustrate the point, let us consider a forecasting exercise with an horizon of three periods and a subset of three input variables (z_1, z_2, z_3) (e.g. exchange rate, external demand, and commodity prices). Let us also assume that z_2 and z_3 are asymmetric in the first and second periods of the horizon. The forecasting error (e.g. of inflation) for the first period, e_{t+1}, will be a function of $(z_{1,t+1}, z_{2,t+1}, z_{3,t+1})$, while e_{t+2} will depend on $(z_{1,t+1}, z_{2,t+1}, z_{3,t+1}, z_{1,t+2}, z_{2,t+2}, z_{3,t+2})$ and e_{t+3} will additionally depend on period $t+3$ variables, $(z_{1,t+3}, z_{2,t+3}, z_{3,t+3})$. When estimating the density of e_{t+1}, there will be only two possible choices of C for this subset of input variables (as indicated above). Without any further temporal restrictions, for the years $t+2$ and $t+3$ the number of possible choices of C would be 15 (corresponding to e_{t+1}). However, the joint distribution of

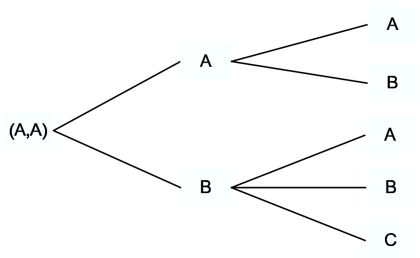

Figure 6 - Tree Diagram - second step (CASE 1)

$(z_{1,t+1}, z_{2,t+1}, z_{3,t+1})$ becomes known when estimating e_{t+1} and it would not be consistent to reestimate a different joint distribution of these three variables when estimating e_{t+2} and e_{t+3}. Therefore, matrix C is restricted to the following formats, respectively for the first, second and third periods:

$$[C_{11}], \quad \begin{bmatrix} C_{11} & 0 \\ C_{21} & C_{22} \end{bmatrix}, \quad \begin{bmatrix} C_{11} & 0 & 0 \\ C_{21} & C_{22} & 0 \\ C_{31} & C_{32} & C_{33} \end{bmatrix}$$

where block C_{11} is the same in all three matrices and the blocks C_{21} and C_{22} are the same in the second and third matrices. A slightly modified tree diagram can also be constructed to list the possible choices for blocks $\begin{bmatrix} C_{21} & C_{22} \end{bmatrix}$, given the previous choice of C_{11}. Figures 6 and 7 illustrate the tree diagrams corresponding to two different choices of C_{11}. As to the third period, the three blocks, C_{31}, C_{32} and C_{33} will be null because we are assuming that none of the input variables is asymmetric at this horizon.

3.3 Summing up: the practical implementation of the alternative approach

In order to clarify how the statistical methodology that we propose can be implemented in practice, let us sum up the successive steps implied by the results presented so far. As a convenient simplification of the notation used in subsection 2.1, we will continue to denote by e the overall forecasting error for a given variable of interest (e.g. inflation or output), at a given forecasting horizon, and by z the corresponding $N \times 1$ vector of input variables. We

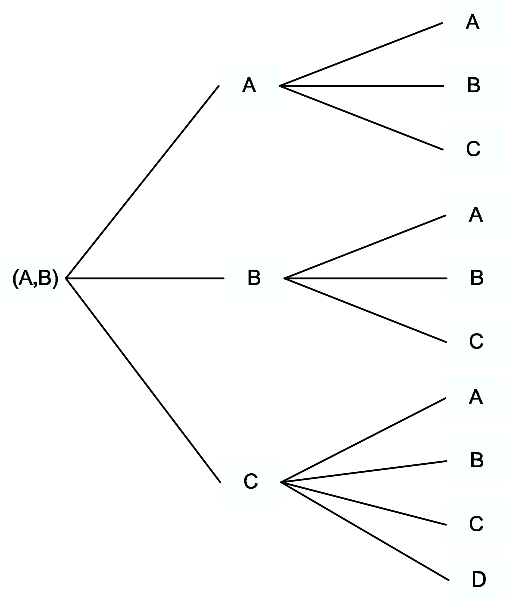

Figure 7 - Tree Diagram - second step (CASE 2)

will assume that $e = \alpha' z$ is a reasonable local linear approximation to the forecasting error generating process.

The first step of our approach consists of computing the parameters of the marginal *sgn* distributions for all the input variables. We admit that the forecaster, besides estimates for α, also has available estimates for the standard deviation and the mode quantile of each input variable z_n $(n = 1, ..., N)$. Given the mode (set to zero), standard deviation and mode quantile, the three parameters of each marginal *sgn* distribution can be easily obtained from a simple algorithm, taking into account the expression of the *sgn* pdf.

The second step of the procedure will consist of partitioning the set of inputs in K subsets such that the input variables belonging to different subsets are assumed to be independent. Accordingly, we may break down e and write $e = e_{(1)} + ... + e_{(k)} + ... + e_{(K)}$, where $e_{(k)} = \alpha'_{(k)} z_{(k)}$ denotes the component of e associated with the input variables belonging to the k-th subset (which contains $N_{(k)}$ input variables, with $N_{(1)} + ... + N_{(K)} = N$). The partition should be based as much as possible on economic reasoning. As suggested above, one of the simplest partitions that is possible to envisage separates the set of input variables in only two subsets, the first including the errors in the conditioning variables and the second including the pure forecasting errors either in their reduced form or their structural form. After deciding on the partition to be used, the forecaster needs to accordingly choose the correlation matrix of the input variables belonging to each subset.

The third step corresponds to checking, for each subset of input variables, all the possible PSD conditions. One should start with $M = 1$ and, if needed, increase the value of M and explore all the possible choices of C (as defined in Theorem 2) associated with each M, until the PSD condition is fulfilled. For a given matrix C, checking the PSD condition requires computing the eigenvalues of the corresponding $\bar{\Omega}$.

If for all subsets of input variables we were able to fulfill the PSD condition, Theorems 1 or 2 provide the formula for the first three moments of $e_{(1)}, ..., e_{(k)}, ..., e_{(K)}$. As these variables are independent by construction, we can obtain the first three moments of e in a straightforward manner:

$$E(e) = \sum_{k=1}^{K} E(e_{(k)}); \; V(e_{t+h}) = \sum_{k=1}^{K} V(e_{(k)}); \; T(e_{t+h}) = \sum_{k=1}^{K} T(e_{(k)}) \quad (32)$$

Taking these moments, the final step will consist of approximating the distribution of e by the *sgn* distribution that has the same relevant moments.

In the benchmark case of symmetric z_n distributions, by construction we have $E(z_n) = M(z_n) = 0$ $(n = 1, \ldots, N)$ and $E(e) = \sum_{n=1}^{N} \alpha_n E(z_n) = 0 = M(e)$. However, these equalities will not stand anymore if the distributions of some inputs are skewed. In the latter case, we will have $E(e) = \sum_{n=1}^{N} \alpha_n E(z_n) \neq 0$, because $E(z_n) \neq 0$ for those inputs with skewed distributions. $M(e)$ becomes also non-null but, unlike the mean, it cannot be expressed as a linear combination of the z_n modes. The effect on the distribution of e of considering risks to some inputs are twofold: (i) the location measures, mode and mean, are shifted, but by different factors; (ii) the distribution will become skewed. Therefore, when evaluating the effect of input risks on the variable of interest, we need to take into account both the "mode effect", defined as the shift of the modal forecasts, and the "skewness effect", measured by the mode quantile of the new distribution of e. A synthetic indicator of both effects is the "baseline quantile", which corresponds to the probability (under the new skewed distribution) of an outcome below zero, i.e., the probability of having one outcome for the endogenous variable below its baseline modal forecast.

An alternative synthetic indicator is the "mean effect" defined as the shift of the mean of the distribution of e relative to the symmetric case. The overall mean effect can be expressed as the linear aggregation of the mean effects of the input variables, making it particularly appropriate if we are interested in performing sensitivity analysis to different risk scenarios. However, the mean effect is less convenient than the mode effect if we want to keep the interpretation of point forecasts as modal forecasts. It should be noted that the mean effect in our approach only differs from the one computed by the BoE approach because we are aggregating *sgn* instead of *tpn* distributions. In addition, the BoE's *tpn* distribution of e will have zero mode, whereas the mode of the *sgn* distribution generated from our approach will reflect the aforementioned mode effect. As a result, confidence intervals (fan charts) under the two approaches will be different, particularly if the skewness is significant.

One may argue that in most cases the "owner" of the forecast will adjust the baseline scenario if the risks to the inputs are judged to be too large, ensuring that the BoE approach will not produce significantly biased risk assessments. Otherwise, the credibility of the point modal baseline would be at stake. In our opinion, however, there may be relevant situations when a large skewness of the distribution of some inputs is warranted without necessarily requiring changes in the baseline for those inputs. For instance, that may happen if the baseline paths for some conditioning variables are set exogenously to the owner of the forecast. In the case of the forecasts produced by central banks, the assumptions made on the paths of fiscal variables constitute an example. Indeed, the fiscal paths are typically set by central banks at the official targets announced by the government, irrespectively of how credible they are. The team of forecasters or the institutional body that owns the forecast may want to attach significant upside or downside risks to those exogenously defined conditioning paths. This situation illustrates that it clearly pays to use a robust statistical procedure to aggregate the risks to the input variables, and the limitations of the BoE approach raise doubts about its appropriateness. In this respect, stochastic simulation methods may be seen as an alternative to our procedure, but in our opinion their practical usefulness in live/real uncertainty and risk assessement is still limited.

3.4 An approximation of last resort

The methodology outlined in the previous subsection may not provide the forecaster with a solution. This will happen whenever the strict PSD condition as well as the generalized PSD conditions of orders up to M are all violated for at least one subset of input variables. In such a case, there are two possible ways to overcome the situation. A first way out, which should be the preferred one if acceptable to the forecaster in economic terms, consists of redefining the initial assumptions. This revision can take several forms: downward revisions (in absolute terms) of the degrees of skewness considered in the marginal distributions of the input variables, multiplication of sample correlations by a shrinking factor (between 0 and 1) and/or a finer partition of the input variables. In the latter case, by working with more subsets of input variables, the number of independence restrictions will increase and, at least in principle, it will become easier to fulfill the PSD condition for each subset.

When all the forms of the PSD condition are violated for some subsets of inputs and further revisions of assumptions appear as rather controversial on economic grounds, we suggest using an "approximation of last resort" which consists of picking the distribution of the extended *sgn* family that in some sense can be considered closer to the set of solutions. The basic result for this approximation is formally stated in the following theorem:

Theorem 3 *Let $\bar{\Omega}$ be a $N \times N$ symmetric matrix (but not necessarily positive semidefinite); $\Lambda = diag(\lambda_1, \ldots, \lambda_n, \ldots, \lambda_N)$ be the matrix of eigenvalues of $\bar{\Omega}$; Q be the (corresponding) orthogonal $N \times N$ matrix of eigenvectors, such that $\bar{\Omega} = Q\Lambda Q'$; and $\tilde{\Lambda} = diag(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n, \ldots, \tilde{\lambda}_N)$ with $\tilde{\lambda}_n = \max\{0, \lambda_n\}$ $(n = 1, 2, \ldots, N)$. Then,*

$$\tilde{\Omega} = Q\tilde{\Lambda}Q'$$

is a symmetric positive semidefinite matrix which best approximates $\bar{\Omega}$, in the sense that it minimizes

$$tr[(\Omega - \bar{\Omega})(\Omega - \bar{\Omega})'] = tr[(\Omega - \bar{\Omega})^2],$$

(that is, in the sense that it minimizes the sum of squares of the $\Omega - \bar{\Omega}$ elements), where Ω represents any symmetric positive semidefinite $N \times N$ matrix.

Let us consider a subset of inputs for which all the PSD conditions are violated because there is no matrix C associated with a positive semidefinite $\bar{\Omega}$.[14] This means that all matrices $\bar{\Omega}$ defined according to both Theorems 1 and 2 have at least one negative eigenvalue. To use the suggested approximation, we should first find the matrix $\bar{\Omega}$ for which we have the smallest (in absolute terms) negative eigenvalue. Let $\bar{C}$ be the associated $N \times M$ matrix as defined in Theorem 2. Notice that by construction,

[14] In practice, this is the relevant condition. In all empirical situations that we have tested, when $\bar{\Omega}$ is tranformed into a positive semidefinite matrice $\tilde{\Omega}$ according to Theorem 3, the diagonal elements of the latter matrix are all positive.

$V(z) = DRD = \bar{\Omega} + 2^{-2/3}\Theta_3\bar{C}\bar{C}'\Theta_3$. The approximation works by replacing $\bar{\Omega}$ by $\tilde{\Omega}$, or equivalently, replacing $V(z)$ by $\tilde{V}(z) \equiv \tilde{\Omega} + 2^{-2/3}\Theta_3\bar{C}\bar{C}'\Theta_3$. Therefore, in intuitive terms, the suggested approximation implies expanding the variances of the input variables until the PSD condition becomes fulfilled. And the smaller in absolute terms the negative eigenvalues, the smaller the required expansion of variances.

A word of caution is needed. The degree of approximation must be carefully scrutinized. The standard deviations of the forecast errors, resulting from such approximation, must be compared with the corresponding "exact" standard deviations (computed using the linear combination $\alpha' z$, the standard deviations for the input variables and the correlation coefficients). The quality of the approximation will depend on the magnitude of the largest (in absolute terms) negative eigenvalue. If the ratio between the corresponding exact and approximate standard deviations is close to one, we may safely use the distribution resulting from the approximation. Otherwise, a revision of the assumptions is unavoidable.

4 Empirical Illustration

In this section, we apply both the methodology suggested above and the BoE's approach to the estimation of the densities of inflation and output for the Euro area. The purpose of this empirical illustration is simply to show how these methodologies work in practice and not to provide the reader with a realistic empirical framework to assess the uncertainty and risks of future inflation and output in the Euro area.[15]

In particular, we will not discuss the economic reasonability of the response coefficients or the historical standard deviations and correlations of the input variables.

[15]Unlike some stochastic simulation approaches, our approach can easily handle a relatively large number of conditioning and endogenous variables. We only took two endogenous variables to simplify the empirical illustration.

4.1 Data and illustrative model

Although, as argued earlier, most institutions do not rely on a single model to assess economic developments, our illustration will be directly based on a very simple econometric model to obtain the local linear approximation and the "pure" forecasting errors. We modeled the Euro area real Gross Domestic Product, $ln(Y_t)$, and the inflation rate, $\Delta ln(P_t)$, using a vector autoregressive (VAR) system, with annual data and in first differences, conditional on three variables: the effective exchange rate, $ln(EE_t)$, a commodity prices index (including oil), $ln(PC_t)$, and the extra-euro-area world output, $ln(WY_t)$.[16] One single lag was chosen after checking the traditional selection criteria. In order to illustrate the multiequational approach, some type of structural form of the model was needed, so we considered a Cholesky identification scheme, assuming that shocks to the inflation rate do not affect output contemporaneously. With this assumption and given a set of sample forecasting errors, the "pure" endogenous forecasting errors were obtained. The errors used as input variables were obtained, but for world output, under the hypothesis of random walk, i.e., one-, two- and three-periods ahead, forecasts are given by the last observed values. With regards to world output, the forecast errors were generated using a AR(2) model in first differences.[17]

In terms of the notation presented in section 2.1, we have (after reparametrizing the model): $e_{t+h} = (e_{t+h}^{ln(Y)}, e_{t+h}^{\Delta ln(P)})'$, $x_{t+h} = (ln(EE_{t+h}),$ $ln(WY_{t+h}),\ ln(PC_{t+h}))'$, and $\nu_{t+h} = (\nu_{t+h}^{ln(Y)},\ \nu_{t+h}^{\Delta ln(P)})'$. The associated matrices of response coefficients are:

$$\Gamma_0 = \begin{bmatrix} -.0259 & .8610 & .0252 \\ .0501 & -.1406 & .0426 \end{bmatrix}; \qquad \Psi_0 = \begin{bmatrix} .0089 & 0 \\ .0026 & .0067 \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} -.0293 & .0560 & .0017 \\ -.0489 & .3880 & -.0010 \end{bmatrix}; \qquad \Psi_1 = \begin{bmatrix} .0100 & -.0013 \\ .0045 & .0065 \end{bmatrix} \tag{33}$$

[16] We used the database built by Fagan, Henry and Mestre (2001) for the estimation of the European Central Bank's (ECB) Area Wide Model (AWM), available at `http://www.ecb.int/pub/wp/ecbwp042.zip`.

[17] The model was: $\Delta ln(WY_t) = 0.01 + 0.42\Delta ln(WY_{t-1}) - 0.35\Delta ln(WY_{t-2})$.

$$\Gamma_2 = \begin{bmatrix} .0045 & -.0679 & .0005 \\ -.0054 & .0037 & .0004 \end{bmatrix}; \qquad \Psi_2 = \begin{bmatrix} .0098 & -.0016 \\ .0047 & .0062 \end{bmatrix}.$$

With the pure forecasting errors and the response coefficient matrices, we can apply both our and the BoE's method of uncertainty and risk analysis. When applying our methodology, we will assume that there are two blocks of independent random variables, namely the block of errors in the conditioning variables and the block of endogenous pure forecasting errors, which in terms of the notation above corresponds to the hypothesis that $x_i \perp \nu_j$, for any (i, j). The BoE approach extends the above independence hypothesis to all variables regardless of which block they belong to and which time period. In the exercises below, we choose not to intervene these errors. However, as with the BoE approach, it is possible to incorporate judgment on the degree of uncertainty by adjusting the observed standard deviations and/or the correlations.

4.2 A limited set of risks

We start by analyzing an example which considers only downward risks in the underlying conditioning path of extra-area world output. In the first period, $t + 1$, it is assumed that there is a 60% probability that the observed world output will fall below the modal path. This probability falls then to 55% in the second period, $t + 2$. For comparison purposes, we run the exercise for each of the four methodologies suggested – BoE's, the *tpn* method of moments of section 2.4 assuming independence of the input variables, an *sgn* variant of the latter method of moments, and, finally, our alternative *sgn* aggregation methodology. The differences between these methods are reflected in the results presented in Table 1. In this table, $E_0 = M_0 = 0$ stand for the mean and mode of the baseline, respectively. They are equal due to the symmetry – no risks – implicit in the baseline assumptions. M_1 and E_1 denote, respectively, the mode and the mean of the distributions after accounting for the risks to the inputs.

First, let us notice that the relaxation of the independence assumption, columns $\perp$ and $\not\perp$, has significant impacts on some of the computed standard deviations (see, for example, second period inflation).

The set of columns labeled $M_1 - M_0$ expresses the mode effect. While under the BoE approach the mode is kept unchanged, that is not the case with our approach. For instance, in period $t+1$, the modal forecast of output will be below the baseline forecast in 0.085%. In the same period, the modal forecast of inflation is shifted upwards by 0.036%.

The first set of columns in the 'Outputs' section of the table, $E_1 - E_0$, represents what we term mean effect. It also highlights the differences due to the parametric hypotheses (*tpn* vs. *sgn*), because E_1 is the linear aggregation of the means of the underlying inputs. Typically, the *sgn* assumption translates into somewhat stronger mean effects than the *tpn* assumption.

The M_1 quantiles provide estimates of the skewness effect. Our method yields slightly less skewness than the *sgn*-based method of moments, owing to the correlations between the input variables. When comparing the BoE and our approaches, besides the differences induced by the correlations, we observe that the former overestimates the degrees of skewness. These biases do not depend on the *tpn* assumption (see *tpn*-based MM), but they are due to the assumption of unchanged modes taken by the BoE approach.

Table 1: Uncertainty and Risks: Comparing alternative methodologies

	$E_1 - E_0$		$M_1 - M_0$		St. Dev.		Quantiles $P[z \leq 0]$						
Inputs	*tpn*	*sgn*	*tpn*	*sgn*									
WY_{t+1}	-0.237	-0.262	0	0	0.73		60.0						
WY_{t+2}	-0.233	-0.277	0	0	1.43		55.0						
	$E_1 - E_0$		$M_1 - M_0$		St. Dev.		Quantiles $P[e \leq M_i]$						
	BoE	Our	BoE	Our			BoE	Method of Moments				Our	
	(tpn)	(sgn)	(tpn)	(sgn)			(tpn)	(tpn)		(sgn)		(sgn)	
Outputs					$\perp$	$\not\perp$	M_0	M_1	M_0	M_1	M_0	M_1	M_0
$ln(Y_{t+1})$	-0.204	-0.225	0	-0.085	0.79	0.81	58.2	56.1	58.7	55.2	59.2	54.4	58.9
$\Delta ln(P_{t+1})$	0.033	0.037	0	0.036	0.78	0.72	48.7	50.0	48.3	50.0	48.1	50.0	48.0
$ln(Y_{t+2})$	-0.214	-0.253	0	-0.065	1.43	1.52	54.7	53.5	55.0	53.5	55.3	53.1	54.9
$\Delta ln(P_{t+2})$	-0.059	-0.063	0	-0.062	1.23	0.98	51.5	50.1	51.9	50.2	52.0	50.0	52.5
$ln(Y_{t+3})$	0.003	0.002	0	0.002	2.94	2.83	50.0	50.0	50.0	50.0	50.0	50.0	50.0
$\Delta ln(P_{t+3})$	-0.091	-0.108	0	-0.089	1.74	1.59	51.6	50.2	52.0	50.3	52.4	50.3	52.6

Table 2: 70% Confidence Intervals

Minimum range method	BoE	sgn
$ln(Y_{t+1})$	$[-0.942,\ 0.677]$	$[-0.921,\ 0.671]$
$\Delta ln(P_{t+1})$	$[-0.787,\ 0.830]$	$[-0.712,\ 0.784]$
$ln(Y_{t+2})$	$[-1.623,\ 1.345]$	$[-1.637,\ 1.399]$
$\Delta ln(P_{t+2})$	$[-1.313,\ 1.236]$	$[-1.079,\ 0.955]$
$ln(Y_{t+3})$	$[-3.048,\ 3.052]$	$[-2.927,\ 2.931]$
$\Delta ln(P_{t+3})$	$[-1.866,\ 1.747]$	$[-1.743,\ 1.552]$

A synthetic indicator of the location-shift and asymmetry effects is given by the baseline quantile. In the "Quantiles" section of Table 1, the columns labeled 'M_0' give us this measure. In spite of different mode and skewness effects, we observe closer baseline quantiles for both methods, but always more assymetric in our approach.

Another form of looking at potential differences between our approach and the BoE's is given in Table 2, which presents 70% minimum range confidence intervals. That is, given the estimated forecast densities, it presents the ranges of the 70% "most likely" outcomes for the forecasting errors under each methodology.[18]

4.3 A larger set of risks

Let us now assume that, in the first and second forecasting periods, there are downside risks to world output and to shocks to euro-area GDP, while the price of commodities and the pure error of inflation have upward risks. The latter can be thought as resulting from factors that cannot be captured by usual determinants of inflation (risks of higher inflation induced, for example,

[18] In the sense that the range is the smallest that covers 70% of the support of the density. Notice that the baselines of all input variables are set to zero, resulting in confidence intervals that always include zero. These intervals can, however, be adjusted by a location shift to reflect the underlying forecast. For instance, if period $t+1$ inflation rate forecast is 1.5%, then the 70% confidence intervals are [0.71, 2.33] and [0.79, 2.28] under the BoE and *sgn* approaches, respectively.

Table 3: Assumed uncertainty and risks of input variables

Exog. Shocks	$E_1 - E_0$	$M_1 - M_0$	St. Dev.	$P[z \leq 0]$
$ln(EE_{t+1})$	0	0	6.23	50%
$ln(WY_{t+1})$	-0.44	0	0.73	**70%**
$ln(PC_{t+1})$	9.63	0	16.15	**30%**
$ln(EE_{t+2})$	0	0	9.07	50%
$ln(WY_{t+2})$	-0.51	0	1.43	**60%**
$ln(PC_{t+2})$	8.50	0	23.75	**40%**
$ln(EE_{t+3})$	0	0	10.50	50%
$ln(WY_{t+3})$	0	0	3.17	50%
$ln(PC_{t+3})$	0	0	32.52	50%
Endog. Shocks	$E_1 - E_0$	$M_1 - M_0$	St. Dev.	$P[z \leq 0]$
$\nu_{t+1}^{ln(Y)}$	-0.18	0	0.93	**55%**
$\nu_{t+1}^{\Delta ln(P)}$	0.13	0	0.67	**45%**
$\nu_{t+2}^{ln(Y)}$	0	0	1.43	50%
$\nu_{t+2}^{\Delta ln(P)}$	0	0	1.03	50%
$\nu_{t+3}^{ln(Y)}$	0	0	1.61	50%
$\nu_{t+3}^{\Delta ln(P)}$	0	0	1.15	50%

Table 4: Uncertainty and risks of the endogenous variables

Endog. Vars.	$E_1 - E_0$	$M_1 - M_0$	St. Dev.	$P[e \leq M_0]$		$P[e \leq M_1]$
				sgn	BoE	*sgn*
$ln(Y_{t+1})$	-0.13	0.01	0.81	53.93%	55.67%	54.72%
$\Delta ln(P_{t+1})$	0.47	0.13	0.72	25.91%	29.18%	35.42%
$ln(Y_{t+2})$	-0.19	-0.02	1.52	53.45%	54.70%	52.83%
$\Delta ln(P_{t+2})$	0.25	-0.09	0.98	44.84%	44.81%	40.59%
$ln(Y_{t+3})$	0.02	0.02	2.83	49.78%	49.77%	50.00%
$\Delta ln(P_{t+3})$	-0.20	-0.17	1.59	54.80%	53.33%	50.54%

Table 5: 70% Confidence Intervals

	Minimum range method		Percentile method	
	BoE	sgn	BoE	sgn
$ln(Y_{t+1})$	$[-0.904,\ 0.720]$	$[-0.819,\ 0.765]$	$[-0.964,\ 0.665]$	$[-0.931,\ 0.668]$
$\Delta ln(P_{t+1})$	$[-0.455,\ 1.103]$	$[-0.341,\ 0.840]$	$[-0.286,\ 1.329]$	$[-0.165,\ 1.129]$
$ln(Y_{t+2})$	$[-1.624,\ 1.344]$	$[-1.596,\ 1.454]$	$[-1.713,\ 1.261]$	$[-1.719,\ 1.341]$
$\Delta ln(P_{t+2})$	$[-1.140,\ 1.404]$	$[-0.860,\ 0.893]$	$[-1.061,\ 1.489]$	$[-0.665,\ 1.162]$
$ln(Y_{t+3})$	$[-3.036,\ 3.064]$	$[-2.913,\ 2.945]$	$[-3.027,\ 3.073]$	$[-2.913,\ 2.945]$
$\Delta ln(P_{t+3})$	$[-1.925,\ 1.685]$	$[-1.825,\ 1.465]$	$[-2.001,\ 1.612]$	$[-1.848,\ 1.442]$

by situations such as the foot-and-mouth disease). How do these risks reflect themselves in the risks to euro-area output and inflation forecasts? Table 3 describes how the asymmetry hypotheses are reflected in terms of the input variables. Table 4 summarizes the implications for the distribution of the forecasting errors of inflation and output of our hypotheses.

In comparing our results with the BoE, it is worth highligthing that whenever the distribution is close to symmetric, the two methodologies produce similar results. However, when the skewness is stronger, as in the case of consumer price inflation, results may be quite distinct. In those cases, it is less likely that the BoE's zero mode effect will approximately compensate the bias in skewness.

Table 5 reports the 70% confidence intervals[19] for the two methodologies, conforming the differences already apparent in Table 4. In particular, the confidence interval for the errors in inflation are quite different.

Tables 4 and 5 results used the approximation of last resort alluded to earlier. That is, neither of the possible combinations of risk sources was capable of yielding a PSD matrix necessary to apply either Theorem 1 or 2. As reflected in terms of the standard deviations, the adjustments are, however, not very significant, maybe with the exception of the variance expansion

[19]We report the confidence intervals using two methods of construction – minimum range and percentile –, which naturally produce different ranges, to illustrate that the choice of the method does not affect the qualitative conclusions drawn about the two aggregation methods.

Table 6: Uncertainty and risks after correlation adjustments

Endog. Vars.	$E_1 - E_0$	$M_1 - M_0$	St. Dev.	$P[e \leq M_0]$		$P[e \leq M_1]$
				sgn	BoE	*sgn*
$ln(Y_{t+1})$	-0.13	0.01	0.81	53.93%	55.67%	54.72%
$\Delta ln(P_{t+1})$	0.47	0.13	0.72	25.91%	29.18%	35.42%
$ln(Y_{t+2})$	-0.23	0.03	1.46	53.65%	54.70%	54.56%
$\Delta ln(P_{t+2})$	0.25	-0.03	1.10	44.52%	44.81%	43.45%
$ln(Y_{t+3})$	0.02	0.02	2.85	49.75%	49.77%	50.00%
$\Delta ln(P_{t+3})$	-0.20	-0.17	1.65	54.68%	53.33%	50.49%

of output in the second period, which standard deviation increased by 23% relative to the 'exact' standard deviation. The full set of adjustment factors, defined as the ratio of the expanded to the 'exact' standard deviations, preserving Table 4 order, is: {1, 1, 1.228, 1.028, 1.128, 1.008}.

If one wants to avoid having to use the last resort approximation, we need to reassess the initial assumptions. To illustrate one of the possible strategies, we chose to introduce adjustments in the correlation matrix. Given that there were no adjustments to period $t+1$ outcomes, we reduced in 50% the data correlations for all auto- and cross-correlations, with the exception of those involving only period one variables. This adjustment lies between the cases of independence and a sample-based correlation matrix. The PSD condition is still not fulfilled, but the adjustment factors are now very close to one: { 1, 1, 1.032, 1.028, 1.000, 1.002}, thereby, playing down the importance of the *sgn* approximation. Table 6 displays the results, which despite the new adjustment factors, show minor changes in skewness relatively to Table 4.[20] Therefore, the analysis drawn in Table 4 still applies.

[20] Notice that the first year results are the same due to the fact that we did not introduce judgment in the first year input variables.

5 Conclusion

In this article we criticized two aspects of the statistical methodology used by the Bank of England, the Riksbank and several other central banks to produce their fan charts. The first issue relates to a statistical shortcut taken by the BoE approach that implicitly considers that the mode of the linear combination of random variables is the (same) linear combination of the modes of those variables. The second issue deals with the assumption of independence among all covariates, which may be restrictive.

We proposed an alternative methodology to estimate the density of the conditional forecasts that addresses both issues. An illustration of our methodology was presented and its results were compared with those obtained using the BoE approach.

In our approach, when we consider risks to some input variables, both the mean and the mode of the distribution are shifted relatively to the benchmark of full symmetry. The size of the mean shift will be larger (in absolute terms) than the size of the mode shift, and the distribution of the forecasted variables will also become skewed. Therefore, in our approach the effects on the distributions of the forecasted variables, resulting from the consideration of risks to the input variables, are twofold: a "mode effect" and a "skewness effect". In the BoE approach, the mean is allowed to vary, but the mode is kept unchanged, implying that the skewness of the distribution is incorrectly measured. In some cases, this bias does not generate a significant distortion when computing confidence intervals, because it compensates the mode shift of the distribution. However, in other cases, this compensation does not materialize. Therefore, the BoE approach has the potential to distort the risk assessment. In practice, the likelihood of obtaining significant differences when comparing the output of both methods will depend on the degree of skewness of the distributions. For only slightly skewed distributions, the BoE and our approaches will generate similar outputs, but the differences may become important when the skewness is strong.

References

Bean, C. (2001), "The Formulation of Monetary Policy at the Bank of England", *Bank of England Quarterly Bulletin*, Winter 2001, 434-441.

Blix, M. & P. Sellin (1998), "Uncertainty Bands For Inflation Forecasts", *Riksbank Working Paper*, 65.

Blix, M. & P. Sellin (1999), "Inflation Forecasts With Uncertainty Intervals", *Riksbank Quarterly Review*, 2/1999, 12-28.

Blix, M. & P. Sellin (2000), "A Bivariate Distribution for Inflation and Output Forecasts, *Riksbank Working Paper*.

Britton, E., P. Fisher & J. Whitley (1998), "The Inflation Report Projections: Understanding the Fan Chart", *Bank of England Quarterly Bulletin*, 38, 30-37.

Cogley, T., S. Morozov & T. Sargent (2003), "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System", *mimeo*.

Fagan, G., J. Henry, & R. Mestre (2001), "An Area-Wide Model (AWM) for the Euro Area", *European Central Bank Working Paper*, 42.

Ihaka, R. & R. Gentleman (1996), "R: A Language For Data Analysis and Graphics", *Journal of Computational and Graphical Statistics*, 5(3), 299-314.

John, S. (1982), "The Three Parameter Two-Piece Normal Family of Distributions and Its Fitting", *Communications in Statistics - Theory and Methods*, 11(8), 879-885.

Johnson, N., S. Kotz & N. Balakrishnan (1994, 2nd Edition), "Continuous Univariate Distributions", Volume 1, Wiley.

Mardia, K. (1970), "Families of Bivariate Distributions", Griffin, London.

Robertson, J., E. Tallman & C. Whiteman (2002), "Forecasting Using Relative Entropy", Working Paper, 2002-22, University of California, Davis.

Wallis, K. (1999), "Asymmetric Density Forecasts of Inflation and the Bank of England's Fan Chart", *National Institute Economic Review*, January, 106-112.

Appendix

Proof of Theorem 1 *Let us consider* $N+1$ *independent random variables*

$$w \sim N(\mu, \Sigma) \text{ and } s \sim 2^{1/3} \exp[2^{1/3}(s + 2^{-1/3})]$$

with $\mu = B'\theta_1$, B *an orthogonal matrix and* Σ *a positive semidefinite diagonal matrix such that (for given* $\bar{\Omega} = DRD - 2^{-2/3}\theta_3\theta_3'$*)* $B\Sigma B' = \bar{\Omega}$.

(1) Let $y = Bw$. *We have*

$$y \sim N(B\mu, B\Sigma B') = N(\theta_1, \bar{\Omega})$$

In particular,

$$y_n \sim N(\theta_{1n}, \omega_{nn}),$$

where ω_{jj} *is the* j*-th diagonal element of* $\bar{\Omega}$. *But, from the expression of* $\bar{\Omega}$:

$$\omega_{nn} = \left(\sqrt{\theta_{2n}^2 + 2^{-2/3}\theta_{3n}^2}\right)^2 - 2^{-2/3}\theta_{3n}^2 = \theta_{2n}^2 > 0.$$

Therefore,

$$y_n \sim N(\theta_{1n}, \theta_{2n}^2)$$

(2) As $y_n = b_{n1}w_1 + \cdots + b_{nN}w_N$ *and* s *are independent and* $\theta_2 > 0$, *we have*

$$z_n^* = y_n + \theta_3 s \sim S(\theta_{1n}, \theta_{2n}, \theta_{3n}).$$

(3) We also need to show that the correlation matrix of z^* *is* R. *Since* $z^* = Bw + \theta_3 s$,

$$\begin{aligned} V(z^*) &= B\Sigma B' + 2^{-2/3}\theta_3\theta_3' = \bar{\Omega} + 2^{-2/3}\theta_3\theta_3' = \\ &= DRD - 2^{-2/3}\theta_3\theta_3' + 2^{-2/3}\theta_3\theta_3' = DRD \end{aligned}$$

As the elements of the diagonal of D *are the standard deviations of the corresponding elements of* z^*, R *is the correlation matrix of* z^*.

(4) As to the distribution of $\alpha' z^*$, it is straightforward to derive the first three moments. The distribution is sgn *because $\alpha' z^*$ may be expressed as the combination of two independent variables: a non-degenerate normal and a shifted exponencial with parameter $2^{1/3}$*

Q.E.D.

Proof of Theorem 2 *Since Theorem 2 is an extension of Theorem 1, the proof is very similar to the proof above.*

Proof of Theorem 3 *If Ω is symmetric positive semi-definite, then there is a matrix R $(N \times N)$ such that*

$$\Omega = R'R.$$

Let

$$\begin{aligned} f(R) &= tr[(R'R - \bar{\Omega})(R'R - \bar{\Omega})'] \\ &= tr[(R'R)^2 - 2\bar{\Omega}R'R + \bar{\Omega}^2]. \end{aligned}$$

The minimization problem may be expressed as:

$$\min_R f(R).$$

The necessary first order conditions are

$$\begin{aligned} & R(R'R - \bar{\Omega}) = 0 && (a) \\ \Leftrightarrow\ & RR'R = RQ\Lambda Q' && (b) \\ \Leftrightarrow\ & (RR')(RQ) = (RQ)\Lambda && (c) \\ \Leftrightarrow\ & (RR')\nu_n = \nu_n \lambda_n, && (d) \end{aligned} \tag{34}$$

where ν_n the n-th column of RQ.

Since

(i) As RR' is positive semi-definite by construction, it cannot have negative eigenvalues, which means that ν_n has to be null for $\lambda_n < 0$;

(ii) *For* $\lambda_n \geq 0$, *it is the case that* λ_n *is a eigenvalue of* RR' *and* $R'R$, *from where we conclude that the positive eigenvalues of* $\bar{\Omega}$ *are also eigenvalues of* RR';

(iii) *Since* Q *is non-singular, the ranks of* RR' *and* RQ *are identical. As, by (i),* ν_n *is null when* $\lambda_n < 0$, *we conclude that all non-null eigenvalues of* RR' *and* $R'R$ *are identical to the positive eigenvalues of* $\bar{\Omega}$.

Thus, we will show that if there are several solutions to the system of first order conditions, (34), *they will correspond to the same value of the objective function, which means that any solution that satisfies the first order condition is an optimal solution. As such, let* $\hat{R}$ *be such a solution of* (34). *Then,*

$$\begin{aligned} f(\hat{R}) &= tr[(\hat{R}'\hat{R})^2 - 2\bar{\Omega}\hat{R}'\hat{R} + \bar{\Omega}^2] \\ &= tr[(\hat{R}'\hat{R})^2 - 2\hat{R}'\hat{R}\hat{R}'\hat{R} + \bar{\Omega}^2] \\ &= tr(\bar{\Omega}^2) - tr[(\hat{R}'\hat{R})^2] \\ &= tr(\bar{\Omega}^2) - tr[(\hat{R}\hat{R}')^2] \\ &= \sum_{n=1}^{N} \lambda_n^2 - \sum_{\lambda_n > 0} \lambda_n^2 = \sum_{\lambda_n < 0} \lambda_n^2 \end{aligned} \tag{35}$$

To complete the demonstration, we will show that one solution to (34) *is given by*

$$\tilde{R} = GQ', \qquad \text{with } G = diag(g_1, \ldots, g_n, \ldots, g_N) \tag{36}$$

where

$$g_n = \begin{cases} 0, & \text{if } \lambda_n \leq 0 \\ \sqrt{\lambda_n}, & \text{if } \lambda_n > 0 \end{cases} \tag{37}$$

Indeed, this solution respects the first order condition ((34)(*c*)) *as*

$$[(GQ')(GQ')'](GQ')Q = (GQ')Q\Lambda \tag{38}$$

which by (36) *is equivalent to* $G^3 = G\Lambda$. *On the other hand, given that* $\tilde{R}\tilde{R}' = G^2 = \tilde{\Lambda}$, *we have*

$$\begin{aligned} f(\tilde{R}) &= tr(\bar{\Omega}^2) - tr[(\tilde{R}\tilde{R}')^2] \\ &= \sum_{n=1}^{N} \lambda_n^2 - \sum_{\lambda_n > 0} \lambda_n^2 = \sum_{\lambda_n < 0} \lambda_n^2, \end{aligned} \tag{39}$$

which respects (35)*. As such*

$$\tilde{\Omega} = \tilde{R}'\tilde{R} = QG^2Q' = Q\tilde{\Lambda}Q' \tag{40}$$

is a solution to the problem.

Q.E.D.