ASSESSING INFLATION UNCERTAINTY IN THREE TRANSITIONAL CENTRAL AND EAST EUROPEAN COUNTRIES: THE WEIGHTED SKEW NORMAL DISTRIBUTION APPROACH

WOJCIECH CHAREMZA*, CARLOS DÍAZ **
AND SVETLANA MAKAROVA***

VERY PLEMILINARY! PLEASE DO NOT QUOTE

December 2015

* University of Leicester, UK and AFiBV, Warsaw, Poland
** University of Leicester, UK
*** University College London, UK

KEYWORDS: skew normal distributions, ex-post uncertainty, inflation forecasting, monetary policy in Central and Eastern Europe,

JEL codes: E17, C46, E52, E37

ACKNOWLEDGEMENT

Financial support of the ESRC/ORA project RES-360-25-0003 *Probabilistic Approach to Assessing Macroeconomic Uncertainties* is gratefully acknowledged. This research used the ALICE High Performance Computing Facility at the University of Leicester. We are solely responsible for all remaining deficiencies.

ABSTRACT

It is conjectured that the *ex-post* inflation forecast uncertainty can be used as a proxy for macroeconomic uncertainty in countries with moderate and high inflation. We argue that the current practice of assessing macroeconomic uncertainty by a count of articles with words related to uncertainty in newspapers might not be adequate. In this study, we have computed *ex-post* inflation uncertainty for Poland, Russia and Ukraine, with U.S. as a benchmark. The data for annual inflation are monthly and cover the period from 1994-until 2014. We derive measures of inflation uncertainty from a skew normal distribution fitted to *ex-post* (pseudo out of sample) forecast errors. It is shown, by Monte Carlo experiments, that different types of skew-normal distributions fitted to data might give similar results, making identification of the true distribution difficult. We suggest the weighted skew-normal (WSN) distribution as the approximation to the distribution of *ex-post* inflation uncertainty. The advantage of this approach is that the parameters of the estimated WSN distribution are directly interpretable in the context of monetary policy. Our results show that WSN provides a good fit for Poland and the U.S. but not for Russia and Ukraine. Furthermore, the estimates of the parameters suggest the prevalence of anti-inflationary policy over the output-stimulating policy for Poland, while the opposite seems to be true for the U.S.

1. INTRODUCTION

The paper tackles some issues related to the evaluation of the *ex-post* inflation forecast uncertainty in Poland, Russia and Ukraine, for the period preceding the current conflict in Ukraine. In a less direct way, it relates inflation uncertainty, observed *ex-post*, through the characteristics of forecast errors, to a wider issue of the unobserved, in principal, macroeconomic uncertainty. In this context we take a sceptical view of the current practice of assessing macroeconomic uncertainty for the economies with moderate and high inflation by a count of articles with words related to uncertainty in a single newspaper. Such practice might be justified for countries with effective monetary policy and, currently, for the resource-importing countries. For the others, it should be applied with caution,. We argue that, in cases where inflation is reasonably high, it might be better to assess macroeconomic uncertainty from the probabilistic characteristics of the distribution of forecast errors of inflation.

2. MEASURING OF UNCERTAINTY IN POST-SOVIET COUNTRIES

Appropriate measurement of macroeconomic uncertainty has recently become a contested issue. Although the research on the relationship between the uncertainty and growth has a long history (see e.g. Knight, 1921), the measurement question was somewhat neglected. It was usually assumed that some measures of variability of GDP or prices, conditional or unconditional, observed over a reasonably long period, were giving sufficient insight into their uncertainty. And quite rightly so; if the probability of an unexpected event is reasonably high and the magnitude of outcomes of such event is substantial, the effects can by observed reasonably frequently. Hence, a variation of the outcomes would represent uncertainty, understood as the absolute magnitude of whatever we cannot forecast, regarding mean.

More recently, however, the scene for observation of outcomes of uncertainty has changed. The unexpected events happen less frequently. They might not even be observed in the historical sample. However, if they happen, the magnitude of outcome might be substantial. In Europe, an evident example is that of the Greek crisis; the default did not happen, the levels of European macroeconomic indicators like prices and GDP have not changed much in the short term, but there has been a substantial increase in trading insurance costs, shortening transactions' horizons, affecting foreign trade, *etc*., which might eventually lead to a slowdown in GDP growth. It shows that simple observation of the historical volatility of the series might not give a right picture prompting the development of more complex measures. Among these measures, the widest coverage recently has the Economic Policy Uncertainty index, EPU (Baker, Bloom and Davis 2013). The EPU index is regularly published for some countries and areas, including China and Russia. For US and some other countries it is a three-component index, consisting of the (1) frequency of newspapers' articles containing words/phrases related to uncertainty, (2) US temporary tax code provisions, often unexpectedly extended and regarded as an additional source of uncertainty and (3) some forecast disagreement measures based on surveys of professional forecasters, as forecast disagreement is often regarded as a measure of uncertainty (see e.g. Bomberger, 1999, Giordani, P. and P. Söderlind, 2003, Lahiri and Sheng, 2010). For most other countries, like China and Russia, only the first of these components is taken into account. For the US, results of implementing the index as an auxiliary tool for forecasting are generally positive and outperforms other measures like the term spread (see e.g. Karnizova and Li, 2014).

For other countries and regions, however, application of such measure, which is based only on the component (1) that is on the newspaper use of the uncertainty-related phases in their articles can be more problematic, especially in international comparison. So far, the newspaper coverage of countries like Russia and China is limited to only one newspaper in each country, *Komersant* for Russia and *South China Morning Post* published in English in Hong Kong, for China. In such cases, an obvious criticism of such measures is that, in countries with some degree of state control over the media, it can be easily manipulated. Additionally, there are problems related to the idiosyncratic sociological and cultural nature of reporting and absorbing the news, linguistic subtleties, misusing particular terms, evaluation of language over time, reporting fads and fashions, cascading (repeating the same news), and other related issues that could make both cross country and in-time comparisons unrealistic. Constructing comparable uncertainty measures that contain all three elements of the full EPU listed above for a wide spectrum of countries seems to be a formidable task. Unresolved issues relate to adjusting the measure to different tax systems, in the second component above, and problems with homogeneity and dependence regarding the 'uncertainty by disagreement' approach, which is the base for constructing the third component see *e.g*. Bowles *et al*., 2007; Andrade and Bihan, 2013; Makarova, 2014).

Concerns raised above prompted us to revert, as an alternative, to measures of uncertainty based on the probabilistic characteristics of the distribution of point forecast errors. In this paper, we concentrate on the evaluation of uncertainty using the forecast errors of inflation (*ex-post* inflation forecast uncertainty, see Clements, 2014). This might be an appropriate measure for countries with moderate to high inflation and without significant price control. Clearly, such measure might be questionable for countries with stable and low-level inflation.

For empirical investigation we have selected Poland, Russia and Ukraine. Because of the possibly distorted effect of the military conflict in Ukraine, our data sample ends in February 2014. The main reason for selecting these countries was that they differ substantially in the nature of their inflation. Poland conducts reasonably successful inflation targeting policy; Russia has experimented with the exchange rate stabilisation under conditions of natural resource economy and 'dirty float' while Ukraine's monetary policy is essentially undefined. Data on inflation are relatively homogeneous and reasonably long for these countries so that the distribution of inflation can be estimated with a high degree of accuracy. The practice of using past forecast errors for approximating forecast uncertainty has often been popular before, especially among the central banks' practitioners (see e.g. Kowalczyk, 2013; for comparison with the survey of forecasters' approach see Clements, 2014; Lahiri, Peng and Sheng, 2014). It developed from comparing the distributions of point forecast errors with the distributions used in probabilistic forecasting (Hall and Mitchell, 2007, Dowd, 2008).

For countries where more complex series of the EPU index data are available, that is based on information from more than one newspaper, there is usually a significant correlation between EPU index and a naïve measure of uncertainty based on squares of *ex-post* inflation forecast errors for periods of moderate of high inflation. For some countries, such correlation reaches 50% (see Charemza, Díaz and Makarova, 2015). If a slightly more complex measure of uncertainty based on moving standard deviation of the logs of squares of such errors is applied, its correlation with the three-components EPU index is markedly higher, exceeding, in some cases, 80% (see Charemza, Díaz and Makarova, 2014).

However, for countries where only one-component EPU index where information from one newspaper is used, like Russia and China, it does not seem to be any relation whatsoever between the EPU index and *ex-post* forecast uncertainty. This is illustrated by the simple rank correlation between the EPU data and a naïve uncertainty measure given by logs of squares of the *ex-post* ARMA 4-step ahead forecast errors for the annual inflation measured monthly for the period from July 2001-February 2013 (more details on computation of these forecast errors are given in Section 3).

Table 1: Rank correlation between log of EPU index and logs of squares of ARMA forecast errors

	July 2001-February 2013	
	rank corr. coeff	*p*-vals
U.S.	0.201	0.007
Russia	-0.078	0.189

Legend: EPU index: http://www.policyuncertainty.com; *p*-values have been obtained by bootstrapping the sample with 1,000 drawings.

Table 1 shows that for U.S., for the fully developed three-components EPU index, there is significant, albeit not high, correlation with a naïve measure of inflation uncertainty for the period before February 2013. However, for Russia, with the one-component EPU index based on data from one newspaper, there is no such correlation. A similar result of low correlation of EPU index with ARMA forecast errors for China has been obtained by Wu (2015). Figure 1compares time paths of the both measures for U.S. and Russia.

Figure 1: Logarithm of EPU index and naïve measure of inflation forecast uncertainty

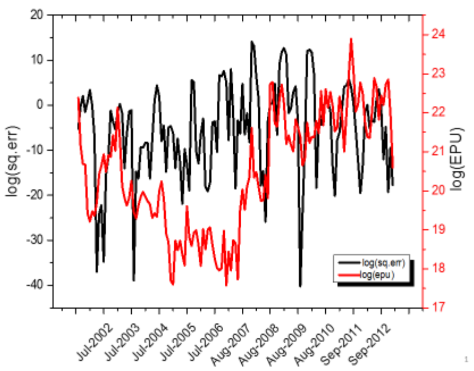
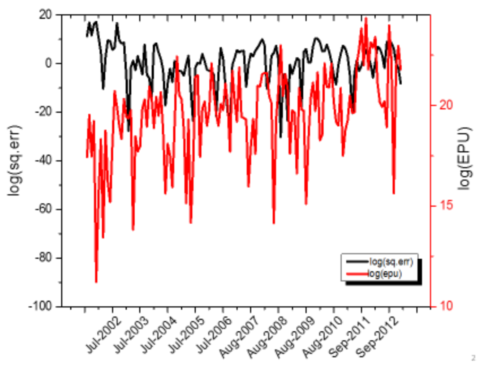
July 2001-February 2013	
U.S.	Russia
	

Figure 1 shows the deviation of the EPU index and naïve inflation forecast uncertainty for the period between the January 2003 and August 2007. For the periods before January 2003 and after August 2007 the dynamics of the both measures is much more similar.[1] For Russia, however, the picture is different. There are visible discrepancies in the dynamics of both series (EPU index has an increasing trend and naïve inflation uncertainty is constant over time. Moreover, the series of EPU index data exhibits a number of drops, suggesting substantial negative skewness, in addition to possible nonstationarity.

The above observation suggests the rationale for using *ex-post* inflation uncertainty measures as a proxy for macroeconomic uncertainty for countries with some inflation variability, in the absence of more complex and perhaps less intuitive methods. Long series of data for consumers' price index, used for computing inflation, are available for most countries and usually regarded as reasonably realistic. In the case of East European and post-Soviet economies, such series have not been manipulated, at least more than in other countries. The rationale for using ARMA forecast errors is supported by the fact that the ARMA model is

[1] For the period after February 2013, there is no positive correlation between EPU and forecast errors of inflation in the US, due to low variability of inflation.

widely regarded as one of the best methods, regarding accuracy and robustness. It was shown that it encompasses forecasts from more complex multivariate models (see e.g. Mitchell, Robertson and Wright, 2014). Even if there are other, more accurate, forecasting methods and models, ARMA inflation forecast is intuitively well understood and widely used, so that any deviations from point forecast can be understood as caused by uncertainty. They are also easily comparable between countries. Consequently, it can be conjectured that, if for countries with well-developed EPU indices there is high and significant correlation between these indices and uncertainty measures developed on the basis of inflation *ex-post* forecast errors in periods where there was some variability of inflation, it is reasonable to assume that such correlation would have also appeared for East European and post-Soviet countries, if fully developed EPU indices were available.[2]

3. SKEW NORMAL DISTRIBUTIONS AND UNCERTAINTY

Observations on the *ex-post* forecast uncertainty are the pseudo-out of sample forecast errors (see Stock and Watson, 2007). Let $\hat{\pi}_{t+h|t}$ be the baseline *h*-step ahead point forecast of inflation π_t made in time *t* for $t+h$. *Ex-post* forecast errors are computed in recursions; that is as:

$$e_{t+h|t} = \pi_{t+h} - \hat{\pi}_{t+h|t} \ , \qquad , \qquad\qquad t = t_0,\ t_0+1, \dots, T-h,$$

where:

t_0 - data used for the initial model estimation;

T - total length of data.

In each recursion the model is re-estimated, the forecast for the horizon *h* is made and forecast error computed. Then, the estimation sample is updated by one observation, model re-estimated, another forecast error computed, *etc*.. Under the assumption of independence of the subsequent forecast errors and ergodicity, the sequence of forecast errors of the length $T-h-t_0+1$ constitutes the set of observations on the *ex-post* forecast uncertainty. According to the terminology of Clements and Hendry (1998), this is the 'what we don't know that we don't know' uncertainty; the 'what we know that we don't know uncertainty can by obtained by scaling $e_{t+h|t}$ by some forecasts of conditional dispersion (e.g. GARCH standard deviations).

Once the data on the *ex-post* forecast errors have been collected, there is a problem of finding an appropriate statistical distribution that best describes the data. Although computing rudimentary measures of uncertainty from empirical distribution, without any assumptions regarding the population, is possible, it would substantially limit the applicability. Inference into the distributions of forecast errors enables to compute the probabilities of particular events, e.g. deflation or hitting inflationary targets. It also enables more complex inference regarding interrelations between uncertainties of particular indicators and between countries. Among the numerous possible distributions that might be fitted to our data we concentrate on the large family of skew-normal distributions.

[2] It should be noted that high correlation of uncertainty measures based on inflation errors is also high for other measures of macroeconomic uncertainty. For instance correlation within the range of 0.53-0.77 (depending on forecast horizon) has been obtained for moving standard deviations of squares of US inflation forecast errors and the measure of uncertainty constructed by Jurado, Ludvigson and Ng (2015); see Charemza, Díaz and Makarova, (2015).

During the last decade a substantial development of the theory and applications of such distributions, which contain normal distribution as their special symmetric case, can be observed. The first distribution of this kind applied in empirical macroeconomics was probably the so-called two-piece normal (or split normal) distribution, TPN, originated by John (1982) and developed further by Kimber (1985). It gained popularity among the practitioners; in particular it has been widely used by economic forecasters for constructing probabilistic forecasts of inflation (see e.g. seminal paper by Wallis, 1999). A further breakthrough was made by Azzalini (1985, 1986), who developed a theory of univariate, and then multivariate, skew-normal distributions. These distributions have been recently subject of substantial generalisations. Most notably, the Balakrishnan skew normal distribution has been proposed by Sharafi and Behboodian (2008), generalized Balakrishnan skew normal distribution, GBSN, by Yagedari, Gerami and Khaledi (2007), and developed further by Hasanalipour and Sharafi (2012), Fujisawa and Abe (2012), Mameli and Musio (2013), and others. For further extension and generalization see Cohoy (2015).

Such plethora of distributions provides a practitioner with a dilemma of which one to choose. Most intuitively, the distribution to be selected is the one with the best fit to the data. However, what if the distributions considered fit the data equally good (or equally bad) so that they are statistically indistinguishable?

There are three types of skew-normal distributions which we consider in this paper: (1) weighted skew normal; see Charemza, Díaz and Makarova (2014), WSN; (2) two-piece normal, TPN, and (3) the Yagedari, Gerami and Khaledi (2007) generalized Balakrishnan skew-normal distribution, GBSN. Appendix A gives their basic probabilistic characteristics and suggests an interpretation of some of the parameters. Generally, these distributions can be characterized as follows:

1. WSN is a six-parameters distribution, with parameters denoted as $\alpha, \beta, \tau_{low}, \tau_{up}, \rho, \sigma$, where $\tau_{low} < \tau_{up}$; $\alpha, \beta \in \mathbb{R}$; $\sigma^2 \in \mathbb{R}^+$; and $|\rho| \leq 1$. It becomes standard normal if $\alpha = \beta = 0$ and $\sigma = 1$. In the case where it is fitted to the inflationary forecast errors, that is to differences between observed inflation and the baseline inflation forecast, and the monetary policy makers have access to additional experts forecasts (assumed not be observable by the model maker), Charemza, Díaz and Makarova (2014) offer the following interpretation of the parameters of WSN distribution.

   α, β: measure effects of the respective anti-inflationary (α) and pro-inflationary (β) monetary policy on the uncertainty,

   τ_{low}, τ_{up}: represent policy thresholds; if additional experts' forecast is between the thresholds, no policy action is undertaken,

   σ^2: is the variance of the distribution of the uncertainty as if the monetary policy was impotent,

   ρ: is the coefficient explaining the degree of expertise (knowledge) of the additional experts.

   More details regarding the derivation and interpretation of this setting are given in Appendix A.

2. TPN is a three-parameters distribution with parameters $\sigma_1, \sigma_2 \in \mathbb{R}^+$ and $\mu \in \mathbb{R}$. It becomes standard normal if $\sigma_1 = \sigma_2 = 1$ and $\mu = 0$. Skewness of this distribution is

decided by the ratio σ_1 / σ_2. According to Wallis (2004), the deviation of TPN from normality represents the balance of risks of the overvalued and undervalued forecasts.

3. GBSN is a there parameters distribution with parameters n, m, δ, where n and m are the non-negative integers and $\delta \in \mathbb{R}$ $\mu = 0$. It becomes normal if $n = 1$ and $\delta = m = 0$ or $n = m = 0$. Its parameters do not have a direct interpretation. Hasanalipour and Sharafi (2012) showed that it fits well to the distributions of some non-economic data.

Figure 2 compares the *pdf*'s of WSN, TPN and GBSN with approximately identical first three moments, namely with means equal to 0.78, variances equal to 0.39 and the coefficient of skewness equal to 0.85. For comparison, the normal distribution with mean equal to 0.78 and variance of 0.86 is plotted in the background.

Figure 2: *pdf*'s of WSN, TPN and GBSN with identical first three moments mean= 0.78, variance=0.39, coef. of skewness=0.85.

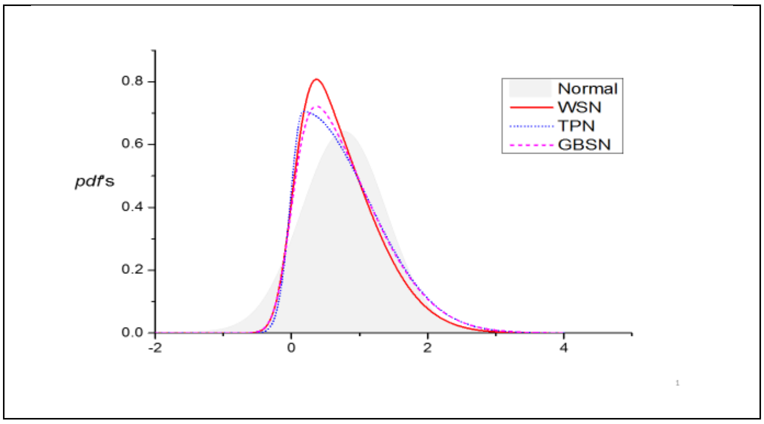
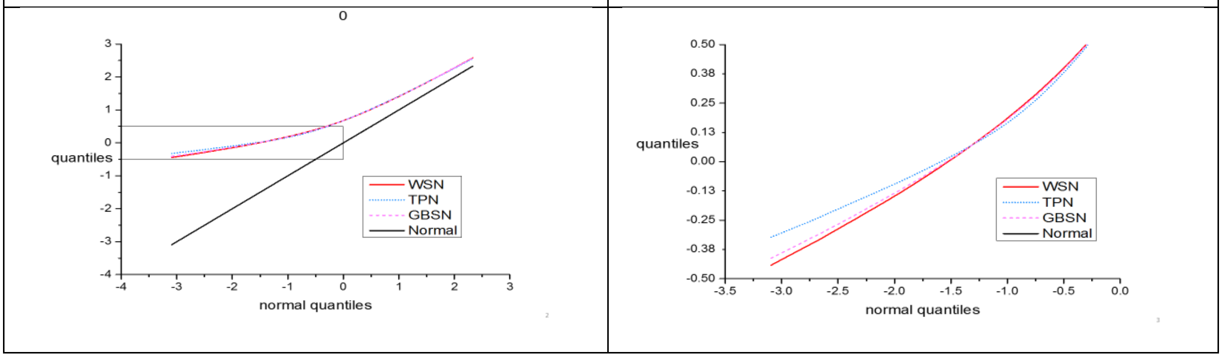


Figure 3 shows the Q-Q diagram that is a scatter diagram of the quantiles of each of the distributions depicted in Figure 1 against the normal distribution. Figure 3a gives the Q-Q plot for the entire range of quantiles, from 0.01 to 0.99, and Figure 3b gives a close-up of Figure 3a for the left tails of the distributions.

Figure 2: Q-Q plots of WSN, TPN and GBSN distributions

Fig 2a: full range of quantiles	Fig. 2b: a close-up of the left tails



Figures 2-3 illustrate potential problems in distinguishing between the distributions. The GBSN and TPN have, in the case illustrated, nearly identical modes, with TPN having a slightly thicker right tail. The corresponding quintiles are nearly identical, except the left-hand side quantiles, where the GBSN are slightly lower than the other corresponding quantiles, relatively to the identical quantile of the normal distribution. Nevertheless, these differences are small, which suggest practical problems in discrimination between skew-normal distributions.

4. FIT OF TRUE AND FALSE MODELS

To check whether using the best fit criterion for selecting the best type of a skew normal distribution might lead to choosing a false one, we have set up three data generating processes (*DGP*'s, or 'true models') and fitted all three distributions to the generated data.

The *DGP*'s are:

DGP 1: WSN with $\alpha = -2.0$, $\beta = -0.5$, $\sigma^2 = 1$, $\tau_{up} = -\tau_{low} = 1$ and $\rho = 0.75$.

DGP 2: TPN with $\sigma_1 = 1.5$, $\sigma_2 = 0.5$, $\mu = 0.4$.

DGP 3: GBSN with $n = 2$, $k = 1$ and $\delta = -0.3$.

All three *DGP*'s have similar first three moments, as given in Table 2:[3]

Table 2: Mean, st. deviation and skewness of *DGP*'s

	Mean	st. dev.	Skewness
DGP 1	-0.363	1.069	-0.628
DGP 2	-0.398	1.113	-0.695
DGP 3	-0.207	0.925	-0.687

1,000 samples have been generated for each *DGP*, and for each sample size of 100, 150, 200, 250, 300, 350, 400, 450 and 500. For each simulated sample, we have fitted all three distributions using the *SMDE* method outlined in Appendix A. Problem in comparison arises because WSN is a 6-parameters distribution and TPN and GBSN have 3 parameters each. To allow for a fair comparison, we have decided to keep the ρ parameter constant (that is, $\rho = 0.75$) and $\sigma = 1$. Also, we are keeping the threshold parameters, τ_{up} and τ_{low} constant in estimation, albeit in two different variations. In the first variation, denoted by WSN(0), we keep the thresholds fixed as in the *DGP* 1, that is $\tau_{up} = -\tau_{low} = 1$. In the second variation, we made the thresholds dependent on σ in such way that $\tau_{up} = \sigma$ and $\tau_{low} = -\tau_{up}$. We denote this as WSN(1). Hence, we are left with three parameters to estimate: α, β, and σ. In such settings the skewness in WSN is induced only through the differences between the parameters α and β. Such settings are close to that applied in the empirical models which are analysed in Section 5.

As a simple, naïve, misspecification measure, we use the frequency of cases when $d_0^{(i)} > d_1^{(i)}$, where $d_0^{(i)}$ denotes the minimum distance measure computed for the estimated properly specified distribution in the i^{th} replication, and $d_1^{(i)}$ denotes the minimum distance measure computed for one of the misspecified distribution estimated using the same generated data. That is, we do the comparison in pairs, comparing the properly specified distribution with the falsely specified one. By the properly specified distribution we understand the distribution of the same type as used for generating the sample. The distance criterion used here is the twice squared Hellinger distance, *HD* (results for other criteria are available on request; they do not differ much from these presented in this paper).

[3] Computing moments of GBSN requires numerical integration over an infinite interval. The algorithms applied here are that of Sikorski and Stenger (1984), named inthp1and inthp2 in GAUSS 13 and later versions.

Another misspecification measure is based on bootstrapping the ratios distance measures for two alternative distributions fitted to the same sample. We have used methodologies developed originally for comparing variances: simple bootstrap and Efron bootstrap (see e.g. Sun, Chernick and LaBudde, 2011).

The algorithm for the simple bootstrap is the following:

Step 1: Draw M pairs with replacement of $\{d_0^{(k)}, d_0^{(j)}\}$, $k,j = 1,\ldots, 1{,}000$, $k \neq j$, $M = 10{,}000$;

Step 2: Compute the ratio of distance measures $r_0^h = \dfrac{d_0^{(k)}}{d_j^{(k)}}$, $h = 1,2,\ldots,M$;

Step 3: Compute the 95[th] quantile of the distribution of r_0^h denoted as $q_{0.95}$;

Step 4: Check the simulated bootstrap criterion for the case where $d_0^{(i)} > d_1^{(i)}$ as:

$$\frac{d_1^{(i)}}{d_0^{(i)}} > q_{0.95} \quad .$$

The frequency of cases where the above inequality is fulfilled tells about the probability of undertaking the right decisions regarding the distribution by rejecting the wrong one. The higher is this ratio, the false distribution is chosen less often. Efron bootstrap is similar, except that in Step 1 drawing of pairs is made from $\{d_0^{(i)}, d_1^{(i)}\}$, rather than from $d_0^{(i)}$ alone. Results in this case are more robust, as the equality of the distance measures is explicit under the null.

Tables 3, 4 and 5 present respectively the naïve misspecification measure and also those based on the simple and Efron bootstraps and for twice squared Hellinger minimum distance criterion. Results for other criteria and different sample sizes are available on request.

Table 3: Frequency of cases where $d_0^{(i)} > d_1^{(i)}$

Sample size	*DGP* 1 (WSN)		*DGP* 2: (TPN)			*DGP* 3: (GBSN)		
	TPN	GBSN	WSN(1)	WSN(0)	GBSN	WSN(1)	WSN(0)	TPN
100	0.380	0.479	0.504	0.507	0.396	0.314	0.287	0.269
250	0.261	0.229	0.37	0.442	0.091	0.304	0.299	0.228
500	0.258	0.06	0.186	0.341	0.005	0.377	0.353	0.233

Table 4: Frequency of cases where $d_1^{(i)} / d_0^{(i)} > q_{0.95}$, simple bootstrap

Sample size	*DGP* 1 (WSN)		*DGP* 2: (TPN)			*DGP* 3: (GBSN)		
	TPN	GBSN	WSN(1)	WSN(0)	GBSN	WSN(1)	WSN(0)	TPN
100	0.088	0.045	0.047	0.046	0.026	0.186	0.207	0.219
250	0.135	0.099	0.073	0.046	0.049	0.201	0.191	0.267
500	0.148	0.216	0.137	0.071	0.183	0.152	0.168	0.254

Table 5: Frequency of cases where $d_1^{(i)} / d_0^{(i)} > q_{0.95}$, Efron bootstrap

Sample size	*DGP* 1 (WSN)		*DGP* 2: (TPN)			*DGP* 3: (GBSN)		
	TPN	GBSN	WSN(1)	WSN(0)	GBSN	WSN(1)	WSN(0)	TPN
100	0.085	0.069	0.045	0.038	0.074	0.071	0.093	0.086
250	0.117	0.112	0.086	0.061	0.156	0.09	0.081	0.116
500	0.131	0.161	0.124	0.103	0.174	0.076	0.082	0.113

Tables 3-5 show that results of fitting WSN and TPN to data generated from GBSN behave differently to that fitted to data generated from WSN or TPN distributions. Let us first concentrate on evaluating the misspecification in case where data are generated by WSN and TPN; it is clearly difficult to distinguish between these two distributions. Using *HD* criterion for the small sample size, it is practically impossible to find out which statistic is systematically smaller than the other, regardless of the data generating process. In particular, if data are generated from TPN, there is a virtually equal chance that WSN would fit better than the true TPN distribution. However, with the increase in sample size the frequencies of cases where the *HD* statistics for the 'true' distribution is smaller than for the 'false' one increase, suggesting the consistency of choice based on the *HD* criterion. This is confirmed by the bootstrap results. The empirical power of the tests based on the *HD* statistics is, in absolute terms, not high. Even for samples of size 500 it is not reaching 20%. In another words, it is in practice problematic to distinguish between the WSN and TPN distributions.

Nevertheless, some differences between the fits given by WSN and TPN can be observed here. Generally TPN is more often falsely well approximated by WSN, particularly in the case when $\tau_{up} = \sigma$ that is for WSN, than WSN by TPN. Also, for middle-sized samples (250 observations) chances for proper identification of WSN against TPN by rejecting the null of identical *MD* statistics are visibly higher than otherwise, albeit still small in absolute terms.

For data generated by WSN and TPN, the danger of misspecification by falsely selecting GBSN is visibly smaller. Except for small samples of data generated by WSN, *HD* statistics for GBSN are usually bigger than for two remaining distributions in this case than the corresponding WSN and TPN statistics, reducing the chance of distributional misspecification. Also, the empirical power of the *HD* ratio test rises relatively quickly with an increase in sample size exceeding, in some cases, 20% for large samples.

In contrast to WSN and TPN, data generated by GBSN exhibit different patterns. In terms of power of the bootstrap tests, they can also be easily confused with two other distributions as the power of the *HD* ratio test is low. However, the power of the test is not visibly increasing with the increase of sample size, causing doubts regarding the consistency. On the positive side, the naïve misspecification benchmark based on the differences between the *HD* statistics for the true and false distributions is less often false than in the case of data generated from WSN and TPN.

To sum up, one would expect the confusions in deciding which skew normal distribution is the best one o be relatively frequent, if the selection is based on the goodness of fit measures alone.

5. EMPIRICAL RESULTS: DISTRIBUTIONS OF INFLATION EX-POST UNCERTAINTY FOR POLAND, RUSSIA, UKRAINE AND U.S.

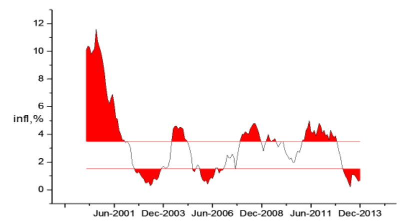
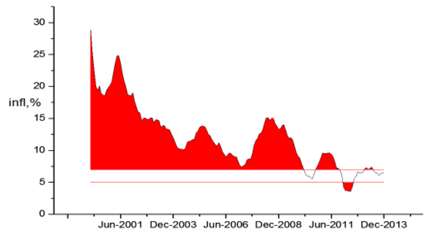
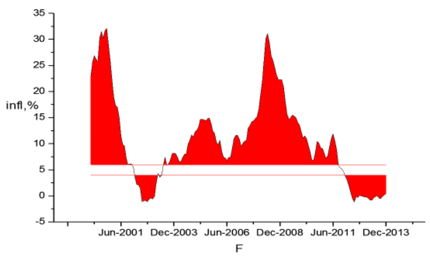
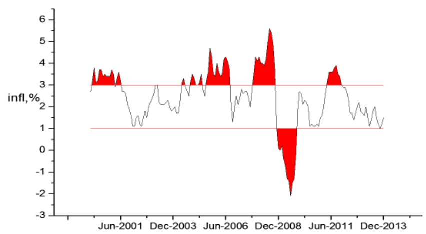
The distributions discussed above have been fitted to the *ex-post* forecasts uncertainties of three selected post-Soviet countries, of different origins and history of monetary policy, namely Poland, Russia and Ukraine. For the sake of comparison, we have also fitted these distributions to the U.S. data as well. As the benchmark test, we have used the estimated distributions for computing the probabilities that inflation is to be within a pre-defined target band. For all these countries, we have decided to use data ending in February 2014. This has been done for the sake of comparability, to avoid the effects of the Ukrainian crisis, which might distort the comparison. Last 12 data points have been retained for checking the *ex-post* forecast properties of the model, so that the distributions have been estimated using data until February 2013.

Economies of these countries differ substantially among themselves, both in term of economic growth, the average level of inflation and types of conducted monetary policy. Among these countries, only Poland conducted a reasonably successful monetary policy since 1998, with a clearly defined inflation target set at 2.5%, with ±1% band, since 2004. Russia, although officially pursuing inflation stabilization, was, in fact, targeting exchange rate stabilisation, until 2010, with some indication that it might earlier target monetary aggregates, at least until

2002 (see Esanov, Merkl and de Souza, 2004). It resulted in inflation fuelled by a ‘dirty float’ (see Vdovichenko and Voronina, 2009). Since 2010, it has been implementing inflation targeting more efficiently. For 2013, the year of the forecast, official inflation target was 5%-6%. Nevertheless, to allow for a comparison with other countries, we have evaluated the probability of inflation being within the ±1% band around its upper limit, that of around 6%. Ukraine monetary policy was the least transparent. It first had implemented the exchange rate targeting that was followed, since 2000, by the exchange rate pegging, relaxed after 2008 (International Monetary Fund,, 2011). In 2013, it announced the transition to inflation targeting, with indications that a likely target will be 5%. Consequently, we have assumed for Ukraine the target band of ±1% around 5%. For the analysis of the development of monetary policy in these three countries see e.g. Égert and MacDonald (2008). In the U.S., the nominal target of 2% has been announced by the Federal Open Market Committee in early 2012. Before that, the unofficial target was within the range of 1.7%-2%. In line with the assumptions made for other countries, for U.S. we are assuming the target of 2%, with the band of ±1%. Figure 4 shows the inflation data for all four countries with the inflation target bands depicted. It depicts the differences in inflation dynamics in the analysed countries and also the differences in the frequencies the inflation target bands for 2013 were crossed in the past.

We have computed the probabilities of inflation being within the bands using the distributions of 4-step ahead forecast errors obtained for forecasts made before February 2013, approximated by the skew normal distributions discussed above, that is WSN, TPN and GBSN.

Figure 4: Annual inflation in Poland, Russia, Ukraine and U.S., 2000-2013, and the assumed target bands for 2013

Poland	Russia
	
Ukraine	U.S.
	

After checking for the order of seasonal and non-seasonal integration by the Taylor (2003) test that takes into account the possibility of the presence of unit roots at frequencies other than tested, we have estimated the seasonal ARIMA (SARIMA) model for inflation data y_t:

$$\phi(L)\Psi(L^s)\Delta^\kappa\Delta^D y_t = \theta(L)\Theta(L^s)u_t \ ,$$

where L is the lag operator, κ is the order of integration of the regular part of y_t, $\Delta^\kappa = (1-L)^\kappa$ is the regular difference operator of order κ, $\Delta^D = (1-L^s)^D$ is the seasonal difference operator of order s for a seasonal $I(D)$ process and u_t is the error term. Polynomials ϕ, θ, Ψ and Θ are based on regular (L) and seasonal (L^s) lag operators correspondingly. Their orders have been obtained using the Gómez and Maravall (1998) procedure that is based on an automatic lag selection criterion that leads to a minimum of Ljung-Box the autocorrelation statistic. The entire data span, for the annual inflation recorded monthly in percentages, is from August 1994 (September 1994 for Russia) to February 2012.[4] Due to data availability, for Ukraine we have used a slightly shorter data span, from January 1995 to February 2012. As described in Section 3, forecasts have been computed recursively, starting from the initial period and updating the sample by one observation in each recursion. The initial (for the first recursion) period for estimation has been defined as a maximum of the first 80 observations of the series. Basic descriptive statistics of the recursive 4-step ahead uncertainty is given in Table 6. All the results given below relate to the 4-step ahead uncertainty. Results for uncertainties for other forecast horizons (up to 12) are available on request.

Table 6: Basic characteristics of 4-step ahead uncertainty

	Poland	Russia	Ukraine	U.S.
Full data span	07-1994 02-2013	08-1994 02-2013	01-1995 02-2013	07-1994 02-2013
total no. of observation in sample.	222	221	217	222
no. of obs. for estimation of the densities	138	137	133	138
mean.	-0129	-0.187	0.106	0.034
std. dev.	1.037	1.977	3.560	1.379
skewness	0.211	0.793	0.106	0.533

All three skew normal distributions discussed here, which is WSN, TPN and GBSN, have been fitted to the forecast errors data. As in Section 4, we have estimated three parameters of WSN: α, β, and σ, keeping $\rho = 0.75$, $\tau_{up} = \sigma$ and $\tau_{low} = -\tau_{up}$. Table 7 presents the results of estimation of the parameters of the density functions.

Table 7: Results of empirical estimation of different skew normal distributions

	Parameters	Poland	Russia	Ukraine	U.S.
WSN	α	-1.817 (0.1803)	-2.281 (0.1303)	-3.320 (0.1262)	-1.887 (0.1051)
	B	-1.223 (0.1792)	-2.358 (0.6400)	-3.478 (0.372)	-2.241 (0.0042)
	σ	1.000 (0.001)	0.999 (0.0.002)	0.999 (0.004)	0.851 (0.0290)
	HD	2.629	21.05	47.42	1.19

[4] Data are from the official statistical agencies of each country, available at http://www.tradingeconomics.com

	Parameters	Poland	Russia	Ukraine	U.S.
TPN	σ_1	1.169 (0.1101)	1.400 (0.617)	1.995 (0.0144)	0.902 (0.0533)
	σ_2	0.8422 (0.367)	2.000 (0.001)	2.000 (0.001)	1.306 (0.0712)
	μ	0.088 (0.482)	-0.7222 (0.4990)	-0.153 (0.275)	-0.336 (0.303)
	HD	3.10	1.11	51.52	6.26
GBSN	*N*	5	5	5	5
	M	2	1	5	2
	δ	-0.061 (0.0510)	-0.084 (0.0110)	0.000 (0.09800)	-0.976009 (0.0229)
	HD	4.71	72.80	251.7	8.90

As in the previous section, for each distribution three parameters have been estimated by the *SMDE* (see Appendix A for the details). As parameters *m* and *n* of the GBSN distribution are integers, their standard errors have not been computed. For the non-integer parameters, standard errors are given in brackets below the estimates.

The distance measure criterion suggests the choice of different distributions for particular countries. For Poland the best fit is that of TPN, followed closely by WSN, and for U.S.A the best fit is that by WSN. As concluded in Section 4, there is a high chance of distributional misspecification between WSN and TPN. With this is mind and taking into account that, for Polish data, differences in *MD*'s in fitting WSN and TPN are not negligible, we can interpret parameters of WSN in the light of monetary policy outcomes. For Poland, the positive difference between the absolute values of the estimates of α and β in WSN indicates footprints of the prevalence of anti-inflationary policy over the output-stimulating policy. For U.S., however, where such difference is negative, there is some evidence of signs of output-stimulating policy. For Russia, TPN gives the best fit, and for Ukraine all minimum distance statistics are rather large, suggesting poor fit of all the distributions considered.

In order to discriminate between the distributions further, we have tested which of these distributions fits better to the observed data with the use of the probability integral transform (*pit*) test (see Diebold, Gunther and Tai, 1998; for application to evaluation of inflation probabilistic forecast see Clemens, 2004, and Galbraith and van Norden, 2012); for other similar approaches and applications to inflation modelling see Mitchell and Hall (2005). The probability integral transform is defined as the probability of observing values of a random variable not greater than its realized value. If the forecasted density is close enough to the true but unknown density, *pit's* will be uniform on the interval from zero to one. If several *pit*'s (that is, for different forecasts) are available, one can test their accuracy by checking whether their values are uniformly distributed using well-known 'goodness-of-fit' tests. Figure 4 give the scatter diagram of *pit*'s for all three distributions and countries analysed and Table 7 gives the results of the Cramer-von Mises test for uniformity of *pit*'s.[5]

[5] Another test often used for evaluating the uniformity of pit's is that of Berkovitz (2001). Hovever, we have decided not to use it, as this test is proved to be biased in evaluation of multi-step forecasts (Dowd, 2007).

Figure 4: *pit*’s for fitted skew normal distributions

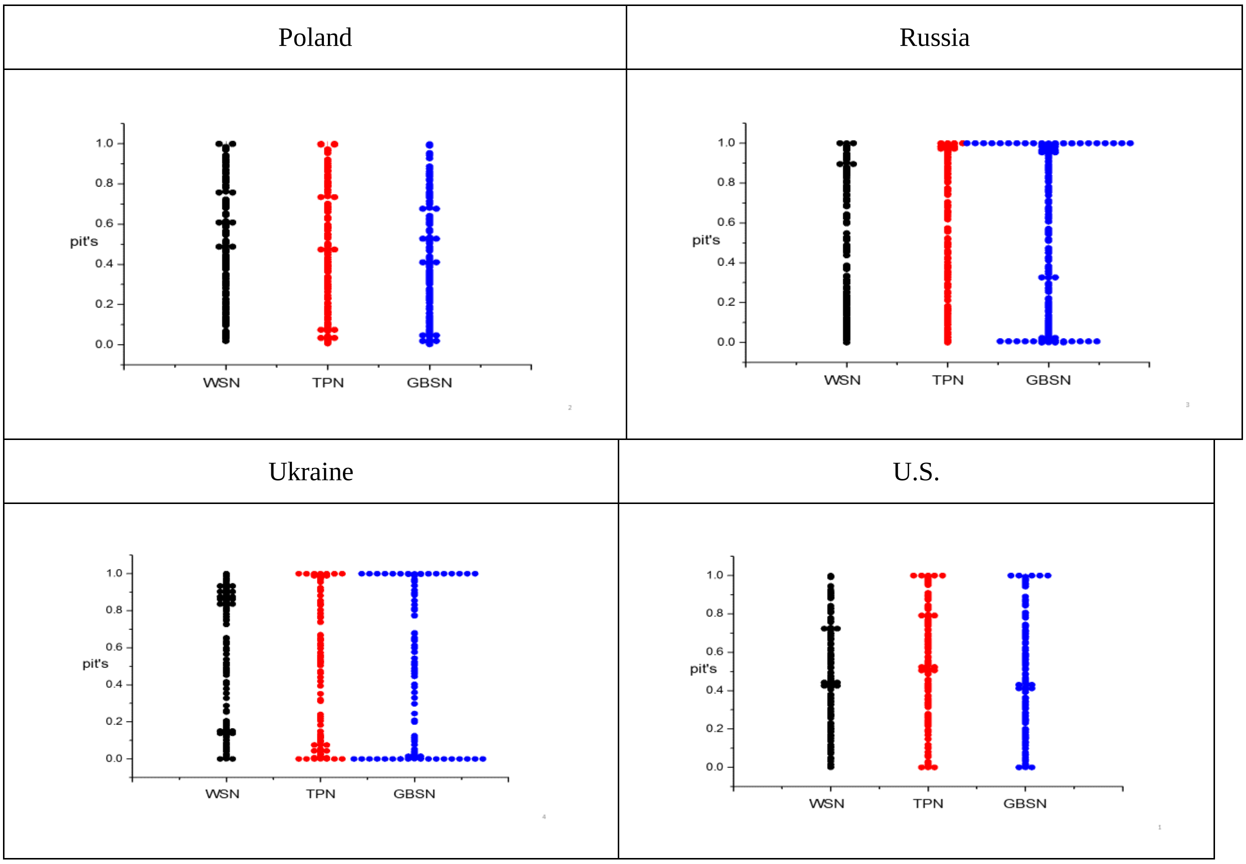


Table 8: Test statistic: Cramer-von Mises statistics for testing uniformity
significant statistics are marked by * for 10% significance and ** for 5% significance

	WSN	TPN	GBSN
Poland	0.034	0.100	0.611**
Russia	0.671**	0.328	2.492**
Ukraine	0.739**	1.649**	2.633**
U.S.	0.095	0.279	0.068

Both Figure 4 and Table 8 show that all distributions give a better fit for the Poland and U.S.A rather than for Russia and Ukraine. In particular, the uniformity of GBSN estimates is questionable for last two countries, with *pit*’s concentrated close to zero and one.

Finally, Table 9 gives the probabilities that inflation, in July 2013, that is four month's periods after the end of the sample data, will be within the target intervals. They were obtained by computing numerical integrals of the respective estimated *pdf*’s over the target interval. It also presents *pit*’s of the realisations of headline inflation in July 2013 (for the interpretation of such *pit*’s in relation to forecast uncertainty see Rossi and Sekhposyan, 2014). They describe the probability of observing values of the random variable not greater than the observed headline inflation.

Table 9: Probabilities of hitting inflation target bands, headline inflation in July 2013 and its *p*-values

	Target interval, %	Infl. in July 2013	WSN		TPN		GBSN	
			(1)	(2)	(1)	(2)	(1)	(2)
Poland	1.5 - 3.5	1.1	0.60	0.25	0.59	0.24	0.63	0.18
Russia	5 – 7	6.5	0.43	0.38	0.79	0.56	0.43	0.26
Ukraine	4 – 6	0.0	0.07	0.34	0.40	0.21	0.00	0.32
U.S.	1 - 3	2.0	0.60	0.56	0.62	0.64	0.68	0.56

Legend: (1): probabilities of inflation being within the target in June 2013 according to the particular distribution;
(2): *pit*'s of observed headline annual inflation in June 2013 according to the particular distribution.

As expected, Table 9 does not show much difference between the estimated probabilities of hitting the inflation target, especially for Poland and U.S., where all examined skew normal distributions fit well. More substantial differences can be noticed for Russia, where the results for TPN, which is the only distribution for which the null hypothesis of the uniformity of *pit*'s is not rejected, shows different probabilities of hitting inflation target and *pit* for the observed inflation distinctively. A positive conclusion that can be drawn here is that more than one skew normal distribution might fit well to the data, and it might not matter much which one is used. However, only the parameters of the estimated WSN distribution can be sensibly interpreted in the context of monetary policy.

6. CONCLUSIONS

Results of the paper indicate striking similarities between the overall efficiency of monetary policy and fit of skew-normal distributions to the *ex-post* forecast uncertainty data. Although no direct comparison is available, it might be prudent to assume that, at least because of an absence of drastic monetary policy changes, monetary policy efficiency, monetary policy efficiency of U.S. and Poland can be classified as similar, as it is in the case of Poland and the West of Europe (see Jarociński, 2010). For these countries, the fit of the skew-normal distributions is much better than that for Russia and Ukraine. Although, for these countries, it is difficult to decide between the two-piece normal and weighted skew normal distributions, the estimates of the latter provide a clear interpretation of the imprints of monetary policy on the uncertainty. For Russia and Ukraine, however, not much can be read from the estimates.

Regarding the statistical side of the results, the general message is somewhat pessimistic. We showed that it might be difficult to tell one skew normal distribution from another by the best fit, especially if the sample size is not very large. As the number of potential skew normal candidates for fitting to data is substantial (especially in the light of the fact that there are other propositions in the literature not considered in this paper) it seems to be sensible to decide on the type of distribution not by the best fit but rather on the basis of interpretation of its parameters, especially if there is not much difference in the closeness of the fit of the competing distributions.

It is worth noting that the difficulty in deciding on the type of skew normal distribution is deepened by the fact that there are no operational statistics developed for testing the degree of disparities between distance measures (or other characteristics) of these distributions. The bootstrap procedure used in this paper suggests a way for further investigation.

APPENDIX A:

THREE SKEW NORMAL DISTRIBUTIONS, INTERPRETATION AND ESTIMATION

WEIGHTED SKEW NORMAL DISTRIBUTION (WSN)

The random variable Z with WSN distribution, as defined by Charemza, Díaz and Makarova, (2014), is described by:

$$Z = X + \alpha \cdot Y \cdot I_{Y>\tau_{up}} + \beta \cdot Y \cdot I_{Y<\tau_{low}} \quad , \tag{A1}$$

where:

$$I_{Y>\tau_{up}} = \begin{cases} 1 & \text{if } Y > \tau_{up} \\ 0 & \text{otherwise} \end{cases} , \quad I_{Y<\tau_{low}} = \begin{cases} 1 & \text{if } Y < \tau_{low} \\ 0 & \text{otherwise} \end{cases} , \quad (X,Y) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \right),$$

$\tau_{low} < \tau_{up}$; $\alpha, \beta \in \mathbb{R}$; $\sigma^2 \in \mathbb{R}^+$; and $|\rho| \leq 1$.

Z is a random variable described fully by six parameters: $\alpha, \beta, \tau_{low}, \tau_{up}, \rho, \sigma$. It is shown in Charemza, Díaz and Makarova (2015), that the probability density function (*pdf*) of the standard weighted skew normal distribution, that is for $\sigma = 1$, is given by:

$$f_{\mathrm{WSN}_1}(t) = \frac{1}{\sqrt{A_\alpha}} \varphi\left(\frac{t}{\sqrt{A_\alpha}} \right) \Phi\left(\frac{B_\alpha t - mA_\alpha}{\sqrt{A_\alpha (1-\rho^2)}} \right) + \frac{1}{\sqrt{A_\beta}} \varphi\left(\frac{t}{\sqrt{A_\beta}} \right) \Phi\left(\frac{-B_\beta t + kA_\beta}{\sqrt{A_\beta (1-\rho^2)}} \right)$$
$$+ \varphi(t) \cdot \left[\Phi\left(\frac{m - \rho t}{\sqrt{1-\rho^2}} \right) - \Phi\left(\frac{k - \rho t}{\sqrt{1-\rho^2}} \right) \right] \quad ,$$

where φ and Φ denote respectively the density and cumulative distribution functions of the standard normal distribution, and $A_\tau = A(\tau) = 1 + 2\tau\rho + \tau^2$, $B_\tau = B(\tau) = \tau + \rho$.

Suppose that the WSN is fitted to a set of inflation forecast errors of the same horizon. It is convenient to assume that the baseline (point) forecast can be in turn improved by second stage forecast, given by *Y* in (A1). Under these settings the parameters of WSN have the following interpretation:

(i) α and β represent the marginal effect of the stimulative and contractionary economic policies on uncertainty respectively;

(ii) τ_{low} and τ_{up} represent the thresholds deciding about the relevance of the second stage forecast information for the policy decisions;

(iii) σ^2 that is variance of *X* and *Y*, represent the uncertainty related to the second stage forecast information used for improving the baseline forecast outcome;

(iv) ρ, that is the correlation coefficient between *X* and *Y*, describes the accuracy of the second stage forecasts.

TWO PIECE NORMAL DISTRIBUTION (TPN)

A random variable with TPN distribution is defined by its *pdf*:

$$f_{TPN}(t; \sigma_1, \sigma_2, \mu) = \begin{cases} A \exp\left\{ -(t-\mu)^2 / 2\sigma_1^2 \right\} & \text{if } t \leq \mu \\ A \exp\left\{ -(t-\mu)^2 / 2\sigma_2^2 \right\} & \text{if } t > \mu \end{cases} , \quad t \in \mathbb{R} \quad ,$$

where $A = \sqrt{2/\pi} \cdot (\sigma_1 + \sigma_2)^{-1}$. Three parameters to be estimated are $\sigma_1, \sigma_2 \in \mathbb{R}^+$ and $\mu \in \mathbb{R}$.

It is often interpreted in the context of forecast uncertainty as a representation of the balance of risks the over-and underestimated forecasts (see Wallis, 2004).

GENERALIZED SKEWED BALAKRISHNAN DISTRIBUTION (GBSN)

The third distribution considered here, the GBSN, is given by the following *pdf*:

$$f_{GBSN}(t;n,m,\delta) = \frac{1}{C(n,m,\delta)}[\Phi(\delta t)]^n[1-\Phi(\delta t)]^m \varphi(t) \quad , \quad t \in \mathbb{R} \quad ,$$

where $C(n,m,\delta) = \sum_{i=0}^{m} \binom{m}{i} (-1)^i \int_{-\infty}^{\infty} [\Phi(\delta t)]^{n+i} \varphi(t) dt$, Φ and φ are respectively the *cdf* and *pdf* of the standard normal distribution and the non-negative integers *n* and *m* and $\delta \in \mathbb{R}$ are the parameters. The GSBN includes the Balakrishnan skew normal distribution for *m* = 0, and the original Azzalini skew normal distribution (with the probability density function $f_{SN}(t;\lambda) = 2\varphi(t)\Phi(\lambda t)$) for *n* = 1 and *m* = 0). Azzalini distribution is also a special case of the WSN for, $\lambda = -2\rho$, $\tau_{up} = \beta = 0$ and $\sigma^2 = 1$. All three distributions can be reduced to a standard normal: WSN for $\alpha = \beta = 0$ and $\sigma^2 = 1$; TPN for $\sigma_1 = \sigma_2 = 1$ and $\mu = 0$; GBSN for $n = 1$ and $\delta = m = 0$ or $n = m = 0$. So far, the parameters of the GBSN distribution have not been given any particular interpretation.

ESTIMATION AND GENERAL SETTINGS

Estimation of WSN, TPN and GBSN distributions by the maximum likelihood or the generalized method of moments is numerically awkward. This problem is particularly well discussed for the Azzalini distribution (see e.g. Azzalini and Capitanio, 1999, Sartori, 2006, Franceschini and Loperfido, 2014), and is evident also for all three families of distributions considered here. For this reason we have resorted to simulation-based estimation methods. These methods are particularly attractive as it is straightforward to derive random number generators for all three distributions. For WSN given by **Error! Reference source not found.** it is described in Charemza, Díaz and Makarova (2014), for TPN in Nakatsuma (2003) and for GBSN in Yagedari, Gerami and Khaledi (2007). With the use of these generators and inspired by Greco (2011) we have applied the simulated minimum distance estimators method (*SMDE*, see Charemza *et al*., 2012), which consists of fitting the approximated by simulation density function to empirical histograms of data and applying a minimum distance criterion.

The version of *SMDE* applied here can be defined as:

$$\hat{\omega}_n^{SMDE} = \arg\min_{\omega \in \Omega} \left\{ \mu_w \left(d(g_n \, , \, f_{t,\omega}) \right)_{t=1}^{T} \right\} \quad ,$$

where $f_{t,\omega}$ is the approximation of the *pdf* , f_ω, of a random variable obtained by generating *t* = 1,…,*T* replications (drawings) from a distribution with parameters ω ($\omega \in \Omega \subset \mathbb{R}^k$), g_n denotes the density of empirical sample of size *n*, μ_w is an operator based on *T* replications, which deals with the problem of the 'noisy' criterion function (median, in this case), and $d(\bullet,\bullet)$ is the distance measure. The minimum distance measures, *MD*, applied here are that of the Cressie and Read (1984) power divergence disparities family given by:

$$d(g_n\ ,\ f_{t,\omega}) = \frac{1}{\lambda_{CR}(\lambda_{CR}+1)} \sum_{i=1}^{m+1} g_n(i) \left[\left(\frac{g_n(i)}{f_{t,\omega}(i)} \right)^{\lambda_{CR}} - 1 \right] \quad , \qquad \text{(A2)}$$

where *m* denotes the number of cells in which data are organized. For $\lambda_{CR} = 1$, (A2) gives the Pearson χ^2 measure, for $\lambda_{CR} = -1/2$ the (twice squared) Hellinger distance (*HD*) and for $\lambda_{CR} = -2$ the Neyman χ^2 measure. For $\lambda_{CR} \to 0$ and $\lambda_{CR} \to -1$ the continuous limits of the right-hand side expression in **Error! Reference source not found.** are respectively the likelihood disparity (*LD*) and the Kullback-Leibler divergence statistics. Cressie and Read (1984) advocate optimal setting $\lambda_{CR} = 3/2$. For a detailed discussion and alternatives see Basu, Shioya and Park (2011). More details and the properties of the MSD are discussed in Charemza et al (2012).

References

Andrade, P. and H. Le Bihan (2013), 'Inattentive professional forecasters' *Journal of Monetary Economics* **60**, 967-982.

Azzalini, A., (1985), 'A class of distribution which includes the normal ones', *Scandinavian Journal of Statistics* **12**, 171-178.

Azzalini, A. (1986), 'Further results on a class of distributions which includes the normal ones', *Statistica*, **46**, pp. 199-208.

Azzalini, A. and A. Capitanio (1999), 'Statistical applications of the multivariate skew normal distributions', *Journal of Royal Statistical Society Series B* **61**, 579-602.

Baker, S. R., N. Bloom and S.J. Davis (2013), 'Measuring economic policy uncertainty', Stanford University and University of Chicago Booth School of Business, http://cep.lse.ac.uk/seminarpapers/04-06-13-SD.pdf.

Basu, A., H. Shioya and C. Park (2011), *Statistical inference: the minimum distance approach*, CRC Press.

Berkowitz, J. (2001), 'Testing density forecasts, with applications to risk management' *Journal of Business and Economic Statistics* **19**, 465-474.

Bomberger, W. A. (1999), 'Disagreement and uncertainty', *Journal of Money, Credit and Banking* **31**, 273-276.

Bowles, C., R. Fritz, V. Genre, G. Kenny, A. Meyer and T. Rautanen (2007), 'The ECB Survey of Professional Forecasters: a review after eight years' experience', Occasional paper No 59, European Central Bank.

Charemza, W., Z. Fan, S. Makarova and Y.Wang (2012), 'Simulated minimum distance estimators in macroeconomics', paper presented at the conference *Computational and Financial Econometrics*, Oviedo.

Charemza, W., C. Díaz Vela and S. Makarova (2014), 'Term structure of inflation forecast uncertainties and skew normal distributions', University of Leicester, Department of Economics, Working Paper No. 14/01.

Clements, M. P. (2004), 'Evaluating the Bank of England density forecasts of inflation', *The Economic Journal* **114**, 844-866.

Clements, M.P. (2014), 'Forecast uncertainty-ex ante and ex post: U.S. inflation and output growth', *Journal of Business and Economic Statistics* **32**, 206-216.

Clements, M.P. and D. Hendry (1998), *Forecasting economic time series*. Cambridge University Press.

Cohoy, D.O. (2015), 'Some skew-symmetric distributions which include the bimodal ones', *Communications in Statistics – Theory and Methods* **44**, 554-563.

Cressie, N. and T.R.C. Read (1984), 'Multinomial goodness-of-fit tests', *Journal of Royal Statistical Society Series B*, **46**, 440-464.

Diebold, F.X., T.A. Gunther and A. S. Tay (1998), 'Evaluating density forecasts: with applications to financial risk management', *International Economic Review* **39**, 863–83

Dominicy, Y. and D. Veredas (2013), 'The method of simulated quantiles', *Journal of Econometrics* **172**, 233-247.

Dowd, K. (2008), 'The Swedish inflation fan charts: an evaluation of the Riksbank's inflation density forecasts', *The Journal of Risk Model Validation* **2**, 73-87.

Dowd, K. (2007), 'Backtesting the RPIX inflation fan charts', CRIS Discussion Paper Series, 2007.V

Égert, B. and R. MacDonald (2008), 'Monetary transmission mechanism in Central and Eastern Europe: surveying the surveyable', OECD Economics Department Working Paper No. 654.

Esanov, A., C Merkl and L.V de Souza (2004), 'A preliminary evaluation of monetary policy rules for Russia', Kiel Institute for World Economics, Working paper No. 1201.

Franceschini, C. and N. Loperfido (2014), 'Testing for normality when the sampled distribution is extended skew-normal', in (M. Corazza and C. Pizzi, eds.): *Mathematical and statistical methods for actuarial sciences and finance,* Springer.

Fujisawa, H. and T. Abe (21012), 'A family of skew-unimodal distributions with mode-invariance through transformation of scale', paper presented at the CFE 2012 conference, Oviedo.

Galbraith, J.W. and S. van Norden (2012), 'Assessing gross domestic product and inflation probability forecasts derived from Bank of England fan charts', *Journal of the Royal Statistical Society, Series A* **175**, 713-727.

Giordani, P. and P. Söderlind (2003), 'Inflation forecast uncertainty', *European Economic Review* **47**, 1037-1059.

Gómez, V. and A. Maravall (1998), 'Automatic modelling methods for univariate series', Banco de España-Servicio de Estudios. Documento de trabajo no. 9808.

Greco, L. (2011), 'Minimum Hellinger distance based inference for scalar skew-normal and skew-t distributions', *Test* **20**, 120-137.

Hall, S.G. and J. Mitchell (2007), 'Combining density forecasts', *International Journal of Forecasting* **23**, **1-13.**

Hasanalipour, P., and M. Sharafi (2012), 'A new generalized Balakrishnan skew-normal distribution', *Statistical Papers* **53**, 219-228.

International Monetary Fund (2011), 'Ukraine: ex-post evaluation of exceptional access under the 2008 stand-by agreement', IMF Country Report No 11/325.

Jarociński, M., (2010), 'Responses to monetary policy shocks in the East and the West of Europe: a comparison', *Journal of Applied Econometrics* **25**, 833-868.

John, S. (1982), 'The three parameter two-piece normal family of distributions and its fitting', *Communications in Statistics: Theory and Methods* **14**, 235-245.

Jurado, K., S. C. Ludvigson and S. Ng (2015), 'Measuring uncertainty', *American Economic Review* **105**, 1177-1216.

Kimber, A.C. (1985), 'Methods for the two-piece normal distribution', *Communications in Statistics: Theory and Methods* **14**, 235-245.

Knight, F.H. (1921), *Risk, uncertainty and profit*, Sentry Press.

Kowalczyk, H. (2013), ‘Inflation fan charts and different dimensions of uncertainty’, National Bank of Poland Working Paper No. 157.

Lahiri, K., H. Peng and X. Sheng (2014), ‘Measuring uncertainty of a combined forecast’, presented at the workshop ‘Uncertainty and Economic Forecasting’, UCL, London.

Lahiri, K. and X. Sheng (2010), ‘Measuring forecast uncertainty by disagreement: the missing link’, *Journal of Applied Econometrics* **25**, 514-538.

Mameli, V. and M. Musio (2013), ‘A generalization of the skew-normal distribution: the beta skew-normal’, *Communications in Statistics – Theory and Methods* **42**, 2229-2244.

Mitchell, J., D. Robertson and S. Wright (2014), ‘What univariate models can tell us about multivariate macroeconomic models’, Warwick Business School, www.bbk.ac.uk/ems/faculty/wright/MRW14.12.2014.pdf .

Monti, A.C. (2003), ‘A note on estimation of the skew normal and the skew exponential power distributions’, *METRON – International Journal of Statistics* **20**, 205-219.

Nakatsuma, T. (2003), ‘Bayesian analysis of two-piece normal regression models’, presented at Joint Statistical Meeting, San Francisco.

Rossi, B. and T. Sekhposyan (2014), ‘Alternative test for correct specification of conditional predictive densities’, paper presented at the *Seminar on Economic Modelling and Forecasting*, Warwick Business School.

Sartori, N. (2006), ‘Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions’, *Journal of Statistical Planning and Inference* **136**, 4259-4275.

Sharafi, M. and J. Behboodian (2008), ‘The Balakrishnan skew-normal density’, *Statistical Papers* **49**, 769-778.

Sikorski, K. and F. Stenger (1984), ‘Optimal quadratures in H_p spaces’, ACM Transactions on Mathematical Software 10, 140-151.

Stock, J.H. and M.W. Watson (2007), ‘Why has U.S. inflation become harder to forecast?’, *Journal of Money, Credit and Banking* **39** (Supplement), 1-33.

Sun, J., Chernick, M. and A. LaBudde (2011), ‘A bootstrap test for comparing two variances: simulation of size and power in small samples’, *Journal of Biopharmaceutical Statistics* **21**, 1079-1093.

Taylor, A.M.R. (2003), ‘Robust stationarity tests in seasonal time series processes’, *Journal of Business and Economic Statistics* **21**, 156-163.

Vdovichenko, A. and V. Voronina (2009), ‘Monetary policy rules and their application in Russia’, EERC Working Paper No 04/09, Economic Research Network, Russia and CIS.

Wallis, K.F. (1999), ‘Asymmetric density forecasts of inflation and the Bank of England’s fan chart’, *National Institute Economic Review*, January, 106-112.

Wallis, K.F. (2004), ‘An assessment of Bank of England and National Institute inflation forecasts uncertainties’, *National Institute Economic Review* **189**, 64-71.

Yadegari, I., A. Gerami and M.J. Khaledi (2007), ‘A generalization of the Balakrishnan skew-normal distribution’, *Statistical and Probability Letters* **78**, 1165-1167.