University of Leicester

DEPARTMENT OF ECONOMICS

Too many skew normal distributions? The practitioner's perspective

Wojciech Charemza, University of Leicester, UK
Carlos Díaz Vela, University of Leicester, UK
Svetlana Makarova, UCL, SSEES, UK

Working Paper No. 13/07
April 2013

TOO MANY SKEW NORMAL DISTRIBUTIONS? THE PRACTITIONEER'S PERSPECTIVE

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

April 2013

* University of Leicester, UK
** University of Leicester, UK
*** University College London, UK

KEYWORDS: skew normal distribution, simulated minimum distance estimators, inflation uncertainties, monetary policy in Eastern Europe

JEL codes: C15, 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

The paper tackles the issue of possible misspecification in fitting skew normal distributions to empirical data. It is shown, through numerical experiments, that it is easy to choose a distribution which is different from this which actually generated the sample, if the minimum distance criterion is used. It is suggested that, in case of similar values of distance measures obtained for different distributions, the choice should be made on the grounds of parameters' interpretation rather than the goodness of fit. This is supported by empirical evidence of fitting different skew normal distributions to the estimated monthly inflation uncertainties for Belarus, Poland, Russia and Ukraine.

1. INTRODUCTION

During the last decade a substantial development of the theory of skew normal distributions, that is distributions which contain normal distribution as the special symmetric case, can be observed. The first distribution of this kind 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 substantial popularity among the practitioners; in particular it has been widely used by economic forecasters for describing uncertainties of the probabilistic forecasts of inflation (for a current review see e.g. Kowalczyk, 2012). Further breakthrough was made by Azzalini (1985, 1986), who developed an elegant theory of univariate, and then multivariate, skew normal distributions. These distributions have been recently subject of substantial generalisations. Most notable, 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 Mameli and Musio (2011), Hasanalipour and Sharafi (2012), Fujisawa and Abe (2012) and others.

These distributions, albeit fairly general and elegant, provide the potential user with three practical problems: (*i*) estimation, (*ii*) interpretation of the parameters and (*iii*) possible distributional misidentification. In this paper the problems (*i*) and (*ii*) are considered only indirectly. Regarding (*i*), the identification and numerical problems have been discussed in a number of papers, e.g. in Pewsey (2000), Monti (2003) and Castro, San Martín and Arellano-Valle (2008). Problem (*ii*) can be tackled by developing skew normal distributions with parameters directly related to the particular theory or the phenomenon described. In particular, in Charemza, Díaz and Makarova (2013) we have proposed a skew normal distribution, called weighted skew normal distribution, WSN, which parameters are directly interpretable in the context of macroeconomic density forecasting under inflation targeting. The current paper deals predominantly with (*iii*), that is the possibility of distributional misspecification. After overcoming (or skipping) problems (*i*) and (*ii*), a practitioner faces a dilemma of choosing from a plethora of different skew normal specifications. It seems to be natural that the researcher would choose that one which fits the best to the data. And here the old problem arises: is the distribution which fits to the data in the best way really the true one?

We tackle (*iii*) by putting three skew normal distributions mentioned above, that is TPN, GBSN and WSN, to the goodness of fit contest. In Section 2 we give brief description of the distributions we are considering. Section 3 explains general settings and estimation procedure. Section 4 presents the results of a Monte Carlo study evaluating the probabilities of choosing a wrongly specified skew normal distribution on the basis of its fit. Section 5 shows empirical results of estimation skew normal distributions, for the one-step ahead forecasts errors of monthly inflation in Belarus, Poland, Russia and Ukraine. Section 6 concludes.

2. THREE SKEW NORMAL DISTRIBUTIONS

There are three distributions which we consider in this paper: weighted skew normal, WSN (which we regard as the benchmark one), two-piece normal, TPN, and the Yagedari, Gerami and Khaledi (2007) generalized Balakrishnan skew normal distribution, GBSN.

A random variable Z with WSN distribution is defined by Charemza, Díaz and Makarova, (2013) as:

$$Z = X + \alpha \cdot Y \cdot I_{Y>\tau_{up}} + \beta \cdot Y \cdot I_{Y<\tau_{low}} \quad , \tag{1}$$

where:

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

$\tau_{low} < \tau_{up}$; $\alpha, \beta, \tau_{low}, \tau_{up}, \mu_X, \mu_Y \in \mathbb{R}$, $\sigma_X, \sigma_Y \in \mathbb{R}^+$, and $|\rho| \le 1$.

It is suggested that the parameters α and β are, in the economic context, related to the strength of the 'corrective' monetary policy (it is sensible to restrict them as non-positive under some additional assumptions considered below). The density and moment-generating functions and the main moments of WSN are given in Charemza, Díaz and Makarova (2013).

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

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

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 $\mu \in \mathbb{R}$.

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; n and m are non-negative integers 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 for $n = 1$ and $m = 0$. Azzalini distribution is also a special case of the WSN for $\alpha = -2\rho$, $\tau_{up} = \beta = 0$, $\sigma_x = \sigma_y = 1$ and $\mu_X = \mu_Y = 0$. All three distributions can be reduced to a standard normal: WSN for $\alpha = \beta = \mu_X = 0$ and $\sigma_X = 1$; TPN for $\sigma_1 = \sigma_2 = 1$ and $\mu = 0$; GBSN for $n = 1$ and $\delta = m = 0$ or $n = m = 0$.

3. ESTIMATION AND GENERAL SETTINGS

As mentioned above, 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), and is evident also for all three families of distributions considered here. However, it is straightforward to derive random number generators for all three distributions. For WSN given by (1) it is described in Charemza, Díaz and Makarova (2013), 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 algorithm requires conducting an iterative grid search over the pre-defined range of admissible parameters.

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 aggregation 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 distance measures, *MD*, applied here are that of the Cressie and Read (1984) power divergence disparities family given by:

$$d(g_n \, , \, f_{t,\omega}) = \frac{1}{\lambda_{CR}(\lambda_{CR}+1)} \sum_{i=1}^{m+1} g_n(i) \left[\left(\frac{g_n(i)}{f_{t,\omega}(i)} \right)^{\lambda_{CR}} - 1 \right], \tag{2}$$

where *m* denotes the number of cells in which data are organized. For $\lambda_{CR} = 1$ formula (2) gives the Pearson χ^2 measure, for $\lambda_{CR} = -1/2$ the Hellinger twice squared 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 (2) are respectively the likelihood disparity (*LD*) and the Kullback-Leibler divergence statistics. Cressie and Read (1984) advocate optimal setting $\lambda_{CR} = 3/2$.[1]

4. FIT OF TRUE AND FALSE MODELS

As the main objective of this paper is to decide whether using the best fit criterion for selecting type of a skew normal model might lead to distributional misspecification, we have set up three data generating processes (*DGP*'s, or 'true models') and fitted all three models to data generated by each of them.

The *DGP*'s are:

DGP 1: WSN with $\alpha = -2.0$, $\beta = -0.5$, $\mu_X = \mu_Y = 0$, $\sigma_X = \sigma_Y = \sigma = 1$, $\tau_{up} = -\tau_{low} = 1$ and $\rho = 0.75$. As in this paper we intend to compare three-parameter distributions only, we are keeping $\mu_X = \mu_Y = 0$ and $\tau_{up} = -\tau_{low} = 1$ constant, so that we are hence left with three parameters to estimate: α, β, and σ .

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 1:[2]

Table 1: 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

For each *DGP*, and for sample sizes of 100, 150, 200, 250, 300, 350, 400, 450 and 500, there have been generated *Nrepl* = 1,000 replications. For each simulated sample we have fitted all three distributions using the *SMDE* method outlined in Section 3. We apply TPN, GBSN and two variations of WSN. In the first variant, denoted by WSN(0) we keep the thresholds fixed

1 For a complex discussion and alternatives see Basu, Shioya and Park (2011).

2 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.

as in the *DGP* 1, that is $\tau_{up} = -\tau_{low} = 1$. In the second variant, we made the thresholds dependent on σ in such way that $\sigma\tau_{up} = -\sigma\tau_{low} = 1$. We denote this as WSN(1).

As a simple, naïve, misspecification measure, we use the frequency of cases when $d_0(\xi^i) > d_1(\xi^i)$, where d_0 denotes the minimum distance measure computed for the estimated properly specified distribution in the i^{th} replication ξ^i, and d_1 denotes the minimum distance measure computed for the estimated misspecified distribution in the same sample. 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 Hellinger distance, *HD* (results for other criteria are available on request; they do not differ much from these presented in this paper).

Another misspecification measure is based on the bootstrapping the ratios of two alternative distance measures obtained for 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 simple bootstrap is the following:

Step 1: Draw *M* pairs of $d_0(\xi^k)$, $d_0(\xi^j)$, *k*,*j* = 1,…*Nrepl*, *k* ≠ *j*. *M* should be large, e.g. 10,000;

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

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

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

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

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. It approximates the probability of rejecting the null hypothesis that the distance measures for the true and false distributions are identical with the implicit alternative that the distribution on which $d_1(\xi^i)$ is based is false. Efron bootstrap is similar, except that in Step 1 drawing is made from the set of all $d_0(\xi^k)$, $d_1(\xi^k)$ rather than from $d_0(\xi^k)$ alone. Results in this case are more robust, as the equality of the distance measures is explicit under the null.

Tables 2, 3 and 4 present respectively the naïve misspecification measure and also these based on the simple and Efron bootstraps. Results for other criteria and for different sample sizes are available on request.

Table 2 Frequency of cases where $d_0(\xi^i) > d_1(\xi^i)$

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 3 Simple simulated bootstrap power

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 4 Efron simulated bootstrap power

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 2-4 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 when data are generated by WSN and TPN; it is clearly difficult to distinguish between these two distributions. For the small sample size it is practically haphazard to find out which statistic is smaller 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 MD statistics for the ‘true’ distribution is smaller than for the ‘false’ one increase, suggesting the consistency of choice based on the MD criterion. This is confirmed by the bootstrap results. In Tables 3a, 3b and 4a, 4b frequencies of the rejection of the null that the MD statistics are identical increase with the increase in sample size. Nevertheless, the empirical power of the tests based on the MD 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 WSN(1), than WSN by TPN. Also, for middle-sized samples (150-350 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. It is also worth noting that the differences between particular MD criteria, in terms of power, and frequencies of the false choice based on the minimum of competing statistics, are meaningless.

For data generated by WSN and TPN, the danger of misspecification by falsely fitting GBSN is visibly smaller. Except for small samples of data generated by WSN, MD 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 MD ratio test rises relatively quickly with the 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 are also be easily confused with two other distributions, as the power of the MD 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 misspecfication benchmark based on the differences between the MD statistics for the true and false distributions is less often false than in the case of data generated from WSN and TPN.

5. EMPIRICAL RESULTS: ASSESSING INFLATION UNCERTAINTY IN BELARUS, POLAND, RUSSIA AND UKRAINE

The distributions discussed above have been used for modelling short-run inflation variability uncertainties, approximated by one-step ahead forecast error. As it is discussed in Charemza, Díaz and Makarova (2013), possible skewness of such uncertainties is caused by monetary policy asymmetries, characterised by the thresholds in the short-run (past-independent) inflation forecast and effectiveness of the anti-inflationary and output-stimulating policies. In the context of WSN, the thresholds are represented by τ_{up} and τ_{low}, and monetary policy effectiveness respectively by α (for the anti-inflationary policy) and β (for the output-stimulating policy).

The four East and Central European countries studied here for the period from January 1995 to December 2012, Belarus, Poland, Russia and Ukraine, represent different types and practices of monetary policy. Poland, for the period under investigation, conducted inflation targeting policy. For Belarus and Russia the targets have been less clear. For Belarus it was predominantly the currency stability, although recently a policy of inflation stabilisation has been announced. For Russia the target was, formally, inflation, for most of the period under study, but practically stabilisation of the exchange rate. For both Belarus and Russia which relies on exporting (in case of Belarus, re-exporting) natural resources, it lead to appreciation pressures and 'dirty float' inflationary effects. In Ukraine the targets and instruments have been usually multiple and loosely defined, with the emphasis on controlling bank liquidities, periods of nominal anchoring the currency to the US dollar, direct commercial banking supervision, etc. Russia and Ukraine and, to a lesser extent, Belarus, have been affected by the Russian currency crisis in 1998.

The data used here are on monthly inflation, not de-seasonalised, with 223 observations per country.[3] After checking for the order of seasonal and non-seasonal integration by the Taylor (2003) test which takes into account the possibility of the presence of unit roots at frequencies other than tested, we have identified the variability uncertainty as u_t in the seasonal ARMA (SARMA) model:

$$\phi(B)\Phi(B^s)\Delta^{\kappa}\Delta^{D}y_t = \theta(B)\Theta(B^s)u_t \quad ,$$

where B is the lag operator, $\Delta^{\kappa} = (1-B)^{\kappa}$ is the regular difference operator, κ is the order of integration of the regular part of y_t, $\Delta^{D} = (1-B^s)^{D}$ is the seasonal difference operator for a seasonal I(D) process, $\phi(B) = (1-\phi_1 B - \cdots - \phi_p B^p)$ is the polynomial of order p in the lag operator B and similarly, the seasonal AR operator is defined as $\Phi(B^s) = (1-\Phi_1 B^s - \cdots - \Phi_P B^{sP})$. Regular, $\theta(B)$, and seasonal, $\Theta(B^s)$, moving average polynomials are defined similarly with their orders denoted by q and Q respectively. The orders p, P, q and Q have been are obtained using the Gómez and Maravall (1998) procedure which is based on an automatic lag selection criterion that minimises the Bayesian Information Criteria (BIC) of the residuals. The algorithm applied here is equivalent to the well-known TRAMO-SEATS and X-11 adjustment methods.

---

[3] Data used for computations are from: http://belstat.gov.by/homep/en/specst/price3.htm; http://www.stat.gov.pl/gus/5840_1638_ENG_HTML.htm ; http://stats.oecd.org/ ; http://ukrstat.org/en/operativ/operativ2006/ct/cn_rik/isc/isc_e/isc_m_e.htm

Basic characteristics of the estimates of uncertainties: standard deviations, coefficients of skewness and *p*-values of Box-Pierce (*BP*) portmanteau autocorrelation statistics are given in Table 5.

Table 5: basic characteristics of empirical distributions of uncertainties

	Belarus	Poland	Russia	Ukraine
std. dev.	0.0184	0.0035	0.0209	0.0098
skewness	0.2033	0.0479	6.1632	0.4141
BP (*p*-val)	1.000	0.7869	0.9932	0.2800

Table 6 presents the empirical results. As in the previous section, for each distribution three parameters have been estimated by the *SMDE*. In case of WSN, the range of selection of other parameters: τ_{up} , τ_{low} and ρ has ben sparse, with only few values searched. For this reason, we do not report standard errors for these parameters. For other, non-integer, parameters, standard errors are given in brackets below the estimates.

Table 6: Results of empirical estimation of skew normal distributions

		Belarus	Poland	Russia	Ukraine
	parameters				
WSN(0)	α	-0.8169 (0.4526)	-3.997 (0.00854)	-3.624 (0.6491)	-3.729 (0.3502)
	β	-0.7226 (0.7507)	-3.713 (0.09892)	-3.548 (0.397)	-1.765 (1.532)
	σ	0.03355 (0.4028)	0.001027 (0.000798)	1.875 (0.3758)	0.009011 (0.4804)
	τ_{up}	0	0	0	1
	τ_{low}	0	0	0	-1
	ρ	0.9	0.5	0.9	0.9
	MD	46.47	8.693	55.35	4.611
TPN	σ_1	0.0712 (0.3098)	0.0117 (0.0053)	0.3550 (0.0840)	0.0652 (0.1656)
	σ_2	0.1520 (0.3060)	0.0149 (0.0154)	0.1492 (0.3148)	0.0944 (0.2243)
	μ	0.2408 (1.5152)	-1.1874 (0.9721)	1.3737 (0.9685)	-1.9402 (0.1795)
	MD	14.7413	9.1105	52.2228	17.7917
GBSN	*n*	20.00	16.00	19.00	19.00
	m	4.00	5.000	4.000	4.00
	δ	-2.586 (0.03822)	1.908 (0.8716)	-2.558 (0.1176)	-2.558 (0.1176)
	MD	18.76	27.90	15.56	15.56

The distance measure criterion suggests the choice of different distributions for particular countries. For Belarus, the best fit is that of TPN, closely followed by GBSN, for Poland and Ukraine the best is WSN, albeit for Poland the difference with respect to TPN is slight. For Russia, the best fitted distribution is that of GBSN. These differences can be explained by different types of monetary policy in each country. It is worthwhile to note that there is no systematic relationship between the absolute level of inflation uncertainties, measured by standard deviations, and the type of best-fitted distribution.

For the countries where WSN fits best, that is Poland and Ukraine, differences in the parameters estimated reflects the changes in monetary policy. For Poland, where there are inflation target bands, the thresholds are non-existing and the estimates of α and β are very low and close to each other, reflecting the near symmetricity of the uncertainties, which are, in turn, related to the balance between anti-inflationary and output-stimulating policies imposed by the inflation target bands of the inflationary targeting. For Ukraine, the distance between the estimated thresholds is substantial and the estimated α is markedly lower than the estimated β. This indicates the preference (or better effectiveness) of the anti-inflationary policy over the output-stimulating one.

6. CONCLUSIONS

The general message from this paper is somewhat pessimistic. It might be difficult to tell one skew normal distribution from another on the basis of 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 on the basis of the best fit but rather on the basis of interpretation of its parameters. In the context of inflation uncertainties in countries conducting consistent and reasonably tight monetary policy, the weighted skew normal distribution seems to be a sensible choice. For modelling other phenomena, different distributions can be of a better use.

The difficulty in deciding on the type of skew normal distribution are 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 offers some hope in this respect. However, further studies are needed here.

References

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.

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

Castro, L. M, E. San Martín and R. B. Arellano-Valle (2008), ‘A note on the identification of skewed-normal distributions’, Department of Statistics, Pontificia Universitad de Conceptión, Chile.

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 (2013), ‘Fan charts, inflation targeting and skew normal distributions’, University of Leicester, Department of Economics, Working Paper No. 13/06.

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

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

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.

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.

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

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

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

Kowalczyk, H. (2012), ‘Fan charts of inflation and different dimensions of uncertainty’ (in Polish: original title ‘Wykresy wachlarzowe inflacji a różne wymiary niepewności), Materuały I Studia Nr. 273, National Bank of Poland.

Mameli, V. and M. Musio (2011), ‘A generalization of the skew-normal distribution: the beta skew-normal’, Departamento di Matematica, University of Caligari.

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.

Pewsey, A. (2000), ‘Problems of inference for Azzalini’s skew-normal distribution’, *Journal of Applied Statistics* **27**, 859-870.

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.

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 Statistocs* **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.

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