Central European Journal of Economic Modelling and Econometrics

Parametric Modelling of Income Distribution in Central and Eastern Europe

Michał Brzeziński*

Submitted: 4.10.2013, Accepted: 25.11.2013

Abstract

This paper models income distribution in four Central and Eastern European (CEE) countries (the Czech Republic, Hungary, Poland and the Slovak Republic) in 1990s and 2000s using parametric models of income distribution. In particular, we use the generalized beta distribution of the second kind (GB2), which has been found in the previous literature to give an excellent fit to income distributions across time and countries. We have found that for Poland and Hungary, the GB2 model fits the data better than its nested alternatives (the Dagum and Singh-Maddala distributions). However, for Czech Republic and Slovak Republic the Dagum model is as good as the GB2 and may be preferred due to its simpler functional form. The paper also found that the tails of parametric income distribution in the Czech Republic, Poland and the Slovak Republic have become fatter in the course of transformation to market economy, which provides evidence for growing income bi-polarization in these societies. Statistical inference on changes in income inequality based on parametric Lorenz dominance suggests that, independently of inequality index used, income inequality in the Czech Republic, Poland and the Slovak Republic has increased during transformation. For Hungary, there is no Lorenz dominance and conclusions about the direction of changes in income inequality depend on the cardinal inequality measure used.

Keywords: generalized beta of the second kind (GB2) distribution, parametric modelling, income distribution, Lorenz dominance, Central and Eastern Europe

JEL Classification: C46, D31, P36

*University of Warsaw; e-mail: mbrzezinski@wne.uw.edu.pl

1 Introduction

Parametric statistical models have been used to model income distributions since the times of Vilfredo Pareto (1897). The models applied in the distributional literature have grown in complexity. After the one-parameter Pareto model, the two-parameter models such as the log-normal model (Gibrat 1931), the gamma (Salem and Mount 1974), and the Weibull model (Bartels and Van Metele 1975) were introduced. In the mid-1970s, the three-parameter models appeared, such as the generalized gamma (Taille 1981), Singh-Maddala (Singh and Maddala 1976) and Dagum (Dagum 1977). In 1984, McDonald (1984) introduced the four-parameter models known as the generalized beta of the first and second kind (GB1 and GB2). The GB1 and GB2 models include all of the previously mentioned distributions as special or limiting cases. Parker (1999) has presented a theoretical model in which firm optimizing behaviour under uncertainty leads to wages that follow a GB2 distribution. Empirically, it was shown that the GB2 distribution fits income distribution data better than the alternative models that it encompasses (the Singh-Maddala, Dagum, generalized gamma, log-normal and Weibull) (Bordley et al. 1996, Bandourian et al. 2003, Dastrup et al. 2007, McDonald and Ransom 2008). McDonald and Xu (1995) have proposed a 5-parameter generalized beta (GB) distribution, which encompasses both GB1 and GB2 distributions. However, empirically this distribution does not seem to improve the fit to data. This was also confirmed in our empirical experiments (not reported). Kleiber and Kotz (2003, p. 232) called the GB distribution "a curious theoretical generalization".

Using parametric models of income distribution is associated with several advantages. Fitting parametric models allows one to represent the entire income distribution through means of a small number of estimated parameters (Brachman *et al.* 1996). The estimated parameters may be then used to reconstruct the entire income distribution, if, for example, income distribution data released in future are published in grouped form (Hajargasht *et al.* 2012) or if available micro data are censored or "top coded" (Burkahuser *et al.* 2012). This kind of reconstruction can be also achieved with the help of a reliable parametric model, when for a given income distribution only empirical estimates of poverty and inequality measures are available (as published for example by the Eurostat or other statistical agency), with no direct access to the underlying micro-data (Graf and Nedyalkova 2013). In addition, a reliable parametric model can be used for poverty and inequality analysis in computable general equilibrium micro-simulation models (Boccanfuso et al. 2013).

The parameters of theoretical models often possess also economic interpretation, which allows, for example, to gain insights about the causes of the evolution of income distribution over time or interpret the differences between income distributions across countries. Moreover, once a given parametric model is fitted to a data set, one can straightforwardly compute inequality and poverty measures, which are analytical functions of the parameters of the model. It is also possible to use estimated parameters to perform stochastic dominance testing (Kleiber and Kotz 2003), which allows for robust inference on inequality and welfare differences between distributions. Finally, estimated parameters may be used in empirical modelling of the impact of macroeconomic conditions (e.g. GDP growth, unemployment and inflation rates, etc.) on the evolution of the personal income distribution (Jäntti and Jenkins 2010).

The present paper models income distribution in four Central and Eastern European (CEE) countries (the Czech Republic, Hungary, Poland and the Slovak Republic) using parametric models of income distribution. In particular, we use the GB2 distribution as it has been found in the previous literature to give an excellent fit to income distributions across time and countries. We perform goodness-of-fit and model selection tests to verify if the GB2 model is a better fit to CEE data than the simpler models (the Singh-Maddala and Dagum) that it encompasses. We also compute inequality indices and perform statistical dominance tests using fitted GB2 models to evaluate changes in income inequality in CEE countries in the period of economic transformation to market economy. Moreover, we analyze and interpret economically the evolution of the GB2 parameters estimates over time.

The paper is related to the previous empirical literature on parametric modelling of income distribution in CEE countries. Kordos (1990) argued that the two-parameter log-normal distribution reasonably describes Polish data on wages until 1980. The log-normal model has been also found to be fitted well to the income distribution of the Polish poor in 2003 and 2006 by Jagielski and Kutner (2010). These authors also found that the income distribution of the middle class and the rich is fitted well by the Pareto model. Domanski and Jedrzejczak (2002) have compared several parametric models (the Dagum, Singh-Maddala, gamma and lognormal) using data on Polish wages in 1990s. They found that the Dagum model best described their data. Lukasiewicz and Orlowski (2004) compared the Dagum and Singh-Maddala models for the distribution of individual incomes in Poland in 2000. The Dagum model gave a slightly better fit to data in their study. Dastrup et al. (2007) provided an extensive comparison of parametric models of income distribution for several countries (including Poland as the only CEE country) roughly in the period from 1980s to 1990s and using several "income" concepts: gross (pre-tax and pre-transfer) household income, disposable (post-tax and post-transfer) household income and earnings. The data used were in grouped format. The authors found that in general the GB2 model gives the best fit to Polish data for each of the income definition used. In particular, the GB2 model seemed to describe Polish data better than its nested alternatives (Dagum and Singh-Maddala), although the differences between these models were not always statistically significant.

Bandourian et al. (2003) provided a comparison of parametric models of income distribution for 23 countries (including Poland, Czech Republic, Hungary and Slovak Republic) in the period from 1970s to the mid-1990s. The main income concept used in gross (pre-tax and pre-transfer) household income, grouped in twenty equal probability intervals. In the context of CEE countries, the results of Bandourian's et al. (2003) study suggest that for Czech Republic in 1992 and 1996, Hungary in

1991 and Poland in 1985, 1992 and 1995, the GB2 model gives the best fit. However, the advantage of the GB2 over alternatives is only statistically significant for Czech Republic in 1992 and Poland in 1986. For Slovak Republic the GB1 has a small advantage over the GB2, but the difference is not statistically significant.
Most of the existing studies on parametric modeling of income distributions suffer from some limitations. Many of them use rather grouped data (data in the form of income classes or income proportions) than individual income data. Other studies do not include newer models like the GB2 distribution, or do not test rigorously for goodness of fit or model selection. The present paper removes these drawbacks by using individual income data and by applying rigorous statistical methods to the GB2 model and its closest rivals.
The paper is structured as follows. The next Section presents the definition and statistical properties of the GB2 model, while Section 3 describes statistical methods used for parametric estimation, goodness-of-fit and model selection testing, as well as tools for testing for stochastic dominance with parametric models. Section 4 introduces the data used. Empirical results and discussion follow in Section 5. The last section concludes.

2 The GB2 distribution – definition and properties

The four-parameter (a, b, p, q) GB2 model was introduced by McDonald (1984). The probability density function for the model takes the form:

$$f(x; a, b, p, q) = \frac{ax^{ap-1}}{b^{ap}B(p, q)\left[1 + (x/b)^{a}\right]^{p+q}}, x > 0, \tag{1}$$

where $B(u,v) = \Gamma(u)\ \Gamma(v)/\Gamma(u+v)$ is the Beta function, and $\Gamma(.)$ is the Gamma function. All four parameters are positive with b being the scale parameter and a, p and q being the shape parameters. The a parameter governs the overall shape of the distribution, while p and q affect the shape of, respectively, the left and the right tail. In particular, the larger the value of a, the thinner the both tails of the GB2 density (Kleiber and Kotz 2003). The larger the value of p, the thinner the left tail and the larger the value of q, the thinner the right tail. Therefore, the smaller values of ap and aq increase density at the, respectively, lower and upper tail. When both ap and aq decrease simultaneously, both tails of the GB2 become fatter. In economic terms, this can be interpreted as an evidence in favour of larger income bi-polarization. The concept of polarization, which is related to but different from inequality, aims at capturing separation or distance between clustered groups in a distribution (Esteban and Ray 1994, 2011, Foster and Wolfson 2010). For the GB2 model, we may interpret the simultaneous decrease in the estimates of ap and aq as growing bi-polarization in the sense of tighter clustering around two income poles – the poor and the rich.
The relative values of p and q affect the skewness of the GB2 distribution. The

cumulative distribution function (cdf) of the GB2 distribution does not have an explicit form as it involves an infinite series, but it can be approximated using functions implemented in most of popular statistical packages (see, e.g., Jenkins 2007, Graf and Nedyalkova 2012).

The often used in the income distribution literature three-parameter models of Singh-Maddala and Dagum are the special cases of the GB2 model. In particular, the Singh-Maddala model is the GB2 model with $p = 1$, while the Dagum model is the GB2 model with $q = 1$. Also, the log-normal model can be obtained from the GB2 model assuming that q goes to infinity and a goes to 0, see McDonald and Xu (1995) for a full characterization of families of distributions nested within the GB1 and GB2 models.

The moment of order k (existing for $ap < k < aq$) for the GB2 is defined as follows:

$$E\left(X^k\right) = \frac{b^k B(p + \frac{k}{a}, q - \frac{k}{a})}{B(p, q)}. \tag{2}$$

Parametric modelling of income distributions is often performed in order to make inferences about income inequality. For this purpose, one can use cardinal inequality indices such as the most popular Gini index of inequality (for a review of various inequality measures, see, e.g., Cowell 2000) or one can test for Lorenz dominance, which provides an unambiguous ranking of distribution in terms of their inequality. The relationship of Lorenz dominance is based on the concept of the Lorenz curve (see, e.g., Kleiber 2008), which is a plot of the cumulative income shares against cumulative population shares, with units (e.g., individuals, households) ordered in ascending order of income. If the Lorenz curve for a distribution y_1 lies nowhere below and at least somewhere above the Lorenz curve of the distribution y_2, then y_1 Lorenz dominates y_2. It is worth noting here that the popular Gini index of inequality is equal to the twice the area between the Lorenz curve and the 45% degree line of perfect equality. Any inequality index satisfying popular axioms like anonymity and the Pigou-Dalton transfer principle will in this case display less inequality for the distribution y_1 than for y_2 (Atkinson 1970).

For the GB2 model and its nested models, the relationship between model parameters and popular inequality indices is complex. McDonald (1984) has derived the analytical formula for the Gini coefficient of the GB2, which, however, takes a rather complicated form:

$$\begin{aligned} G \quad = \quad & \frac{2B\left(2p + \frac{1}{a}, \quad 2q - \frac{1}{a}\right)}{pB\left(p, q\right) B\left(p + \frac{1}{a}, \quad q - \frac{1}{a}\right)} \cdot \\ & \cdot \left\{ \tfrac{1}{p} {}_3F_2\left[1, p + q, 2p + \tfrac{1}{a};; p + 1, 2\left(p + q\right);; 1\right] \right. \\ & \left. - \tfrac{1}{p + \frac{1}{a}} {}_3F_2\left[1, p + q, 2p + \tfrac{1}{a};; p + \tfrac{1}{a} + 1, 2\left(p + q\right);; 1\right] \right\}. \end{aligned} \tag{3}$$

The generalized hypergeometric function ${}_3F_2$involves an infinite series and present computational difficulties. For the purposes of the present paper, the Gini coefficient for the GB2 distribution has been implemented in Stata using an algorithm for computing the ${}_3F_2$ function proposed by Wimp (1981).
The Gini index of inequality is most sensitive to income differences around the mode of distribution and therefore is it not suitable to detecting distributional changes that occur in the bottom or in the top of distribution. For this purpose, a family of distribution-sensitive generalized entropy inequality measures GE(γ) has been designed (Shorrocks 1984). The more positive parameter γ is, the more sensitive GE(γ) is to income differences at the top of the distribution; the more negative it is, the more sensitive is GE(γ) to income differences at the bottom of the distribution. The most popular members of the GE family include the mean logarithmic deviation, GE(0), the Theil index, GE(1) and the half the square of the coefficient of variation, GE(2). In this paper, we are especially interested in the GE(2) inequality measure, as it has been shown that inequality measures are particularly sensitive to the presence of extremely large income observations (Cowell and Flachaire 2007). Generalized entropy inequality measures for the GB2 distribution have been recently derived by Jenkins (2009). The GE(2) index for the GB2 model takes the form:

$$GE\left(2\right)=-\frac{1}{2}+\frac{\Gamma\left(\mathrm{p}\right)\Gamma\left(\mathrm{q}\right)\Gamma\left(\mathrm{p}+\frac{2}{\mathrm{a}}\right)\Gamma\left(\mathrm{q}-\frac{2}{\mathrm{a}}\right)}{2\Gamma^{2}\left(\mathrm{p}+\frac{1}{\mathrm{a}}\right)\Gamma^{2}\left(\mathrm{q}-\frac{1}{\mathrm{a}}\right)}. \tag{4}$$

The appropriate expressions for all indices presented above in the cases of the Singh-Maddala and Dagum distributions can be obtained by setting, respectively, the parameter p to 1 and parameter q to 1.
Kleiber (1999) showed that for two GB2 distributions, $X_i \sim \mathrm{GB2}(a_i, b_i, p_i, q_i), i = 1, 2$, if $a_1 \leqslant a_2, a_1p_1 \leqslant a_2p_2$, and $a_1q_1 \leqslant a_2q_2$, then distribution X_2 Lorenz-dominates (is less unequal than) distribution X_1. Notice that Kleiber's conditions are sufficient, but not necessary. Therefore there may be some practical cases in which it will be impossible to verify Lorenz dominance on the basis of these conditions. Necessary conditions for Lorenz dominance were derived by Wilfling (1996): if distribution X_2 Lorenz-dominates (is less unequal than) distribution X_1, then $a_1p_1 \leqslant a_2p_2$, and $a_1q_1 \leqslant a_2q_2$.

3 Methods

3.1 Parameter estimation, goodness of fit and model selection techniques

All models analyzed in this paper were fitted to individual income data using the maximum likelihood estimation (MLE). The expressions for the log-likelihoods of the GB2 and its nested models (the Singh-Maddala and Dagum) are given in Kleiber and

Kotz (2003). MLE methods for the GB2 model with sampling weights is carefully discussed in Graf and Nedyalkova (2013). Hajargasht et al. (2012) developed an optimal GMM estimator for fitting the GB2 and its nested models to grouped data (i.e. data available in n income classes). For fitting models to data, we use Stata programs developed by Stephen Jenkins (Jenkins 2007). The programs maximize the likelihoods numerically using the modified Newton–Raphson algorithm, or optionally Berndt–Hall–Hall–Hausman, Davidon–Fletcher–Powell or Broyden–Fletcher–Goldfarb–Shanno algorithms. For an implementation of GB2 maximum likelihood estimation in R, see Graf and Nedyalkova (2012). Parameter variances are based on the negative inverse Hessian. Inequality and poverty indices implied by a fitted GB2 model, and their associated standard errors computed using the delta method, can be obtained using the gb2dist Stata command developed by the author. The command can be obtained from the author's webpage. The implementation covers also poverty indices for this distribution, which have been recently derived by Chotikapanich et al. (2013).

The plausibility of models' fit to data should be in principle assessed using goodness-of-fit tests like the Kolmogorov-Smirnov (KS) or Anderson-Darling (AD) tests (see, e.g., Stephens 1986), with p-values determined using a nonparametric bootstrap approach. The distributions of the goodness-of-fit tests based on the empirical distribution function (as the KS and the AD tests are) depend on the assumption that the data are drawn from the known (fixed) distributions. In our case, the distributions are fitted by the maximum likelihood procedure and hence they are not fixed. For this reason, the nonparametric bootstrap procedure should be used (see Clauset et al. 2009). However, our experiments have shown that for our data sets the goodness-of-fit tests always reject the hypothesis that the data follow even the best model selected by model selection tests (see below). This is not surprising as it often happens in the literature on fitting parametric models to income distribution data and in other large-sample settings (McDonald 1984), when even small deviations from a model result in model rejection. For this reason, often graphical and numerical methods for assessing goodness of fit are used (see, e.g., Graf and Nedyalkova 2013). The most popular graphical method is the quantile-quantile (q-q) plot, which for a given model plots the theoretical quantiles versus empirical quantiles of a variable. If the estimated model fits the data perfectly, the resulting q-q plot would coincide with the 45-degree line. The numerical approach to assessing goodness of fit relies on comparing the numerical values of theoretical and sample indicators such as the mean, the median, the standard deviation, the Gini index, the poverty rate, and others. In Section 4, we use both graphical and numerical methods in evaluating our fitted models.

In order to compare the fit of the GB2 model and its nested alternatives (the Singh-Maddala and Dagum), we use the likelihood ratio test. The likelihood ratio statistics takes the form:

$$LR = 2\left(\widehat{l_u} - \widehat{l_r}\right) \sim \ \chi^2(h), \tag{5}$$

where $\widehat{l_u}$ and $\widehat{l_r}$ are, respectively, the log-likelihood values corresponding to the unconstrained (GB2) and restricted or nested models (Singh-Maddala and Dagum), and h is the difference in the number of parameters in the two compared models (equal to 1 in our setting). The differences between GB2 and its nested alternatives can be thus compared using a chi-square distribution with one degree of freedom.

3.2 Testing for Lorenz dominance with the GB2 model

As pointed out in Section 2, Kleiber (1999) showed that for two GB2 distributions, $X_i \sim \text{GB2}(a_i, b_i, p_i, q_i), i = 1, 2$, if $a_1 \leqslant a_2, a_1p_1 \leqslant a_2p_2$, and $a_1q_1 \leqslant a_2q_2$,, then distribution X_2 Lorenz-dominates (is less unequal than) distribution X_1. After the GB2 model is fitted to data, the set of conditions implying Lorenz dominance can be tested using parameter estimates and their variances. In order to test equality of the Lorenz curves for two GB2 distributions with vectors of parameters $\theta_i = (a_i, b_i, p_i, q_i)^T$, $i = 1, 2$, we may use the following Wald test (Prieto-Alaiz 2007):

$$W = \left[H\left(\widehat{\theta}_1\right) - H(\widehat{\theta}_2)\right]^T \widehat{\Omega}_{12}^{-1} \left[H\left(\widehat{\theta}_1\right) - H(\widehat{\theta}_2)\right], \tag{6}$$

where $\widehat{\theta}$ is the MLE of θ, $H()$ is the 3×1 vector of nonlinear functions of the GB2 parameters, which state the Lorenz dominance:

$$H(\theta) = [h_1(\theta), h_2(\theta), h_3(\theta)]^T = [a, ap, aq]^T.$$

The W statistics is distributed as chi-square with three degrees of freedom. Assuming independence between compared distributions, $i = 1, 2$, the matrix $\widehat{\Omega}_{12}$ is given by:

$$\widehat{\Omega}_{12} = \left(\hat{D}\widehat{\Sigma}_1\hat{D}^T/n_1\right) + \left(\hat{D}\widehat{\Sigma}_2\hat{D}^T/n_2\right), \tag{7}$$

where n_1 and n_2 are the sample sizes for respective distributions, $\widehat{\Sigma}$ is the covariance matrix of MLE evaluated at $\widehat{\theta}$ and $\hat{D}$ is the (3×4) matrix with elements defined as follows:.

$$\hat{D}_{ij} = \left[\frac{\partial h_i(\theta)}{\partial \theta_j}\right]_{\theta=\widehat{\theta}}, \ i = 1,\ 2,\ 3;; j = 1,\ 2,\ 3,\ 4. \tag{8}$$

If the equality of the Lorenz curves is rejected, then if Kleiber's (1999) conditions are satisfied for a pair of GB2 distributions, $X_i \sim \text{GB2}(a_i, b_i, p_i, q_i), i = 1, 2$, that is if $a_1 \leqslant a_2, a_1p_1 \leqslant a_2p_2$, and $a_1q_1 \leqslant a_2q_2$, then we may conclude that distribution X_2 Lorenz-dominates (is less unequal than) distribution X_1.

4 Data

We use individual income data taken from two sources. For Poland, we use yearly data for the period 1993-2010 coming from the Household Budget Survey (HBS) study

conducted by the Polish Central Statistical Office. Data for other countries analysed in this paper (the Czech Republic, Hungary, the Slovak Republic) was obtained from the Luxembourg Income Study (LIS) database (see `www.lisdatacenter.org` for a detailed description of the LIS database.) LIS data is available in roughly 5-year intervals; this paper uses all data sets available for our choice of countries since the early 1990s to the most recent year available.
The main income variable that is modelled in the paper is disposable (post-tax and post-transfer) household income, equivalized using the square root equivalence scale. In order to obtain personal income distributions, in all our estimations we have used weights defined as a product of the household sampling weights and the number of household members. Income is measured in real (inflation-corrected) national currency units. Observations with negative and zero incomes were excluded from the analysis, but this affected less than 1% of all observations for all of our data sets. Table 1 presents descriptive statistics for the income variable used in our empirical analyses.

Table 1: Descriptive statistics for the real equivalent household disposable income variable

Data set	Mean	Median	Std. Dev.	Max.	No. of households
Czech Republic					
1992	103135.6	95509.62	49028.06	1271468	16234
1996	152586.8	134757.8	87317.75	3741595	28148
2004	177948.3	154467.5	107963	3095899	4351
Hungary					
1991	1209948	1073457	749995.7	8275354	2019
1994	1032074	864613	764597.3	2.03e+07	1936
1999	993708.6	854647	620465.9	7423942	1636
2005	1219921	1042275	859600.5	2.26e+07	2035
Poland					
1993	864.7	750.6	604.2	20127.1	32108
1998	1138.1	1003.0	778.3	21338.6	31745
2004	1102.7	949.4	847.0	27578.9	32214
2010	1503.6	1254.3	1741.1	181072.3	37127
Slovak Republic					
1992	115519.7	108743.8	46462.17	1208909	15990
1996	142847	132141.9	73055.85	1319030	16336
2004	156054.5	140326.3	94531.2	1844909	5147
2010	7299.088	6594.618	4759.55	291874.1	5198

5 Empirical results

5.1 Fitting models to CEE data

Tables 2-9 present our estimates of models' parameters together with their standard errors. We also give the values of log-likelihoods and the results of likelihood ratio tests for the fitted models. Results of the likelihood ratio tests for Poland, presented in Table 3, suggest that the GB2 model for Poland is preferred to the Singh-Maddala and Dagum models for all years under study. The results of model selection for other countries are less straightforward. In the case of the Czech Republic, at least one nested model seems to be as good as the GB2 for each studied year. For Hungary, the GB2 model is a better fit to data in all years except 1999. For the 1999 Hungarian

Table 2: Maximum likelihood estimates of models' parameters for Poland

Parameter estimates	Singh-Maddala	Dagum	GB2
	1993		
a	3.660 (0.031)	3.652 (0.0293)	5.463 (0.1990
b	739.7 (6.058)	769.6 (6.087)	749.4 (4.841)
p	-	0.955 (0.018)	0.575 (0.027)
q	0.951 (0.018)	-	0.564 (0.027)
Log-likelihood	-235121.2	-235121.6	-235047.4
	1998		
a	3.391 (0.028)	3.695 (0.031)	4.673 (0.167)
b	1041.2 (9.672)	1066.6 (8.531)	1044.0 (7.727)
p	-	0.860 (0.016)	0.638 (0.030)
q	1.093 (0.022)	-	0.710 (0.034)
Log-likelihood	-241793.3	-241771.5	-241746.8
	2004		
a	2.991 (0.024)	3.396 (0.029)	4.330 (0.159)
b	1018.6 (10.84)	1040.5 (8.858)	1011.7 (8.189)
p	-	0.814 (0.015)	0.600 (0.028)
q	1.161 (0.024)	-	0.702 (0.035)
Log-likelihood	-246737.2	-246703.7	-246678.7
	2010		
a	3.289 (0.026)	3.220 (0.024)	4.014 (0.131)
b	1226.1 (10.69)	1255.5 (11.03)	1238.7 (9.582)
p	-	1.004 (0.018)	0.752 (0.033)
q	0.946 (0.017)	-	0.726 (0.032)
Log-likelihood	-295975.1	-295979.6	-295954.5

Standard errors are given in parentheses.

sample, the three models are empirically indistinguishable. Similar conclusion applies do the Slovak Republic in 1992, but in 1996 the GB2 fits the data better than the alternatives. For both 2004 and 2010 Slovakian samples, the Dagum model is as good as the GB2. In general, the GB2 model fits the data best in 8 out of 15 analyzed data sets. However, there are stark differences between countries. The GB2 model is clearly the best model for Polish data. It seems also to be the best model for Hungary. For the Czech Republic and the Slovak Republic, the Dagum model is often as good as the GB2 and may be preferred in practical applications due to its simpler functional form.

Table 3: Likelihood ratio test for Poland

Year	Singh-Maddala vs. GB2		Dagum vs. GB2	
	LR	*p*-value	LR	*p*-value
1993	147.6	0.000	148.4	0.000
1998	93.0	0.000	49.4	0.000
2004	117.0	0.000	50.0	0.000
2010	41.1	0.000	50.2	0.000

Table 4: Maximum likelihood estimates of models' parameters for Czech Republic

Parameter estimates	Singh-Maddala	Dagum	GB2
	1992		
a	5.373 (0.064)	4.811 (0.055)	5.823 (0.274)
b	90938.49 (709.664)	91353.39 (877.08)	91574.01 (757.937)
p	-	1.157 (0.034)	0.885 (0.060)
q	0.845 (0.022)	-	0.762 (0.048)
Log-likelihood	-192443.57	-192450.99	-192441.99
	1996		
a	4.146 (0.040)	3.782 (0.033)	3.776 (0.133)
b	129775.2 (1080.466)	128804.7 (1198.58)	128810 (1206.311)
p	-	1.151 (0.026)	1.153 (0.061)
q	0.882 (0.019)	-	1.002 (0.052)
Log-likelihood	-350202.42	-350198.63	-350198.63
	2004		
a	3.902 (0.093)	3.711 (0.083)	3.864 (0.372)
b	152841 (3406.206)	153019.8 (3587.802)	152764.1 (3513.3)
p	-	1.072 (0.060)	1.014 (0.140)
q	0.929 (0.051)	-	0.941 (0.131)
Log-likelihood	-54971.207	-54971.296	-54971.202

Table 5: Likelihood ratio test for Czech Republic

	Singh-Maddala vs. GB2		Dagum vs. GB2	
	LR	p-value	LR	p-value
1992	3.16	0.075	18.0	0.000
1996	7.58	0.006	0.000	1
2004	0.01	0.920	0.188	0.665

Table 6: Maximum likelihood estimates of models' parameters for Hungary

Parameter estimates	Singh-Maddala	Dagum	GB2
	1991		
a	3.176 (0.099)	3.912 (0.135)	5.096 (0.685)
b	1203528 (48852.41)	1221943 (34794.49)	1177459 (33980.01)
p	-	0.725 (0.050)	0.525 (0.088)
q	1.295 (0.109)	-	0.676 (0.123)
Log-likelihood	-29239.886	-29234.724	-29232.439
	1994		
a	2.908 (0.097)	3.314 (0.108)	5.455 (0.912)
b	930855.5 (41540.52)	963970.8 (31761.01)	898641.1 (26572.51)
p	-	0.799 (0.055)	0.445 (0.087)
q	1.148 (0.096)	-	0.489 (0.105)
Log-likelihood	-28132.567	-28128.831	-28123.03
	1999		
a	3.719 (0.146)	3.309 (0.116)	4.005 (0.574)
b	791106.6 (28950.39)	800571 (33804.23)	796783.6 (29573.6)
p	-	1.159 (0.104)	0.896 (0.182)
q	0.833 (0.071)	-	0.756 (0.149)
Log-likelihood	-23594.707	-23595.511	-23594.561
	2005		
a	3.548 (0.117)	3.549 (0.114)	5.065 (0.700)
b	1035360 (34819.89)	1073059 (35182.79)	1049593 (28627.22)
p	-	0.958 (0.071)	0.609 (0.109)
q	0.959 (0.073)	-	0.603 (0.109)
Log-likelihood	-29729.559	-29729.542	-29725.822

Table 7: Likelihood ratio test for Hungary

Year	Singh-Maddala vs. GB2		Dagum vs. GB2	
	LR	p-value	LR	p-value
1991	14.894	0.000	4.57	0.033
1994	19.074	0.000	11.60	0.001
1999	0.292	0.589	1.9	0.168
2005	7.474	0.006	7.44	0.006

Table 8: Maximum likelihood estimates of models' parameters for Slovak Republic

Parameter estimates	Singh-Maddala	Dagum	GB2
	1992		
a	5.351 (0.063)	5.364 (0.065)	5.734 (0.269)
b	108484.6 (897.56)	109062.6 (920.735)	108848.1 (886.55)
p	-	0.987 (0.028)	0.901 (0.061)
q	0.994 (0.028)	-	0.906 (0.060)
Log-likelihood	-190613.23	-190613.16	-190612.06
	1996		
a	3.032 (0.030)	5.107 (0.064)	8.109 (0.432)
b	180820 (2979.55)	166536.4 (1143.753)	153834.3 (1280.921)
p	-	0.488 (0.010)	0.293 (0.017)
q	2.123 (0.073)	-	0.502 (0.035)
Log-likelihood	-203463.67	-203232.53	-203177.56
	2004		
a	3.383 (0.066)	4.107 (0.093)	4.413 (0.389)
b	156102.5 (3777.48)	156104.4 (2832.94)	154814.9 (3048.3)
p	-	0.752 (0.035)	0.686 (0.081)
q	1.301 (0.069)	-	0.898 (0.113)
Log-likelihood	-64425.834	-64421.299	-64420.94
	2010		
a	3.099 (0.058)	4.402 (0.098)	4.811 (0.382)
b	8266.501 (239.84)	7873.509 (122.922)	7724.4 (164.85)
p	-	0.616 (0.026)	0.554 (0.055)
q	1.690 (0.102)	-	0.868 (0.104)
Log-likelihood	-49330.235	-49313.165	-49312.467

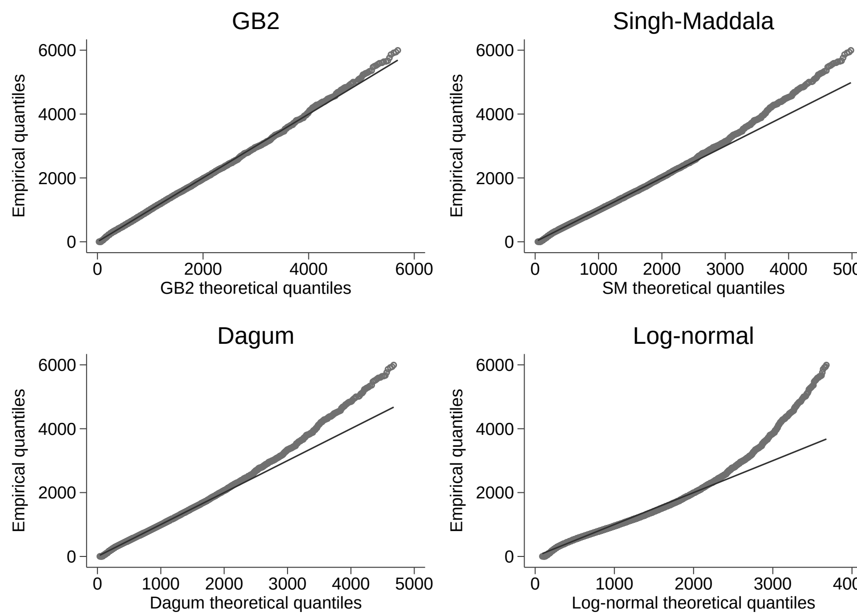
Table 9: Likelihood ratio test for Slovak Republic

Year	Singh-Maddala vs. GB2		Dagum vs. GB2	
	LR	p-value	LR	p-value
1992	2.34	0.126	2.2	0.138
1996	572.22	0.000	109.94	0.000
2004	9.788	0.002	0.718	0.396
2010	35.536	0.000	1.396	0.237

Goodness of fit is assessed using both visual and numerical methods. Figures 1-2 show quantile-quantile plots for Poland in 1993 and 2010. We have also included a log-normal model in Figures 1-2 in order to show how the three-parameter models improve the fit in comparison with a two-parameter model. We do not provide quantile-quantile plots for the Czech Republic, Hungary and the Slovak Republic as the data for these countries were taken from LIS, which is a remote-execution data access system not allowing for producing graphs. It can be easily seen that for

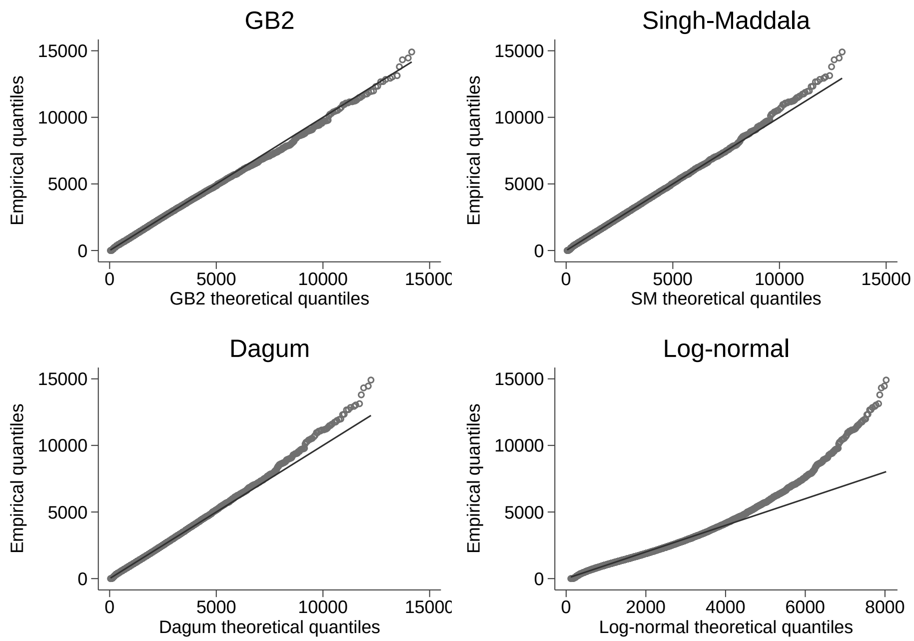
Poland, the GB2 model gives the best fit to data. Other models are visibly worse, especially for higher quantiles. It can be also observed that the two-parameter log-normal model gives a significantly worse fit to Polish data than the three-parameter Singh-Maddala and Dagum models. Goodness of fit is also evaluated numerically

Figure 1: Quantile-quantile plots, Poland, 1993



in Tables 10-13, by comparing the sample values of chosen distributional indicators with their counterparts implied by the fitted models. For brevity, the analyses are performed only for the last available year for each country. The results suggest that for most of the indices, the best fitting models produce indices' values that are often in a close agreement with the corresponding sample values. The two exceptions are the top-sensitive inequality index, GE(2), and the poverty rate. The poverty rate here is defined as the proportion of the population that has an income lower or equal to the 60% of the median income. The GE(2) index for Poland for the best fitting GB2 distribution differs by about as much as 54% from its sample counterpart. For Slovak Republic, the respective difference is also large and reaches about 33%. These facts reflect the high sensitivity of some inequality indices to the presence of extremely large incomes (Cowell and Flachaire 2007). The estimates implied by fitted parametric models seem to be much less sensitive to extreme observations than sample estimates. It is worth stressing here that both types of estimates (the sample estimates and

Figure 2: Quantile-quantile plots, Poland, 2010



estimates implied by the fitted model) for the most popular inequality measure – the Gini index – differ in our analyses by no more than 1.1%. This suggests that the GB2 model is quite successful in describing the inequality of income distribution in the CEE countries, at least if one is focusing on the Gini index.

The differences between sample estimates and estimates implied by fitted models for poverty rates in Hungary and Slovak Republic are also rather big and reach 10-12%. This suggests that, at least in some cases, the parametric distributions may have troubles in modelling also the lower tails of income distributions.

Figure 3 plots the evolution of the estimated GB2 parameters over time. The scale parameter, b, has increased markedly throughout the analyzed period in all countries, except for Hungary, representing the increase in mean income during the transition to market economies. The parameter b is proportional to the mean of the GB2 distribution (see equation 2). There are no visible trends in other parameters' behaviour for Hungary. For the Slovak Republic, the values of all three shape parameters – a, p, and q – have fallen over 1992-2010. This means that both tails of the fitted GB2 distribution have become fatter in the period under study. As

Table 10: Numerical goodness of fit, Czech Republic, 2004

	Empirical value	Percentage difference between empirical value and value implied by a fitted model		
		GB2	Singh-Maddala	Dagum
Mean	177948.3	0.2	0.2	0.3
Std. Dev.	107963	4.8	4.6	6.0
Median	154467.5	-1.6	-1.6	-1.6
Gini index	0.267	0.5	0.4	0.7
GE(2) index	0.184	10.0	8.7	11.2
P90/P10	3.212	1.0	1.1	1.0
P75/P25	1.801	1.1	1.1	0.9
Poverty rate	0.115	-0.9	-0.9	-1.4

P90/P10 and P75/P25 denote, respectively, the ratio of the 90^{th} percentile to the 10^{th} percentile and the ratio of the 75^{th} percentile to the 25^{th} percentile.

Table 11: Numerical goodness of fit, Hungary, 2005

	Empirical value	Percentage difference between empirical value and value implied by a fitted model		
		GB2	Singh-Maddala	Dagum
Mean	1219921	0.1	0.6	1.1
Std. Dev.	859600.5	0.3	8.8	12.6
Median	1042275	-1.0	-1.0	-1.2
Gini index	0.291	0.2	1.1	2.2
GE(2) index	0.248	0.5	16.0	21.8
P90/P10	3.311	-5.4	-6.1	-6.1
P75/P25	1.845	0.6	-1.5	-1.5
Poverty rate	0.125	-11.6	-10.4	-12.1

Table 12: Numerical goodness of fit, Poland, 2010

	Empirical value	Percentage difference between empirical value and value implied by a fitted model		
		GB2	Singh-Maddala	Dagum
Mean	1503.7	0.7	1.0	1.5
Std. Dev.	1741.15	32.4	36.6	39.0
Median	1254.3	-0.1	-0.1	-0.3
Gini index	0.319	1.1	1.9	2.7
GE(2) index	0.670	53.7	59.0	61.7
P90/P10	3.847	-1.1	-1.5	-1.6
P75/P25	1.955	0.3	-1.0	-1.1
Poverty rate	0.157	-0.2	-1.3	-2.1

Table 13: Numerical goodness of fit, Slovak Republic, 2010

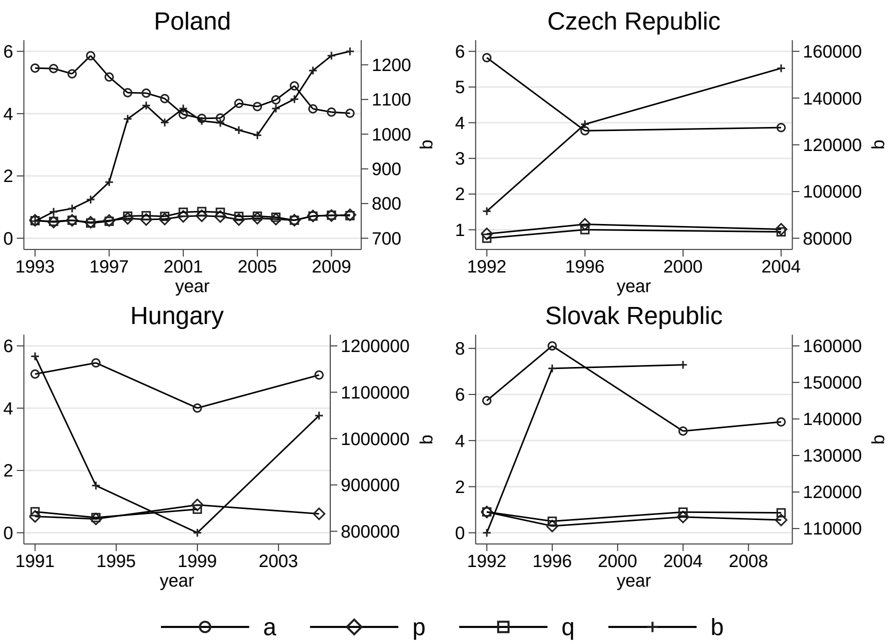
	Empirical value	Percentage difference between empirical value and value implied by a fitted model		
		GB2	Singh-Maddala	Dagum
Mean	7299.088	0.5	0.6	0.7
Std. Dev.	4759.55	18.6	22.6	20.3
Median	6594.618	-1.1	-0.7	-1.1
Gini index	0.265	0.0	0.8	0.5
GE(2) index	0.213	33.3	39.3	35.7
P90/P10	3.253	-4.6	-5.1	-4.7
P75/P25	1.814	-0.4	-2.6	-0.9
Poverty rate	0.134	-14.7	-12.6	-12.1

suggested in Section 2, this can be interpreted as evidence for growing income bi-polarization in the Slovakian society. The bi-polarization process, which concentrates incomes around two distributional poles (grouping the poor and the rich), shrinks the size of the middle class and in this way it can have significant negative consequences for economic growth and social stability. Recent theoretical literature has linked polarization to the intensity of social conflicts (Esteban and Ray 1994, 2011).
There was a notable fall in the value of a parameter for Poland and the Czech Republic. At the same time, the values of p and q for these countries have increased. These trends are similar to those reported for household income in Germany for 1984–93 by Brachmann et al. (1996), and for 1970–1990 for the US family income as reported by Bordley et al. (1996). For Poland and the Czech Republic, the fall in a, which is making both tails of the GB2 distribution fatter is combined with increases in both p and q, which have opposite effects on, respectively, the left and the right tail of income distributions. The conclusions with respect to changes in bi-polarization depend therefore on the joint changes in ap and aq, which is investigated in the next section.

5.2 Inference on changes in income inequality

In this section, we perform statistical tests on Lorenz dominance, which allow to make robust (independent of the choice of inequality measure) inferences on changes in income inequality. Table 14 presents sample estimates of four widely used inequality indices: the Gini index, the GE(2) index, and the two percentile ratios. According to these estimates, income inequality during the transformation to market economy has increased substantially in the Czech Republic, Poland and the Slovak Republic. For Hungary, the Gini and the GE(2) indices suggest that the inequality increased, but the percentile ratios suggest otherwise. The scale of the inequality increase in the Czech Republic, Poland and the Slovak Republic depends on the particular cardinal inequality measure used, but all of them suggest that income inequality has risen.

Figure 3: The evolution of the GB2 parameters over time (b measured on the right axis)



However, we cannot be sure that this conclusion would remain valid for other cardinal inequality indices that could be used. Testing for Lorenz dominance allows one to reach a conclusion that is valid for a wide range of popular inequality measures (see Section 2). Moreover, as shown in Section 3.2, parametric Lorenz dominance can be tested statistically and thus provide a conclusion, which is statistically significant. Statistical inference on inequality changes could be, of course, also conducted using tests based on sampling variances for particular inequality indices. However, such tests would have to be performed for all (possibly many) inequality measures used.

The results of the tests for Lorenz curves equality for chosen pairs of years are presented in Table 15. For Hungary, the fall in both a and q combined with a rise in p implies that the necessary conditions for Lorenz dominance are not satisfied and neither distribution Lorenz-dominates the other one (see Section 2). Therefore, the conclusions about the direction of inequality changes in Hungary depend on a particular cardinal inequality measure applied. It is notable that for Hungary the ap index, which regulates the fatness of the GB2 left tail, has increased over time. It means that the left tail of the Hungarian income distribution has become thinner; this had an inequality-reducing effect according to some inequality indices (including the percentile ratios, see Table 14).

Table 14: Inequality indices for the CEE countries, sample estimates

Data set	Inequality index			
	Gini	GE(2)	P90/P10	P75/P25
Czech Republic				
1992	0.206	0.112	2.360	1.548
1996	0.256	0.163	2.974	1.765
2004	0.267	0.184	3.212	1.801
Hungary				
1991	0.283	0.186	3.355	1.873
1994	0.321	0.273	4.138	1.970
1999	0.292	0.195	3.432	1.888
2005	0.291	0.248	3.311	1.845
Poland				
1993	0.284	0.239	3.312	1.808
1998	0.286	0.220	3.469	1.856
2004	0.313	0.259	4.000	1.981
2010	0.319	0.670	3.847	1.955
Slovak Republic				
1992	0.189	0.081	2.251	1.519
1996	0.250	0.131	3.038	1.716
2004	0.268	0.179	3.286	1.810
2010	0.265	0.213	3.253	1.814

For the Czech Republic, Poland and the Slovak Republic, the conditions of the Lorenz

Table 15: Test results for equality of the Lorenz curves

	Combinations of estimated parameters and test statistics						
	a	p	q	ap	aq	χ^2	p-value
Czech Republic							
1992	5.823	0.885	0.762	5.153	4.437	63.49	0.000
2004	3.864	1.014	0.941	3.918	3.636		
Hungary							
1991	5.096	0.525	0.676	2.6754	3.445	-	-
2005	5.065	0.609	0.603	3.085	3.045		
Poland							
1993	5.463	0.575	0.564	3.141	3.081	114.63	0.000
2010	4.014	0.752	0.726	3.019	2.914		
Slovak Republic							
1992	5.734	0.901	0.906	5.166	5.195	278.24	0.000
2010	4.811	0.554	0.868	2.665	4.176		

p-values in the last column are Sidak-adjusted.

dominance for the GB2 model are fulfilled. In particular, we observe that in these countries a fall in a over time is combined with a fall in both ap and aq. Therefore, income distributions observed in these countries in early 1990s Lorenz-dominate (are less unequal than) income distributions observed in the respective countries in the mid- or late-2000. P-values from the chi-square test confirm that these conclusions are statistically significant. The fall in both ap and aq means also that Poland and the Czech Republic have experienced a rise in income bi-polarization, similar to that occurring in the Slovak Republic. This confirms earlier results on changes in income polarization in Poland, obtained in a non-parametric framework (Kot 2008, Brzezinski 2011).

6 Conclusions

The objective of this paper was to model income distributions in four Central and Eastern European (CEE) countries (the Czech Republic, Hungary, Poland and the Slovak Republic) in 1990s and 2000s using parametric statistical models proposed in the theoretical literature. In particular, we have used the generalized beta distribution of the second kind (GB2) and the models that it encompasses (the Singh-Maddala and Dagum distributions). The models were fitted to micro-data on household incomes using the maximum likelihood estimation. We have found that for Poland, and to somewhat lesser extent for Hungary, the GB2 model fits the data better than the considered alternatives. For the Czech Republic and the Slovak Republic, the Dagum model is often in practice as good as the GB2 and may be preferred in empirical research due to its greater simplicity.

The paper also found that the tails of the fitted GB2 models for the Czech Republic, Poland and the Slovak Republic have become fatter over time. This can be interpreted as an evidence in favour of the view that the process of transformation to market economies in these countries has brought growing income bi-polarization – incomes began to cluster around the poles situated around the tails of the distribution. Our analysis for Hungary suggests that this country is the only one in our sample for which the left tail has become thinner – some of the probability mass has shifted to the middle or to the right tail of the distribution.

We have also provided statistical inference on changes in income inequality based on parametric Lorenz dominance. The results show that for a wide class of popular inequality indices, the period of economic transformation since the early 1990s to the mid- or late-2000s has brought unambiguously an increase in income inequality in the Czech Republic, Poland and the Slovak Republic. There is no Lorenz dominance in case of Hungary – income inequality has increased in this country according to some measures, but decreased according to others. Overall, this paper has shown that parametric modelling is a useful tool to describe the shape and the evolution of income distributions in the CEE countries. The results of this paper concerning the best fitting parametric model for a given country can be used in applying the model to study more specific economic problems involving income distribution – for example, to study the effect of economic reforms on income distribution in general equilibrium modelling.

Acknowledgments

The previous versions of this paper have received helpful comments from the participants of the 2012 Polish Stata Users Group meeting (Warsaw, 19 October 2012) and the 5th Scientific Conference "Modelling and forecasting the national economy" (Sopot, 10-12 June 2013). This work was supported by Polish National Science Centre grant no. 2011/01/B/HS4/02809.

References

[1] Bandourian R., McDonald J.B., and Turley R.S. (2003), A comparison of parametric models of income distribution across countries and over time, *Estadistica* 55, 135-152.

[2] Bartels, C.P.A., Van Metele, H. (1975), Alternative Probability Density Functions of Income, *Research memorandum* 29, Vrije University Amsterdam.

[3] Boccanfuso, D., Richard, P., Savard, L. (2013), Parametric and nonparametric income distribution estimators in CGE micro-simulation modelling, *Economic Modelling*, http://dx.doi.org/10.1016/j.econmod.2013.07.002.

[4] Bordley, R. F., McDonald, J.B., and Mantrala, A. (1996), Something New, Something Old: Parametric Models for the Size Distribution of Income, *Journal of Income Distribution* 6, 91-103.

[5] Brachmann, K., Andrea, S., and Trede, M. (1996), Evaluating Parametric Income Distribution Models, *Allegemeine Statistiches Archiv* 80, 285-98.

[6] Brzezinski, M. (2011), Statistical inference on income polarization in Poland, *Przeglad Statystyczny* 58, 102-113.

[7] Burkhauser, R. V., Feng, S., Jenkins, S. P., Larrimore, J. (2012), Recent trends in top income shares in the United States: reconciling estimates from March CPS and IRS tax return data, *Review of Economics and Statistics* 94, 371-388.

[8] Chotikapanich, D., Griffiths, W., Karunarathne, W. (2013), Calculating Poverty Measures from the Generalised Beta Income Distribution, *Economic Record* 89, 48-66.

[9] Clauset, A., Shalizi, C. R., Newman, M. E. J. (2009), Power-law distributions in empirical data, *SIAM Review* 51, 661-703.

[10] Cowell, F. A. (2000), Measurement of Inequality, [in:] *Handbook of Income Distribution*, Vol. 1, [ed.:] A. B. Atkinson and F. Bourguignon, Elsevier, Amsterdam, 87-166.

[11] Cowell, F. A., Flachaire, E. (2007), Income distribution and inequality measurement: The problem of extreme values, *Journal of Econometrics* 141, 1044-1072.

[12] Dagum, C. (1977), A new model for personal income distribution: specification and estimation, *Economie Appliquée* 30, 413-437.

[13] Dastrup, S. R., R. Hartshorn, J. B. McDonald (2007), The Impact of Taxes and Transfer Payments on the Distribution of Income: A Parametric Comparison, *Journal of Economic Inequality* 5, 353-369.

[14] Domanski, C., Jedrzejczak, A. (2002), Income Inequality Analysis in the Period of Economic Transformation in Poland, *International Advances in Economic Research* 8, 215-220.

[15] Esteban, J. and Ray, D. (1994), On the measurement of polarization, *Econometrica* 62, 819-851.

[16] Esteban, J., Ray, D. (2011), Linking conflict to inequality and polarization, *American Eco-nomic Review* 101, 1345-1374.

[17] Foster, J., Wolfson, M. (2010), Polarization and the decline of the middle class: Canada and the U.S, *Journal of Economic Inequality* 8, 247-273.

[18] Gibrat, R. (1931), *Les Inegalites Economiques*, Sirey, Paris.

[19] Graf, M., Nedyalkova, D. (2012), *GB2: Generalized Beta Distribution of the Second Kind: Properties,Likelihood, Estimation*, R Package Version 1.1.

[20] Graf, M., Nedyalkova, D. (2013), Modeling of Income and Indicators of Poverty and Social Exclusion Using the Generalized Beta Distribution of the Second Kind, *Review of Income and Wealth*, doi: 10.1111/roiw.12031.

[21] Hajargasht, G., Griffiths, W., Brice, J., Rao, D.S.P., Chotikapanich, D. (2012), Inference for Income Distributions Using Grouped Data, *Journal of Business of Economic Statistics* 30, 563-76.

[22] Jagielski, M., Kutner, R. (2010), Study of Households' Income in Poland by Using the Statistical Physics Approach, *Acta Physica Polonica A* 117, 615-618.

[23] Jäntti, M., Jenkins, S. P. (2010), The Impact of Macroeconomic Conditions on Income Inequality, *Journal of Economic Inequality* 8, 221-240.

[24] Jenkins, S. P. (2007), gb2fit: Stata Module to fit Generalized Beta of the Second Kind Distribution by Maximum Likelihood, *Statistical Software Components Archive* S456823, (`http://ideas.repec.org/c/boc/bocode/s456823.html`).

[25] Jenkins, S. P. (2009), Distributionally-sensitive inequality indices and the GB2 income distribution, *Review of Income and Wealth* 55, 392-398.

[26] Kleiber, C. (1999), On the Lorenz Order Within Parametric Families of Income Distributions, *Sankhya* B61, 514-17.

[27] Kleiber, C. (2008), The Lorenz Curve in Economics and Econometrics, [in:] *Advances on Income Inequality and Concentration Measures*, [ed.:] G. Betti, A. Lemmi, Routledge, London, 225-242.

[28] Kleiber, C., Kotz, S. (2003), *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley, Hoboken, NJ.

[29] Kordos, J. (1990), Research on income distribution by size in Poland, [in:] *Income and Wealth Distribution, Inequality and Poverty*, [ed.:] C. Dagum, M. Zenga, Springer, New York, Berlin, London, and Tokyo, 335-351.

[30] Kot, S. M. (2008). *Polaryzacja ekonomiczna. Teoria i zastosowanie*, PWN, Warszawa.

[31] Lukasiewicz, P., Orlowski, A. (2004), Probabilistic Models of Income Distributions, *Physica A* 344, 146-151.

[32] McDonald, J. B. (1984), Some Generalized Functions for the Size Distribution of Income, *Econometrica* 52, 647-63.

[33] McDonald, J. B., Xu, Y. J. (1995), A Generalization of the Beta Distribution with Applications, *Journal of Econometrics* 66, 133-52, (Erratum: *Journal of Econometrics* 69, 427-8, 1995).

[34] McDonald, J. B., Ransom, M. (2008), The Generalized Beta Distribution as a Model for the Distribution of Income: Estimation and Related Measures of Inequality, [in:] *Modeling Income Distributions and Lorenz Curves*, [ed.:] D. Chotikapanich, Springer, New York, 147-66.

[35] Pareto, V. (1897), *Cours d'économie politique*, Ed. Rouge, Lausanne.

[36] Parker, S. C. (1999), The Generalized Beta as a Model for the Distribution of Earnings, *Economics Letters* 62, 197-200.

[37] Prieto-Alaiz, M. (2007), Spanish Economic Inequality and Gender: A Parametric Lorenz Dominance Approach, *Research on Economic Inequality* 14, 49-70.

[38] Salem, A.B., Mount, T. D. (1974), A convenient descriptive model of income distribution: The gamma density, *Econometrica* 42, 1115-1127.

[39] Shorrocks, A. F. (1984), Inequality decomposition by population subgroups, *Econometrica* 52, 1369-1388.

[40] Singh, S.K., Maddala, G.S. (1976), A function for the size distribution of incomes, *Econometrica* 44, 963-970.

[41] Stephens, M. A. (1986), Tests based on EDF statistics, [in:] *Goodness-of-fit techniques*, [ed.:] R. B. D'Agostigno, M. A. Stephens, Marcel Dekker, New York, 95-193.

[42] Taille, C. (1981), Lorenz ordering within the generalized gamma family of income distributions, [in:] *Statistical Distributions in Scientific Work*, vol. 6, [ed.:] C. Taille, G. P. Patil, B. Balderssari, Reidel, Boston, 181-192.

[43] Wilfling, B. (1996), Lorenz ordering of generalized beta-II income distributions, *Journal of Econometrics* 71, 381-388.

[44] Wimp, J. (1981), The computation of ${}_3F_2(1)$, *International Journal of Computer Mathematics* 10, 55-62.