PRZEGLĄD STATYSTYCZNY
R. LVIII – ZESZYT 1-2 – 2011

MICHAŁ BRZEZIŃSKI

STATISTICAL INFERENCE ON CHANGES IN INCOME POLARIZATION IN POLAND

1. Introduction

Over the last years, measurement and estimation of economic polarization has been getting an increased attention among economists and other social scientists. Starting with contributions of Esteban and Ray [14] and Wolfson [35], several researchers proposed different measures of polarization, often developed formally in an axiomatic framework[1]. Major applications of these and other approaches to polarization measurement include analysis for the UK [25], Spain [21, 22], China [36], Uruguay [23], Poland [27], and Denmark [24]. There are also studies devoted to the regional polarization in Russia [20], the European Union [18] and Central and Eastern European countries [19], as well as many papers conducting cross-country analysis [8, 10, 11, 13, 28, 32].

Economic polarization as conceptualized in these studies concerns the distribution of socio-economic characteristics (for example income, consumption, race, education) and refers to the extent to which the society is grouped in a number of clusters. In other words, in a polarized society incomes or other relevant attributes become concentrated around two or more diverging poles. Such a phenomenon is related, but certainly conceptually different from the more familiar concept of inequality. For example, it is intuitively easy to imagine a society converging in incomes to two local poles (for example grouping 'the rich' and 'the poor'). Such a society would display lowering level of inequality as measured by standard inequality indices, since the overall income dispersion would decrease, but it would also become more clustered around the two poles and therefore more bi-polarized[2].

What gives polarization a special appeal, both from the abstract and policy-oriented perspective, is that it has been theoretically linked to the phenomenon of social conflicts. Esteban and Ray [15] argued that it is polarization, and not inequality as commonly perceived, which is correlated with social conflicts like organised, large-scale

[1] The major contributions include Wang and Tsui [34], Chakravarty and Majumder [5], Zhang and Kanbur [36], Duclos et al. [8], and Esteban et al. [13].

[2] The issue of whether polarization and inequality are distinguishable empirically has been a matter of some debate. Ravallion and Chen [31] and Zhang and Kanbur [36] argued that measures of polarization generally do not generate very different results from those of standard measures of inequality. Esteban [12], Duclos et al. [8], and Lasso de la Vega and Urrutia [28] provide evidence that the two sets of indices significantly differ empirically.

protests, strikes, demonstrations, or even revolts and armed unrests. They showed that in the case of bipolarized society the level (or intensity) of social conflicts increases with the magnitude of polarization[3].

From another point of view, polarization has often been associated with the 'disappearing of the middle class' – a phenomenon characteristic for the US and the UK in the 1980s [25, 35]. Indeed, if incomes concentrate around two opposite distributive poles, then the size of the middle class has to decrease. The common opinion and a number of economic arguments suggest that a stable and sizable middle class is a necessary condition for successful economic and political transition and development (see, e.g., [2], pp. 3–4). Therefore, high or increasing level of bi-polarization can pose a serious problem for policymakers.

Recently, Kot [27] has offered an important contribution to both theoretical and empirical literature on economic polarization. He introduced a new class of polarization indices based on the concept of the Lorenz curve. Using Polish Household Budget Survey (HBS) data, he estimated these new indices as well as several existing polarization measures, including those introduced in Wolfson [35] and Esteban et al. [13], for the distribution of household expenditures in Poland during 1993-1997 and 1998-2005[4]. He found that during the first period there was no clear trend in the economic polarization, while in the second period polarization has increased by 0.85 to 8.1%, depending on the polarization index used.

This paper attempts to make two contributions. First, we complement Kot's [27] empirical analysis by estimating another important polarization index proposed by Duclos et al. [8] for the income distribution in Poland during 1998-2007. Beside developing the axiomatic foundation for the index, Duclos et al. [8] have provided also a formula for an asymptotic variance for their polarization index estimated from household survey samples. It is therefore possible to conduct formal statistical inference on this index by calculating confidence intervals and testing hypotheses about the changes in the values of the index over time. Hence, the second contribution of the paper is to test whether estimated changes in income polarization are statistically significant. It allows us to make polarization comparisons for Poland robust to sampling variability and to formulate reliable conclusions on whether income polarization has increased or not.

Section 2 of the paper introduces polarization measure as well as the methods of its sample estimation and tools of statistical inference as offered by Duclos et al. [8]. Micro data from Polish Household Budget Survey study are discussed in section 3. Section 4 reports and discusses empirical results, while section 5 concludes.

---

[3] See also Esteban and Ray [16, 17]. Esteban and Ray ([16], p. 164) argue that empirical evidence on the connection between *inequality* and social conflicts is ambiguous and inconclusive.

[4] The period between 1993 and 2005 is divided into two sub-periods because HBS data for these sub-periods are not directly comparable (see also section 3).

2. MEASURES OF POLARIZATION

The approach to measuring economic polarization introduced in Duclos et al. [8] has two attractive features – one theoretical, the other practical. The theoretical advantage of this formulation is that it is formulated in the so-called identification-alienation framework, introduced in Esteban and Ray [14]. This framework allows for formal analysis of the abovementioned link between polarization and the phenomenon of social conflict [15, 16, 17]. The practical benefit of the approach is that it is carefully operationalized in a statistical framework. Duclos et al.'s [8] methodology offers distribution-free statistical inference procedures, which allows to check whether the resulting estimates are not the effect of sampling noise and measurement errors[5].

Identification-alienation framework, introduced in Esteban and Ray [14], suggests that polarization can be measured as the effect of two interrelated mechanisms: (1) alienation, which is felt by individuals from a given group (defined by income class, religion, race, education etc.) toward individuals belonging to other groups, and (2) identification, which unites members of any given group. This approach assumes that polarization requires that individuals identify with other members of their socio-economic group and feel alienation to members of other groups. The approach of Duclos et al. [8] is proposed for the space of distributions of incomes (or consumption expenditures, wealth, etc.) described by income density functions (f), with integrals of these functions corresponding to various population sizes. It is assumed that identification at income y depends on the density at x, $f(x)$, while alienation between two individuals with incomes y and x is given by $|y-x|$. By imposing a set of axioms on the identification-alienation structure, Duclos et al. [8] derive the following polarization measure (DER index):

$$P_a(f) = \iint f(y)^{1+\alpha} |y - x| \, dydx, \tag{1}$$

where α is an ethical parameter expressing the weight given to the identification part of the framework. The axioms introduced by Duclos et al. [8] require that α must be bounded in the following way: $0.25 \leqslant \alpha \leqslant 1$. When $\alpha = 0$, the DER index is equal to the twice the popular Gini coefficient of inequality, which for density f and all incomes normalized by their mean takes the form

$$G(f) = \frac{1}{2} \iint f(y) |y - x| \, dydx. \tag{2}$$

Therefore, the lowest admissible value for DER index of $\alpha = 0.25$ should produce polarization results close to those of inequality measured by Gini index, while $\alpha = 1$ leads potentially to the highest disparity between the Gini coefficient and the DER index[6].

[5] See also Duclos et al. [9] for the more complete presentation of their statistical framework.

[6] In order to make comparisons between the Gini index and the DER index easier, in our empirical application in section IV we compute all measures for incomes normalized by the mean income and divide DER indices by 2.

Duclos et al. [8] analyze also the problem of the statistical estimation of their index. It can be shown that for every income distribution function F with associated density function f and mean μ, Eq. 1 can be restated in the form

$$P_\alpha(F) = \int_y f(y)^\alpha \alpha(y)\, dF(y), \tag{3}$$

with $a(y) \equiv \mu + y(2F(y) - 1) - 2\int_{-\infty}^{y} x dF(x)$.

Suppose next that the index is to be estimated using a random sample of n i.i.d (independent and identically distributed) sample observations of income yi, i =1,..., n, drawn from the distribution $F(y)$ and ordered such that $y_1 \leqslant y_2 \leqslant ... \leqslant y_n$. It is also assumed that a set of sampling weights is provided with w_i being a sampling weight on observation i and $\bar{w} = \sum_{j=1}^{n} w_j$ being the sum of weights. All incomes are normalized by the sample weighted mean income $\left(\mu = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i\right)$. Using Eq. 3, the DER index can be then estimated by

$$P_\alpha\left(\hat{F}\right) = \bar{w}^{-1} \sum_{i=1}^{n} w_i \hat{f}(y_i)^\alpha \hat{a}(y_i), \tag{4}$$

where $\hat{a}(y_i)$ is given as

$$\hat{a}(y_i) = \mu + y_i\left(\bar{w}^{-1}\left(2\sum_{j=1}^{i} w_j - w_i\right) - 1\right) - \bar{w}^{-1}\left(2\sum_{j=1}^{i-1} w_j y_j + w_i y_i\right), \tag{5}$$

and where $\hat{f}(y_i)^\alpha$ is estimated non-parametrically using kernel density estimation methods[7]. Duclos et al. [8] use a symmetric weighted kernel function $K(u)$, defined such that $\int_{-\infty}^{+\infty} K(u)du = 1$ and K(u) $\geqslant$ 0. The estimator $\hat{f}(y)$ is then defined as

$$\hat{f}(y) \equiv (\bar{w}h)^{-1} \sum_{i=1}^{n} w_i K\left(\frac{y - y_i}{h}\right), \tag{6}$$

where the smoothing parameter h, which is usually called the bandwidth, describes the width of the density window around each point, and kernel function $K(u)$ is specified as the Gaussian kernel, defined by

$$K(u) = (2\pi)^{-0.5} \exp^{-0.5u^2}. \tag{7}$$

[7] Kernel density estimation methods are presented in detail in, for example, Pagan and Ullah ([29], cha. 2) and Silverman ([33], cha. 3). See also Jenkins [25], and Duclos and Araar ([7], cha. 15) for analysis of these methods in the context of income distribution.

The optimal value of the bandwidth ($h*$) is chosen as to minimize the mean square error (MSE) of the estimator (1) for a given sample size (n). Duclos et al. [8, 9] devise computational formulas, which approximate h^*. For a normal distribution with variance and a Gaussian kernel function used in (3), the formula is

$$h^* \cong 4.7 n^{-0.5} \sigma \alpha^{0.1}, \tag{8}$$

where α is again polarization sensitivity parameter. For an income distribution characterized by skewness greater than about 6, a better approximation is given by a more complex formula of the form

$$h^* \cong n^{-0.5} IQ \frac{(3.76 + 14.7\sigma_{\ln})}{(1 + 1.09 \cdot 10^{-4}\sigma_{\ln})^{7268+15323\alpha)}}, \tag{9}$$

where IQ is the interquartile range (the difference between the 75^{th} and the 25th percentile) and is the variance of the logarithms of income[8].

Duclos et al. [8, 9] further show that under standard regularity conditions concerning the finiteness of moments of certain variables, if h in Eq. 6 vanishes when n tends to infinity, then $n^{0.5}\left(P_\alpha\left(\widehat{F}\right) - P_\alpha(F)\right)$ has an asymptotic limiting normal distribution $N(0, V_\alpha)$, with

$$V_\alpha = var_{f(y)}\left((1+\alpha) f(y)^\alpha a(y) + y \int f(x)^\alpha dF(x) + 2 \int_y^\infty (x - y) f(x)^\alpha dF(x)\right). \tag{10}$$

Variance given by Eq. 10 is distribution-free in the sense that it can be computed without knowledge of the form or the parameters of the distribution from which the sample is drawn. The expression (11) is operationalized by the use of the delta method (see, e.g., Rao [30], pp. 385-391) in the DASP software package written for STATA [1, 7]. In this paper, asymptotic variance and standard errors for the DER index are computed with the help of the DASP package.

Using estimates of the DER index and its asymptotic variance, it is possible to construct confidence intervals for the index and to test hypotheses about the changes in its value. The statistical inference in this paper is based on the asymptotic t-type statistic computed using Eq. 3 and the variance estimate of the DER index computed on the basis of Eq. 10. As we are mainly interested in changes in income polarization in time, the relevant hypothesis states that two different income distributions have the same value of the DER index. Assuming that we have independent samples drawn from two distributions (e.g. income distributions in periods t_0 and t_1), we can compute estimates $P_\alpha\left(\hat{F}_{t_0}\right)$ and $P_\alpha\left(\hat{F}_{t_1}\right)$ from the two samples using formula (3), and variance estimates $\hat{V}_\alpha(t_0)$ and $\hat{V}_\alpha(t_1)$ using Eq. 10[9]. For a null hypothesis that $P_\alpha(F_{t_0}) = P_\alpha(= F_{t_1})$, a

[8] The skewness of income distribution estimated from household survey data may often be greater than 6 (cf. Kot [27], p. 78), especially if extreme incomes are not eliminated, recoded or dealt with in other way.

[9] Cf. Davidson and Flachaire [6] for asymptotic inference for differences in inequality indices.

t-type statistic is

$$W = \frac{P_\alpha\left(\hat{F}_{t_0}\right) - P_\alpha\left(\hat{F}_{t_1}\right)}{\left(\hat{V}_\alpha\left(t_0\right) + \hat{V}_\alpha\left(t_1\right)\right)^{0.5}}. \tag{11}$$

Finally, using expression (11) we compute asymptotic *P* values based on the standard normal distribution or on the Student distribution with *n* degrees of freedom.

3. DATA

This paper uses data from Household Budget Survey (HBS) study conducted yearly by the Polish Central Statistical Office (CSO). We use yearly HBS micro-data for the period 1998-2007. An important modification of the HBS occurred experimentally in 1997, and definitively in 1998, when in order to fulfil Eurostat recommendations, new definitions of some core concepts (i.e. disposable income) were implemented. Due to this change it is rather difficult to construct fully comparable series of household disposable incomes for the period before 1998 and after this year. Therefore, in this paper we use HBS data from 1998 to 2007 (the last available year). The HBS sample design has changed several time during the period under study[10]. From the point of view of the present paper, it is important to notice that during 1996-2000 the design assumed sampling of primary sampling units (PSU's) for a period of four years, while since 2001 every year a new subsample is drawn for use during two years. Therefore, year-to-year samples are correlated and the statistic *W* given by Eq. 11 should take into account also the covariance of the two estimated polarization indices. To avoid this complication we study statistical inference on the DER index using independent samples for the years 1998, 2001, 2004 and 2007.

Household net disposable income (i.e. post-tax-and-transfer income) is the main income concept used. It includes cash wages and salaries, self-employment income (including farm income), cash property income, social transfers (including social insurance, social assistance) and other income. Income taxes, mandatory payroll taxes and gifts donated to other households are not included. As it is standard in income distribution literature, we consider the individual as the main unit of analysis. In order to obtain personal income distributions, all household observations are weighted by the product of household weights provided in the HBS and household size[11]. All incomes are divided by equivalence scales defined as $h^{0.5}$, where *h* is household size, to adjust for the size and composition of households. Finally, to obtain real equivalent disposable income we use CPI deflator to express all incomes in December 2007 price levels.

[10] The detailed description of the HBS design and its other features can be found in Kordos et al. [26] and Central Statistical Office [4].

[11] HBS weights are non-response weights adjusting sample data for the differential non-response rates of different types of households. The method of estimating these weights has changed several times between 1998 and 2004. See Kordos et al. [26] and Central Statistical Office [4] for details.

4. RESULTS

Figure 1 shows trends in the values of the DER index for $\alpha \in \{0.25, 0.5, 0.75, 1\}$ during the period under study. To verify whether income polarization as measured by the DER indices is indeed behaving differently from income inequality as measured by the Gini index, which is equal to the DER index with $\alpha = 0$, the changes in the latter are presented as well. As expected, changes in the DER index with $\alpha = 0.25$ are very similar in quantity and direction to the changes in the Gini index. The same can be said for other values of the parameter α except $\alpha = 1$. It can be observed that in this case income polarization as measured by the DER index is changing in a markedly different ways from inequality. For example, while the Gini index is clearly decreasing over 2004-2006, DER (α =1) is slightly increasing. Such opposite tendencies are displayed by the indices for 2002-2003 and 1998-1999 periods as well. Our analysis confirms, therefore, the conclusion of Duclos et al. [8] that the DER index, at least for the highest admissible value of α, is empirically distinct from the Gini index.

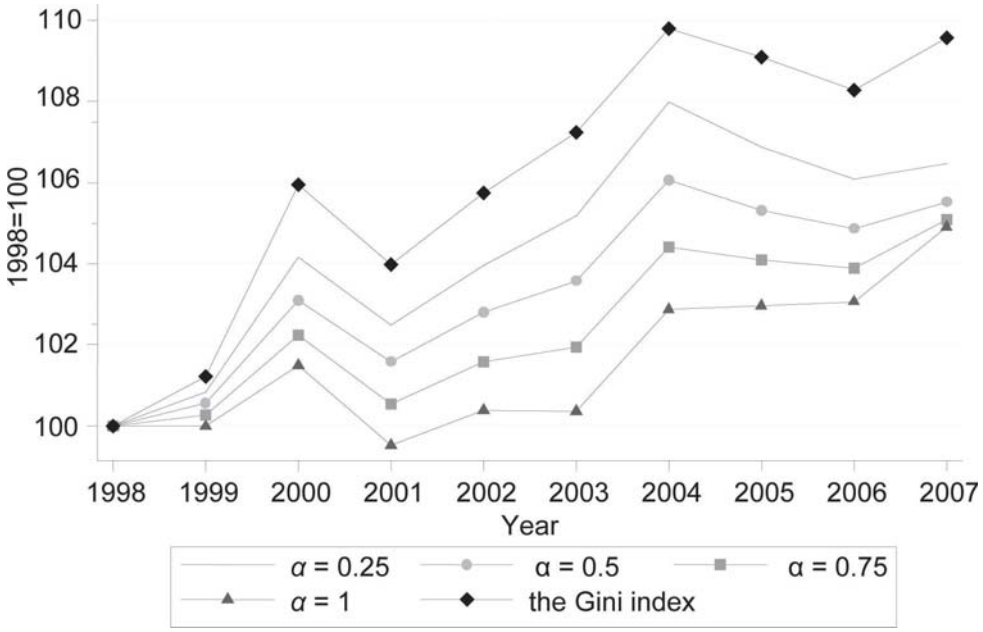

Figure 1. Changes in income polarization (DER index) in Poland, 1998-2007

Throughout the period under study, there was an increase in income polarization ranging from 4.9 to 6.5%, depending on the value of α parameter. This finding is consistent with the slight increase in the polarization of consumption expenditures over the 1998-2005 period found by Kot [27]. Table 1 reports values of the DER index for $\alpha = 1$, which gives the highest weight to the identification part of the identification-alienation framework in the Duclos et al.'s [8] account. Moreover, for our datasets, this value of the α is the one for which polarization is empirically clearly different from

inequality. Table 1 shows also the values of the standard error for the estimate of the index along with 95% and 99% confidence intervals. For the comparison between 1998 and 2007 both types of confidence intervals are non-overlapping and therefore we may conclude that the difference in income polarization as measured by the DER index with α =1 is statistically significant. Similarly, there is a statistically significant increase in income polarization between 2001 and 2007. However, for 4 out of 6 paired comparisons under study (i.e. 1998 and 2001, 1998 and 2004, 2001 and 2004, 2004 and 2007) both types of confidence intervals are overlapping and therefore they cannot be used to state unambiguously whether income polarization has changed. Hence, to provide statistical inference on these changes we turn to hypotheses testing.

Table 1

Income polarization according to the DER index (α = 1), Poland, 1998-2007

	DER index	SE	95% CI		99% CI	
			LB	UB	LB	UB
1998	0.1614	0.0010	0.1594	0.1635	0.1587	0.1641
1999	0.1614	0.0018	0.1578	0.1650	0.1566	0.1661
2000	0.1638	0.0015	0.1610	0.1667	0.1600	0.1675
2001	0.1606	0.0009	0.1589	0.1623	0.1584	0.1629
2002	0.1620	0.0013	0.1594	0.1646	0.1586	0.1654
2003	0.1620	0.0009	0.1603	0.1637	0.1597	0.1642
2004	0.1660	0.0013	0.1634	0.1687	0.1626	0.1695
2005	0.1662	0.0011	0.1641	0.1683	0.1635	0.1689
2006	0.1664	0.0010	0.1645	0.1682	0.1639	0.1688
2007	0.1693	0.0012	0.1669	0.1717	0.1662	0.1725

Notes: Own calculations using HBS data. Column 1 reports point estimates of the DER index calculated on the basis of the household data with weights equal to the product of the household weight and the size of the household. Column 2 reports standard errors of the DER index. Columns 3 and 4 report lower and upper bounds (LB and UB) for, respectively, 95% and 99% confidence intervals.

Table 2 shows, first, the change (D) in income polarization as measured by the DER index (α = 1) for every pair among the analyzed years (1998, 2001, 2004, 2007) with index for the later year always being the minuend. This is followed by standard error (SE) for the estimate of the measured change in the DER index, and the probability (P value) that such a change is greater than zero (P value = Pr($D > 0$). If D is less than 0, then $1-P$ value = Pr($D < 0$). Among the paired comparisons, the change from 1998 to 2001 is not statistically significant at the conventional 5% level with P value equal to 0.28. The change from 2004 to 2007 is significant only at the 5% level, but not at the 1% level. All other changes are statistically significant at standard levels.

The main conclusion from these results is that short-period observed changes in income polarization in Poland, which are usually quantitatively rather small, can often reflect rather sampling variability than the real changes in the underlying income distribution. Therefore, such short-period changes, including year-to-year changes, should be treated cautiously especially is they are to be used as a basis for policy recommendations. On the other hand, for the HBS data even moderate changes in the DER index, as observed during the 6- and 9-year time spans among our paired comparisons, appear to be statistically significant.

Table 2

Hypotheses tests for changes in the DER index (α = 1)

		2001	2004	2007
1998	*D*	–0.000792	0.004614	0.007921
	SE	0.001362	0.001697	0.001611
	P value	0.2804	0.0033	0.0000
2001	*D*		0.005406	0.008713
	SE		0.001602	0.001511
	*P*value		0.0004	0.0000
2004	*D*			0.003307
	SE			0.001819
	P value			0.0345

Notes: *D* denotes differences in measured values of DER indices over time, SE is the standard error of *D*, and *P* value is the probability that *D* is greater than 0.

5. CONCLUSIONS

This paper has provided estimates of the DER index of economic polarization, calculated on the basis of income distribution HBS data for Poland between 1998 and 2007. It has also offered statistical inference on the changes in income polarization by estimating confidence intervals and testing statistical hypotheses. The main conclusion is that for the entire period under study there was a moderate rise in income polarization in the range of 4.9 to 6.5%, depending on the polarization-sensitivity parameter α. For the period 1998-2007, the changes in the DER index with α = 1, for which polarization is empirically distinguishable from inequality, are statistically significant at conventional levels. The paper has found, however, that for smaller increases, which usually occur over shorter periods, it may be difficult to conclude whether observed changes are due to the changes in the income distribution in the population or due to the sampling variability of the estimates.

Finally, a caveat is in order concerning the methods used to estimate the variance of the DER index from the HBS data. These methods, introduced in Duclos et al. [8], assume that income observations are i.i.d., whereas HBS is a complexly designed survey with stratification, clustering and weighting of sample observations. While weighting can be rather easily incorporated in the statistical inference procedures, the effects

of stratification and clustering are harder to dealt with. However, accounting for the complex sample design would almost certainly increase variance estimates for the DER index[12]. This would lead to wider confidence intervals, larger P values in our empirical application, and could force us to reverse the direction of inference in at least some cases. We leave the issue of accounting for complexity of survey design in inference for polarization measures for future research.

REFERENCES

[1] Araar A. and Duclos J.-Y., *User Manual for Stata Package DASP: Version 2.0*, PEP, CIRPEE, World Bank and WIDER, 2009.

[2] Banerjee A and Duflo E., What is middle class about the middle classes around the world? *Journal of Economic Perspectives*, 22(2) (2008), pp. 3-28.

[3] Bhattacharya D., Inference on Inequality from Household Survey Data, *Journal of Econometrics*, 137(2) (2007), pp. 674-707.

[4] Central Statistical Office, *Household Budget Surveys in 2007*, CSO, Warszawa, 2008.

[5] Chakravarty S.R. and Majumder A., Inequality, polarization and welfare: theory and applications. *Australian Economic Papers*, 40(1) (2001), pp. 1-13.

[6] Davidson R., Flachaire E., Asymptotic and bootstrap inference for inequality and poverty measures, *Journal of Econometrics*, 141(1) (2007), pp. 141-166.

[7] Duclos J.-Y. and Araar A., *Poverty and Equity: Measurement, Policy, and Estimation with DAD*, Springer Publishing Company, New York, 2006.

[8] Duclos J.-Y., Esteban J. and Ray D., Polarization: concepts, measurement, estimation. *Econometrica*, 72(6) (2004), pp. 1737-1772.

[9] Duclos J.-Y., Esteban J. and Ray D., Polarization: concepts, measurement, estimation. in C. Barrett (ed.), *The Social Economics of Poverty: Identities, Groups, Communities and Networks*, Routledge, London, 2005, pp. 56-105.

[10] Duro J.A., Another look to income polarization across countries. *Journal of Policy Modeling*, 27(9) (2005), pp. 1001-1007.

[11] Duro J.A., International income polarization: a note. *Applied Economics Letters*, 12(12) (2005), pp. 759-762.

[12] Esteban J., *Economic polarization in the Mediterranean Basin*, Els Opuscles del CREI #10, Barcelona, 2002.

[13] Esteban J., Gradín C. and Ray D., An extension of a measure of polarization, with an application to the income distribution of five OECD countries. *Journal of Economic Inequality*, 5(1) (2007), pp. 1-19.

[14] Esteban J. and Ray D., On the measurement of polarization. *Econometrica*, 62(4) (1994), pp. 819-851.

[15] Esteban J. and Ray D., Conflict and distribution. *Journal of Economic Theory*, 87(2) (1999), pp. 379-415.

[16] Esteban J. and Ray D., Polarization, fractionalization and conflict. *Journal of Peace Research*, 45(2) (2008), pp. 163-182.

[17] Esteban J., Ray D., Linking conflict to inequality and polarization. UFAE and IAE Working Paper 766.09, 2009.

[12] The usual empirical finding is that the joint effect of stratification and clustering in samples design is to increase standard errors of the Gini index by considerable margin (see, e.g., [3]).

[18] Ezcurra R., Pascual P., Regional polarisation and national development in the European Union. *Urban Studies*, 44(1) (2007), pp. 99-122.
[19] Ezcurra R., Pascual P. and Rapún M., The dynamics of regional disparities in Central and Eastern Europe during transition. *European Planning Studies*, 15(10) (2007), pp. 1397-1421.
[20] Fedorov L., Regional inequality and regional polarization in Russia, 1990-99. *World Development*, 30(3) (2002), pp. 443-456.
[21] Gradín C., Polarization by sub-populations in Spain. *Review of Income and Wealth*, 46(4) (2000), pp. 457-474.
[22] Gradín C., Polarization and inequality in Spain: 1973-91. *Journal of Income Distribution*, 11(1) (2002), pp. 34-52.
[23] Gradín C. and Rossi M., Income distribution and income sources in Uruguay. *Journal of Applied Economics,* 9(1) (2006), pp. 49-69.
[24] Hussain M. A., The sensitivity of income polarization: time, length of accounting periods, equivalence scales, and income definitions, *Journal of Economic Inequality*, 7(3) (2008), pp. 207-223.
[25] Jenkins S. P., Did the middle class shrink during the 1980s? UK evidence from kernel density estimates. *Economics Letters*, 49(4) (1995), pp. 407-413.
[26] Kordos J., Lednicki B., Zyra M., The household sample surveys in Poland, *Statistics in Transition*, 5(4) (2002), pp. 555-589.
[27] Kot S. M., *Polaryzacja ekonomiczna. Teoria i zastosowanie*, Warszawa, PWN, 2008.
[28] Lasso de la Vega M. C. and Urrutia A. M., An alternative formulation of the Esteban-Gradín-Ray extended measure of polarization. *Journal of Income Distribution*, 15 (2006), pp. 42-54.
[29] Pagan A. and Ullah A., *Nonparametric Econometrics*, Cambridge University Press, New York, 1999.
[30] Rao C. R, *Linear Statistical Inference and Its Applications*, Second Edition, Wiley, New York, 1973.
[31] Ravallion M. and Chen S., What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review*, 11(2) (1997), pp. 357-382.
[32] Seshanna S. and Decornez S., Income polarization and inequality across countries: an empirical study. *Journal of Policy Modeling*, 25(4) (2003), pp. 335-358.
[33] Silverman B.W., *Density Estimatrion for Statistics and Data Analysis. Monographs on Statistics and Applied Probability*, Chapman and Hall, London, 1986.
[34] Wang Y.-Q. and Tsui K.-Y., Polarization orderings and new classes of polarization indices. *Journal of Public Economic Theory*, 2(3) (2000), pp. 349-363.
[35] Wolfson M. C., When inequalities diverge. *American Economic Review*, 84(2) (1994), pp. 353-358.
[36] Zhang X. and Kanbur R., What difference do polarisation measures make? An application to China, *Journal of Development Studies*, 37(3) (2001), pp. 85-98.

STATISTICAL INFERENCE ON CHANGES IN INCOME POLARIZATION IN POLAND

Summary

This paper estimates a popular measure of economic polarization (DER index) using Polish income micro-data from the Household Budget Survey (HBS) study for the period from 1998 to 2007. Using asymptotically distribution-free statistical inference we test whether the changes in the values of the estimated indices are statistically significant. Results show that during the period under study DER index has increased in the range from 4.9 to 6.5%, depending on the polarization-sensitivity parameter α. For the period 1998-2007, the changes in the DER index with $\alpha = 1$, for which polarization is empirically

distinguishable from inequality, are statistically significant at conventional levels. It is found, however, that for some sub-periods changes in the DER index are not statistically significant.

Keywords: polarization, income distribution, statistical inference, Poland

WNIOSKOWANIE STATYSTYCZNE O ZMIANACH POLARYZACJI DOCHODOWEJ W POLSCE

Streszczenie

W artykule dokonano estymacji popularnej miary polaryzacji ekonomicznej (indeks DER) dla danych o dochodach z polskich Badań Budżetów Gospodarstw Domowych w okresie 1998-2007. Używając niezależnych od rozkładu metod wnioskowania statystycznego przeprowadzono testy istotności różnic w wartościach estymowanych wskaźników. Wyniki pokazują, że w badanym okresie wartość indeksu DER wzrosła pomiędzy 4,9 a 6,5% w zależności od wartości parametru α mierzącego wagę przykładaną do polaryzacji. Zmiany indeksu DER dla $\alpha = 1$, dla którego polaryzacja zachowuje się empirycznie w sposób odmienny od nierówności, są w badanym okresie statystycznie istotne. W artykule pokazano jednak, że niektóre ze zmian indeksu DER dla krótszych okresów nie są istotne statystycznie.