

*Alina Jędrzejczak**

INCOME INEQUALITY IN SOCIOECONOMIC GROUPS IN THE PERIOD OF ECONOMIC TRANSFORMATION

1. INTRODUCTION

Household income is the main factor determining personal standard of living. In the period of economic transformation from a centrally-planned to a market economy we observed many important changes concerning income distributions in Poland. The changes were concentrated within socio-economic groups as well as in the relations between the groups.

We can distinguish six socio-economic groups of households taking into consideration the way of gaining their income:

1. Households of employees,
2. Households of employees-farmers,
3. Households of farmers,
4. Households of self-employed,
5. Households of pensioners and retirees,
6. Households maintained from non-earned sources.

The source of information on income distributions by socio-economic groups was individual data coming from the Household Budgets Survey conducted by the Central Statistical Office in the years 1999–2003. This survey is based on a random sample of households. Sample elements are selected by means of two-stage stratified sampling with rotation. Primary sampling units are selected with different probabilities. Overall sample size, that is the number of household examined annually, exceeds 30 thousand. Table 1 presents the changes in the structure of samples in the period 1999–2002.

* Ph. D., Chair of Statistical Methods, University of Łódź.

Table 1. Structure of Household Budgets sample in the years 1999–2002

Year	Total	Households					
		Employees	Employees-farmers	Farmers	Self-employed	Retirees and pensioners	Non-earned sources
1999	31429	13366	2878	1604	1997	10981	930
2000	36163	14509	2934	1710	2359	13004	1647
2001	31847	12611	2518	1318	2058	11711	1631
2002	32342	12449	2382	1273	2034	12160	2044

S o u r c e: Own calculations for all tables.

In the paper we focus the analysis on theoretical income distributions which approximate the empirical ones. Using a theoretical distribution instead of the empirical one can be useful for many reasons. From the point of view of income distributions the most important is that a good theoretical model can flatten irregularities resulting from the way of gathering information (nonresponses, misclassifications etc.).

2. CHANGES IN THE AVERAGE LEVEL OF INCOME IN POLAND IN THE PERIOD 1999–2002

Arithmetic mean is the measure most often used in comparisons of income and wage distributions in time and in space. The analysis of average level of income in different socio-economic groups can be the first step in the researches concerning income distributions in Poland in the period of economic transformation.

The results of the calculations are presented in the tables 2 and 3. The table 2 shows changes in the average household income by socio-economic groups in the period 1999–2002 while the table 3 includes the similar measures calculated for the real household income (the comparative period is 1999). It can be easily noticed that the average household income increased by 18.5% during the years 1999–2002 (see: Table 2). The highest growth was observed for the households of farmers and maintained from non-earned sources (38.8% and 37,73% respectively). Except for the households of farmers and pensioners the highest rates of growth was noticed in the year 2000. Income of farmers increased by 10% in the years 2000 and 2001 and by 15% in 2002. If is worth mentioning that the incomes of pensioners and retirees increased by 9% in 2001 while in the other years the rates of growth were very small (about 3%).

Comparing the results of the calculations presented in the table 3 with the results included in the previous table one can notice some important differences. Only for the households of farmers and maintained from non-earned sources the real income increased significantly in the period under consideration. What is important also the year average growth rate was positive only for these two groups of households. The average real income of the remaining household groups decreased year by year during the considered period,

Table 2. Average household income in socio-economic groups

Households	1999 = 100			Preceding year = 100			Year average growth rate
	2000	2001	2002	2000	2001	2002	
Total	108.94	115.00	118.52	108.94	105.56	103.06	4.3
Employees	110.96	116.33	117.85	110.96	104.84	101.31	4.2
Employees farmers	110.31	116.28	116.60	110.31	105.41	100.27	3.9
Farmers	110.68	120.77	138.80	110.68	109.11	114.93	8.5
Self-employed	110.99	112.88	117.77	110.99	101.71	104.33	4.2
Retirees and pensioners	103.48	112.98	117.19	103.48	109.17	103.73	4.1
Non-earned sources	120.70	129.09	137.73	120.70	106.95	106.69	8.3

Table 3. Average household income in socio-economic groups in constant prices

Households	1999 = 100			Preceding year = 100			Year average growth rate
	2000	2001	2002	2000	2001	2002	
Total	98.68	98.73	99.96	98.68	100.06	101.24	-0.02
Employees	100.15	99.61	99.03	100.15	99.47	99.42	-0.02
Employees farmers	99.74	100.22	99.17	99.74	100.48	98.95	-0.02
Farmers	99.90	104.21	118.18	99.90	104.32	113.41	+4.3
Self-employed	101.17	98.00	100.43	101.17	96.86	102.48	+0.1
Retirees and pensioners	93.74	96.80	98.28	93.74	103.43	101.53	-0.45
Non-earned sources	108.94	110.84	115.42	108.94	101.74	104.13	+5.0

3. CHANGES IN INCOME INEQUALITY WITHIN AND BETWEEN SOCIO-ECONOMIC GROUPS

The standard of living of a particular household depends not only on the average family income in a group but also on the level of income inequality. To complete the analysis it is necessary to take into consideration the following problems:

- changes in the level of concentration of income in socio-economic groups (inequality within distributions)
- comparison of economic situation of different socio-economic groups (inequality between distributions).

On the basis of the empirical data coming from the Household Budgets Survey the parameters of selected theoretical distributions were estimated. As theoretical distributions the lognormal and the Dagum models were applied. In the paper we analysed two kinds of income distribution: the first one was the distribution of family income and the second one the distributions of per-capita income.

The lognormal distribution has been widely used in wage and income distribution analysis for many years. The advantage of this distribution is its simplicity; a disadvantage, however, is its poor fitting to the data, especially in the tails. We say that a random variable Y is log-normally distributed when the logarithm of this variable is normally distributed. Hence the density function of Y is the following:

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\ln y - \mu)^2\right\}, \quad y > 0 \quad (1)$$

where:

μ – expected value of the logarithms of a random variable Y ,

σ – standard deviation of the logarithms of a random variable Y .

The formula (1) describes two-parameter lognormal distribution. The statistical characteristics of this distribution can be expressed by means of the parameters μ and σ . Besides the well known two-parameter model three and four parameter lognormal distributions are used in wage and income analysis. The third parameter τ determines the lowest level of positive income ($y > \tau$). The four-parameter distribution is characterised by four parameters: μ, σ, τ and θ , where θ determines the highest level of income ($\tau < y < \theta$).

Unlike the lognormal the Dagum model was based on empirical observations of income distributions in many countries. Dagum and Lemini (1977) noted that the function describing income elasticity of a cumulative distribution function of income is convex, decreasing and bounded. It can be described by the following differential equation:

$$\varepsilon(y, F(y)) = \frac{d \ln F(y)}{d \ln y} = \beta_1 [1 - [f(y)^{\beta_2}]], \quad (2)$$

for $y \geq 0$, $\beta_1, \beta_2 > 0$.

The cumulative distribution function of the Dagum model was the solution of the equation given by formula (2). It can be written as follows:

$$F(y) = \begin{cases} (1 + \lambda y^{-\delta})^{-\beta}, & y > 0 \\ 0, & y \leq 0 \end{cases} \quad (3)$$

for: $\beta, \lambda, \delta > 0$,

where: $\beta = 1/\beta_1$, $\delta = \beta_1\beta_2$, $\lambda = \exp c$,

c – constant of integration resulting from the solution of equation (2).

The density function corresponding to the cumulative distribution function given by (3) has the form:

$$f(y) = \begin{cases} \beta\lambda\delta y^{-\delta-1}(1 + \lambda y^{-\delta})^{-\beta-1}, & y > 0 \\ 0, & y \leq 0 \end{cases} \quad (4)$$

The parameter λ is a scale parameter of the Dagum model while β and δ are shape (inequality) parameters.

The three-parameter distribution given by (3) and (4) is called the Dagum type-I model. For data with null and negative or strictly positive incomes the four – parameter models proposed by Dagum should be useful. The distribution function of such the models can be expressed by the formula:

$$F(y) = a + (1 - \alpha)(1 + \lambda y^{-\delta})^{-\beta}, \text{ for } \beta, \lambda, \delta > 0 \quad (5)$$

The type of a distribution depends on the value of the fourth parameter α : when $\alpha = 0$ we obtain type I model for $0 < \alpha < 1$ type II model and for $\alpha < 0$ type III model. Parameter α taking values in the interval $(0,1)$ can be interpreted as an unemployment rate.

The lognormal model was chosen to the analysis of income distributions in Poland for its simplicity and clear economic interpretation of parameters. The Dagum model was chosen because of its properties giving better fitting to the data. Existence of null and negative values of observed family income was the reason for the application of the four-parameter models.

The parameters of the theoretical distributions were estimated by means of the maximum likelihood method. In order to find the maximum of log likelihood function an individual numerical procedure was applied.

The results of the estimation are presented in tables 4–9. The goodness-of-fit of the theoretical distributions with the empirical ones was measured by means of coefficients of distributions similarity (W_p) and standard deviations of squared errors (Sd). The best results were obtained for the four-parameter

Dagum model – most coefficients of distributions similarity exceed 0.99 for both total and per capita family income (see Tables 4, 6 and 8). The fourth parameter α usually took values below zero so we obtained the type III distributions. Only for the households of farmers and employees – farmers the parameter α takes values in the interval (0, 1) what is characteristic for the type II Dagum distribution. It is connected with significant amount of negative incomes observed in these household groups.

It is worth mentioning that the goodness-of-fit for the three-parameter Dagum distribution is also very high. For per capita incomes the differences between the compared distributions are negligible. The consistency of the lognormal model with the empirical distributions of per capita income is not satisfactory (Tables 6 and 8). Coefficients of distributions similarity usually do not exceed 0.95. The goodness-of-fit is extremely poor for the groups with high level of income inequality and heavy right tail (farmers).

The parameters of the theoretical model proposed by Dagum, with presents high consistency with empirical data, were then used to the estimation of concentration coefficients and economic distance ratios. Table 10 presents concentration coefficients proposed by Gini (1912), Zenga (1990) and Bonferroni (1930). The Gini ratio, based on the Lorenz curve, is the most popular concentration measure. It is said to be the best synthetic measure of income inequality while the coefficient derived by Bonferroni is more sensitive to changes in lower income groups. The Zenga measure corresponds to the concept of point concentration.

One can notice that almost all the concentration measures placed in the table 10 increased in the period under consideration. The highest level of income inequality was observed for the households of farmers ($G_{2003}=0.48$, $Z_{2003}=0.60$, $B_{2003}=0.55$). The level of income inequality is relatively low for households of retirees ($G_{2003}=0.24$) an pensioners ($G=2003_{=0.29}$). Figure 1 shows the Zenga curves, which can be regarded as point concentration measures calculated for per capita family income in the years 1999 and 2003. The presented concentration curves confirm that income inequality increased at each point of the distribution (The area under the Zenga curve can be interpreted as concentration area).

To complete the analysis, economic distance ratios D_1 for socio-economic groups were calculated. Economic distance d_1 (see Dagum, 1980) is defined as the weighted sum of income differences ($y - x$) given that $E(Y) > E(X)$. The weighting factor is a joint probability density function $f(y)f(x)$. Economic distance ratio D_1 is derived by normalizing d_1 in the unit interval. It is worth mentioning that D_1 measures not only differences in the average level of income but also is sensitive to any changes in the shape of the compared distributions. For example, the economic situation of the households of self-employed is by 28.67% better than the situation of the households of employees and by 70.25% better in comparison with the households of employees – farmers (see Table 11).

Table 4. Distributions consistency measures calculated for total household income

Year	Lognormal distribution		Dagum distribution			
			three-parameter		four-parameter	
	W_p	S_d	W_p	S_d	W_p	S_d
1999	0.9769	0.0051	0.9713	0.0077	0.9974	0.0005
2000	0.9747	0.0052	0.9762	0.0061	0.9932	0.0016
2001	0.9705	0.0061	0.9696	0.0077	0.9912	0.0020
2002	0.9722	0.0057	0.9724	0.0075	0.9979	0.0004
2003	0.9778	0.0043	0.9705	0.0070	0.9882	0.0032

Table 5. Parameters of the four-parameter Dagum model estimated for total household income

Year	Dagum distribution parameters			
	α	λ	β	δ
1999	-0.0334	11.7903	0.5803	3.5180
2000	-0.0268	11.8744	0.6152	3.3344
2001	-0.0308	23.0367	0.5413	3.6375
2002	-0.0290	16.7222	0.5852	3.4688
2003	-0.0210	11.9443	0.6874	3.1440

Table 6. Distributions consistency measures calculated for per capita income

Year	Lognormal distribution		Dagum distribution			
			three-parameter		four-parameter	
	W_p	S_d	W_p	S_d	W_p	S_d
1999	0.9532	0.0135	0.9879	0.0033	0.9851	0.0045
2000	0.9492	0.0132	0.9859	0.0037	0.9823	0.0049
2001	0.9639	0.0094	0.9888	0.0024	0.9929	0.0014
2002	0.9617	0.0104	0.9893	0.0027	0.9946	0.0012
2003	0.9600	0.0085	0.9904	0.0019	0.9806	0.0046

Table 7. Parameters of four-parameter Dagum model estimated for per capita income

Year	Dagum model parameters			
	α	λ	β	δ
1999	-0.0624	0.3024	0.5881	3.4444
2000	-0.0440	0.3813	0.6000	3.4181
2001	-0.0616	0.6484	0.5058	3.5576
2002	-0.0689	0.7256	0.04781	3.4854
2003	-0.0908	0.8625	0.4724	3.3548

Table 8. Distributions consistency measures for per capita income by socio-economic groups

Households	Year	Longnormal model		Dagum model			
				three-parameter		four-parameter	
		W_p	S_d	W_p	S_d	W_p	S_d
Employees	1999	0.9322	0.0165	0.9894	0.0025	0.9863	0.0039
	2000	0.9405	0.149	0.9887	0.0027	0.9630	0.0109
	2001	0.9509	0.0106	0.9868	0.0028	0.9695	0.0081
	2002	0.9525	0.0107	0.9760	0.0059	0.9830	0.0046
	2003	0.9456	0.0120	0.9784	0.0044	0.9658	0.0090
Employees- farmers	1999	0.9049	0.0291	0.9397	0.0164	0.9866	0.0027
	2000	0.9177	0.0230	0.9873	0.0029	0.9882	0.0027
	2001	0.9187	0.0203	0.9748	0.0051	0.9793	0.0045
	2002	0.9298	0.0208	0.9810	0.0041	0.9810	0.0042
	2003	0.9316	0.0159	0.9872	0.0029	0.9886	0.0026
Farmers	1999	0.8460	0.0441	0.9789	0.0046	0.9800	0.0046
	2000	0.8532	0.0414	0.9810	0.0035	0.9842	0.0031
	2001	0.8480	0.0338	0.9504	0.0106	0.9594	0.0083
	2002	0.8491	0.0419	0.9821	0.0033	0.9832	0.0032
	2003	0.8450	0.0359	0.9664	0.0063	0.9703	0.0060
Self-employed	1999	0.9273	0.0146	0.9758	0.0048	0.9732	0.0055
	2000	0.9420	0.0111	0.9739	0.0051	0.9724	0.0056
	2001	0.9489	0.0093	0.9564	0.0099	0.9731	0.0052
	2002	0.9446	0.0087	0.9570	0.0079	0.9570	0.0088
	2003	0.9147	0.0162	0.9476	0.0101	0.9656	0.0078
Retirees	1999	0.9230	0.0224	0.9640	0.0095	0.9748	0.0068
	2000	0.9306	0.0185	0.9715	0.0066	0.9795	0.0050
	2001	0.9292	0.0171	0.9809	0.0042	0.9886	0.0026
	2002	0.9276	0.0164	0.9844	0.0033	0.9845	0.0036
	2003	0.9300	0.0168	0.9866	0.0026	0.9867	0.0028
Pensioners	1999	0.9284	0.0214	0.9864	0.0030	0.9893	0.0023
	2000	0.9215	0.0256	0.9891	0.0028	0.9910	0.0022
	2001	0.9439	0.0172	0.9672	0.0080	0.9871	0.0026
	2002	0.9236	0.0245	0.9874	0.0025	0.9842	0.0037
	2003	0.9560	0.0139	0.9839	0.0033	0.9883	0.0024
Non-earned sources	1999	0.7635	0.0743	0.9550	0.0112	0.9748	0.0051
	2000	0.7468	0.0816	0.9859	0.0030	0.9878	0.0029
	2001	0.7798	0.0658	0.9427	0.0151	0.9502	0.0139
	2002	0.8297	0.0514	0.9589	0.0091	0.9627	0.0082
	2003	0.8423	0.0462	0.9655	0.0084	0.9706	0.0072

Table 9. Parameters of the four-parameter Dagum model estimated for per capita income by socio-economic groups

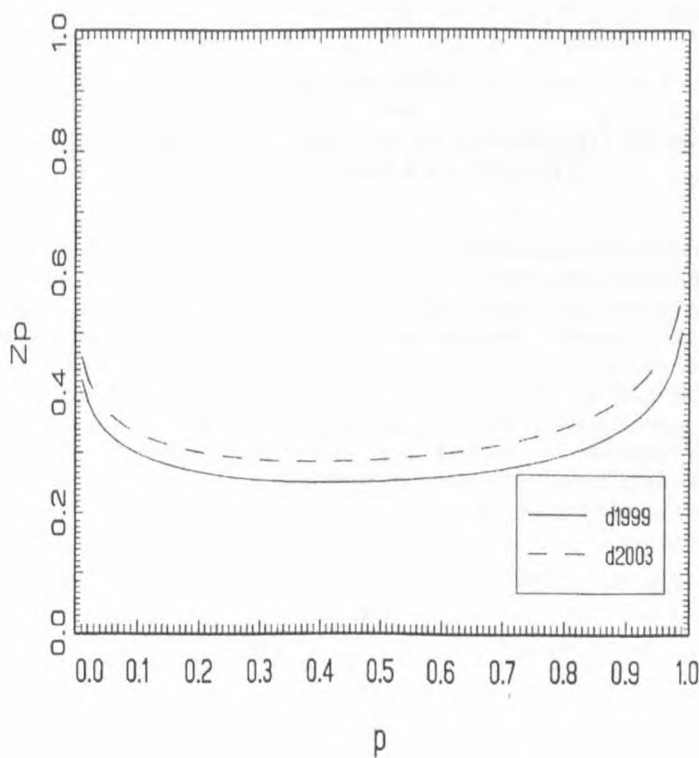
Households	Year	Dagum model parameters			
		α	λ	β	δ
Employees	1999	-0.6708	0.6732	0.2097	3.3739
	2000	-0.6352	0.9463	0.2093	3.2971
	2001	-0.2152	0.8585	0.3740	3.3127
	2002	-0.3842	1.2701	0.2540	3.4108
	2003	-0.3679	1.5904	0.2673	3.2050
Employees- farmers	1999	-0.9999	0.1606	0.2415	3.1982
	2000	0.0161	0.1030	0.9005	3.6050
	2001	-0.9497	0.3495	0.1823	3.5571
	2002	-0.0264	0.1187	0.8156	3.4353
	2003	0.0056	0.1000	1.2663	3.2754
Farmers	1999	-0.0929	0.1580	0.9235	2.0650
	2000	0.1789	0.1077	2.0003	2.3664
	2001	-0.9999	0.4009	0.2899	1.9757
	2002	-0.1447	0.4188	0.4310	2.4709
	2003	-0.4284	0.1503	0.8078	1.8248
Self-employed	1999	-0.0210	0.3478	1.0094	2.7249
	2000	0.0107	0.2853	1.4236	2.5475
	2001	-0.0663	0.8802	0.6420	2.9738
	2002	-0.1413	2.1786	0.3729	3.1705
	2003	-0.1990	1.5155	0.4302	2.5176
Retirees	1999	0.0119	0.1875	1.0866	4.5700
	2000	0.0085	0.2705	0.9782	4.3857
	2001	0.0060	0.5805	0.8148	4.35865
	2002	0.0063	0.5563	0.9142	4.31349
	2003	0.0005	0.5865	0.9493	4.2752
Pensioners	1999	0.0342	0.1245	0.7628	4.0586
	2000	-0.0067	0.2422	0.4973	4.3793
	2001	-0.2208	0.7127	0.2493	4.5681
	2002	-0.0343	0.4910	0.3878	4.5585
	2003	-0.0210	0.4521	0.5382	4.0533
Non-earned sources	1999	-0.9998	0.1000	0.2993	2.4635
	2000	-0.7161	0.1842	0.2199	2.7204
	2001	-0.5526	0.5881	0.1494	3.5352
	2002	-0.8597	0.2955	0.2232	2.6032
	2003	-0.4614	0.4222	0.2726	2.7018

Table 10. Gini, Zenga and Bonferroni concentration measures for per capita income by socio-economic groups

Households	Year	Dagum model parameters		
		G	Z ₁	B
Total	1999	0.3150	0.2948	0.4291
	2000	0.3216	0.3060	0.4374
	2001	0.3214	0.3057	0.4386
	2002	0.3323	0.3244	0.4515
	2003	0.3390	0.3357	0.4551
Employees	1999	0.3277	0.3178	0.4324
	2000	0.3384	0.3367	0.4433
	2001	0.3371	0.3332	0.4457
	2002	0.3456	0.3483	0.4545
	2003	0.3641	0.3824	0.4713
Employees-farmers	1999	0.3071	0.2817	0.4215
	2000	0.2902	0.2543	0.4012
	2001	0.2953	0.2628	0.4003
	2002	0.2983	0.2672	0.4074
	2003	0.2926	0.2588	0.3949
Farmers	1999	0.4715	0.5823	0.5790
	2000	0.4381	0.5149	0.5629
	2001	0.4843	0.6114	0.5850
	2002	0.4563	0.5483	0.5807
	2003	0.4810	0.6025	0.5849
Self-employed	1999	0.3602	0.3745	0.4692
	2000	0.3704	0.3928	0.4808
	2001	0.3547	0.3643	0.4649
	2002	0.3784	0.4058	0.9952
	2003	0.4396	0.5223	0.5485
Retirees	1999	0.2209	0.1528	0.3173
	2000	0.2336	0.1700	0.3336
	2001	0.2342	0.1712	0.03373
	2002	0.2501	0.1933	0.3548
	2003	0.2377	0.1756	0.3357
Pensioners	1999	0.2787	0.2371	0.3934
	2000	0.2842	0.2468	0.4014
	2001	0.2957	0.2645	0.4141
	2002	0.2929	0.2608	0.4108
	2003	0.2902	0.2554	0.4072
Non-earned sources	1999	0.4135	0.4703	0.5400
	2000	0.4228	0.4875	0.5497
	2001	0.4109	0.4573	0.5301
	2002	0.4100	0.4636	0.5282
	2003	0.4188	0.4794	0.5420

Table 11. Economic distance ratios between socio-economic groups in 2003

Households	Employees	Retirees	Farmers	Pensioners	Non-earned sources	Employees farmers
Self-employed	0.2863	0.2959	0.5316	0.5841	0.6038	0.7025
Employees		0.0350	0.3394	0.3584	0.5115	0.5330
Retirees			0.3912	0.4624	0.5822	0.6353
Farmers				0.0692	0.1576	0.2121
Pensioners					0.2604	0.2363
Non-earned sources						0.0712

**Fig. 1.** Zenga cocentration curves for per capita income in the years 1999 and 2003

REFERENCES

- Bonferroni C. E. (1930), *Elementi di Statistica Generale*, Liberia Seber, Firenze.
- Dagum C. (1977), *A New Model of Personal Income Distribution. Specification and Estimation*, "Economie Appliquee" XXX(3), pp. 413–436.
- Dagum C. (1980), *Inequality Measures Between Income Distribution with Application*, "Econometrica" 48, pp. 1790–1803.
- Dagum C., Lemmi A. (1989), *A Contribution to the Analysis of Income Distribution and Income Inequality and a Case Study: Italy*, [in:] D. I. Slottjee, *Advances in Econometrics*, Jai Press Greenwich.
- Gini C. (1912), *Variabilita e Mutabilita*, Studi Economicogiuridici, Universita di Cagliari III, 2a.
- Gini C. (1939), *Memorie di Metodologia Statistica*, Milano, pp. 397, 474–520.
- Jędrzejczak A. (1993), *Application of the Dagum Distribution in the Analysis of Income Distributions in Poland*, "Acta Universitatis Lodzianensis, Folia Oeconomica" 131, pp. 103–111.
- Zenga M. (1990), *Concentration Curves and Concentration Indexes Derived from them*, [in:] *Income and Wealth Distribution, Inequality and Poverty*, Springer-Verlag, Berlin, pp. 94–110.

Alina Jędrzejczak

**NIERÓWNOŚCI DOCHODOWE W GRUPACH SPOŁECZNO-ZAWODOWYCH
W POLSCE W OKRESIE TRANSFORMACJI**

Okres transformacji gospodarczej, który rozpoczął się w Polsce w 1989 r., wiąże się z narastaniem nierówności dochodowych w różnych przekrojach. Powoduje to m.in. zwiększanie się dystansu ekonomicznego pomiędzy niektórymi grupami społeczno-ekonomicznymi oraz wzrost koncentracji płac i dochodów wewnątrz tych grup. W efekcie tych zmian poziom nierówności rozkładu dochodów w Polsce jest obecnie bardzo wysoki w porównaniu z innymi krajami Wspólnoty Europejskiej. W artykule przedstawione zostały miary koncentracji dochodów w grupach społeczno-zawodowych w ujęciu dynamicznym oraz współczynniki dystansu ekonomicznego pomiędzy grupami. Miary te były szacowane na podstawie parametrów rozkładów teoretycznych dopasowanych do empirycznych rozkładów dochodów pochodzących z Badania Budżetów Gospodarstw Domowych.