

DOCHODY INDYWIDUALNE A DOCHODY GOSPODARSTW DOMOWYCH W POLSCE W LATACH 2000 – 2010

Piotr Łukasiewicz, Krzysztof Karpio, Arkadiusz Orłowski
Katedra Informatyki
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mails: piotr_lukasiewicz@sggw.pl, krzysztof_karpio@sggw.pl,
arkadiusz_orlowski@sggw.pl

Streszczenie: W pracy przedstawiono wyniki badań dotyczących zależności pomiędzy rozkładami dochodów indywidualnych i dochodów gospodarstw domowych w Polsce w okresie 2000–2010. W USA rozkład dochodów gospodarstw domowych można wyrazić poprzez splot rozkładów dochodów indywidualnych członków rodziny. Jak wynika z wcześniejszych badań, w Polsce nie obserwujemy takiej zależności. W zasadzie, mamy bardzo małą wiedzę o rodzaju stosunków między tymi kategoriami dochodów naszym kraju. Przypuszcza się, że w USA dochody osobiste członków rodzin są zmiennymi niezależnymi. Z drugiej strony, w Polsce są na pewno zmiennymi zależnymi. W tej pracy stosujemy technikę splotów, a rozkłady dochodów przybliżamy trójparametrycznym modelem Daguma.

Słowa kluczowe: rozkład dochodów, splot rozkładów, model Daguma

WSTĘP

Z badań Budżetów Gospodarstw Domowych z lat 2000-2010 wynika, że w grupie polskich gospodarstw domowych rodziny pozostające na utrzymaniu dwóch osób (w tym większość małżeństw) stanowią ok. 50%-52% wszystkich rodzin, a gospodarstwa będące na utrzymaniu jednej osoby ok. 28%-30%. Łącznie te dwie grupy stanowią przeważającą, ok. 80% grupę wszystkich rodzin. Interesujące jest zbadanie związku pomiędzy rozkładami dochodów tych dwóch grup gospodarstw domowych. Metoda badawcza przyjęta w niniejszej pracy wynika z prostej obserwacji. Dochód z gospodarstw domowych pozostających na utrzymaniu dwóch osób jest sumą dochodów indywidualnych x i y . Przyjmijmy, że

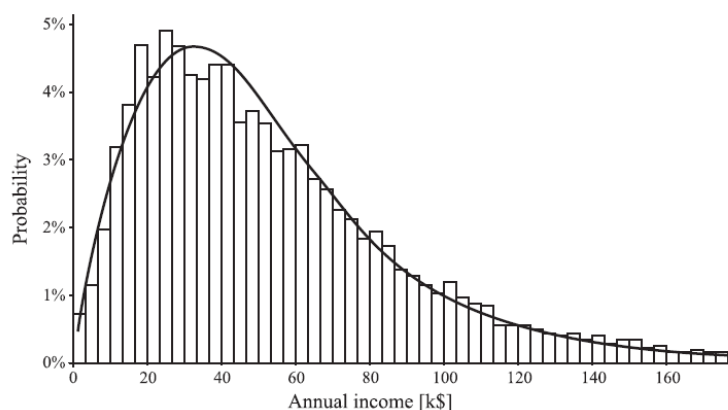
Z jest zmienną losową o wartościach z , a X i Y zmiennymi losowymi odpowiednio o wartościach x i y . Wówczas $Z = X + Y$. Przyjmijmy dalej, że f_X i f_Y są funkcjami gęstości zmiennych losowych X , Y i że zmienne te są niezależne. Wówczas, funkcja gęstości f_Z zmiennej losowej Z jest splotem funkcji f_X i f_Y oraz wyraża się wzorem [Plucińska, Pluciński 2000]

$$f_Z(z) = (f_X * f_Y)(z) = \int_R f_X(t) f_Y(z-t) dt. \quad (1)$$

Zauważmy, że równość $f_Z = f_X * f_Y$ nie implikuje niezależności zmiennych losowych, natomiast brak równości implikuje zależność tych zmiennych.

W przypadku dochodów amerykańskich obserwujemy, że rozkład dochodów rodzin z dwoma osobami pracującymi jest splotem rozkładów dochodów indywidualnych. Pokazano to w pracy [Drăgulescu, Yakovenko 2001] oraz potwierdzono w [Łukasiewicz, Orłowski 2004] dla danych dochodowych z roku 2000 – Rys. 1. Związek ten obserwuje się także w kolejnych latach, co wskazuje, że można go uważać za stałą cechę dochodów amerykańskich. Jednocześnie w pracy [Łukasiewicz, Orłowski 2004] pokazano, że w Polsce taka zależność nie istnieje. Model (1) jest wyraźnie przesunięty względem rozkładu empirycznego w kierunku wyższych dochodów – Rys. 2.

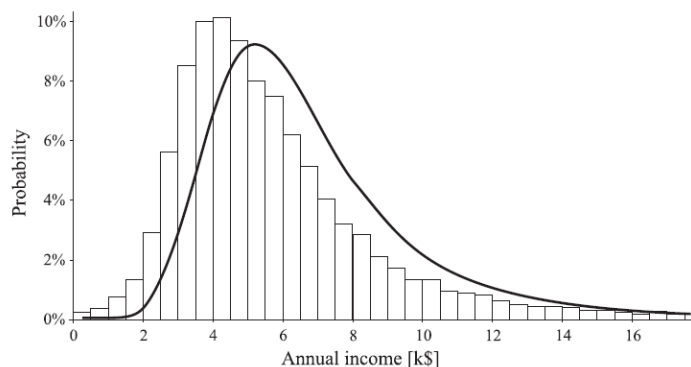
Rysunek 1. Empiryczny rozkład dochodów rodzin z dwiema osobami pracującymi w USA w roku 2000 oraz splot rozkładów indywidualnych (linia ciągła)



Źródło: [Łukasiewicz, Orłowski 2004]

W niniejszej pracy przedstawiono wyniki badań dla polskich gospodarstw domowych dla lat 2000–2010. Przeprowadzono analizy w dwóch zakresach: 1) pełnym, w odniesieniu do wszystkich rodzin utrzymywanych przez dwie osoby, 2) wąskim, dotyczącym tylko małżeństw bez dzieci.

Rysunek 2. Empiryczny rozkład dochodów rodzin z dwiema osobami pracującymi w Polsce w roku 2000 oraz splot rozkładów indywidualnych (linia ciągła)



Źródło: [Łukasiewicz, Orłowski 2004], przyjęto 1 USD = 4,143 zł

DANE I ORGANIZACJA BADANIA

Analizy oparto na danych dochodowych pochodzących z badań Budżetów Gospodarstw Domowych (BGD) z okresu 2000–2010. Zmiany w rozkładach dochodów obserwowane w kolejnych latach są niewielkie. Dane pobrano z odstępem dwuletnim, biorąc pod uwagę sześć lat parzystych. Dostępne dane dochodowe to miesięczne dochody rozporządzalne gospodarstw domowych, tzn. dochody netto, które przeznaczone są na wydatki oraz przyrost oszczędności. Dochód rozporządzalny każdego gospodarstwa jest równy sumie dochodów netto wszystkich jego członków. Dane pozwalają na analizowanie łącznego dochodu gospodarstwa (ewentualnie przeliczonego na osobę); w bazach danych BGD dochody poszczególnych członków gospodarstwa nie są udostępniane. Przed analizą dochody miesięczne przeliczone zostały na dochody roczne.

Gospodarstwa podzielono na grupy przyjmując za jedyne kryterium liczbę osób wnoszących dochód do gospodarstwa. Osoby te to nie tylko osoby pracujące, ale również osoby posiadające dochody w formie emerytury, renty, zasiłków, dochodów z wynajmu lokali, itp. W prezentowanym badaniu rozpatrywano następujące zbiory gospodarstw domowych:

- $R^{(2)}$ – rodziny utrzymywane przez dwie osoby;
- M – małżeństwa bez dzieci, w których każdy z małżonków osiąga dochody;
- $R^{(1)}$ – rodziny utrzymywane przez jedną osobę;
- S – gospodarstwa jednoosobowe.

Oczywiście $M \subset R^{(2)}$ oraz $S \subset R^{(1)}$.

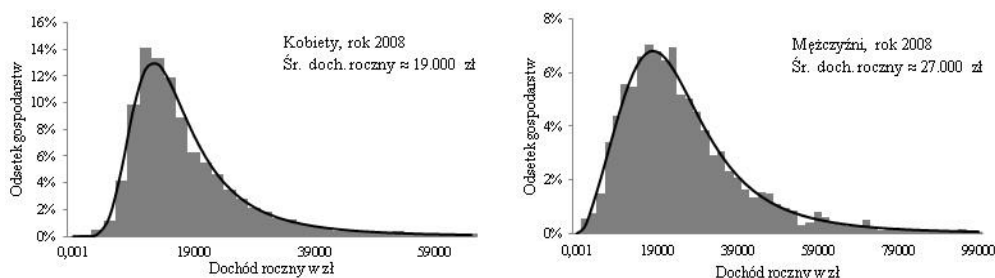
Zbiory $R^{(1)}$ i S posłużyły konstrukcji rozkładów dochodów indywidualnych. Ponieważ rozkłady dochodów kobiet i mężczyzn wykazują we wszystkich latach duże różnice (Rys. 3) dokonano, więc dodatkowych podziałów wg płci:

- $R^{(1)} = R_k^{(1)} \cup R_m^{(1)}$, gdzie $R_k^{(1)}$, $R_m^{(1)}$ – rodziny utrzymywane odpowiednio przez kobiety, mężczyznę;
 - $S = S_k \cup S_m$, gdzie S_k , S_m – samotne kobiety, samotni mężczyźni.
- Liczebności wszystkich prób zamieszczono w Tabeli 1.

Tabela 1. Liczebności prób

Rok	$R_k^{(1)}$	$R_m^{(1)}$	$R^{(2)}$	S_k	S_m	M
2000	5.638	4.661	18.357	3.905	1.185	5.854
2002	5.215	4.438	15.460	3.671	1.377	5.080
2004	5.277	4.466	15.578	3.654	1.336	5.008
2006	6.102	4.840	18.095	4.377	1.644	6.248
2008	6.176	4.468	18.111	4.775	1.774	6.938
2010	6.106	4.804	18.472	4.804	1.856	7.507

Źródło: obliczenia własne

Rysunek 3. Rozkłady dochodów kobiet i mężczyzn w roku 2008 (rozkłady dochodów gospodarstw utrzymywanych przez kobiety $R_k^{(1)}$ i mężczyznę $R_m^{(1)}$). Linia ciągła – wykres funkcji aproksymującej (2)

Źródło: opracowanie własne

Przebieg badania dla każdego roku był następujący.

- 1) W oparciu o zbiory $R_k^{(1)}$, $R_m^{(1)}$ skonstruowano unormowane rozkłady dochodów kobiet i mężczyzn. Rozkłady te aproksymowano następnie funkcją gęstości Daguma, która określona jest wzorem

$$f(x) = \frac{abc}{x^{b+1}(1+ax^{-b})^{c+1}}, \quad (2)$$

gdzie $a > 1$, $b > 0$, $c > 0$ są parametrami modelu, a $x > 0$ jest dochodem.

Model (2) opisany jest w pracach [Dagum 1977] i [Łukasiewicz, Orłowski 2003]. Charakteryzuje się bardzo wysoką zgodnością z empirycznymi rozkładami dochodów. Stosowany jest często w empirycznych badaniach dochodów, por. [Dagum, Lemmi 1987], [Stoppa 1995], [Bandourian et al 2002], [Quintano, D'Agostino 2006]. Parametry funkcji (2) oszacowano metodą największej wiarygodności w oparciu o dane indywidualne [por.

Łukasiewicz, Orłowski 2003]. Otrzymano w ten sposób teoretyczne rozkłady dochodów indywidualnych kobiet i mężczyzn: f_k oraz f_m . Stopień zgodności z rozkładami empirycznymi zmierzono za pomocą wskaźników SSE i W_r [Bandourian et al 2002]:

$$SSE = \sum_{i=1}^k \left[\frac{n_i}{n} - p_i \right]^2, \quad (3)$$

$$W_r = \left(1 - \frac{1}{2} SAE \right) \times 100\%, \quad SAE = \sum_{i=1}^k \left| \frac{n_i}{n} - p_i \right|, \quad (4)$$

gdzie n_i / n oznacza częstości empiryczne, p_i częstości teoretyczne, natomiast k jest liczbą przedziałów histogramu.

- 2) Wyznaczono spłot rozkładów f_k i f_m zgodnie z formułą (1), który w przypadku funkcji Daguma i przy założeniu $x > 0$ przybiera postać

$$(f_k * f_m)(x) = \int_0^x \frac{a_1 b_1 c_1 \cdot a_2 b_2 c_2}{t^{b_1+1} (x-t)^{b_2+1} (1+a_1 t^{-b_1})^{c_1+1} (1+a_2 (x-t)^{-b_2})^{c_2+1}} dt, \quad (5)$$

gdzie a_1, b_1, c_1 oraz a_2, b_2, c_2 to parametry oszacowanych modeli, a $x > 0$ jest dochodem. Całkę (5) obliczano numerycznie.

- 3) Model $f_k * f_m$ porównano z empirycznym rozkładem dochodów rodzin grupy $R^{(2)}$. Wykorzystano w tym celu wskaźniki SSE i W_r .
- 4) Powtórzono 1) i 2) dla grup S_k, S_m , a otrzymany spłot porównano z rozkładem dochodów małżeństw bez dzieci.

WYNIKI

Rozkłady dochodów indywidualnych przybliżono za pomocą modeli Daguma (2). Łącznie oszacowano 24 funkcje, po 4 dla każdego roku. Nie zamieszczamy szczegółowych informacji o poszczególnych modelach, nadmienimy jedynie, że w prezentowanym badaniu wartości wskaźnika SSE obserwowano w granicach 0,00042 – 0,0013, a wskaźnika W_r w granicach od 93% do 96%. Oszacowane modele charakteryzują się wysoką i bardzo wysoką zgodnością z rozkładami empirycznymi.

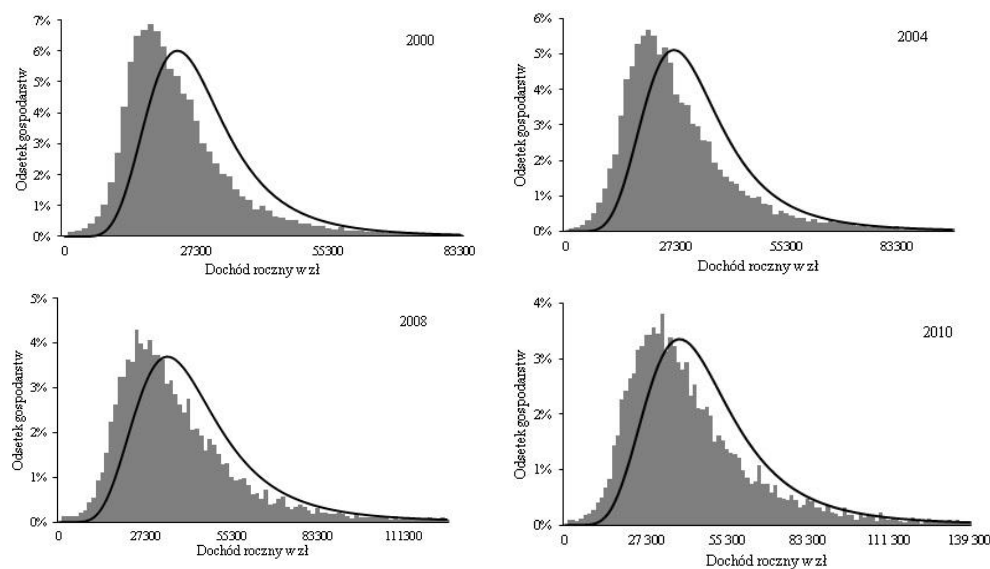
Wykorzystując formułę (5) dokonano spłotu funkcji Daguma przybliżających rozkłady dochodów kobiet i mężczyzn (grupy $R_k^{(1)}, R_m^{(1)}$). Otrzymano w ten sposób modele dochodów rodzin będących na utrzymaniu dwóch osób. Wyniki porównano z rozkładami empirycznymi (grupa $R^{(2)}$). Rezultaty dla wybranych lat pokazano na Rys. 4. W Tabeli 2 zamieszczono wartości

wskaźników zgodności SSE i W_r , których zmiany w czasie pokazano również na Rys. 6.

We wszystkich latach mamy do czynienia ze znacznym przesunięciem rozkładów teoretycznych w kierunku wyższych dochodów, natomiast kształt wykresów w przybliżeniu odpowiada rozkładowi empirycznemu. Ogólnie można powiedzieć, że dochody teoretyczne rodzin będących na utrzymaniu dwóch osób są wyższe od empirycznych. Jednocześnie, w całym okresie 2000-2010 obserwujemy spadek wskaźnika SSE i wzrost W_r , co wskazuje na systematyczne zmniejszanie się różnicy pomiędzy splotem, a rozkładem empirycznym.

W drugim etapie badań dokonano zawężenia danych, wykluczając dodatkowe, nieprzynoszące dochodu osoby w gospodarstwach. Dochody rodzin z jedną osobą pracującą zastąpiono dochodami rodzin jednoosobowych kobiet i mężczyzn (grupy S_k , S_m). Splot funkcji Daguma tych rozkładów porównano z rozkładem dochodów małżeństw bez dzieci. Wyniki przedstawiono na Rys. 5, w Tabeli 2 i na Rys. 7. W przypadku dochodów małżeństw również obserwujemy przesunięcie rozkładu teoretycznego w kierunku wyższych dochodów, ale znacznie mniejsze niż w przypadku ogólnym. Wartości wskaźnika SSE są znacznie niższe, a wartości wskaźnika W_r osiągają granicę 90%. Splot rozkładów indywidualnych nie opisuje dokładnie rozkładu dochodów małżeństw, jednak można uznać, że jest do niego bardzo zbliżony. Jednocześnie widzimy, że niezgodność ta maleje w czasie.

Rysunek 4. Rozkłady dochodów rodzin na utrzymaniu dwóch osób oraz splot rozkładów dochodów indywidualnych – model (5)



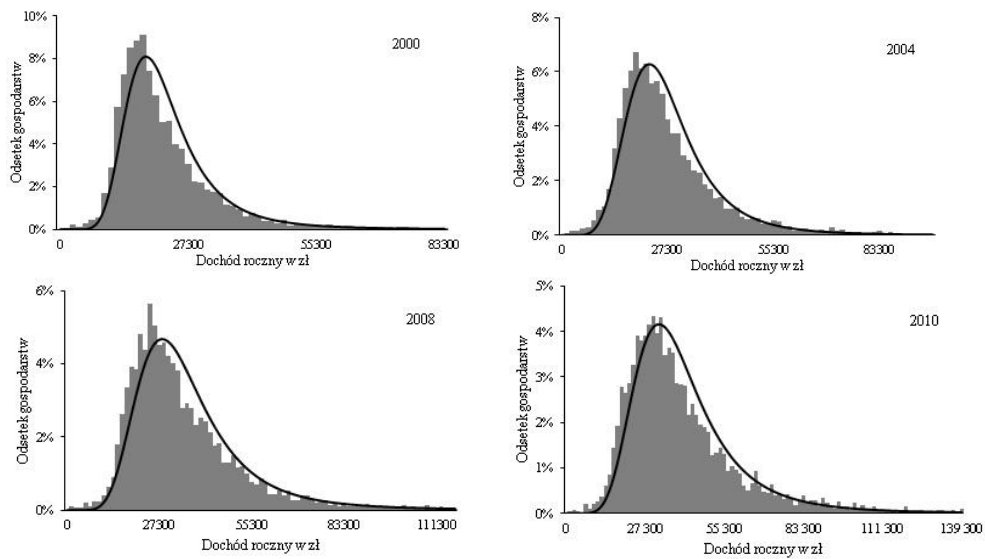
Źródło: opracowanie własne

Tabela 2. Zmiany wskaźników SSE i W_r

Rok	Rodziny na utrzym. dwóch osób		Małżeństwa bez dzieci	
	SSE	W_r	SSE	W_r
2000	0,0089	76,1%	0,0028	88,2%
2002	0,0070	77,9%	0,0020	89,0%
2004	0,0063	78,2%	0,0013	90,4%
2006	0,0058	77,8%	0,0021	88,4%
2008	0,0036	80,8%	0,0015	89,4%
2010	0,0031	81,5%	0,0010	90,1%

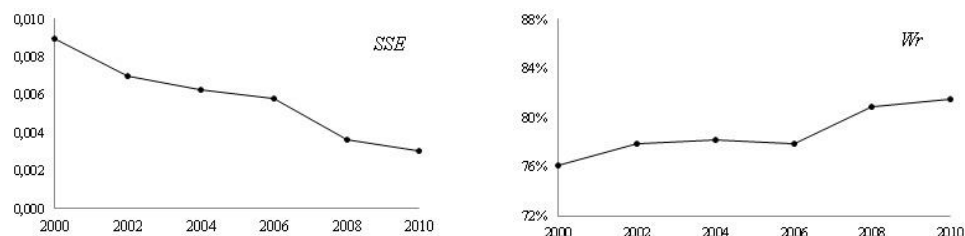
Źródło: obliczenia własne

Rysunek 5. Rozkłady dochodów małżeństw bez dzieci oraz splot rozkładów dochodów indywidualnych – model (5)



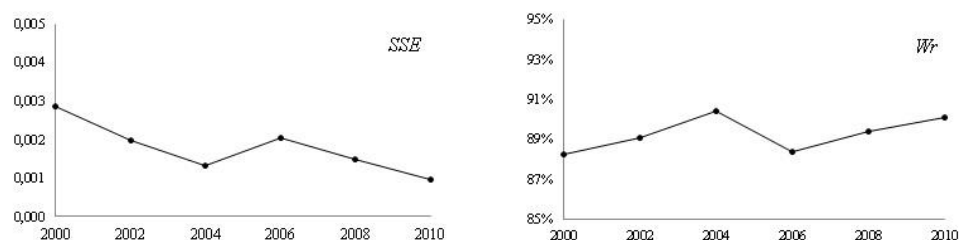
Źródło: opracowanie własne

Rysunek 6. Zmiany wskaźników SSE i W_r . Stopień zgodności rozkładów dochodów rodzin na utrzymaniu dwóch osób z modelem (5)



Źródło: opracowanie własne

Rysunek 7. Zmiany wskaźników SSE i W_r . Stopień zgodności rozkładów dochodów małżeństw bez dzieci z modelem (5)



Źródło: opracowanie własne

PODSUMOWANIE

W pracy porównano dochody gospodarstw domowych utrzymywanych przez dwie osoby ze splotem dochodów gospodarstw z jedną osobą zarabiającą. Zaobserwowano wyraźne, lecz malejące w okresie 2000 – 2010 różnice. Dochody otrzymane w wyniku splotu są wyższe dla wszystkich lat. Posiadanie rodziny skutkuje zmniejszeniem się łącznych dochodów, prowadzi do pewnej degradacji dochodowej. Otrzymane wyniki jednoznacznie wskazują, iż dochody członków rodzin w Polsce są zależne. Pokazano, że stopień tej zależności jednak systematycznie maleje.

Poszukując przyczyny zaobserwowanych różnic dokonano zawężenia danych, wykluczając nieprzynoszące dochodu osoby w gospodarstwach. Dochody małżeństw bez dzieci porównano ze splotem dochodów gospodarstw jednoosobowych. Również w tym przypadku dochody otrzymane dla splotu są wyższe w całym analizowanym okresie czasu. Zaobserwowane różnice są jednak wyraźnie mniejsze. Dodatkowo obserwujemy zmniejszanie się tych różnic na przestrzeni lat.

Otrzymane wyniki wskazują, że zawarcie małżeństwa i założenie wspólnego gospodarstwa domowego skutkuje pewnym zmniejszeniem łącznych dochodów. Ich dalsza degradacja następuje, gdy wzrasta liczebność gospodarstwa domowego.

Z drugiej strony, zaobserwowane różnice dochodowe na przestrzeni lat stopniowo maleją.

BIBLIOGRAFIA

- Bandourian R., McDonald J.B., Turley R.S. (2002) A comparison of parametric models of income distribution across countries and over time, Department of Economics, Brigham Young University.
- Dagum C. (1977) A new model of personal income distribution: Specification and Estimation. *Economic Appliquee* XXX (3), 413-437.
- Dagum C., Lemmi A. (1987) A Contribution to the Analysis of Income Distribution and Income Inequality and a Case Study: Italy, *Econometric Society Meeting*, Chicago.
- Drăgulescu A., Yakovenko V.M. (2001) Evidence for the exponential distribution of income in the USA. *The European Physical Journal B* 20, 585-589.
- Łukasiewicz P., Orłowski A. (2004) Probabilistic models of income distributions, *Physica A* 344, 146-151.
- Łukasiewicz P., Orłowski A. (2003) Probabilistic models of income distributions of polish households, *Metody ilościowe w badaniach ekonomicznych III*, Wyd. SGGW, 122-130.
- Plucińska A., Pluciński E. (2000) *Probabilistyka*, WNT, Warszawa.
- Quintano C., D'Agostino A. (2006) Studying inequality in income distribution of single-person households in four developed countries. *Review of Income and Wealth*, 52, 525-546.
- Stoppa G. (1995) Explicit estimators for income distributions, in: Dagum C., Lemmi, A. (eds.): *Research on Economic Inequality, Vol. 6: Income Distribution, Social Welfare, Inequality and Poverty*. Greenwich, CT: JAI Press, 393-405.

PERSONAL INCOMES VS HOUSEHOLDS INCOMES IN POLAND IN THE YEARS 2000 – 2010

Abstract: In this paper we study dependencies between distributions of personal and households incomes in Poland in 2000 to 2010. For example, the distributions of households incomes in the USA can be obtained by a convolution of distributions of personal incomes of family members. As was derived from our previous studies for 2004, there is no such a simple dependency in Poland. In principle, there is a very little knowledge about a type of relationship between those incomes in Poland. In the case of the USA personal incomes of family members are probably independent of each other. On the other hand, those variables are for sure dependent for Poland. In this paper we study personal and households incomes in Poland, their changes in time and their mutual relationship. Incomes distributions have been described by the three-parameter Dagum function.

Keywords: income distribution, convolution of distributions, Dagum's model