

**Anna Turczak**

---

Zachodniopomorska Szkoła Biznesu w Szczecinie

---

## **Analiza zależności między wartością dochodu na osobę i wielkością gospodarstwa domowego w latach 2006–2013**

### **Streszczenie:**

Celem artykułu jest znalezienie modelu, który najlepiej opisuje zależność między poziomem dochodu rozporządzalnego na osobę a wielkością gospodarstwa domowego. Sformułowano dwie hipotezy badawcze. Pierwsza z nich głosi, że funkcja liniowa adekwatnie obrazuje zależność między rozpatrywanymi zmiennymi. Z kolei druga hipoteza stanowi, iż wzrost liczby osób w gospodarstwie domowym powoduje zazwyczaj stały spadek średniego dochodu rozporządzalnego na osobę. Obydwie hipotezy zostały zweryfikowane pozytywnie.

**Słowa kluczowe:** wielkość gospodarstwa domowego, dochód rozporządzalny, liniowy model przyczynowo-opisowy, model regresji, weryfikacja modelu ekonometrycznego

---

### **Wprowadzenie**

---

Badanie budżetów gospodarstw domowych w Polsce pozwala na analizę poziomu życia poszczególnych grup mieszkańców. W szczególności badanie takie daje możliwość dokonania porównania wartości dochodu na osobę między określonymi grupami ludności oraz przeprowadzenia oceny wpływu różnych czynników na kształtowanie się poziomu i dyspersję tego dochodu. Zatem dochód rozporządzalny *per capita* jest wielkością zróżnicowaną i zależy między innymi od wielkości gospodarstwa domowego mierzonej liczbą osób wchodzących w skład gospodarstwa. Toteż jako główny cel niniejszego artykułu obrano poszukiwanie modelu, który adekwatnie opisywałby zależność między poziomem dochodu rozporządzalnego na osobę a wielkością gospodarstwa domowego. Dzięki temu badaniu można będzie określić, czy zależność między wspomnianymi zmiennymi ma charakter liniowy, czy może funkcja liniowa nie jest tutaj właściwą aproksymantą. Analizę przeprowadzono dla każdego roku oddzielnie od 2006 do 2013 włącznie.

Niewątpliwie proces modelowania ekonometrycznego przeprowadzony w niniejszym artykule należałoby poprzedzić precyzyjnym zdefiniowaniem badanej cechy statystycznej. Otóż dochód rozporządzalny będący tutaj przedmiotem analizy to suma bieżących dochodów gospodarstwa domowego z poszczególnych źródeł pomniejszona o zaliczki na podatek dochodowy od osób fizycznych płacone przez płatnika w imieniu podatnika (od dochodów z pracy najemnej oraz od niektórych świadczeń z ubezpieczenia społecznego i pozostałych świadczeń społecznych), o podatki od dochodów z własności, podatki płacone przez osoby pracujące na własny rachunek, w tym przedstawicieli wolnych zawodów i osób użytkujących gospodarstwo indywidualne w rolnictwie, oraz o składki na ubezpieczenia społeczne i zdrowotne. W skład dochodu rozporzadzalnego wchodzi dochody pieniężne i niepieniężne, w tym spożycie naturalne (tj. towary i usługi konsumpcyjne pobrane na potrzeby gospodarstwa domowego z gospodarstwa indywidualnego w rolnictwie bądź prowadzonej działalności gospodarczej na własny rachunek) oraz towary i usługi otrzymane nieodpłatnie.

Dochód rozporządzalny przeznaczany jest przez gospodarstwo domowe na wydatki oraz przyrost oszczędności<sup>1</sup>.

Na dochód rozporządzalny składają się<sup>2</sup>:

- dochód z pracy najemnej;
- dochód z gospodarstwa indywidualnego w rolnictwie;
- dochód z pracy na własny rachunek poza gospodarstwem indywidualnym w rolnictwie oraz z wykonywania wolnego zawodu;
- dochód z tytułu własności;
- dochód z wynajmu nieruchomości;
- świadczenia z ubezpieczeń społecznych i inne świadczenia społeczne;
- oraz pozostałe źródła dochodu (na przykład dary czy alimenty).

Wszystkie dane niezbędne do przeprowadzenia obliczeń zaczerpnięto z opracowań Głównego Urzędu Statystycznego: *Budżety gospodarstw domowych 2006 r. (2007 r., 2008 r., 2009 r., 2010 r., 2011 r., 2012 r., 2013 r.)*. Co warto podkreślić, badanie budżetów gospodarstw domowych przez GUS prowadzone jest metodą reprezentacyjną, która daje możliwość uogólnienia uzyskanych wyników na wszystkie gospodarstwa domowe w Polsce.

<sup>1</sup> *Budżety gospodarstw domowych 2013 r.* Główny Urząd Statystyczny, Warszawa 2014, s. 19.

<sup>2</sup> tamże, s. 19.

Do zrealizowanych w artykule obliczeń wykorzystano dane dotyczące średniej wartości miesięcznego dochodu rozporządzalnego gospodarstwa domowego przypadającej na jedną osobę wchodzącą w skład tego gospodarstwa. Wspomniana zmienna na potrzeby niniejszego opracowania nazwana została przeciętnym dochodem *per capita*.

Dla poszczególnych lat objętych analizą przeprowadzono procedurę badawczą obejmującą następujące etapy:

1. oszacowanie parametrów strukturalnych liniowego modelu przyczynowo-opisowego z jedną zmienną objaśniającą;
2. zweryfikowanie otrzymanego modelu ekonometrycznego poprzez:
  - 2.1. ocenienie dopasowania modelu do danych empirycznych,
  - 2.2. sprawdzenie istotności parametru strukturalnego stojącego przy zmiennej objaśniającej.

Artykuł niniejszy ma więc charakter badawczy. W artykule tym stawia się dwie hipotezy badawcze. Pierwsza z nich głosi, że funkcja liniowa jest wystarczająco dobrą aproksymantą obrazującą zależność między przeciętnym miesięcznym dochodem na osobę i wielkością gospodarstwa domowego. Z kolei druga hipoteza stanowi, iż wzrost liczby osób w gospodarstwie domowym powoduje zazwyczaj stały spadek średniego dochodu rozporządzalnego na osobę. Pozytywna weryfikacja postawionych hipotez będzie upoważniała do ostatecznej konstatacji, że bez względu na to, czy wzrost wielkości gospodarstwa domowego oznacza zwiększenie liczby jego członków z jednego do dwóch, z dwóch do trzech, z trzech do czterech itd., zawsze taka zmiana o jedną osobę będzie powodowała redukcję przeciętnego dochodu *per capita* w przybliżeniu o stałą wartość.

## 1. Zastosowane narzędzia badawcze

W celu określenia, czy istnieje związek stochastyczny między wybranymi zmiennymi, zebrano dla poszczególnych lat materiał statystyczny obejmujący realizacje zmiennej  $X$  oraz realizacja zmiennej  $Y$ . Wówczas otrzymano następujące pary liczb<sup>3</sup>:

$$(x_i, y_i),$$

gdzie  $x_i$  i  $y_i$  są zaobserwowanymi wartościami odpowiednio zmiennej  $X$  i  $Y$ , a  $i$  jest kolejnym numerem obserwacji odnotowanej w danym roku.

<sup>3</sup> Ch. Heij, P. de Boer, P.H. Franses, T. Kloek, H.K. van Dijk, *Econometric methods with applications in business and economics*. Oxford University Press, New York 2004, s. 79.

Generalnie związek między dwiema zmiennymi może mieć charakter liniowy albo nieliniowy. To, czy związek między zmiennymi ma charakter liniowy czy nie, można stwierdzić na podstawie wzrokowej oceny sporządzonego wykresu punktowego zwanego empirycznym rozrzutem punktów<sup>4</sup>. Ocena takiego wykresu składającego się z punktów o współrzędnych  $(x_i, y_i)$  da więc możliwość określenia, czy występuje związek liniowy, czy nie<sup>5</sup>, a następnie – jeśli nie ma podstaw do założenia związku liniowego – czy powinno się przyjąć związek krzywoliniowy i jaki konkretnie charakter taki nieliniowy związek będzie miał<sup>6</sup>.

Jeżeli wszystkie punkty składające się na wykres punktowy da się ograniczyć elipsą, to spodziewać się należy, iż między analizowanymi zmiennymi istnieje związek liniowy. Warto też wiedzieć, że im mniejszy jest rozrzut punktów względem prostej będącej liniową aproksymantą, czyli im większe jest spłaszczenie elipsy, tym szacowany model będzie lepiej dopasowany do danych empirycznych. Z kolei im większy jest rozrzut punktów względem tej aproksymanty, czyli im mniejsze jest spłaszczenie elipsy, tym gorsze dopasowanie wyznaczonego modelu do danych empirycznych.

Formalnie liniowy model regresji z jedną zmienną objaśniającą można przedstawić w następującej postaci<sup>7</sup>:

$$y_i^* = a_0 + a_1 x_i,$$

gdzie:

$x_i$  – wartości empiryczne zmiennej objaśniającej (czyli niezależnej);

$y_i$  – wartości empiryczne zmiennej objaśnianej (czyli zależnej);

$y_i^*$  – wartości teoretyczne zmiennej objaśnianej;

$i$  – numer obserwacji ( $i = 1, 2, \dots, n$ );

$n$  – liczba obserwacji;

$a_0, a_1$  – oceny parametrów strukturalnych modelu.

Parametry strukturalne modelu ekonometrycznego mającego przedstawioną powyżej budowę można oszacować wykorzystując wzory<sup>8</sup>:

<sup>4</sup> G. Argyrous, *Statistics for research*. SAGE Publications, London 2005, s. 162.

<sup>5</sup> J.E. Freund, B.M. Perles, *Modern elementary statistics*. Pearson Education, New Jersey 2007, s. 399.

<sup>6</sup> F.L. Coolidge, *Statistics. A gentle introduction*. SAGE Publications, Thousand Oaks 2006, s. 168.

<sup>7</sup> B. Guzik, *Podstawy ekonometrii*. Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poznań 2008, s. 46.

<sup>8</sup> *Ekonometria stosowana w przykładach i zadaniach*. Praca zbiorowa pod redakcją naukową J. Hoze-  
ra, Katedra Ekonometrii i Statystyki Uniwersytetu Szczecińskiego, Stowarzyszenie „Pomoc i Rozwój”,  
Szczecin 2007, s. 33.

$$a_1 = \frac{\overline{y \cdot x} - \bar{y} \cdot \bar{x}}{\overline{x^2} - \bar{x}^2} \text{ oraz } a_0 = \bar{y} - a_1 \cdot \bar{x},$$

gdzie:

$\bar{x}, \bar{y}$  – średnia arytmetyczna odpowiednio zmiennej  $X$  i zmiennej  $Y$ .

Weryfikacja statystyczna oszacowanego modelu powinna objąć przede wszystkim ocenę stopnia dopasowania tego modelu do danych empirycznych oraz sprawdzenie istotności parametru strukturalnego towarzyszącego zmiennej objaśniającej. W celu dokonania oceny dopasowania modelu do danych empirycznych można wyznaczyć następujące miary: odchylenie standardowe składnika losowego (zwane średnim błędem szacunku modelu), współczynnik zmienności losowej modelu, współczynnik indeterminacji oraz współczynnik determinacji (zwany też współczynnikiem dopasowania). Wymienione wielkości oblicza się z następujących wzorów:

- Średni błąd szacunku modelu<sup>9</sup>:

$$S_e = \sqrt{\frac{\sum_{i=1}^n (y_i - y_i^*)^2}{n-2}};$$

Średni błąd szacunku modelu wyrażony jest zawsze w jednostce zmiennej objaśnianej<sup>10</sup>.

- Współczynnik zmienności losowej modelu<sup>11</sup>:

$$V_e = \frac{S_e}{|\bar{y}|};$$

Współczynnik zmienności losowej modelu jest wielkością niemianowaną i do interpretacji zazwyczaj wyraża się go w procentach.

- Współczynnik indeterminacji<sup>12</sup>:

$$\varphi^2 = \frac{\sum_{i=1}^n (y_i - y_i^*)^2}{\sum_{i=1}^n (y_i - \bar{y})^2};$$

<sup>9</sup> B. Pułaska-Turyna, *Statystyka dla ekonomistów*. Wydawnictwo „Difin”, Warszawa 2005, s. 272.

<sup>10</sup> J.H. Stock, M.W. Watson, *Introduction to econometrics*. Pearson Education, New Jersey 2007, s. 124.

<sup>11</sup> P. Dittmann, *Prognozowanie w przedsiębiorstwie*. Oficyna Ekonomiczna, Oddział Polskich Wydawnictw Profesjonalnych Sp. z o.o., Kraków 2003, s. 117.

<sup>12</sup> A. Welfe, *Ekonometria. Metody i ich zastosowanie*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2009, s. 41.

Współczynnik indeterminacji jest wielkością unormowaną i przyjmuje wartości wyłącznie z przedziału  $(0; 1)$ . Im współczynnik ten jest mniejszy, tym lepiej, a czym jest większy, tym gorzej<sup>13</sup>.  $\varphi^2 = 0$  świadczy o dokładnym dopasowaniu modelu do danych empirycznych.

■ Współczynnik determinacji<sup>14</sup>:

$$R^2 = \frac{\sum_{i=1}^n (y_i^* - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \varphi^2.$$

Współczynnik determinacji jest wielkością unormowaną i przyjmuje wartości z przedziału  $(0; 1)$ <sup>15</sup>. Im współczynnik ten jest większy, tym lepiej, a czym jest mniejszy, tym gorzej<sup>16</sup>.  $R^2 = 1$  świadczy o dokładnym dopasowaniu modelu do danych empirycznych<sup>17</sup>.

Współczynniki indeterminacji i determinacji są wielkościami niemianowanymi i w celu interpretacji wyraża się je w procentach.

W ramach oceny jakości oszacowanego modelu ekonometrycznego sprawdza się też istotność parametru strukturalnego stojącego przy zmiennej objaśniającej, tj. bada się, czy uzyskany parametr istotnie różni się od zera<sup>18</sup>. W tym celu oblicza się wartość statystyki określonej formułą<sup>19</sup>:

$$t_1 = \frac{|a_1|}{D(a_1)},$$

<sup>13</sup> M. Sobczyk, *Ekonometria*. Wydawnictwo C.H. Beck, Warszawa 2013, s. 40.

<sup>14</sup> *Ekonometria. Metody, przykłady, zadania*. Praca zbiorowa pod redakcją naukową J. Dziechciaż, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2002, s. 134–135.

<sup>15</sup> J. Józwiak, J. Podgórski, *Statystyka od podstaw*. Polskie Wydawnictwo Ekonomiczne, Warszawa 2012, s. 371.

<sup>16</sup> R. Czyżycki, R. Klóska, *Ekonometria i prognozowanie zjawisk ekonomicznych w przykładach i zadaniach*. Wydawnictwo ECONOMICUS, Szczecin 2011, s. 45.

<sup>17</sup> Ch. Dougherty, *Introduction to econometrics*. Oxford University Press, New York 2002, s. 66.

<sup>18</sup> G. Keller, B. Warrack, *Statistics for management and economics*. Brooks/Cole – Thomson Learning, Pacific Grove 2003, s. 620.

<sup>19</sup> J. Sołtysiak, *Podstawy ekonometrii*. Wydawnictwo Gdańskiej Wyższej Szkoły Humanistycznej, Gdańsk 2007, s. 17.

gdzie  $D(a_1)$  jest standardowym błędem szacunku testowanego parametru. W przypadku liniowego modelu przyczynowo-opisowego z jedną zmienną objaśniającą standardowy błąd szacunku  $D(a_1)$  można otrzymać wykorzystując formułę<sup>20</sup>:

$$D(a_1) = \frac{S_e}{\sqrt{n(\overline{x^2} - \bar{x}^2)}} .$$

Aby sprawdzić istotność parametru strukturalnego stojącego przy zmiennej objaśniającej, należy z tablic rozkładu  $t$  Studenta dla  $n - 2$  stopni swobody i zadanego poziomu istotności  $\alpha$  odczytać wartość krytyczną  $t_\alpha$ <sup>21</sup>. I wtedy, jeżeli  $|t_1| > t_\alpha$ , to parametr  $a_1$  jest istotny, czyli zmienna objaśniająca  $X$  istotnie wpływa na zmienną objaśnianą  $Y$ . Z kolei jeśli zachodzi nierówność  $|t_1| \leq t_\alpha$ , to testowany parametr strukturalny jest nieistotny<sup>22</sup>.

## 2. Estymacja parametrów strukturalnych modeli regresji

Niech zmienna  $X$  będzie liczbą osób w gospodarstwie domowym, a zmienna  $Y$  – przeciętnym miesięcznym dochodem rozporządzalnym *per capita*<sup>23</sup>, którego jednostką jest zł/osobę. W badaniu rozpatrywane jest sześć grup gospodarstw domowych, a mianowicie gospodarstwa 1-osobowe, 2-osobowe, 3-osobowe, 4-osobowe, 5-osobowe oraz 6- i więcej osobowe. Poszczególne obserwacje zmiennej niezależnej  $X$  są następujące:  $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5$  oraz wartość  $x_6$  uzyskana w wyniku podzielenia łącznej liczby osób z gospodarstw domowych należących do kategorii „6- i więcej osobowe” przez liczbę takich gospodarstw. Otrzymany w ten sposób iloraz  $x_6$  wynosi: 6,63 w 2006 r., 6,63 w 2007 r., 6,65 w 2008 r., 6,64 w 2009 r., 6,65 w 2010 r., 6,64 w 2011 r., 6,60 w 2012 r. i 6,63 w 2013 r.

W pierwszym etapie badania założono liniowy charakter stochastycznej zależności między wielkością gospodarstwa domowego a przeciętnym miesięcznym dochodem rozporządzalnym na osobę. Wówczas parametry strukturalne oszacowanych modeli regresji mają wartości zaprezentowane w tabeli 1.

<sup>20</sup> E. Nowak, *Zarys metod ekonometrii*. Wydawnictwo Naukowe PWN, Warszawa 2002, s. 38.

<sup>21</sup> D. Asteriou, S.G. Hall, *Applied econometrics*. PALGRAVE MACMILLAN, New York 2007, s. 40.

<sup>22</sup> R. Czyżycki, M. Hundert, R. Klóska, *Wybrane zagadnienia z ekonometrii*. Wydawnictwo ECONOMICUS, Szczecin 2004, s. 64.

<sup>23</sup> Przeciętny miesięczny dochód rozporządzalny *per capita* rozumiany jest tutaj jako średnia miesięczna wartość dochodu rozporządzalnego gospodarstwa domowego przypadająca na jedną osobę wchodzącą w skład gospodarstwa.

**Tabela 1.**Parametry strukturalne  $a_0$  i  $a_1$  oszacowanych modeli

	Lata							
	2006	2007	2008	2009	2010	2011	2012	2013
Parametr $a_0$	1408,31	1510,88	1681,10	1811,85	1965,14	2005,29	2070,89	2113,54
Parametr $a_1$	-151,36	-154,81	-169,33	-185,78	-204,97	-207,03	-213,28	-218,72

Źródło: obliczenia własne na podstawie danych z publikacji GUS: Budżety gospodarstw domowych w 2006, 2007, 2008, 2009, 2010, 2011, 2012 i 2013 r., Warszawa 2006 r.: s. 51–56; 2007 r.: s. 63–68; 2008 r.: s. 77–82; 2009 r.: s. 73–78; 2010 r.: s. 81–86; 2011 r.: s. 84–89; 2012 r.: s. 77–82; 2013 r.: s. 101–106.

Tak więc pierwszy z modeli regresji liniowej po estymacji parametrów ma postać:  $y_i^* = 1408,31 - 151,36x_i$  (jest to model dla 2006 roku). Wartość  $a_1 = -151,36$  oznacza, że pod wpływem wzrostu zmiennej objaśniającej  $X$  o jednostkę zmienna objaśniana  $Y$  maleje o 151,36 jednostek. Zatem wzrost wielkości gospodarstwa domowego o osobę powoduje – w tym konkretnym przypadku – spadek miesięcznego dochodu rozporządzalnego na osobę średnio o około 151,36 zł. Natomiast parametr  $a_0$  nie ma tutaj interpretacji ekonomicznej.

Analiza informacji zawartych w tabeli 1 pozwala również na wyciągnięcie wniosku, że w przypadku wszystkich wziętych pod uwagę lat (2006–2013), wartość estymowanego parametru  $a_1$  jest ujemna. Oznacza to, że dla każdego z ośmiu oszacowanych modeli wzrost liczby osób w gospodarstwie domowym skutkuje spadkiem przeciętnego miesięcznego dochodu na osobę.

### 3. Ocena dopasowania modeli do danych empirycznych

Jak już wspomniano, weryfikacja modeli obejmuje sprawdzenie dopasowania wartości teoretycznych do danych empirycznych oraz zbadanie istotności parametru strukturalnego towarzyszącego zmiennej objaśniającej.

Jako miary służące ocenie dopasowania modelu do danych empirycznych wybrano:

- średni błąd szacunku modelu  $S_e$ ;
- współczynnik zmienności losowej modelu  $V_e$ ;
- współczynnik indeterminacji  $\varphi^2$ ;
- współczynnik determinacji  $R^2$ .

W tabeli 2 zebrano wyniki otrzymane dla wymienionych miar w przypadku poszczególnych modeli.



**Tabela 2.**

Wartości  $S_e$ ,  $V_e$ ,  $\varphi^2$  i  $R^2$  dla oszacowanych modeli regresji liniowej

	Lata							
	2006	2007	2008	2009	2010	2011	2012	2013
Średni błąd szacunku modelu $S_e$	73,009	60,466	61,692	72,374	87,106	81,682	84,952	89,964
Współczynnik zmienności losowej modelu $V_e$	0,085	0,063	0,058	0,063	0,071	0,065	0,065	0,068
Współczynnik indeterminacji $\varphi^2$	0,043	0,028	0,025	0,028	0,033	0,029	0,030	0,031
Współczynnik determinacji $R^2$	0,957	0,972	0,975	0,972	0,967	0,971	0,970	0,969

Źródło: obliczenia własne na podstawie tabeli 1 oraz publikacji GUS: *Budżety gospodarstw domowych w 2006, 2007, 2008, 2009, 2010, 2011, 2012 i 2013 r.*, Warszawa 2006 r.: s. 51–56; 2007 r.: s. 63–68; 2008 r.: s. 77–82; 2009 r.: s. 73–78; 2010 r.: s. 81–86; 2011 r.: s. 84–89; 2012 r.: s. 77–82; 2013 r.: s. 101–106.

Średni błąd szacunku dla pierwszego liniowego modelu przyczynowo-opisowego jest na poziomie 73,009. Oznacza to, że w 2006 roku wartości empiryczne zmiennej objaśnianej  $Y$  różniły się przeciętnie o około  $\pm 73,01$  zł od wartości teoretycznych tej zmiennej (czyli od wartości wyznaczonych na podstawie oszacowanego modelu).

Współczynnik zmienności losowej pierwszego modelu wynosi 0,085. Czyli odchylenie standardowe składnika losowego stanowiło w 2006 roku 8,5% średniej wartości zmiennej zależnej  $Y$ . Świadczy to o bardzo małej zmienności losowej modelu, bowiem wartość rozpatrywanego miernika jest poniżej 10%.

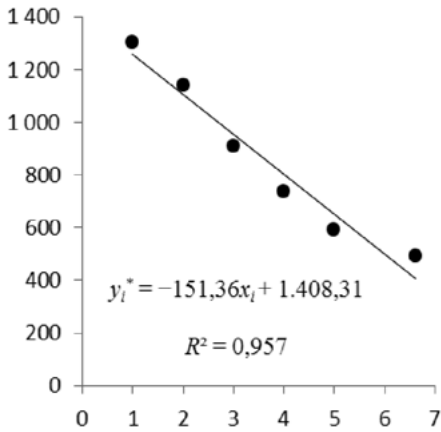
Współczynnik indeterminacji w przypadku liniowego modelu regresji oszacowanego dla 2006 roku opiewa na 0,043. Toteż zmienność badanej zmiennej  $Y$  jedynie w 4,3% nie została wyjaśniona przez oszacowany model. Z kolei wartość współczynnika determinacji otrzymana dla pierwszego modelu wynosi 0,957, czyli zmienność rozpatrywanej cechy  $Y$  aż w 95,7% została wyjaśniona przez oszacowany model.

$V_e$ ,  $\varphi^2$  i  $R^2$  są miarami stosunkowymi niemianowanymi, dzięki czemu istnieje możliwość porównywania ich wartości między latami. Dodatkowo współczynniki  $\varphi^2$  i  $R^2$  są miarami unormowanymi (zawsze mieszczą się w przedziale od 0 do 100% włącznie), toteż bardzo łatwo można ocenić, czy uzyskana wartość jest duża, czy mała. Porównanie wspomnianych współczynników otrzymanych dla poszczególnych lat pozwala stwierdzić, że na przestrzeni okresu 2006–2013 utrzymywały się one na dość stabilnym poziomie. Dodatkowo analiza wartości zebranych w tabeli 3 daje podstawę do wyciągnięcia wniosku, iż wszystkie oszacowane modele przyczynowo-opisowe są

bardzo dobrze dopasowane do danych empirycznych. Oznacza to, że zastosowana postać analityczna modeli prawidłowo opisuje kształtowanie się badanej zmiennej  $Y$ .

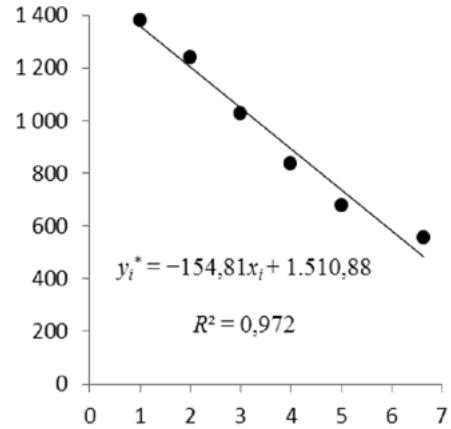
Rysunki 1–8 przedstawiają empiryczny rozrzut punktów dla danych dotyczących liczby osób w gospodarstwie domowym (zaznaczonej na osi odciętych) oraz dochodu rozporządzalnego na osobę (zaznaczonego na osi rzędnych) w przypadku ośmiu kolejnych lat objętych badaniem. Na każdym z rysunków umieszczono także wykres oszacowanej funkcji liniowej, która okazała się najlepszą aproksymantą.

Rysunek 1. Związek między  $X$  i  $Y$  w 2006 r.



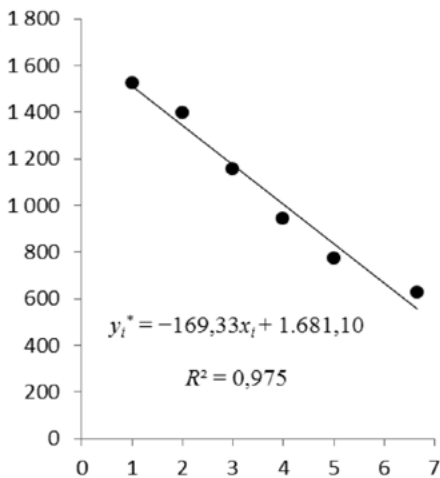
Źródło: opracowanie własne.

Rysunek 2. Związek między  $X$  i  $Y$  w 2007 r.



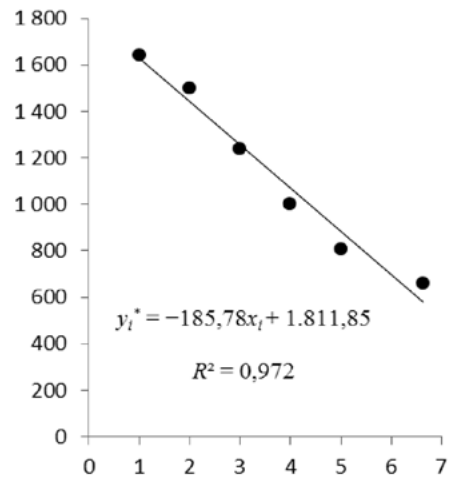
Źródło: opracowanie własne.

Rysunek 3. Związek między  $X$  i  $Y$  w 2008 r.



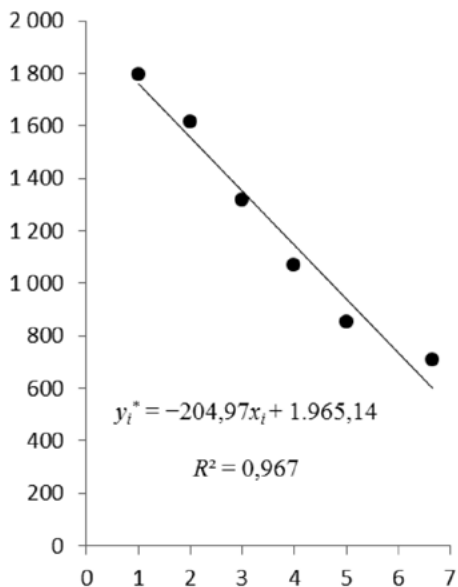
Źródło: opracowanie własne.

Rysunek 4. Związek między  $X$  i  $Y$  w 2009 r.



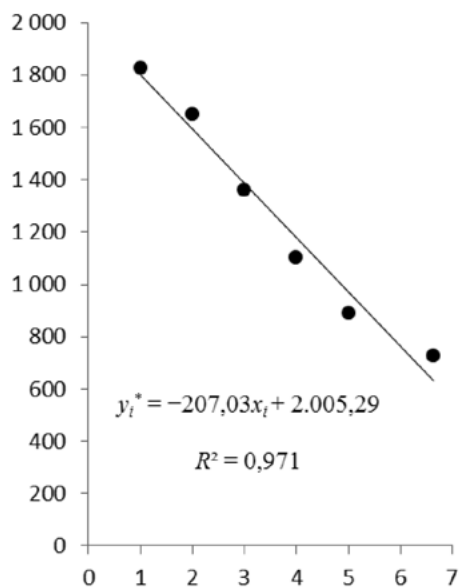
Źródło: opracowanie własne.

Rysunek 5. Związek między  $X$  i  $Y$  w 2010 r.



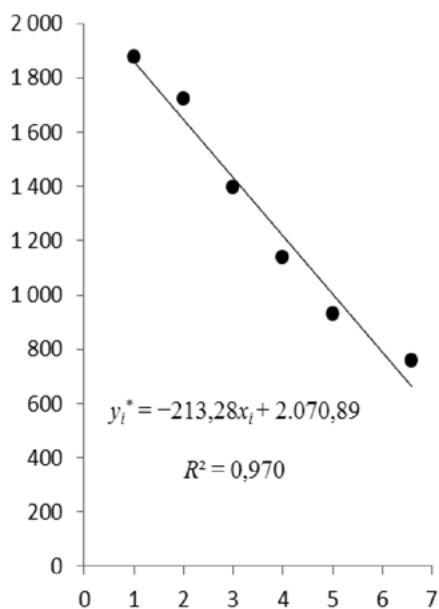
Źródło: opracowanie własne.

Rysunek 6. Związek między  $X$  i  $Y$  w 2011 r.



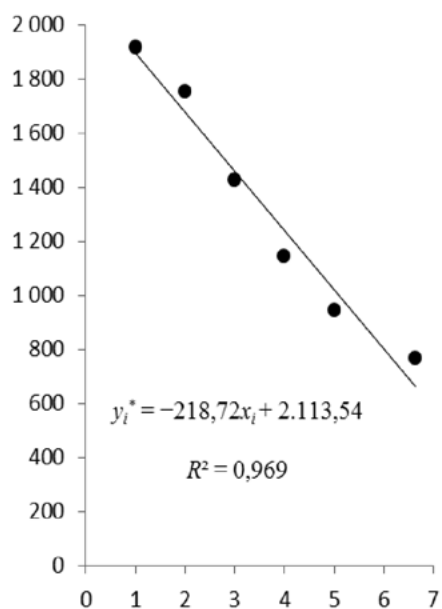
Źródło: opracowanie własne.

Rysunek 7. Związek między  $X$  i  $Y$  w 2012 r.



Źródło: opracowanie własne.

Rysunek 8. Związek między  $X$  i  $Y$  w 2013 r.



Źródło: opracowanie własne.

Na każdym z przedstawionych rysunków punkty składające się na wykres punktowy można ograniczyć elipsą, stąd między analizowanymi zmiennymi istnieje związek liniowy. W każdym z ośmiu przypadków z ułożenia punktów na układzie współrzędnych wynika też, że wzrostowi (spadkowi) wartości zmiennej  $X$  towarzyszy spadek (wzrost) wartości zmiennej  $Y$ , czyli parametr  $a_1$  liniowego modelu przyczynowo-opisowego musi mieć wartość ujemną. W oparciu o rysunki 1–8 można również potwierdzić wyciągnięty już wcześniej wniosek, że zbudowane modele są bardzo dobrze dopasowane do danych empirycznych.

#### 4. Sprawdzenie istotności parametru stojącego przy zmiennej objaśniającej

Wartości standardowego błędu szacunku  $D(a_1)$  oraz statystyki  $t_1$ , które zostały wyznaczone dla każdego z ośmiu oszacowanych modeli, umieszczono w tabeli 3.

Tabela 3.

Wartości  $D(a_1)$  i  $t_1$  dla oszacowanych modeli regresji

	Lata							
	2006	2007	2008	2009	2010	2011	2012	2013
Standardowy błąd szacunku $D(a_1)$	15,950	13,199	13,433	15,767	18,958	17,805	18,630	19,652
Statystyka $t_1$	9,490	11,729	12,606	11,783	10,812	11,628	11,448	11,129

Źródło: obliczenia własne na podstawie tabeli 1 oraz publikacji GUS: *Budżety gospodarstw domowych w 2006, 2007, 2008, 2009, 2010, 2011, 2012 i 2013 r.*, Warszawa 2006 r.: s. 51–56; 2007 r.: s. 63–68; 2008 r.: s. 77–82; 2009 r.: s. 73–78; 2010 r.: s. 81–86; 2011 r.: s. 84–89; 2012 r.: s. 77–82; 2013 r.: s. 101–106.

Dla poziomu istotności 0,1 i 4 stopni swobody odczytana z tablic rozkładu  $t$  Studenta wartość krytyczna  $t_\alpha$  wynosi 2,13. Konfrontacja wyników zebranych w tabeli 3 z wartością krytyczną pozwala orzec, że dla każdego roku obliczona wartość statystyki  $t$  Studenta jest wyższa od wartości  $t_\alpha$ , co uprawnia do stwierdzenia, iż parametr strukturalny towarzyszący zmiennej  $x_i$  istotnie różni się od zera.

## Podsumowanie

Analiza różnic w wielkości dochodu mieszkańców Polski stanowi ważny obszar badawczy, ponieważ łączy się z problematyką niejednorodności w poziomie życia poszczególnych grup społeczeństwa. Dochód nie jest oczywiście jedyną, ale wciąż pozostaje kluczową zmienną różnicującą jakość życia mieszkańców Polski.

Celem niniejszego artykułu było znalezienie stochastycznej zależności między wielkością dochodu rozporządzalnego *per capita* oraz liczbą osób kształtujących gospodarstwo domowe, a następnie udzielenie odpowiedzi na pytanie, czy zależność ta ma charakter liniowy.

W artykule postawiono hipotezę badawczą stanowiącą, że funkcja liniowa to wystarczająco dobra aproksymanta obrazująca zależność między przeciętnym miesięcznym dochodem rozporządzalnym na osobę oraz wielkością gospodarstwa domowego. Dla każdego przeanalizowanego roku oszacowany model ekonometryczny był bardzo dobrze dopasowany do danych empirycznych oraz zmienna objaśniająca istotnie wpływała na zmienną objaśnianą, stąd można uznać, że funkcja liniowa jest odpowiednią aproksymantą.

Dodatkowo w artykule postawiono hipotezę stanowiącą, iż wzrost liczby osób w gospodarstwie domowym powoduje spadek przeciętnego dochodu na osobę i jest to mniej więcej stały spadek. Na pozytywną weryfikację tej hipotezy pozwalał fakt, iż w przypadku wszystkich ośmiu liniowych modeli przyczynowo-opisowych wartość parametru  $a_1$  była ujemna.

Podsumowując więc powyższe rozważania warto skonstatować, że bez względu na to, czy rośnie liczba członków gospodarstwa domowego z jednego do dwóch, z dwóch do trzech, z trzech do czterech itd., to zawsze taki wzrost wielkości gospodarstwa domowego o jedną osobę sprawia, że spada przeciętny dochód *per capita* w przybliżeniu o stałą kwotę. Jest to informacja nieoczywista i z naukowego punktu widzenia dość zaskakująca, co udowadnia wysoką wartość poznawczą niniejszego artykułu.

# **Analysis of the relationship between income *per capita* and the size of household in the years 2006–2013**

## **Summary:**

The aim of the article is to find the model that best describes the relationship between the level of available income *per capita* and the size of the household. Two hypotheses are formulated. The first of them states that the linear function adequately shows the relationship between the two variables under consideration. The second hypothesis states that the increase of the number of people in the household usually results in the constant decrease of average available income *per capita*. Both of these hypotheses are verified positively.

**Keywords:** available income, size of household, linear causal-descriptive model, regression model, verification of econometric model