

dr Henryk Kowgier

Uniwersytet Szczeciński

Kilka uwag na temat statystycznej analizy czasowej emigracji Polaków za granicę w latach 2004-2015

Streszczenie:

W artykule dokonano analizy statystycznej czasowej emigracji Polaków za granicę w latach 2004-2015 wykorzystując do tego celu dane zaczerpnięte z GUS(2017). Do zbadania tych danych zastosowano głównie metodę statystyczną w zakresie analizy korelacji, podstawowych statystyk opisowych jak również elementy wnioskowania statystycznego. Zbudowano także modele liniowe, które łączą emigrację Polaków do niektórych krajów europejskich z emigracją do krajów spoza Unii Europejskiej. Artykuł kończą stosowne wnioski związane z przeprowadzoną analizą.

Słowa kluczowe: analiza statystyczna, regresja liniowa, emigracja Polaków za granicę

Wprowadzenie

W trakcie przemian ustrojowych w Polsce po 1989 roku emigracja Polaków za granicę początkowo miała charakter ekonomiczny czasowy i zarobkowy. Brali w niej udział zwykle ludzie z podstawowym i zasadniczym wykształceniem zawodowym wywodzący się z obszarów Polski objętych wysokim bezrobociem. Przystąpienie Polski do Unii Europejskiej w 2004 roku zmieniło nieco osobę statystycznego emigranta z Polski. W poszukiwaniu pracy zaczęło wyjeżdżać coraz więcej ludzi z wyższym wykształceniem znających języki obce. Często zabierali ze sobą całe rodziny z myślą o pozostaniu na emigracji już na stałe. Dzięki dużej inteligencji i pracowitości Polacy, którzy wyjechali budowali i budują do tej pory dobrobyt krajów, które stały się ich drugą ojczyzną. Jak pokazują dane statystyczne, w latach 2004-2015 nie udało się powstrzymać masowej emigracji Polaków a wręcz odwrotnie – nieustannie w tym okresie rosła. W raporcie KPMG z 2008 roku¹ podano zarówno dodatnie jak i ujemne czynniki związane z czasową emigracją polskich pracowników za granicę. W niniejszym opracowaniu dokonano analizy statystycznej danych zaczerpniętych z Rocznika Statystycznego GUS (2017) w zakresie korelacji, regresji liniowej, wnioskowania statystycznego oraz podstawowych statystyk opisowych dotyczących emigracji Polaków za granicę w latach 2004-2015.

¹Por. Raport KPMG (2008). Migracja pracowników – szansa czy zagrożenie?

Celem artykułu jest statystyczny opis czasowej emigracji Polaków za granicę w latach 2004-2015. Dotyczy on w szczególności następujących problemów:

- Wyznaczenia wielkości współczynników korelacji między emigracją Polaków do poszczególnych krajów i wynikających z tego wniosków.
- Wyznaczenia wielkości podstawowych statystyk opisowych dotyczących czasowej emigracji i wynikających z tego wniosków.
- Wyznaczenia niektórych równań regresji liniowej, które łączą emigrację Polaków do poszczególnych krajów Unii Europejskiej z emigracją do krajów spoza Unii Europejskiej.
- Zweryfikowania hipotezy statystycznej dotyczącej średniej wielkości emigracji Polaków do krajów spoza Unii Europejskiej.
- Wyznaczenia przedziału ufności dla odchylenia standardowego wielkości emigracji Polaków do Szwecji.

Aby ułatwić analizę, dane z GUS-u umieszczono w tabelach 1-2. Dotyczą one wielkości (wyrażonych w tysiącach) emigracji Polaków do krajów Unii Europejskiej, Europy oraz krajów spoza Unii Europejskiej.

L	A	B	C	CZ	D	F	FR	G	H	HI
2004	15,0	13,0	-	-	-	0,4	30,0	13,0	23,0	26,0
2005	25,0	21,0	-	-	-	0,7	44,0	17,0	43,0	37,0
2006	34,0	28,0	-	-	-	3,0	49,0	20,0	55,0	44,0
2007	39,0	31,0	4,0	8,0	17,0	4,0	55,0	20,0	98,0	80,0
2008	40,0	33,0	4,0	10,0	19,0	4,0	56,0	20,0	108,0	83,0
2009	36,0	34,0	3,0	9,0	20,0	3,0	60,0	16,0	98,0	84,0
2010	29,0	45,0	3,0	7,0	19,0	3,0	60,0	16,0	92,0	48,0
2011	25,0	47,0	3,0	7,0	21,0	2,0	62,0	15,0	95,0	40,0
2012	28,0	48,0	2,0	8,0	23,0	2,0	63,0	14,0	97,0	37,0
2013	31,0	49,0	1,0	8,0	25,0	3,0	63,0	12,0	103,0	34,0
2014	34,0	49,0	1,0	9,0	28,0	3,0	63,0	9,0	109,0	32,0
2015	36,0	52,0	1,0	9,0	30,0	3,0	64,0	8,0	112,0	30,0

Tabela1. Czasowa emigracja Polaków za granicę (w. tys.) w latach 2004-2015

Źródło: Rocznik Statystyczny GUS 2017.

Oznaczenia: L – lata, A – Austria, B – Belgia, C – Cypr, CZ – Czechy, D – Dania, F – Finlandia,

FR – Francja, G – Grecja, H – Holandia, HI – Hiszpania, - oznacza brak danych.

L	IR	N	P	SZ	WB	W	NOR	SUE	E	UE
2004	15,0	385,0	0,5	11,0	150,0	59,0	-	20,0	770,0	750,0
2005	76,0	430,0	0,6	17,0	340,0	70,0	-	30,0	1200,0	1170,0
2006	120,0	450,0	1,0	25,0	580,0	85,0	-	60,0	1610,0	1550,0
2007	200,0	490,0	1,0	27,0	690,0	87,0	36,0	65,0	1925,0	1860,0
2008	180,0	490,0	1,0	29,0	650,0	88,0	38,0	67,0	1887,0	1820,0
2009	140,0	465,0	1,0	31,0	595,0	88,0	45,0	75,0	1765,0	1690,0
2010	133,0	440,0	1,0	33,0	590,0	92,0	50,0	78,0	1685,0	1607,0
2011	120,0	470,0	1,0	36,0	625,0	94,0	56,0	85,0	1754,0	1670,0
2012	118,0	500,0	1,0	38,0	637,0	97,0	65,0	96,0	1816,0	1720,0
2013	115,0	560,0	1,0	40,0	642,0	96,0	71,0	102,0	1891,0	1789,0
2014	113,0	614,0	1,0	43,0	685,0	96,0	79,0	112,0	2013,0	1901,0
2015	111,0	655,0	1,0	46,0	720,0	94,0	84,0	115,0	2098,0	1983,0

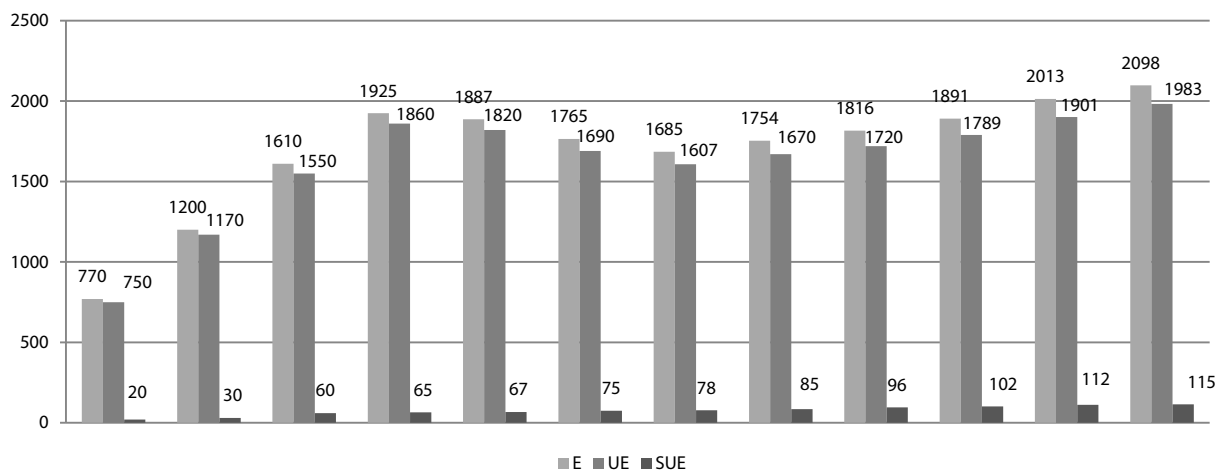
Tabela2. Czasowa emigracja Polaków za granicę w latach 2004 - 2015

Źródło: Rocznik Statystyczny GUS 2017.

Oznaczenia: IR – Irlandia, N – Niemcy, P – Portugalia, Sz- Szwecja, WB – Wielka Brytania, W – Włochy, NOR – Norwegia, SUE – kraje spoza Unii Europejskiej, E – Europa, UE – Unia Europejska, - oznacza brak danych.

Analiza statystyczna danych

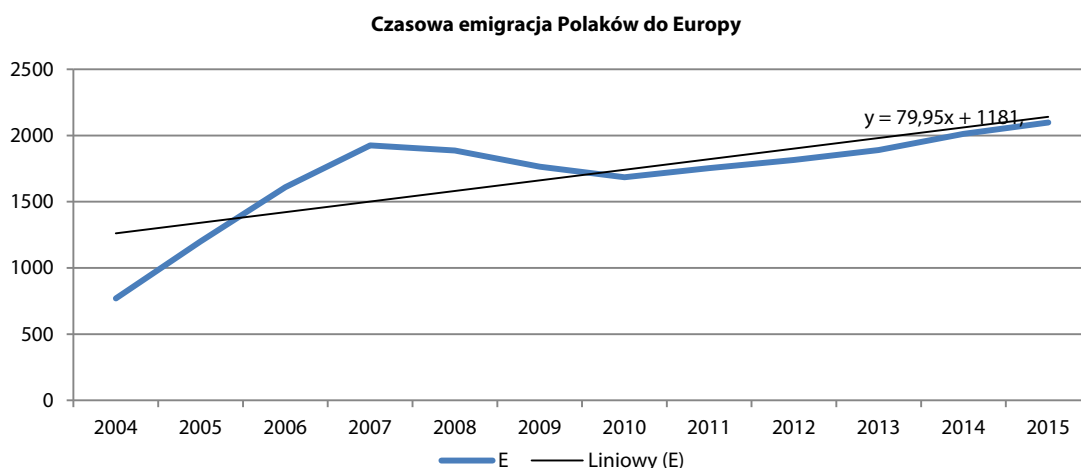
W celu dokonania analizy danych zaczerpniętych z GUS-u poniżej sporządzono dodatkowo dziewięć rysunków i dwie tabele 3-4 w oparciu o rzeczywiste dane umieszczone w tabelach 1-2.



Rysunek 1. Porównanie czasowej emigracji Polaków (w tys.) w latach 2004-2015 do krajów: Unii Europejskiej, Europy oraz spoza Unii Europejskiej

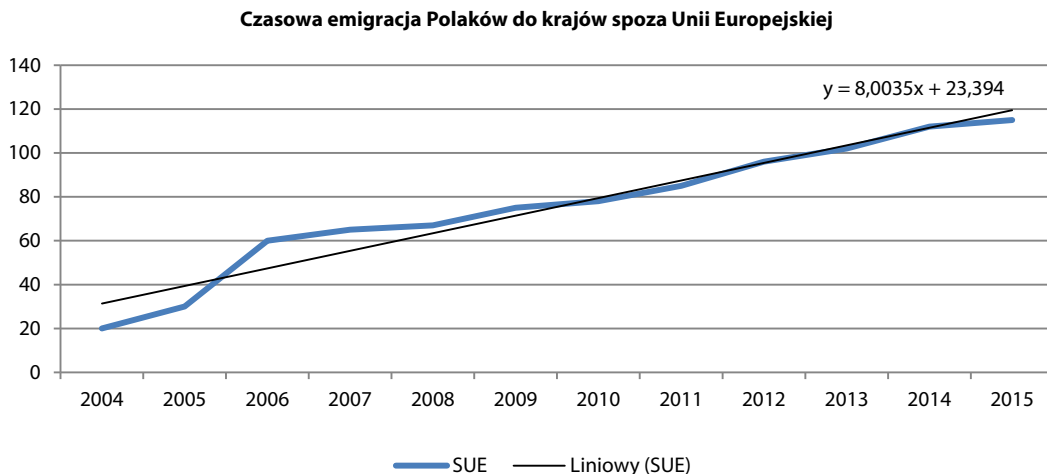
Źródło: opracowanie własne na podstawie tabeli 2.

Z danych GUS-u przedstawionych na rysunku 1 wynika, że czasowa emigracja Polaków w okresie 2004-2015 dotyczyła głównie Europy (najwyższe słupki w każdym roku) a w szczególności Unii Europejskiej. Emigracja poza Unię Europejską w poszczególnych latach mimo, że nieustannie rosła od 20 tysięcy w 2004 roku do 115 tysięcy w 2015 roku była stosunkowo niewielka na tle emigracji do Europy. Emigracja w 2004 roku do krajów spoza Unii Europejskiej stanowiła 2,6% emigracji do Europy oraz 2,7% emigracji do Unii Europejskiej. W 2015 roku emigracja do krajów spoza Unii Europejskiej stanowiła 5,5% emigracji do Europy oraz 5,8% emigracji do Unii Europejskiej. Ogółem na przestrzeni lat 2004-2015 emigracja do krajów spoza Unii Europejskiej stanowiła 4,4 % emigracji do Europy oraz 4,6% emigracji do krajów Unii Europejskiej. Emigracja do krajów Unii Europejskiej stanowiła aż 95,57% emigracji do Europy.



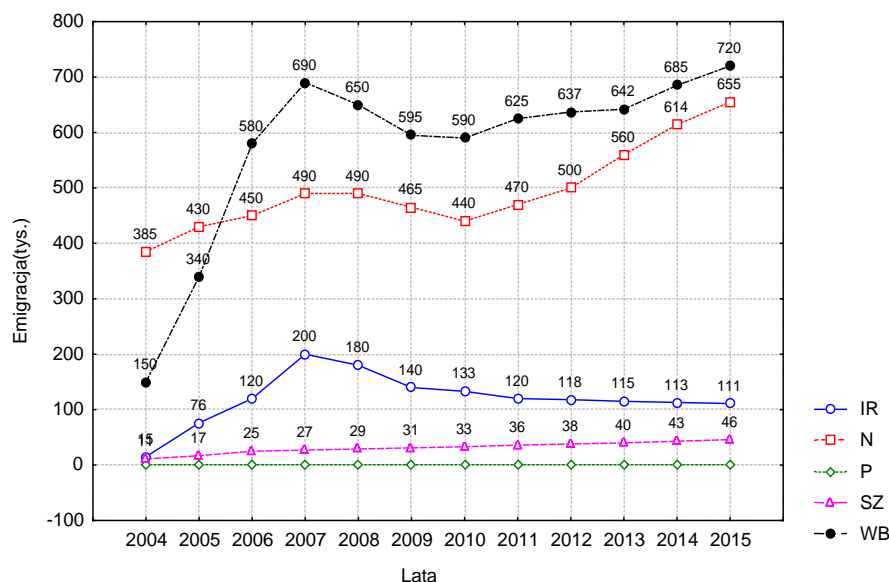
Rysunek 2. W przybliżeniu rosnący trend liniowy czasowej emigracji Polaków (w tys.) do Europy w latach 2004-2015

Źródło: opracowanie własne na podstawie tabeli 2 z wykorzystaniem pakietu Excel 2007.



Rysunek 3. W przybliżeniu rosnący trend liniowy czasowej emigracji Polaków (w tys.) do krajów spoza Unii Europejskiej w latach 2004-2015

Źródło: opracowanie własne na podstawie tabeli 2 z wykorzystaniem pakietu Excel 2007.



Rysunek 4. Porównanie czasowej emigracji Polaków (w tys.) do niektórych krajów europejskich w latach 2004-2015

Źródło: opracowanie własne na podstawie tabeli 2 z wykorzystaniem pakietu Statistica 6.0.

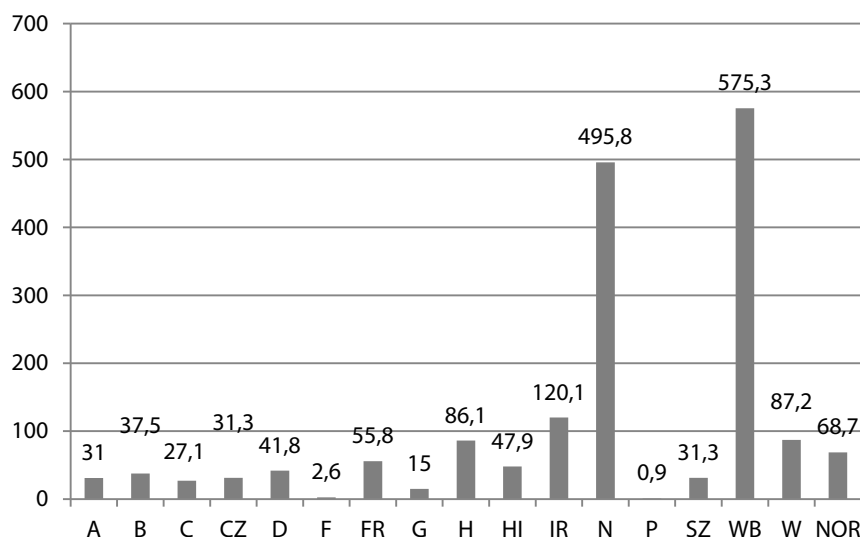
Jak wynika z danych GUS-u przedstawionych na rysunku 4 w okresie 2004-2015 Polacy najchętniej wyjeżdżali do Wielkiej Brytanii na drugim miejscu plasują się Niemcy, zaś na trzecim Irlandia. Łączna wartość czasowej emigracji do Niemiec w latach 2004-2015 wynosiła 5949 tys., i stanowiła 86,17% łącznej wartości czasowej emigracji Polaków do Wielkiej Brytanii – 6904 tys. W okresie 2004-2015 oprócz wymienionych krajów Polacy często wyjeżdżali do Włoch i Holandii. Łączna wielkość (po zsumowaniu) emigracji do Włoch wynosiła 1046 tys., natomiast do Holandii – 1033 tys. W sumie wielkość emigracji do Niemiec, Wielkiej Brytanii i Irlandii stanowiła 70,02% wielkości czasowej emigracji Polaków do Europy w okresie 2004-2015. Najrzadziej w wymienionym okresie Polacy wyjeżdżali do Portugalii oraz nieco chętniej jak wynika z danych GUS-u do Szwecji.

L	Śr	Mediana	Minimum	Maksimum	O.Standardowe	Skośność	Kurtoza
A	31	32,5	15	40	7,10	-0,92	0,95
B	37,5	39,5	13	52	12,68	-0,64	-0,7
C	27,08	3	1	44	44,59	1,32	-0,33
CZ	31,25	9	7	100	41,47	1,32	-0,33
D	41,83	24	17	100	35,28	1,28	-0,38
F	2,59	3	0,4	4	1,13	-0,89	0,23
FR	55,75	60	30	64	10,22	-1,73	2,85
G	15	15,5	8	20	4,05	-0,35	-0,69
H	86,08	97,5	23	112	29,05	-1,38	0,68
HI	47,92	38,5	26	84	21,59	1,05	-0,6
IR	120,08	119	15	200	46,41	-0,56	2,07
N	495,75	480	385	655	78,02	0,92	0,39
P	0,93	1	0,5	1	0,18	-2,12	3,13
SZ	31,33	32	11	46	10,35	-0,58	-0,08
WB	575,33	631	150	720	164,95	-2,02	3,76
W	87,17	90	59	97	11,54	-1,73	2,58
NOR	68,67	68	36	100	24,09	0,09	-1,49
SUE	75,42	76,5	20	115	29,67	-0,54	-0,23
E	1701,17	1790,5	770	2098	372,00	-1,73	3,0
UE	1625,83	1705	750	1983	346,19	-1,77	3,19

Tabela 3. Ważniejsze statystyki opisowe

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 6.0.

Aż w czternastu przypadkach badane rozkłady zmiennych wykazywały lewostronną asymetrię rozkładu (ujemna wartość współczynnika skośności) natomiast w sześciu: *C*, *CZ*, *D*, *HI*, *N*, *NOR* prawostronną asymetrię rozkładu (dodatnia wartość współczynnika skośności)². Dodatni współczynnik kurtozy występował w jedenastu przypadkach: *A*, *F*, *FR*, *H*, *IR*, *N*, *P*, *WB*, *W*, *E*, *UE*. Świadczy to o tym, że w przypadku tych zmiennych wartości cechy były bardziej skoncentrowane wokół wartości średniej, niż ma to miejsce w przypadku rozkładu normalnego. W dziewięciu przypadkach: *B*, *C*, *CZ*, *D*, *G*, *HI*, *SZ*, *NOR*, *SUE* współczynnik kurtozy był ujemny i wartości cechy w wymienionych przypadkach były mniej skoncentrowane wokół wartości średniej niż ma to miejsce dla rozkładu normalnego.

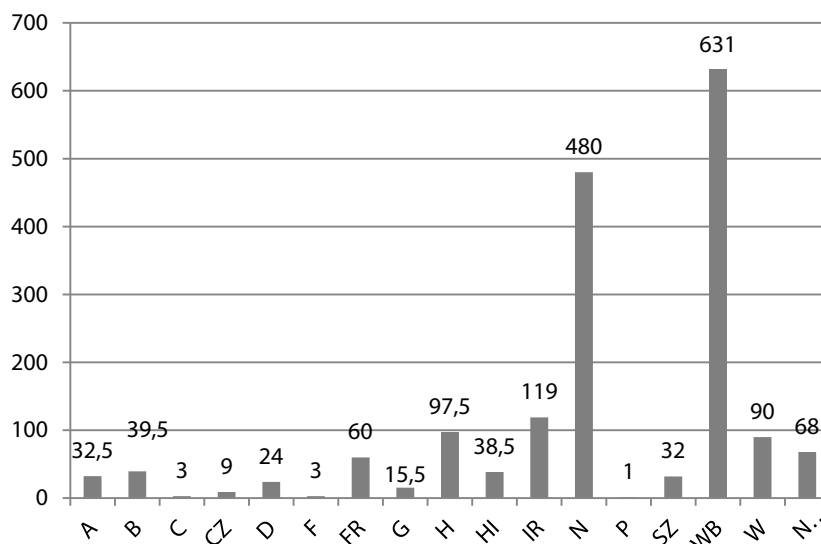


Rysunek 5. Porównanie średnich wartości (w tys.) czasowej emigracji Polaków do poszczególnych krajów europejskich w latach 2004-2015

Źródło: opracowanie własne na podstawie tabeli 3 z wykorzystaniem pakietu Excel 2007.

² Por. Statystyka. Opis statystyczny, (red.) J. Hozer US, Stowarzyszenie Pomoc i Rozwój, Szczecin 1998, s. 106-112.

Największa średnia wartość czasowej emigracji Polaków w latach 2004-2015 równa 575,33 tys., przypadła na Wielką Brytanię (nie licząc zmiennych: *E, UE*). Największą wartość odchylenia standardowego odnotowano również w przypadku Wielkiej Brytanii - 164,95 tys. Analogicznie najmniejsze średnie wielkości emigracji przypadły na Portugalię - 0,9 tys., oraz Finlandię - 2,59 tys. Podobnie najmniejsze odchylenia standardowe odnotowano w przypadku Portugalii - 0,18 tys., oraz Finlandii - 1,13tys.



Rysunek 6. Porównanie median (w tys.) czasowej emigracji Polaków w latach 2004-2015 do poszczególnych krajów europejskich

Źródło: opracowanie własne na podstawie tabeli 3 z wykorzystaniem pakietu Excel 2007.

Największe wartości median odnotowano dla Wielkiej Brytanii - 631 tys., i Niemiec -480 tys., a najmniejsze dla Portugalii - 1tys., i Finlandii oraz Cypru - 3 tys.

	WB	N	IR	H	W	HI	FR	,B
WB	1,00	0,70	0,79	0,92	0,94	0,32	0,92	0,80
N	0,70	1,00	0,25	0,73	0,67	-0,20	0,69	0,75
IR	0,79	0,25	1,00	0,72	0,62	0,77	0,60	0,35
H	0,92	0,73	0,72	1,00	0,90	0,33	0,93	0,84
W	0,94	0,67	0,62	0,90	1,00	0,12	0,98	0,93
HI	0,32	-0,20	0,77	0,33	0,12	1,00	0,15	-0,18
FR	0,92	0,69	0,60	0,93	0,98	0,15	1,00	0,93
B	0,80	0,75	0,35	0,84	0,93	-0,18	0,93	1,00

Tabela 4. Kształtowanie się współczynników korelacji dotyczące czasowej emigracji Polaków do wybranych krajów Unii Europejskiej w latach 2004-2015. Oznaczone współczynniki korelacji są istotne z prawdopodobieństwem $p < 0,05$

Źródło: opracowanie własne na podstawie tabel 1-2 z wykorzystaniem pakietu Statistica 6.0. Istotne statystycznie współczynniki korelacji oznaczono kursywą i boldem.

Z tabeli 4 wynika, że w zdecydowanej większości przypadków otrzymano duże, istotne statystycznie³ i dodatnie współczynniki korelacji. Świadczy to o tym, że najczęściej wzrost emigracji Polaków do danego kraju powodował wzrost emigracji do innego kraju. Największe korelacje dodatnie odnotowano między emigracjami Polaków do Francji i Włoch (0,98), Belgii i Francji (0,93), Francji i Holandii

³ Por. H. Kowgier, Elementy rachunku prawdopodobieństwa i statystyki na przykładach z ekonomii, WNT, Warszawa 2011, s. 115-116.

(0,93), Wielkiej Brytanii i Włoch (0,94), Wielkiej Brytanii i Holandii (0,92) oraz Wielkiej Brytanii i Francji (0,92). Zależność liniowa występowała między emigracją Polaków: do Europy i krajów spoza Unii Europejskiej, do Niemiec i krajów spoza Unii Europejskiej, do Szwecji i krajów spoza Unii Europejskiej, do Włoch i krajów spoza Unii Europejskiej, do Francji i krajów spoza Unii Europejskiej, do Unii Europejskiej i krajów spoza Unii Europejskiej przy założeniu, że zmienna UE jest zmienną objaśnianą. Odpowiednie równania regresji, błędy standardowe szacunku parametrów strukturalnych i stopnie dopasowania tych modeli do danych rzeczywistych ukazano poniżej:

$$E = 869,22 + 11,033 \cdot SUE, R^2 = 0,77; E = -44,44 + 1,07 \cdot UE, R^2 = 0,99 - \text{nieistotny statystycznie wyraz}$$

$$\text{wolny}; SUE = -87,36 + 0,32 \cdot N, R^2 = 0,74; SUE = -131,66 + 2,37 \cdot W, R^2 = 0,85;$$

$$SUE = 14,01 + 2,85 \cdot SZ, R^2 = 0,99; UE = 869,23 + 10,03 \cdot SUE, R^2 = 0,73;$$

$$SUE = -73,79 + 2,68 \cdot FR, R^2 = 0,84.$$

Jak wynika z powyższych modeli regresji w każdym przypadku występował relatywnie duży współczynnik determinacji tych modeli do danych rzeczywistych. Małą niespodzianką, która wyszła w trakcie obliczeń jest to, że w modelu łączącym emigrację do Unii Europejskiej z emigracją do Europy nieistotny statystycznie okazał się wyraz wolny.

Między emigracją w latach 2004-2015 Polaków do Niemiec, Wielkiej Brytanii i krajów leżących poza Unią Europejską przy założeniu, że SUE jest zmienną objaśnianą występowała zależność liniowa z istotnymi parametrami stojącymi przy zmiennych N oraz WB oraz istotnym wyrazie wolnym:

$$SUE = -73,74 + 0,2 \cdot N + 0,08 \cdot WB, R^2 = 0,86. \quad (1)$$

W nawiasach zwykłych podano standardowe błędy szacunku parametrów stojących przy zmiennych N , WB oraz wyrazie wolnym. Aby zbadać istotność parametrów strukturalnych występujących w modelu ekonometrycznym, należy sprawdzić, czy parametry te istotnie różnią się od zera. Do tego celu wykorzystano test istotności t – Studenta. Zweryfikowano hipotezę zerową $H_0: |\alpha_i| = 0$ przy hipotezie alternatywnej $H_1: |\alpha_i| \neq 0$ ($i = 0, 1, 2$). Sprawdzianem testu istotności jest statystyka t - Studenta liczona według wzoru $|t_{\hat{\alpha}_i}| = \frac{|\hat{\alpha}_i|}{D(\hat{\alpha}_i)}$, gdzie $\hat{\alpha}_i$ - ocena i -tego parametru strukturalnego, $D(\hat{\alpha}_i)$ - standardowy błąd szacunku parametru strukturalnego $\hat{\alpha}_i$. W przypadku modelu (1) otrzymano:

$$|t| = \frac{73,74}{24,17} = 3,05 > t_{0,05;9} = 2,26, |t_N| = \frac{0,2}{0,06} = 3,33 > t_{0,05;9} = 2,26, |t_{WB}| = \frac{0,08}{0,03} = 2,66 > 2,26,$$

gdzie $t_{0,05;9} = 2,26$ - wartość krytyczna statystyki t - Studenta obliczona dla poziomu istotności 0,05 oraz 9 stopni swobody. Uzyskane powyżej nierówności: $3,05 > 2,26$; $3,33 > 2,26$; $2,66 > 2,26$ świadczą o tym, że hipotezę H_0 na poziomie istotności 0,05 we wszystkich przypadkach należy odrzucić na korzyść hipotezy H_1 . Zatem parametry występujące w modelu (1) są istotne statystycznie i są potrzebne przy prognozowaniu szacowanej wartości zmiennej SUE . Stopień dopasowania modelu do danych rzeczywistych wynosi 86%.

Względne błędy oszacowań parametrów przyjmują wartości:

$$\frac{24,17}{73,74} \cdot 100\% = 32,77\% < 50\%, \frac{0,06}{0,2} \cdot 100\% = 30\% < 50\%, \frac{0,03}{0,08} \cdot 100\% = 37,5\% < 50\%.$$

Kwantyle statystyki d_L i d_U testu Durбина-Watsona dla $\alpha = 0,05$, $k = 2$ – liczba zmiennych objaśniających oraz $n = 12$ obserwacji przyjmują wartości $d_L = 0,812$, $d_U = 1,579$. Wykorzystując funkcję statystyczną Reglinp w Excelu i inne obliczenia otrzymano, że statystyka Durбина-Watsona przyjmuje wartość:

$$d = \frac{\sum_{t=2}^{12} (e_t - e_{t-1})^2}{\sum_{t=1}^{12} e_t^2} = \frac{1061,0}{1341,27} = 0,79 < 2,$$

gdzie e_1, e_2, \dots, e_{12} – zaobserwowane reszty modelu.

Stawiając hipotezę $H_0 : \rho = 0$ przy hipotezie alternatywnej $H_1 : \rho > 0$ otrzymano, że $d = 0,79 < d_L = 0,812$. Zatem hipotezę $H_0 : \rho = 0$ mówiącą o braku autokorelacji składnika losowego należy odrzucić na korzyść hipotezy alternatywnej mówiącej o istnieniu autokorelacji dodatniej. Wariancja składnika losowego przyjmuje wartość:

$$S_e^2 = \frac{\sum_{t=1}^{12} e_t^2}{n - k - 1} = \frac{1341,27}{9} = 149,03$$

n – liczba obserwacji, k – liczba zmiennych objaśniających.

Odchylenie standardowe składnika losowego przyjmuje wartość:

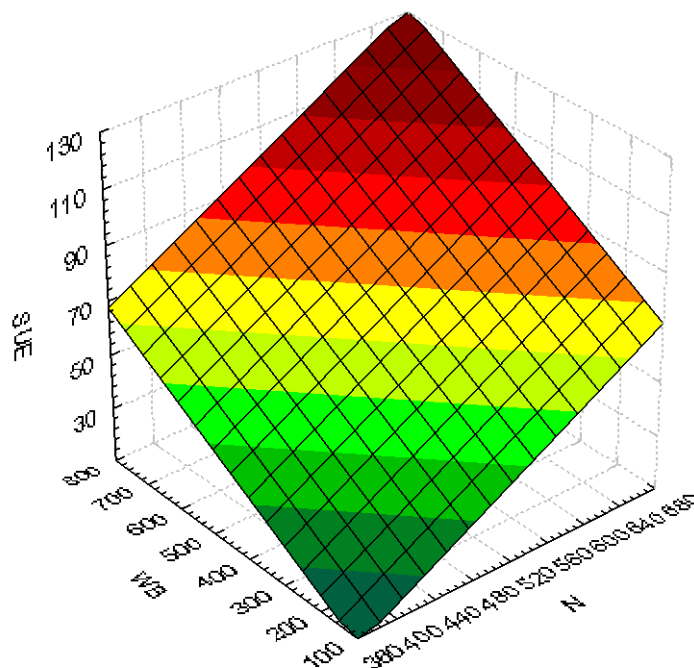
$$S_e = 12,2.$$

Wartości empiryczne odchyłały się średnio od teoretycznych o 12,2 tys. Współczynnik zmienności losowej ma wartość:

$$V_s = \frac{S_e}{SUE} \cdot 100\% = \frac{12,2}{75,41} \cdot 100\% = 16,17\%.$$

Otrzymany wynik świadczy o stosunkowo wysokiej zmienności losowej stanowiącej ponad 16,17% wartości średniej arytmetycznej zmiennej SUE .

Wartość statystyki F – Snedecora obliczona numerycznie za pomocą funkcji Reglinp w Excelu spełnia relację: $F = 27,99 > F_{0,05;2;9} = 4,26$, gdzie $F_{0,05;2;9} = 4,26$ jest wartością krytyczną testu F – Snedecora przy poziomie istotności 0,05; dwóch zmiennych objaśniających i 9 stopniach swobody. Niepożądaną rzeczą występującą w modelu jest autokorelacja dodatnia składnika losowego. Dlatego mimo istotności statystycznej parametrów występujących w modelu, zachodzącej relacji $F = 27,99 > F_{0,05;2;9} = 4,26$, oraz dobrego dopasowania modelu do danych rzeczywistych wynoszącego 86% nie powinno raczej używać się tego modelu do celów prognostycznych.



Rysunek 7. Ilustracja graficzna modelu liniowego (płaszczyzny) emigracji Polaków do krajów spoza Unii Europejskiej na tle emigracji do Niemiec i Wielkiej Brytanii sporządzona w oparciu o relację $SUE = -73,74 + 0,2 \cdot N + 0,08 \cdot WB$

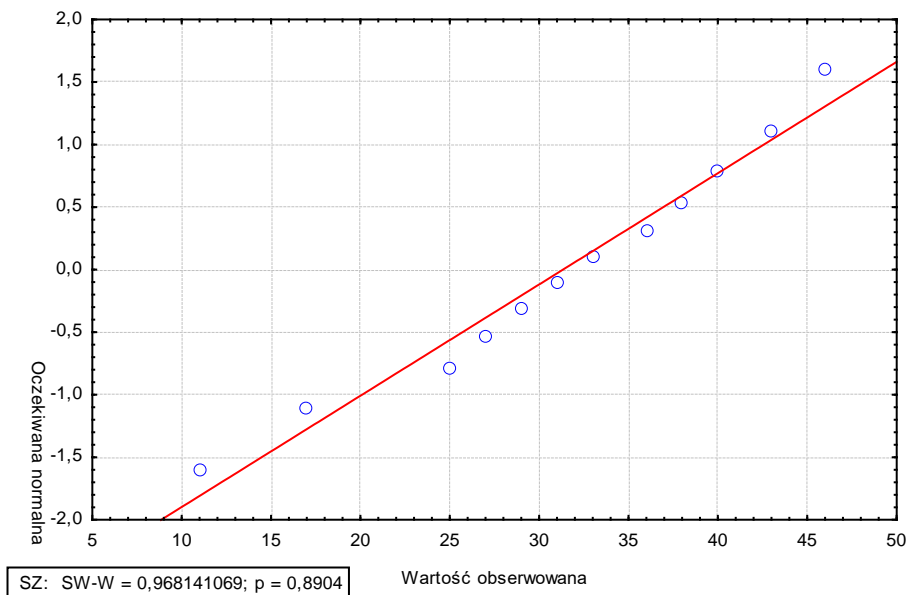
Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 6.0.

Rozpatrując model regresji gdzie zmienną objaśnianą jest SUE a zmiennymi objaśniającymi: N , WB , W otrzymano:

$$SUE = -200,69 + 0,19 \cdot N - 0,08 \cdot WB + 2,62 \cdot W, R^2 = 0,98, \quad (2)$$

(17,04) (0,02) (0,02) (0,02)

który jest liniowy, ponieważ względne błędy oszacowań parametrów są w każdym przypadku mniejsze od 50%. Ponadto jak łatwo sprawdzić wszystkie parametry są istotne statystycznie wobec tego, że wartość krytyczna statystyki t –Studenta przy poziomie istotności 0,05 oraz 8 stopniach swobody przyjmuje wartość $t_{0,05;8} = 2,306$. Model ma również własności prognostyczne gdyż wartość statystyki F – Snedecora obliczona za pomocą funkcji Reglinp w Excelu spełnia relację $F = 189,27 > F_{0,05;3;8} = 4,066$. Jak pokazały liczne symulacje dokonane za pomocą funkcji statystycznej Reglinp (które wobec objętości artykułu opuszczono) nie wszystkie modele regresji o trzech, czterech, pięciu, sześciu, siedmiu i ośmiu zmiennych objaśniających są liniowe jeżeli traktujemy SUE jako zmienną objaśnianą a emigrację do poszczególne krajów jako zmienne objaśniające. Na koniec przytoczono przykłady zastosowania elementów wnioskowania statystycznego do zmiennych SZ oraz SUE .



Rysunek 8 . Ilustracja wykorzystania testu Shapiro –Wilka do zbadania normalności rozkładu zmiennej SZ

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 6.0.

Stosując test Shapiro – Wilka przyjęto, że zmienna SZ ma w przybliżeniu rozkład normalny. Wariancja z próby dotyczącej zmiennej SZ ma wartość $s^2 = 107,12$.

Założmy, że próba SZ pochodzi z populacji o rozkładzie normalnym. Przyjmując współczynnik ufności $1 - \alpha = 0,9$; $\frac{\alpha}{2} = 0,05$ oraz $1 - \frac{\alpha}{2} = 0,95$ z tablic rozkładu χ^2 , dla $1 - \frac{\alpha}{2} = 0,95$ i 11 – stopni swobody odczytano, że $z_1 = 4,575$. Podobnie $z_2 = 19,675$, dla $\frac{\alpha}{2} = 0,05$ i 11-stopni swobody.

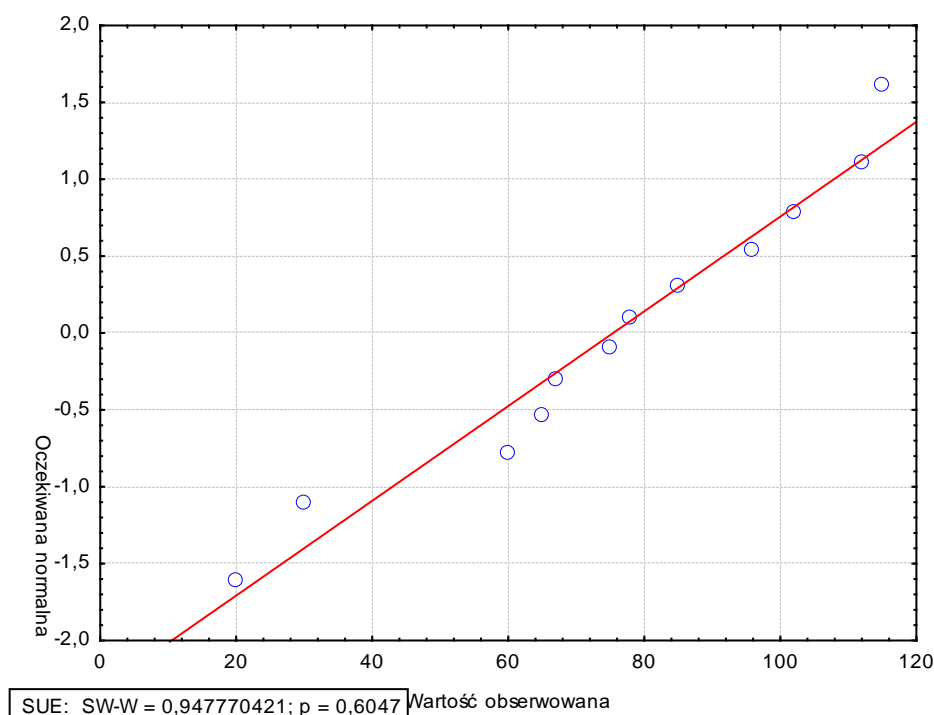
Zatem dla małej próby 12 elementowej zachodzi:

$$\frac{12 \cdot 107,12}{19,675} < \sigma^2 < \frac{12 \cdot 107,12}{4,575} \Leftrightarrow 65,33 < \sigma^2 < 280,97$$

Podobnie, rozpatrując przedział ufności dla odchylenia standardowego, i pierwiastkując lewą i prawą stronę ostatniej nierówności dostajemy:

$$\sqrt{65,33} < \sigma < \sqrt{280,97} \Leftrightarrow 8,08 < \sigma < 16,76$$

Zatem przedział (8,08; 16,76) wyrażony w tysiącach z ufnością 0,9 pokrywa prawdziwą wartość odchylenia standardowego emigracji Polaków do Szwecji w latach 2004-2016.



Rysunek 9. Ilustracja wykorzystania testu Shapiro –Wilka do zbadania normalności rozkładu zmiennej SUE

Źródło: opracowanie własne z wykorzystaniem pakietu Statistica 6.0.

Stosując test Shapiro-Wilka przyjęto, że zmienna *SUE* ma w przybliżeniu rozkład normalny. Wartość średnia z próby dotyczącej emigracji Polaków do krajów spoza Unii Europejskiej wynosi $\bar{x} = 75,42$ tys., zaś odchylenie standardowe $s = 29,67$ tys. Załóżmy, że próba *SUE* pochodzi z populacji o rozkładzie normalnym. Przyjmując hipotezę zerową mówiącą o tym, że wartość średnia z próby wynosi 77 tys., osób tzn. $H_0 : m_0 = 77$ tys., przy hipotezie alternatywnej $H_1 : m \neq m_0$ mamy do czynienia z małą próbą i obustronnym obszarem krytycznym. Do weryfikacji użyto testu *t* – Studenta:

$$t = \frac{\bar{x} - m_0}{s} \sqrt{n-1} = \frac{75,42 - 77}{29,67} \sqrt{11} = -0,1768.$$

Z tablic rozkładu *t* – Studenta dla poziomu istotności 0,05 i 11 stopni swobody otrzymano $t_{0,05} = 2,201$. Ponieważ $|t| = 0,1768 < t_{0,05} = 2,201$ więc nie ma podstaw do odrzucenia hipotezy zerowej mówiącej o tym, że przeciętna wielkość emigracji Polaków do krajów spoza Unii Europejskiej w latach 2004 -2016 wynosiła 77 tysięcy.

Podsumowanie

Założone cele pracy podane w wprowadzeniu artykułu zostały zrealizowane. Przeprowadzona analiza pokazała, że w latach 2004-2015 występowała dość masowa emigracja Polaków za granicę. Istotnym powodem tego było między innymi wejście Polski do Unii Europejskiej co wiązało się w dalszej kolejności z otwarciem granic i uwolnieniem rynku pracy. Rosnący w przybliżeniu trend liniowy dotyczył emigracji Polaków do Europy oraz krajów spoza Unii Europejskiej. Najwięcej naszych rodaków wyjechało czasowo w okresie 2004-2015 do Wielkiej Brytanii, Niemiec, Irlandii i Holandii. Badane rozkłady zmiennych wykazywały głównie lewostronną asymetrię rozkładu. W przypadku jedenastu na dwadzieścia badanych zmiennych występował dodatni współczynnik kurtozy a w przypadku dziewięciu – ujemny współczynnik kurtozy. Jak pokazuje tabela 4, między emigracją Polaków do poszczególnych krajów (wybrano kraje gdzie wyjechało najwięcej Polaków) występowały

najczęściej wysokie, dodatnie i istotne statystycznie korelacje tzn. wzrost emigracji do jednego kraju powodował wzrost emigracji do innego badanego kraju. W przypadku czterech krajów Unii Europejskiej odnotowano regresje liniowe, które łączyły emigrację Polaków do tych krajów i emigrację do krajów spoza Unii Europejskiej. Regresja liniowa występowała również w przypadku emigracji do Niemiec, Wielkiej Brytanii i krajów spoza Unii Europejskiej (zmienna objaśniana), jak również w przypadku emigracji do Niemiec, Wielkiej Brytanii, Włoch i krajów spoza Unii Europejskiej (zmienna objaśniana). Model liniowy nie występował w przypadku trzech najbardziej licznych emigracji Polaków: do Niemiec, Wielkiej Brytanii, Irlandii (zmienna objaśniana) i krajów spoza Unii Europejskiej (zmienna objaśniana). Model liniowy występował w przypadku emigracji do Europy (zmienna objaśniana) i krajów spoza Unii Europejskiej (zmienna objaśniana), jak również w przypadku Unii Europejskiej (zmienna objaśniana) oraz krajów spoza Unii Europejskiej (zmienna objaśniana). Ponadto znaleziono przedział ufności dla wariancji i odchylenia standardowego czasowej emigracji Polaków do Szwecji oraz zweryfikowano hipotezę o przeciętnej wielkości emigracji Polaków do krajów spoza Unii Europejskiej za pomocą testu istotności t-Studenta. Do ujemnych skutków czasowej emigracji Polaków za granicę należy zaliczyć utratę siły roboczej oraz wykształconej kadry. Ponadto długookresowa rozłąka niejednokrotnie powodowała destrukcyjny wpływ na więzi rodzinne. Dodatkowo skutki emigracji to: tworzenie powiązań Polski z zagranicą, możliwość realizacji aspiracji zawodowych, transfer wiedzy, lepsza znajomość języków. Po 2015 roku jak wynika z informacji środków masowego przekazu trend dotyczący czasowej emigracji Polaków za granicę zaczął się stopniowo odwracać. Część emigrantów powróciła do Polski co wiąże się z coraz bardziej dynamicznie rozwijającą się gospodarką Polski i budzi pewien optymizm na przyszłość.

Some remarks on the statistical analysis of temporary emigration of Poles abroad in the years 2004-2015

Summary:

The article presents a statistical analysis of the temporary emigration of Poles abroad in the years 2004-2015 using for this purpose data from the Central Statistical Office (2017). The statistical method for correlation analysis, basic descriptive statistics as well as statistical inference elements were used to examine these data. Linear models were also built that connect the emigration of Poles to some European countries with emigration to countries outside the European Union. The article ends with relevant conclusions.

Keywords: statistical analysis, linear regression, emigration of Poles abroad