

Barbara M. Cieślik

Instytut Ekonometrii, Kolegium Analiz Ekonomicznych
Szkoła Główna Handlowa w Warszawie
barbara.cieslik@sgh.waw.pl
ORCID: 0000-0002-0492-9488

Ewa M. Syczewska

Instytut Ekonometrii, Kolegium Analiz Ekonomicznych
Szkoła Główna Handlowa w Warszawie
ewams@sgh.waw.pl
ORCID: 0000-0003-1401-4844

Sezonowe wahania umieralności w wybranych państwach Europy Środkowo-Wschodniej i Zachodniej

Streszczenie

Artykuł poświęcono analizie sezonowości umieralności w wybranych państwach Europy Środkowo-Wschodniej oraz Zachodniej w latach 1999–2019 na podstawie danych Eurostatu. Do statystycznego opisu sezonowości obserwowanej w danych wykorzystano miary proponowane w literaturze, a występujące wahania zwizualizowano za pomocą wykresów. We wszystkich badanych populacjach stwierdzono (typową dla państw rozwiniętych na półkuli północnej) wyraźną nadumieralność w okresie zimowym. Określono podobieństwa i różnice we wzorcach tej sezonowości oraz w przeciętnym poziomie umieralności między badanymi państwami. Do modelowania stwierdzonej sezonowości liczby zgonów zostały wykorzystane modele deterministycznej sezonowości ze zmiennymi binarnymi dla poszczególnych miesięcy. Artykuł stanowi wstęp do dalszych badań, w tym do oceny wpływu pandemii COVID-19 na umieralność w poszczególnych państwach.

Słowa kluczowe: sezonowość umieralności, modele ze zmiennymi zero-jedynkowymi, modelowanie sezonowości

Wprowadzenie

Występowanie sezonowych różnic umieralności, w szczególności nadwyżka zgonów dla krajów półkuli północnej w miesiącach zimowych, było wielokrotnie analizowane w literaturze przedmiotu (Bull, Morton, 1978; Grut, 1987). Zimowy wzrost umieralności jest tłumaczony występowaniem niskich temperatur, wzrostem zanieczyszczenia powietrza drobnymi pyłami (smogiem) oraz przypadkami grypy i infekcji grypopochodnych. Badacze wskazują, że możliwe jest zniwelowanie tak dużych różnic przez zwiększenie wydatków na służbę zdrowia, wzrost PKB *per capita* i zmniejszanie nierówności ekonomicznych (Healy, 2003), a także zwracają uwagę na to, że amplituda wahań sezonowych może być związana z poziomem życia (Mielecka-Kubień, 2018).

Początek pandemii COVID-19 i wprowadzanych ograniczeń nastąpił w lutym–marcu 2020 r., czyli w środku okresu, w którym zazwyczaj stwierdzano sezonowy wzrost zachorowań na grypę i wzrost liczby zgonów z różnych przyczyn. W wielu państwach nastąpił znaczący wzrost liczby zgonów – w Europie szczególnie dramatycznymi przykładami są Wielka Brytania, Hiszpania, Włochy, Szwecja i Francja¹. Jednocześnie wstępne dane z niektórych państw, w tym z Polski, wskazują, paradoksalnie, na obniżenie liczby zgonów w tym okresie, poprzez ograniczenie liczby infekcji grypowych (dzięki podwyższeniu reżimu sanitarnego) oraz liczby przypadków powikłań pooperacyjnych (szpitale odwołały w tym okresie większość planowych operacji)². Być może rozprzestrzenianie się COVID-19 i ostre objawy tej infekcji skutkujące podwyższoną śmiertelnością będą bardziej nasilone w miesiącach zimowych (Lauc, Markotic, Gornik, Primorac, 2020). Jednocześnie można się spodziewać przesuniętego w czasie względem początku pandemii wzrostu liczby zgonów wynikającego z zaniedbań w diagnostyce i leczeniu wielu chorób, szczególnie nowotworowych i kardiologicznych. Do późniejszej analizy wpływu obecnej pandemii na umieralność

¹ Według danych na 04.09.2020, czyli po około pół roku trwania epidemii, w tych państwach europejskich natężenie zgonów na COVID-19 przekracza już 40 na 100 000 osób, en.wikipedia.org/wiki/COVID-19_pandemic_death_rates_by_country (dostęp: 06.09.2020).

² Informacja pobrana ze strony internetowej www.medonet.pl/koronawirus/koronawirus-w-polsce,jak-epidemia-wplynela-na-statystyki-zgonow-w-polsce-,artykul,93145387.html (dostęp: 06.09.2020).

w poszczególnych populacjach konieczne jest poznanie charakterystyk obecnie obserwowanych sezonowych wzorców umieralności w każdej z tych populacji.

Niniejsze badanie oparto na miesięcznych danych Eurostatu dotyczących liczby zgonów dla Polski, Czech, Słowacji, Węgier, Litwy i Łotwy oraz – dla porównania – Francji, Niemiec i Włoch. Przedmiotem badania jest sezonowość umieralności, a jego celem jest identyfikacja, pomiar i modelowanie tej sezonowości w wybranych populacjach, w tym testowanie istotności nadwyżek umieralności w okresach zimowych, analiza różnic i podobieństw wzorca sezonowego pomiędzy krajami.

Opis wykorzystanego zbioru danych

Zastosowano dane miesięczne dotyczące liczby zgonów oraz dane roczne dotyczące wielkości populacji pochodzące z bazy Eurostatu według stanu na 25.08.2020. Zamiarem auterek było porównanie wybranych sześciu państw Europy Środkowo-Wschodniej z wybranymi trzema państwami Europy Zachodniej. Wybór państw był podyktowany kompletnością i wiarygodnością danych.

Podstawową zmienną jest miesięczna liczba zgonów dla każdego z badanych państw w okresie 01.1999–06.2019 ogółem – $Deaths_{Country\ t}$, gdzie *Country* oznacza symbol kraju: *PL* – Polska, *CZ* – Czechy, *HU* – Węgry, *LT* – Litwa, *LV* – Łotwa, *DE* – Niemcy, *IT* – Włochy, *FR* – Francja, *SV* – Słowacja, a indeks *t* – numer kolejny obserwacji. Pominięto przypadki zgonów nieprzypisanych do żadnego miesiąca. Liczby zgonów wyrównano (co jest stosowane powszechnie w literaturze w celu zapewnienia porównywalności umieralności w poszczególnych miesiącach³), przeliczając je na miesiące 30-dniowe według wzoru:

$$Deaths30_{Country\ t} = Deaths_{Country\ t} \cdot 30/n, \quad (1)$$

gdzie *n* – liczba dni w danym miesiącu, *t* – numer kolejny obserwacji; $t = 1, 2, \dots, 82$, *Country* – skrót nazwy państwa.

Następną zmienną jest liczebność populacji danego państwa – $PopulationY_{Country\ t}$ – podawana w Eurostatie w ujęciu rocznym, według stanu na 1 stycznia. Zgodnie z powszechnie stosowaną praktyką została dokonana interpolacja wartości tej zmiennej z danych rocznych na miesięczne według wzoru:

³ Sumaryczna hipotetyczna roczna liczba zgonów dla miesięcy 30-dniowych jest o około 1% niższa od rzeczywistej rocznej liczby zgonów, gdyż rok kalendarzowy ma 365 lub 366 dni, a nie 360.

$$PopulationM_{Country\ t} = PopulationY_{Country\ t} + \frac{i-1}{12} (PopulationY_{Country\ t+12} - PopulationY_{Country\ t}), \quad (2)$$

gdzie i – numer miesiąca ($i = 1, \dots, 12$).

Zauważono skokowe korekty wartości zmiennej $PopulationY_{Country\ t}$ dla niektórych państw. Przykładowo w 2011 r. liczby ludności Litwy i Łotwy zostały w danych Eurostatu skorygowane (odpowiednio o $-2,84\%$ i $-2,16\%$), a spis powszechny w Niemczech wykazał w tymże roku spadek liczby ludności w stosunku do roku 2010 o 1,58 mln osób (zmiana względna na poziomie $-1,93\%$)^{4,5}. Obliczenia bazują na tych danych, bez stosowania wygładzania (poza wspomnianą interpolacją).

Na podstawie powyższych zmiennych obliczono wartości współczynnika zgonów: $d_rate_{Country\ t}$, czyli natężenie zgonów w poszczególnych miesiącach w przeliczeniu na 1000 osób danej populacji w skali roku (przy przyjęciu miesiący 30-dniowych), według wzoru:

$$d_rate_{Country\ t} = (Deaths30_{Country\ t} / PopulationM_{Country\ t}) \cdot 1000 \cdot 12, \quad (3)$$

co zapewnia porównywalność danych z państw o populacjach różnej wielkości przez uzyskanie miary określającej, ile osób na 1000 z danej populacji umarłoby w ciągu jednego roku, gdyby natężenie umieralności przez cały rok (12 miesięcy) było takie, jak obserwowane w danym miesiącu⁶. Na podstawie wartości tego współczynnika przeprowadzono analizę sezonowości zaprezentowaną w dalszej części artykułu.

Analiza wzorców sezonowych

Do analizy podobieństw i różnic w szeregach czasowych współczynników zgonów dla analizowanych państw można wykorzystać pojęcie wzorca sezonowego (*seasonal pattern*) znane z analizy szeregów czasowych. Według Hyndmana i Athanasopouloasa (2018) wzorzec sezonowy w szeregu czasowym występuje, gdy na szereg wpływają

⁴ Por. www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/bevoelkerungsforschreibung-2017.pdf?_blob=publicationFile (dostęp: 22.09.2020).

⁵ We Włoszech w danych Eurostatu wystąpił wzrost liczby ludności w 2014 r. o 1,84% w porównaniu z 2013 r. Największa względna zmiana dla Polski to spadek liczby ludności w 2000 r. o 1,04%.

⁶ Por. *Biuletyn Statystyczny* (2020), 3, s. 44 – podane współczynniki zgonów na 1000 ludności dla poszczególnych miesięcy są tam przeskalowywane (przemnożone przez 12), co ułatwia ich porównywanie z analogicznymi współczynnikami dla danych rocznych. Analogiczny zabieg można zastosować dla danych kwartalnych (mnożnie razy 4) czy tygodniowych (razy 52).

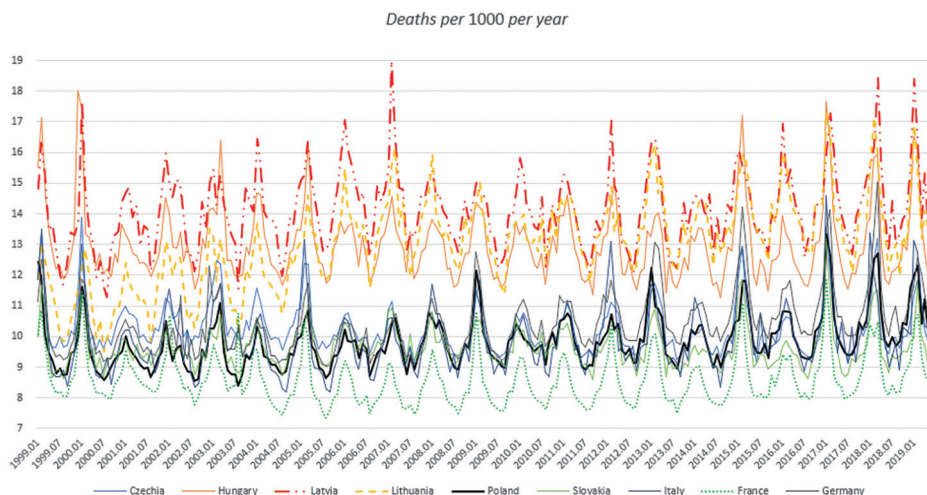
czynniki sezonowe (np. kwartał roku, miesiąc lub dzień tygodnia). Sezonowość ma zawsze określony i znany okres.

Porównując szeregi czasowe współczynników zgonów dla analizowanych państw pod względem podobieństw i różnic wzorców sezonowych, można brać pod uwagę kształt wahań sezonowych przejawiający się w długości okresu, rozłożeniu wahań oraz w ich amplitudzie.

Na rysunku 1 zobrazowano wartości współczynnika zgonów dla badanych dziewięciu państw w poszczególnych miesiącach całego analizowanego okresu. Wyraźnie widać powtarzający się wzorec sezonowy o okresie dwunastomiesięcznym, charakteryzujący się istotnie podwyższoną umieralnością w miesiącach przypadających na okres zimowy. Charakterystyczny jest nie tylko wyższy średni poziom umieralności w tym okresie, ale też wyższa wariancja. Zauważamy, że analizowane państwa dzielą się na dwie grupy, istotnie różniące się przeciętnym poziomem umieralności – do państw o niższej umieralności należą wszystkie analizowane państwa Europy Zachodniej oraz Polska, Czechy i Słowacja. Do państw o wyższej umieralności należą Węgry, Łotwa oraz Litwa, w której na początku badanego okresu umieralność była na poziomie zbliżonym do pierwszej grupy państw. Znacząco wyróżnia się też Francja jako państwo o najniższej średniej umieralności w całym analizowanym okresie.

Rysunek 1. Liczba zgonów na 1000 ludności w przeliczeniu na rok w okresie 01.1999–06.2019

Figure 1. Deaths per 1000 per year, since January 1999 to June 2019.



Źródło: obliczenia własne na podstawie danych Eurostatu.

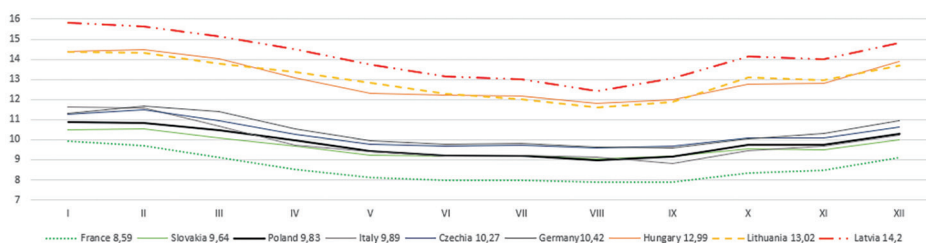
Source: own computations based on the Eurostat data.

Na rysunku 2 przedstawiono wieloletnie średnie miesięczne współczynniki zgonów (na 1000 osób w ciągu roku) w poszczególnych państwach. W legendzie tego rysunku państwa są uszeregowane rosnąco według przeciętnej wartości współczynnika zgonów w całym badanym okresie. Ta wartość jest najniższa dla Francji, gdzie w ciągu roku umiera mniej niż 9 osób na 1000, a najwyższa dla Łotwy (ponad 14 osób na 1000 w skali roku).

Rysunek 1 stanowi też potwierdzenie tego, że w Polsce wystąpiła podwyższona częstość zgonów w latach 2016–2018 – zimowe piki w tych latach są dość wyraźnie wyższe niż w poprzedzającej ten okres dekadzie (Wojtyniak, Rabczenko, Czarkowski, Skotak, 2018). Bliższa analiza (której tu nie przytaczamy) wykazała, że analogiczne wzrosty miały miejsce też w innych państwach środkowoeuropejskich oraz w Niemczech.

Rysunek 2. Średnie wieloletnie miesięcznej umieralności w wybranych państwach na 1000 ludności na rok w okresie 01.1999–06.2019

Figure 2. Average deaths rate by country per 1000, since January 1999 to June 2019



Źródło: obliczenia własne na podstawie danych Eurostatu.

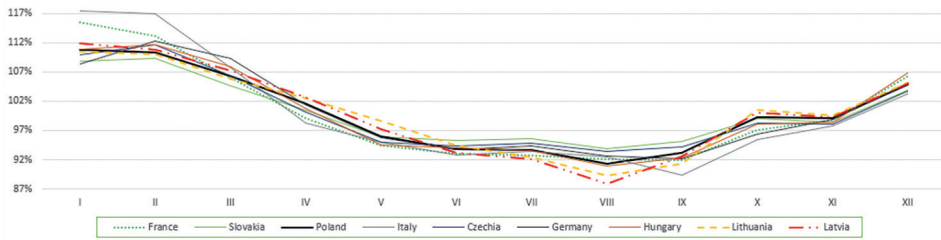
Source: own computations based on the Eurostat data.

Rysunek 3 przedstawiający relację umieralności dla danego miesiąca do średniej rocznej w danej populacji ilustruje występowanie wyraźnych podobieństw wzorca sezonowego w badanej grupie państw. Przede wszystkim widać, że kształt poszczególnych łamanych oraz względna amplituda obserwowanych wahań sezonowych są zbliżone. Opisana już podwyższona częstość zgonów w miesiącach przypadających na okres zimowy mieści się tu w przedziale 108,4–117,4% średniej rocznej w styczniu i w lutym oraz 104,6–109,3% w marcu, przy czym efekt wyższej umieralności w styczniu i lutym jest najsilniejszy we Włoszech. Styczeń jest najgorszym miesiącem pod względem średniej wartości współczynnika umieralności dla Włoch, Francji, Łotwy, Litwy i Polski, a dla pozostałych państw współczynnik ten jest najwyższy w lutym, co jest szczególnie silnie widoczne w przypadku Niemiec. Dla większości analizowanych państw miesiącem o najniższej umieralności jest sierpień, co

jest szczególnie wyraźne w przypadku Łotwy i Litwy. Ponadto widać (na przykładzie Włoch) potwierdzenie opisanego w literaturze (Falagas i in., 2009) fenomenu września w państwach rejonu śródziemnomorskiego, dla których to ten miesiąc charakteryzuje się najniższą umieralnością. Także dla Niemiec i Francji wrzesień okazał się pod tym względem lepszy od sierpnia, ale nieznacznie.

Rysunek 3. Relacja miesięcznej umieralności do średniej umieralności dla danego państwa w wybranych państwach w okresie 01.1999–06.2019

Figure 3. Monthly *d_rate* to average *d_rate* ratio, by country, since January 1999 until June 2019



Źródło: obliczenia własne na podstawie danych Eurostatu.

Source: own computations based on the Eurostat data.

W celu dokładniejszego zbadania zaobserwowanej sezonowości, dla każdego państwa obliczono trzy miary sezonowości zdefiniowane w literaturze (Rau, 2007). Pierwszą z nich jest wskaźnik zima/lato (*winter/summer ratio*) oznaczony symbolem φ_1 , wyrażający względną nadwyżkę umieralności w okresie zimowym (I, II i III) w porównaniu z okresem letnim (VII, VIII, IX), kiedy obserwowana umieralność jest na ogół najniższa:

$$\varphi_1 = \frac{\sum_{i=I}^{III} Deaths_i}{\sum_{j=VII}^{IX} Deaths_j}, \quad (4)$$

gdzie $\sum_{i=I}^{III} Deaths_i$ oznacza sumaryczną liczbę zgonów w pierwszym kwartale,

a $\sum_{j=VII}^{IX} Deaths_{30j}$ analogiczną sumę dla trzecich kwartałów. W literaturze są też spo-

tykane inne wersje tej miary – np. porównanie sumarycznej liczby zgonów w XII, I i II

do liczby zgonów w VI, VII i VIII (Nobbe, 2017) lub wskaźnik CSVM⁷ – *coefficient of seasonal variation in mortality* (Healy, 2003) – ale w przypadku wszystkich analizowanych w tym artykule populacji miesiącami o najwyższej umieralności są I, II i III.

Ponieważ analizowane tu dane obejmowały okres 20,5 roku, czyli niepełny cykl, a ponadto przedmiotem porównań są współczynniki zgonów obliczone według wzoru (3), wzór ten został na potrzeby tego artykułu zmodyfikowany do postaci, której istotą jest porównanie średniej umieralności w miesiącach zimowych do średniej umieralności w miesiącach letnich w całym analizowanym okresie:

$$\varphi_{1m} = \overline{d_rate_{I_III}} / \overline{d_rate_{VII_IX}}. \quad (5)$$

We wzorze (5) w liczniku jest średnia wartość współczynników zgonów we wszystkich pierwszych kwartałach analizowanego okresu (było ich 21), a w mianowniku analogiczna średnia dla trzecich kwartałów (których było 20).

Wskaźnik ten nie uwzględnia roli pozostałych miesięcy w obserwowanej sezonowości, co może mieć znaczenie w populacjach o innym wzorcu miesięcznej zmienności umieralności, dlatego stosuje się też inne miary. Do ich wyznaczenia należy obliczyć udziały zgonów przypadających na dany miesiąc w całkowitej liczbie zgonów przypadających na rok, co może być też interpretowane jako prawdopodobieństwo tego, że zgon losowo wybranej osoby z danej populacji nastąpi w określonym miesiącu. Wartości te zostały obliczone według wzoru:

$$p_i = \overline{d_rate_i} / \overline{d_rate}, \quad (6)$$

gdzie $i \in \{I, II, \dots, XII\}$, a średnie ponownie są obliczone dla całego analizowanego okresu.

Druga miara sezonowości umieralności omawiana przez Rau (2007, s. 43) to miara nierównomierności rozkładu udziału zgonów z poszczególnych miesięcy w ogólnej liczbie zgonów, zdefiniowana, na wzór współczynnika Giniego, jako połowa sumy wartości bezwzględnych odchyłeń prawdopodobieństw zgonu w poszczególnych miesiącach od wartości charakteryzujących rozkład jednostajny. Symbolem tej miary jest φ_2 :

$$\varphi_2 = \frac{1}{2} \sum_{i=I}^{XII} \left| p_i - \frac{1}{12} \right|. \quad (7)$$

⁷ Definiowany jako nadwyżka zgonów w okresie grudzień–marzec (miesiące „zimowe”) nad połową zgonów w pozostałych miesiącach (tj. kwiecień–listopad) odniesiona do połowy zgonów w okresie kwiecień–listopad.

Jedną z wad tego wskaźnika jest jego niski zakres zmienności. Nietrudno wykazać, że jego wartości nie mogą przekroczyć 0,91667, lecz według Rau (2007, s. 44) oscylują zwykle wokół wartości 0,03.

Trzecią miarą jest wskaźnik φ_3 , mierzący poziom entropii, czyli niepewności. Jego konstrukcja odpowiada założeniu, że gdyby umieralność we wszystkich miesiącach była jednakowa, to φ_3 miałyby wartość 1; gdyby natomiast wszystkie zgony następowały w jednym z 12 miesięcy, to wartość φ_3 byłaby równa 0. Wzór (8) podany w (Rau, 2007, s. 44):

$$\varphi_3 = \frac{\log_2(n) - \frac{1}{n} \sum_{i=1}^{12} n_i \log_2(n_i)}{\log_2 12}, \quad (8)$$

gdzie n_i to liczba zdarzeń (tu zgonów) w miesiącu i , a n to liczba wszystkich zdarzeń w ciągu roku, został na potrzeby tego artykułu zmodyfikowany do postaci⁸:

$$\varphi_{3m} = \frac{-\sum_{i=1}^{XII} p_i \log_2(p_i)}{\log_2 12}. \quad (9)$$

Wskaźnik ten również charakteryzuje się niewielką zmiennością – jego wartości są zwykle bardzo bliskie 1; im wyższa rola sezonowości, tym niższa wartość φ_{3m} . Wartości wszystkich miar dla każdego z analizowanych państw zawiera tabela 1.

Zastosowane miary sezonowości dają spójny obraz sytuacji – w tabeli 1 nie ma zmiany ani na pierwszym miejscu (największe sezonowe wahania umieralności można zaobserwować we Włoszech), ani na ostatnich trzech (państwa o najniższej roli sezonowości to Słowacja, Czechy i Polska). Dla państw ze środka rankingów wartości miar φ_{1m} , φ_2 i φ_{3m} są bardzo zbliżone. Potwierdza to obserwację o podobieństwie wzorca sezonowego w badanych populacjach. Także współczynniki korelacji liniowej Pearsona wyznaczone dla wartości współczynnika zgonów (tabela 2) potwierdzają podobieństwo wzorca sezonowego w badanej grupie państw. Należy tu pamiętać, że współczynnik korelacji liniowej Pearsona bada związek między odchyleniami obu korelowanych zmiennych od swoich średnich, czyli w tym przypadku pozwala stwierdzić podobieństwo szeregów czasowych wynikające ze zgodności kierunku i siły tych odchyżeń, a nie z wartości średniej każdego z nich. Najsilniejsze podobieństwo

⁸ Nietrudno sprawdzić, że wzory te są sobie tożsame, gdy $p_i = n_i/n$, bowiem różnica logarytmów dwóch liczb jest równa logarytmowi ich ilorazu. Dodatkowo można zauważyć, że w każdym z tych wzorów można zmienić podstawę logarytmu (np. zastosować logarytm naturalny lub dziesiętny), co nie wpłynie na wynik.

pod tym wzgłędem wykazują szeregi dla Czech i Węgier (wyraźnie różniące się za to pod wzgłędem średniego poziomu umieralności), najsłabsze – dla Czech i Litwy. Polska wykazuje najsilniejsze podobieństwo z Niemcami.

Tabela 1. Rankingi analizowanych państw według wartości miar sezonowości umieralności

Table 1. Rank of the analysed countries, based on seasonal measures of d-rates

φ_{1m}	Oznaczenie kraju Country symbol	Państwo Country	φ_2	Oznaczenie kraju Country symbol	Państwo Country	φ_{3m}	Oznaczenie kraju Country symbol	Państwo Country
1,243	IT	Włochy	0,038	IT	Włochy	0,99840	IT	Włochy
1,210	LV	Łotwa	0,034	FR	Francja	0,99875	FR	Francja
1,209	FR	Francja	0,032	HU	Węgry	0,99889	LV	Łotwa
1,196	LT	Litwa	0,031	LV	Łotwa	0,99898	HU	Węgry
1,193	HU	Węgry	0,030	DE	Niemcy	0,99906	LT	Litwa
1,182	DE	Niemcy	0,029	LT	Litwa	0,99908	DE	Niemcy
1,178	PL	Polska	0,028	PL	Polska	0,99915	PL	Polska
1,162	CZ	Czechy	0,027	CZ	Czechy	0,99924	CZ	Czechy
1,134	SV	Słowacja	0,022	SV	Słowacja	0,99946	SV	Słowacja

Źródło: obliczenia własne na podstawie danych Eurostatu.

Source: own computation based on the Eurostat data.

Tabela 2. Współczynniki korelacji między szeregami czasowymi współczynników zgonów

Table 2. Pearson correlation coefficients between time series of deaths coefficients

	CZ	HU	LV	LT	PL	SV	IT	FR	DE
Czechy	1								
Węgry	0,897	1							
Łotwa	0,719	0,749	1						
Litwa	0,427	0,510	0,748	1					
Polska	0,763	0,783	0,776	0,779	1				
Słowacja	0,847	0,867	0,761	0,611	0,786	1			
Włochy	0,811	0,809	0,713	0,612	0,821	0,780	1		
Francja	0,859	0,825	0,702	0,475	0,778	0,768	0,884	1	
Niemcy	0,757	0,731	0,695	0,726	0,895	0,739	0,809	0,794	1

Źródło: obliczenia własne na podstawie danych Eurostatu.

Source: own computation based on the Eurostat data

Testowanie istotności sezonowości

Przedmiotem analizy w tej części są wartości współczynników zgonów dla poszczególnych populacji, określone wzorem (3) i oznaczone jako $d_rate_{Country\ t}$.

Wobec braku wyraźnego trendu⁹ w analizowanych szeregach, dla zbadania, czy zaobserwowane wahania sezonowe są statystycznie istotne można wykorzystać najprostszy model deterministycznej sezonowości – tu został wykorzystany model z wyrazem wolnym i zero-jedynkowymi zmiennymi sezonowymi¹⁰. Model ten jest tożsamy z proponowanym przez Eurostat modelem określanym jako *simplest deterministic seasonal model* (Mazzi, 2018, s. 60) i ma postać:

$$d_rate_{Country\ t} = \alpha_0 + \sum_{i=1}^{11} \alpha_i Dmi_t + \varepsilon_t, \quad (10)$$

gdzie α_i oznaczają parametry regresji liniowej, $i=0, 1, 2, \dots, 11$, ε_t – składnik losowy, a sezonowe zmienne zero-jedynkowe są określone wzorem:

$$Dmi = \begin{cases} 1 & \text{w miesiącu } i\text{-tym,} \\ 0 & \text{w pozostałych miesiącach.} \end{cases} \quad (11)$$

Wyniki estymacji modelu podano w tabeli 3. Pokazano, oprócz ocen parametrów, wartości współczynnika determinacji i skorygowanego współczynnika determinacji, kryteriów informacyjnych AIC Akaike (1974) i BIC Schwarza (1978). Błędy estymacji są w wersji odpornej na występowanie autokorelacji i heteroskedastyczności, HAC – *heteroskedasticity and autocorrelation consistent* (zob. Newey, West, 1987).

W specyfikacji tego modelu uwzględniono wyraz wolny, a pominięto zmienną $Dm12$ dla grudnia, zatem oceny parametrów przy zmiennych zero-jedynkowych interpretuje się jako przeciętną nadwyżkę współczynnika zgonów w danym miesiącu nad wartością przeciętną z grudnia w okresie objętym próbą, ocena zaś wyrazu wolnego odpowiada uśrednionemu poziomowi tej zmiennej dla grudnia i stanowi punkt odniesienia do porównań dla pozostałych miesięcy¹¹.

⁹ W modelach trendu z wahaniami miesięcznymi trend był nieistotny.

¹⁰ Natomiast do celów opisu zjawiska oraz do prognoz, ze względu na niestacjonarność zmiennych i stacjonarność ich przyrostów, właściwszy będzie model dynamiczny SARIMA, w którym zmienną objaśnianą są przyrosty lub przyrosty sezonowe badanej zmiennej.

¹¹ Oceny parametrów wariantu modelu bez wyrazu wolnego, a z dwunastoma zmiennymi zero-jedynkowymi łatwo obliczyć; pozostawiamy jednak powyższą postać modelu, gdyż ułatwia zauważenie wzrostów i spadków umieralności w miesiącach zimowych i letnich, odpowiednio.

Tabela 3. Wyniki estymacji deterministycznych modeli sezonowości, ze zmiennymi zero-jedynkowymi dla miesięcy I–XI
 Table 3. Deterministic seasonality models estimations with binary variables for months I–XI

Zmienna/Kraj	CZ	HU	PL	SV	LV	LT	IT	FR	DE
Wyraz wolny	10,644 ***	13,884 ***	10,296 ***	10,0056 ***	14,8212 ***	13,7064 ***	10,2132 ***	9,12 ***	10,9332 ***
Dm1	0,6396 **	0,5184 **	0,5892 ***	0,4776 ***	0,9840 ***	0,6792 **	1,3956 ***	0,7980 ***	0,3564 **
Dm2	0,8232 ***	0,6144 *	0,5436 ***	0,5292 ***	0,8268 ***	0,6216 **	1,3524 ***	0,5976 ***	0,7548 ***
Dm3	0,2808	0,1416	0,1524	0,0744	0,3156	0,0744	0,4392 ***	10,644	0,4536 *
Dm4	-0,3588 ***	-0,7956 ***	-0,3204 **	-0,3264 ***	-0,3312 *	-0,3552	-0,4992 ***	-0,6192 ***	-0,3780 **
Dm5	-0,8868 ***	-1,6020 ***	-0,8640 ***	-0,7632 ***	-1,0968 ***	-0,0721 ***	-0,0683 ***	-0,0842 ***	-0,0803 ***
Dm6	-0,9564 ***	-1,6728 ***	-1,0848 ***	-0,8244 ***	-1,6632 ***	-1,4196 ***	-1,0380 ***	-1,1328 ***	-1,1736 ***
Dm7	-0,9168 ***	-1,7280 ***	-1,1040 ***	-0,7932 ***	-1,8108 ***	-1,6764 ***	-0,9768 ***	-1,1604 ***	-1,1064 ***
Dm8	-1,0476 ***	-2,0712 ***	-1,3284 ***	-0,9588 ***	-2,4120 ***	-2,0820 ***	-1,0752 ***	-1,2216 ***	-1,2828 ***
Dm9	-0,9756 ***	-1,8876 ***	-1,1436 ***	-0,8424 ***	-1,7472 ***	-1,8228 ***	-1,3800 ***	-1,2228 ***	-1,3380 ***
Dm10	-0,5544 ***	-1,1340 ***	-0,5412 ***	-0,4656 ***	-0,6900 ***	-0,6132 ***	-0,7704 ***	-0,7872 ***	-0,8904 ***
Dm11	0,5724 ***	-1,1016 ***	-0,5580 ***	-0,5148 ***	-0,8184 ***	-0,7344 ***	-0,5400 ***	-0,6336 ***	-0,6396 ***
R ²	0,616	0,614	0,558	0,563	0,673	0,373	0,693	0,695	0,519
R ²	0,598	0,596	0,537	0,542	0,658	0,3435	0,68	0,681	0,497
AIC	-838,30	-647,71	-775,546	-900,985	-654,462	-429,453	-754,51	-888,78	-688,26
BIC	-796,24	-605,64	-733,481	-858,921	-612,398	-387,39	-712,45	-846,71	-646,20

Źródło: obliczenia własne na podstawie danych Eurostatu; *** istotność na poziomie 0,01; ** istotność na poziomie 0,05; * istotność na poziomie 0,10.
 Source: own computation based on the Eurostat data; *** significance on the level 0,01; ** significance on the level 0,05; * significance on the level 0,1.

Ocena parametru przy zmiennej D_{mi} , $i = 1, 2, \dots, 11$ jest zatem miernikiem uśrednionego po wszystkich latach w próbie odchylenia (dodatniego lub ujemnego) wartości współczynnika zgonów w danym miesiącu od jego średniego poziomu dla grudnia.

Oceny parametrów dla stycznia i lutego są dodatnie i statystycznie istotne, oznacza to wyższe niż w grudniu poziomy umieralności. Oceny parametrów dla marca także są dodatnie (poza Francją), ale w większości nieistotne (poza Włochami i Niemcami). Nawiązuje to do wspomnianego już we wprowadzeniu spostrzeżenia, że pierwszy kwartał w krajach półkuli północnej cechuje się wzrostem umieralności. Z kolei oceny parametrów przy zmiennych D_{m4} do D_{m11} są ujemne i statystycznie istotne (poza kwietniem dla Litwy), co oznacza, że poziom umieralności od kwietnia do listopada jest przeciętnie niższy niż w grudniu, zatem w cieplejszych miesiącach następuje obniżenie umieralności.

Analogiczne modele deterministycznej sezonowości dla zmiennych $Deaths_{Country t}$ mają inne wartości ocen parametrów, jednak układ znaków dodatnich lub ujemnych oraz schemat istotności zmiennych jest taki jak dla modeli przytoczonych w tabeli 3.

Podsumowanie

Na podstawie miesięcznych danych Eurostatu z okresu 01.1999–06.2019 wyznaczono dla współczynników zgonów w badanych populacjach trzy miary sezonowości opisane w literaturze przedmiotu (Rau, 2007) oraz współczynniki korelacji liniowej Pearsona. Stwierdzono zróżnicowanie średniej liczby zgonów na 1000 mieszkańców – wyraźny podział na dwie grupy państw: Polska, Czechy i Słowacja wraz z Francją, Niemcami i Włochami to państwa o niższej umieralności, oraz Węgry, Łotwa i Litwa – państwa z wyższą umieralnością. Na podstawie oszacowanego modelu deterministycznej sezonowości potwierdzono istotność zmian sezonowych w badanych krajach, w tym zimowego wzrostu umieralności.

W dalszych badaniach jest planowane wykorzystanie modeli SARIMA do oceny łącznej nadwyżki zgonów w okresie pandemii w porównaniu z typowymi wahaniami sezonowymi w poszczególnych państwach. Ważnym kierunkiem dalszych badań nad sezonowością zgonów jest uwzględnienie struktury wiekowej populacji, zwłaszcza rosnącego udziału osób starszych, bardziej podatnych na zachorowania, szczególnie w okresie grypowym (Wróblewska, 2019). W przypadku części państw warto też zwrócić uwagę na większe ruchy migracyjne mające wpływ na strukturę wiekową populacji oraz możliwe różnice etniczne w podatności na zachorowania i umieralności (Rau, 2007, s. 24).

Bibliografia

- [1] Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control* AC, 19, s. 716–723. DOI: 10.1109/TAC.1974.1100705.
- [2] Bull, G.M., Morton, J. (1978). Environment, Temperature and Death Rates, *Age and Ageing*, 7, s. 210–224. DOI: 10.1093/ageing/7.4.210. pubmed.ncbi.nlm.nih.gov/727071/.
- [3] Falagas, M.E., Karageorgopoulos, D.E., Moraitis, L.I., Vouloumanou, E.K., Rousos, N., Peppas, G., Rafailidis, P.I. (2009). Seasonality of mortality: The September phenomenon in Mediterranean countries, *CMAJ*, 181(8), s. 484–486. DOI: 10.1503/cmaj.090694. www.ncbi.nlm.nih.gov/pmc/articles/PMC2761439/.
- [4] Grut, M. (1987). Cold-related death in some developed countries, *The Lancet*, 8526: 212. DOI: 10.1016/S0140-6736(87) 90019-5.
- [5] Healy, J.D. (2003). Excess winter mortality in Europe: a cross country analysis identifying key risk factors, *J Epidemiol Community Health*, 57, s. 784–789, www.ncbi.nlm.nih.gov/pmc/articles/PMC1732295/.
- [6] Hyndman, R.J., Athanasopoulos, G. (2018). *Forecasting: principles and practice* (2nd ed.). OTexts: Melbourne, Australia. OTexts.com/fpp2 (dostęp: 15.10.2020).
- [7] Lauc, G., Markotic, A., Gornik, I., Primorac, D. (2020). Fighting COVID-19 with water, *Journal of Global Health*, Vol. 10, jogh.org/documents/issue202001/jogh-10-010344.pdf.
- [8] Ljung, G.M., Box, G.E.P. (1978). On a measure of lack of fit in time series models, *Biometrika*, Vol. 65, Iss. 2, s. 297–303. DOI: 10.1093/biomet/65.2.297.
- [9] Mazzi, G.L. (ed.). (2018). *Handbook on seasonal adjustment, 2018 edition*, Luxembourg: Publications Office of the European Union, ec.europa.eu/eurostat/documents/3859598/8939616/KS-GQ-18-001-EN-N.pdf.
- [10] Mielecka-Kubień, Z. (2018). Sezonowość zgonów w Polsce w latach 2012–2015, *Studia Ekonomiczne. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach*, 375.
- [11] Newey, W.K., West, K.D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, s. 703–708.
- [12] Nobbe, K. (2017). Seasonality of Mortality, *Reflections RGA's Global Magazine*, Vol. 40, s. 8–13, www.rgare.com/knowledge-center/media/articles/seasonality-of-mortality.
- [13] Rau, R. (2007). *Seasonality in Human Mortality: A Demographic Approach*. Springer.
- [14] Schwarz, G.E. (1978). Estimating the dimension of a model, *Annals of Statistics*, 6(2), s. 461–464. DOI: 10.1214/aos/1176344136.
- [15] Wojtyński, B., Rabczenko, D., Czarkowski, M., Skotak, K. (2018). *Ocena przyczyn podwyższonej umieralności ludności Polski na początku 2017 roku*. Warszawa: Narodowy Instytut Zdrowia Publicznego – Państwowy Zakład Higieny, www.pzh.gov.pl.
- [16] Wróblewska, W. (2019). Zmiana liczby zgonów i trwania życia w latach 2016–2018. Przyczynek do analizy zgonów w Polsce, *Studia Demograficzne*, 1(175), s. 9–25, econ-journals.sgh.waw.pl/SD/article/view/2335/2066. DOI: 10.33119/SD.2019.1.1.

Baza danych Eurostatu

- [1] ec.europa.eu/eurostat/databrowser/view/demo_mmonth/default/table?lang=en (dostęp: 25.08.2020).
- [2] ec.europa.eu/eurostat/databrowser/view/demo_gind/default/table?lang=en (dostęp: 25.08.2020).

Seasonal mortality patterns in Central and East European countries and their comparison with selected Western countries

Abstract

The report concerns analysis of seasonal patterns of mortality for selected countries of Central and Eastern Europe, compared with France, Italy, and Germany, based on monthly Eurostat data from 1999 to 2019. In the literature it is known that the increase of mortality rates in winter is typical for the North Hemisphere. We show that there are similarities and differences in the pattern – both in the averages and volatility – of this increase in particular countries. To this aim the φ_1 , φ_2 and φ_3 seasonality measures, known in the literature, were applied. The results are visualised in a graphical and tabular form. In addition, the significance of this seasonal variation is tested based on the regression with binary seasonal variables.

Keywords: seasonality of mortality, binary variables model, modeling of mortality