

Kornelia Kozaczewska

e-mail: kozaczewskakornelia@gmail.com

ORCID: 0009-0002-1084-0284

Uniwersytet Ekonomiczny we Wrocławiu

Analiza interwencji na przykładzie badania wpływu pandemii COVID-19 na dzietność za pomocą modeli SARIMA

DOI: 10.15611/2024.76.5.05

JEL: C22, I1, J13

© 2024 Kornelia Kozaczewska

Praca opublikowana na licencji Creative Commons Uznanie autorstwa-Na tych samych warunkach 4.0 Międzynarodowe (CC BY-SA 4.0). Skrócona treść licencji na <https://creativecommons.org/licenses/by-sa/4.0/deed.pl>

Cytuj jako: Kozaczewska, K. (2024). Analiza interwencji na przykładzie badania wpływu pandemii COVID-19 na dzietność za pomocą modeli SARIMA. W: A. Stanimir (red.), *Współczesne problemy społeczno-ekonomiczne w ujęciu analitycznym* (s. 69-85). Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu.

Streszczenie: Wystąpienie pandemii mogło różnie wpłynąć na liczbę urodzeń – zależało to od wielu czynników społecznych, ekonomicznych i zdrowotnych. Problemem badawczym w niniejszym artykule jest określenie, w jakim stopniu pandemia COVID-19 wpłynęła na dzietność w Polsce. Efekty takiego zdarzenia mogą być trwałe, ale mogą też mieć krótkoterminowe skutki. W pracy dokonano analizy wpływu pandemii COVID-19 na dzietność w Polsce, wykorzystując do tego modele SARIMA. Dane dotyczące liczby urodzeń pochodzą z Głównego Urzędu Statystycznego (GUS). Model I lepiej przewidywał liczbę urodzeń, co sugeruje, że COVID-19 nie wpłynął tak mocno na wcześniej podjęte decyzje reprodukcyjne. Jednocześnie dla modelu II wartości zaobserwowane były wyższe niż prognozowane. Badanie należałoby powtórzyć, wykorzystując bardziej skomplikowane metody, np. transmitancję operatorową (*transfer function*). Pełne skutki wpływu pandemii COVID-19 na dzietność mogą być jeszcze nieznanne, co sugeruje potrzebę dalszej analizy.

Słowa kluczowe: dzietność, liczba urodzeń, pandemia, COVID-19, modelowanie, SARIMA, prognozowanie, analiza interwencji, szeregi czasowe

1. Wstęp

Pandemia COVID-19, która wybuchła na początku 2020 roku, wywarła ogromny wpływ na wiele aspektów życia społecznego i gospodarczego na całym świecie. Dotychczasowe badania pokazują, że doprowadziła ona do spadku liczby urodzeń pod koniec 2020 roku i na początku roku 2021 (Aassve i in., 2021; Sobotka i in., 2021), ale ten trend odwrócił się w niektórych krajach w drugiej połowie 2021 roku (UNFPA, 2021). Reakcje społeczeństwa na zmieniające się warunki mają istotny wpływ na

kierunek i dynamikę badanych zjawisk nie tylko w krótkim, ale również w długim okresie. Na kształtowanie się zjawisk społecznych ma wpływ szereg różnorodnych czynników. Dodatkowo pandemia COVID-19 komplikuje analizę, wprowadzając zaburzenia, które mogą skutkować zmianą dynamiki i trendów dzietności.

Badania sugerują, że normy dotyczące rodzicielstwa, polityki łączenia pracy z rodzicielstwem oraz bezrobocie mają wpływ na wybory kobiet w zakresie płodności, całkowitej wielkości rodziny oraz różnicę między nimi (Beaujouan i Berghammer, 2019). Natomiast inne badania podkreślają istnienie zawitych zależności między ekonomią i kulturą w kwestii dzietności, gdzie większy wpływ na normy społeczne mają warunki ekonomiczne niż odwrotnie (Snopkowski i Kaplan, 2015).

Przyczynowość jest szczególnie trudna do zbadania w naukach społecznych ze względu na skomplikowane relacje między zmiennymi. W takim przypadku badanie wpływu zdarzeń jest bardziej zasadne niż badanie bezpośredniej przyczynowości (Morrison, 2012). Przeszłość szeregu czasowego może pomóc w wyjaśnieniu jego bieżącej wartości i w ten sposób uchwycić jego endogeniczną dynamikę. Zmiany, które wpływają na szereg, mogą mieć charakter endogeniczny lub egzogeniczny – ten wpływ może być tymczasowy lub długotrwały (Vasileiadou i Vliegenthart, 2014). W takich sytuacjach często korzysta się ze skomplikowanych modeli, które mimo wszystko nie są skuteczniejsze od prostych modeli prognostycznych (Armstrong i Green, 2015).

Problemem badawczym jest ocena wpływu pandemii COVID-19 na trendy urodzeń przy użyciu narzędzi analizy szeregów czasowych oraz wybranych modeli prognostycznych. Analiza ma na celu zrozumienie, jak kształtują się trendy demograficzne w obliczu trudnych i jedynych w swoim rodzaju okoliczności, jakie przyniosła pandemia.

Artykuł składa się z sześciu punktów. W punkcie drugim opisano dzietność oraz zjawiska z nią związane, czynniki, które na nią wpływają, oraz rolę, jaką odegrała pandemia COVID-19. W punkcie trzecim zaprezentowano dane. W czwartym omówiono analizę interwencji, modele SARIMA oraz szacowanie i metody oceny jakości modeli, a w piątym przedstawiono wyniki analizy liczby urodzeń w Polsce. Zakończenie zawiera podsumowanie rozważań przeprowadzonych w artykule.

2. Dzietność

2.1. Współczynnik dzietności

Współczynnik dzietności teoretycznej TFR (*Total Fertility Rate*) określa liczbę urodzonych dzieci przypadających na jedną kobietę w wieku rozrodczym (15-49 lat). W Polsce od lat 90. wskaźnik ten przyjmuje wartości, poniżej których gwarantowana jest prosta zastępowalność pokoleń. Niekorzystne proporcje między liczbą dzieci, dorosłych i osób starszych wpływają na organizację systemu emerytalnego, rynku pracy i ogólnego funkcjonowania społeczeństwa. Te proporcje są możliwe do odwrócenia właśnie poprzez zwiększenie liczby urodzeń (Kotowska, 2019).

W 2019 roku tylko dwie osoby na sto nie wyrażały chęci posiadania dzieci (CBOS, 2019), ale w 2023 roku takich osób było już 8%. Jednocześnie jedynie 30% ma tyle dzieci, ile chciałoby mieć, a 57% ma mniej dzieci, niż by chciało (CBOS, 2023). Zjawisko to oznacza to, że część planów prokreacyjnych nie jest realizowana. Takie rozbieżności określa się mianem *fertility gap* – jest to różnica między faktyczną a pożądaną liczbą dzieci, która wynika z szeregu powiązanych ze sobą czynników.

2.2. Czynniki wpływające na dzietność

Czynniki, które mają wpływ na liczbę urodzeń, to między innymi czynniki kulturowe, które są związane z dynamiką w związku oraz oczekiwaniami wobec partnera i rodziny. Badania pokazują, że egalitarny podział obowiązków domowych istotnie wpływa na podejmowanie decyzji o posiadaniu potomstwa (Okun, 2016). Natomiast tradycyjne poglądy nie muszą oznaczać większej liczby dzieci – w rodzinach Norwegów o nowoczesnych poglądach średnia liczba dzieci nie była mniejsza niż w rodzinach Polaków o tradycyjnych poglądach (Kocot-Górecka, 2015), gdzie częściej obserwuje się model rodziny z podwójnym obciążeniem kobiet (*dual earner – double burden of women*) (Matysiak, 2006).

Dodatkowo na podejmowanie decyzji reprodukcyjnych wpływ mają czynniki polityczno-prawne związane ze wsparciem na rynku pracy oraz polityką rodzinną, a które są silnie związane z sytuacją zawodową matek. Ponad dwie trzecie kobiet uważa, że powiększenie rodziny wpłynie negatywnie na ich sytuację na rynku pracy, natomiast prawie połowa mężczyzn uważa, że taka decyzja wpłynie pozytywnie na ich sytuację. Najczęściej wymieniane czynniki, które pomagają kobietom w powrocie do pracy, to elastyczne godziny pracy, wsparcie w zapewnieniu opieki nad dziećmi, a także możliwość wykonywania jej zdalnie (Ministerstwo Rodziny i Polityki Społecznej, 2021). Praca z domu pomaga kobietom zmniejszyć koszty i czas związany z pracą (np. oszczędzanie czasu na dojazdach) (Osiewalska i in., 2022). Te czynniki są powiązane z czynnikami kulturowymi – często w sytuacji, gdy kobieta wraca do pracy po urlopie macierzyńskim, opieka nad dzieckiem jest powierzana rodzinie (Styrc, 2007). Elastyczny czas pracy oraz lepsze możliwości opieki poza domem nad dziećmi w wieku do 7 lat są najbardziej preferowanymi rozwiązaniami problemów wynikających z łączenia rodzicielstwa i pracy zawodowej (Strzelecki i in., 2013).

2.3. Pandemia COVID-19 a dzietność

Widoczny jest wpływ pandemii na czynniki ekonomiczne, które z kolei oddziałują na dzietność. W pierwszych miesiącach pandemii wiele osób zostało zwolnionych z pracy, a grupy, które zostały najbardziej poszkodowane, to ludzie młodzi oraz kobiety (Smit i in., 2023). W Polsce stanowią one większość pracowników w sektorach, które zostały najsilniej dotknięte pandemią (Krause, 2021).

Również rodzicielstwo zostało dotknięte konsekwencjami pandemii. To głównie kobiety przejmowały nowe obowiązki, które spadały na rodzinę (np. pomoc dzieciom w nauce w trakcie godzin szkolnych) (Hays Poland, 2021). Na świecie kobiety na niepłatną pracę (zajmowanie się rodziną) oraz na wykonywanie obowiązków domowych poświęcają aż trzy razy więcej czasu niż mężczyźni (Committee for the Coordination of Statistical Activities, 2020). Poczucie odpowiedzialności za wszelkie skutki, które przyniosła pandemia, było również przyczyną niepewności oraz strachu. Stwierdzono też, że to kobiety częściej miały negatywne wizje przyszłości związane z pandemią (Drozdowski i in., 2020) – co wynika z silnych powiązań między czynnikami kulturowymi oraz społecznymi.

Pandemia wpłynęła również na czynniki zdrowotne – kobiety oraz ludzie młodzi deklarowali wzrost postrzeganego u siebie stresu (Kowal i in., 2020). Pomimo osłabienia restrykcji poziom stresu u kobiet nie był istotnie różny od tego odczuwanego na początku pandemii (Rokosz i in., 2021). Wydłużający się kryzys wyniszcza organizm – najbardziej na ten stan są narażeni ludzie młodzi, którzy jeszcze nie mają wypracowanych mechanizmów radzenia sobie z trudnymi sytuacjami (Markiewicz, 2019). Najczęściej wymieniane obawy związane z pandemią to niepewna sytuacja w kraju, brak stabilności zatrudnienia i strach przed utratą pracy, a także utrudniony dostęp do lekarzy. Wspominano również o restrykcjach w szpitalach, a zwłaszcza na oddziałach położniczych (Ministerstwo Rodziny i Polityki Społecznej, 2021).

3. Opis danych

Do analizy został wykorzystany szereg czasowy liczby urodzeń żywych (tys.) w Polsce. Dane były kompletne. Pobrano je z „Biuletynu Statystycznego” nr 11/2023 Głównego Urzędu Statystycznego (GUS)¹, a dokładniej z tablicy 7. *Stan i ruch naturalny ludności*. Rejestrowane dane są miesięczne, podane w tysiącach. Do analizy został użyty okres od stycznia 2013 roku do października 2022 roku włącznie.

4. Metodyka

4.1. Analiza interwencji

Aby ocenić wpływ czynników zewnętrznych na szereg czasowy, można wykorzystać analizę interwencji ITSA (*Intervention Time Series Analysis*). Jest to metoda badawcza służąca do oceny wpływu jednorazowego zdarzenia na zachowanie szeregu. Interwencje mogą przyjmować różne formy, takie jak zmiany polityki, wydarzenia społeczne czy kryzysy ekonomiczne, czas wystąpienia (jednorazowa lub ciągła) oraz mieć różną skalę – od lokalnej po globalną.

¹ Dane do pobrania ze strony GUS: <https://stat.gov.pl/obszary-tematyczne/inne-opracowania/informacje-o-sytuacji-spoeczno-gospodarczej/biuletyn-statystyczny-nr-112023,4,144.html>.

Analiza interwencji znalazła zastosowanie do zbadania rozkładu radioaktywności po katastrofie w Czarnobylu oraz jego wpływu na łańcuch pokarmowy człowieka i procesy gospodarcze (Hultkrantz i Olsson, 1997; Smith i in., 2000) czy do sprawdzenia siły wpływu dostępności heroiny w Australii w latach 2001-2002 (Degenhardt i in., 2005).

Jednym ze sposobów przeprowadzenia analizy interwencji jest analiza różnic między okresem przed interwencją a okresem po interwencji. Skutki takiego zdarzenia można oszacować, stosując modele do określenia wzrostów lub spadków średniego poziomu lub zmiany trendu (Gilmour i in., 2006). W tym celu utworzono dwa modele dla szeregu czasowego. Pierwszy opiera się na danych zebranych przed wystąpieniem interwencji i na jego podstawie zaprognozowano kolejne 12 miesięcy po interwencji. Jego celem jest analiza zachowania szeregu czasowego i próba określenia jego kierunku, gdyby interwencja nigdy nie wystąpiła. Drugi model uwzględnia dane z 12 miesięcy po wystąpieniu interwencji. Na podstawie tych danych również wykonano prognozę na kolejne 12 miesięcy. Celem drugiego modelu jest zbadanie ewentualnych zmian w zachowaniu szeregu czasowego spowodowanych interwencją. Oba modele oraz ich rezultaty zostały porównane ze sobą, co umożliwi ocenę wpływu interwencji na badany szereg czasowy.

Stosując metodę ITSA, należy mieć na uwadze dokładne zdefiniowanie nieplanowanego zdarzenia. Data wystąpienia interwencji ma kluczowe znaczenie – bez jasnej daty zakres możliwych modeli poszerza się, co może prowadzić do nieścisłości.

4.2. Modele SARIMA

Modele ARIMA są uważane za najbardziej ogólny rodzaj analizy szeregów czasowych (Boef i Keele, 2008) i nadają się szczególnie dla szeregów czasowych o długości większej niż 50 obserwacji (Yanovitzky i VanLear, 2008).

Model ARIMA składa się z trzech elementów: składnika AR (p) – procesu autoregresyjnego, którego każda wartość jest kombinacją liniową poprzednich wartości, składnika MA (q) – procesu średniej ruchomej, który działa podobnie jak proces AR, a także uwzględniają stopień integracji – I (d). Stopień integracji odnosi się do liczby różnicowania szeregu potrzebnego do uzyskania jego stacjonarności.

Końcowy model, uwzględniający opóźnienia, można zapisać w następujący sposób:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

gdzie: y'_t – zróżnicowany szereg, $y_{t-1} + \dots + y_{t-p}$ – wartości szeregu w czasie od $t-1$ do $t-p$, ϕ – parametr siły wpływu poprzedniej wartości na wartość bieżącą (dla składnika AR), θ – parametr siły wpływu starszych zaburzeń szeregu na bieżącą wartość (dla składnika MA), ε_t – składnik losowy w czasie t .

Rozszerzeniem modelu ARIMA (p, d, q) jest model SARIMA, zapisywany jako ARIMA $(p, d, q)(P, D, Q)_m$, gdzie dodatkowo jest brany pod uwagę składnik sezonowy S oraz liczba obserwacji w ciągu roku – m . Część sezonowa modelu zawiera takie same elementy jak niesezonowa, ale obejmuje one sezonowe opóźnienia. W modelu SARIMA sezonowe i niesezonowe składniki są przez siebie mnożone:

$$\begin{aligned} (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^m - \dots - \Phi_P B^{mP}) \\ (1 - B)^d (1 - B^m)^D y_t = (1 + \theta_1 B + \dots + \theta_q B^q) \\ (1 + \Theta_1 B^m + \dots + \Theta_Q B^{mQ}) \varepsilon_t, \end{aligned}$$

gdzie: ϕ_p – parametr AR dla składowych niesezonowych, Φ_p – parametr AR dla składowych sezonowych, θ_q – parametr MA dla składowych niesezonowych, Θ_Q – parametr MA dla składowych sezonowych, B – operator różnicowy (zwykle jest to opóźnienie o jeden okres).

Wartości parametrów modeli ARIMA można określić, analizując wykresy funkcji autokorelacji ACF (*Autocorrelation Function*) i funkcji autokorelacji cząstkowej PACF (*Partial Autocorrelation Function*). ACF i PACF są również przydatne do analizy struktury szeregu czasowego.

4.3. Szacowanie oraz ocena jakości modelu

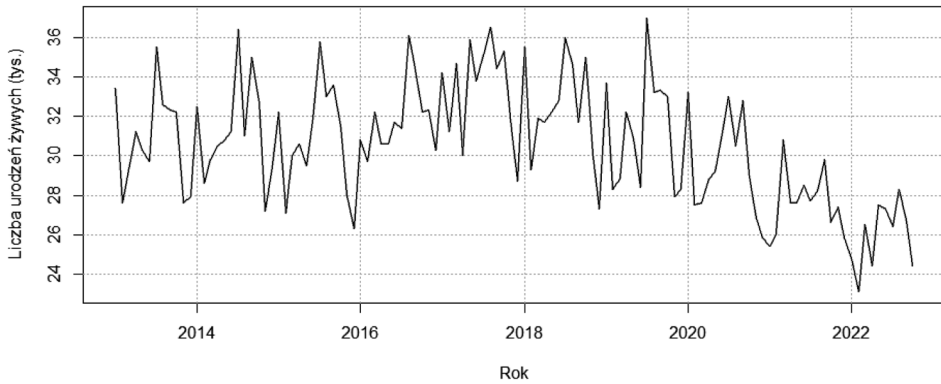
Modele zostały oszacowane za pomocą metody największej wiarygodności (przy wykorzystaniu funkcji `auto.arima()` dla modeli SARIMA oraz `ets()` dla modeli ETS). Porównywane są ze sobą kryteria informacyjne – im niższa wartość, tym lepsze dopasowanie modelu do danych. Ostatecznie wybierany jest model z najniższym AIC (*Akaike's Information Criterion* – kryterium informacyjne Akaike). Zarówno AIC (Akaike, 1973), jak i BIC (*Bayesian Information Criterion* – kryterium bayesowskie Schwartz'a) (Schwarz, 1978) są najważniejszymi kryteriami doboru modeli w ekonometrii i statystyce (Hossain, 1998). Często sprawdza się również AIC_c (AIC z poprawką) (Hurvich i Tsai, 1989), które jest szczególnie polecane dla małych prób.

W celu oceny jakości modelu porównuje się zbiór testowy (*test set*) (w tym przypadku 12 ostatnich obserwacji szeregu) z prognozą wykonaną na podstawie pozostałych obserwacji – zbioru uczącego (*training set*), za pomocą błędów *ex post*. Taka metoda umożliwia ocenę predykcyjną modelu. Te same miary można stosować, aby sprawdzić, czy model dobrze wykonał prognozę. Błędy *ex post* opierają się na resztach modelu, czyli różnicy między wartością empiryczną a teoretyczną. Oceny jakości predykcji modelu dokonuje się poprzez porównanie ze sobą błędów *ex post*. Niższa wartość błędu oznacza dokładniejszą prognozę.

5. Liczba urodzeń w Polsce

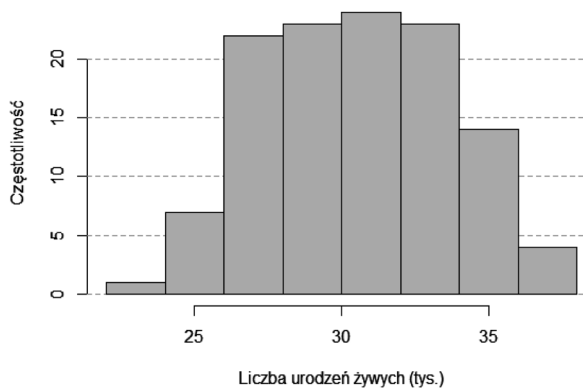
5.1. Charakterystyka liczby urodzeń przed pandemią i po jej zakończeniu

Rysunki 1 oraz 2 przedstawiają wykres oraz histogram liczby urodzeń żywych. Od 2016 do 2018 roku występuje widoczny wzrost liczby urodzeń, po którym następuje spadek – przez okres pandemii aż do 2022 roku (rys. 1). Najczęściej występowały wartości od 30 do 32 tys. urodzeń. Dane są dość symetryczne, z lekką lewostronną asymetrią (rys. 2).



Rys. 1. Wykres liniowy liczby urodzeń żywych w Polsce (w tys.)

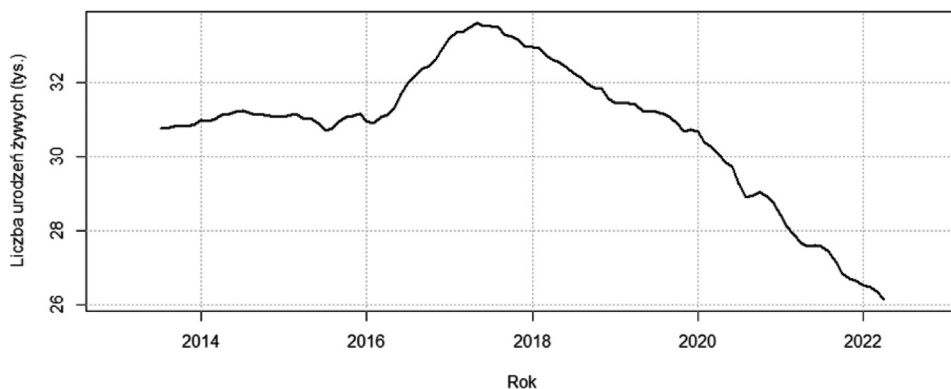
Źródło: obliczenia własne na podstawie danych GUS.



Rys. 2. Histogram liczby urodzeń żywych w Polsce (w tys.)

Źródło: obliczenia własne na podstawie danych GUS.

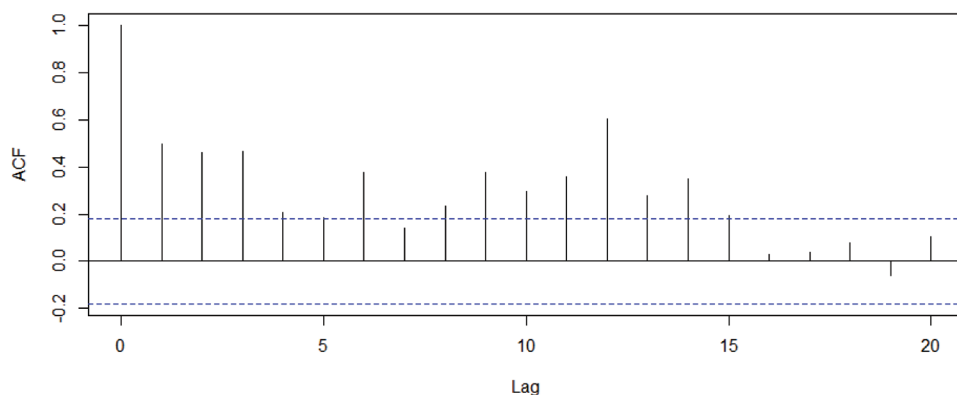
Trend liczby urodzeń był spadkowy już przed 2018 rokiem. Przed 2016 rokiem trend był stabilny. Okres wzrostu został zarejestrowany od roku 2016 do czwartego kwartału 2017 roku (rys. 3).



Rys. 3. Wykres linii trendu liczby urodzeń żywych w Polsce (w tys.)

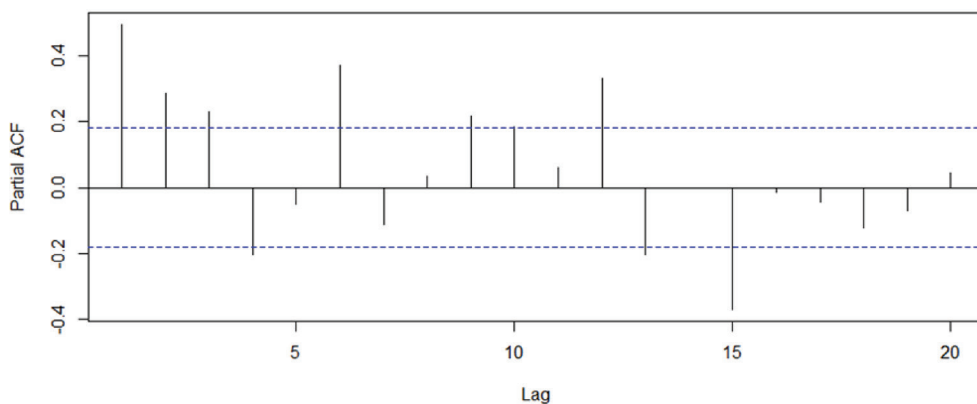
Źródło: obliczenia własne na podstawie danych GUS.

W zdecydowanej większości opóźnień występuje dodatnia korelacja. Największa wartość jest przyjmowana dla 12-tego opóźnienia, co wskazuje na sezonowość roczną (rys. 4). Korelacje między obserwacją a konkretnymi opóźnieniami są zarówno dodatnie, jak i ujemne. Największa ujemna korelacja jest dla opóźnienia równego 1 oraz 15, natomiast największa dodatnia korelacja jest dla 1., 6. oraz 12. miesiąca (rys. 5).



Rys. 4. Wykres ACF liczby urodzeń żywych w Polsce (w tys.)

Źródło: obliczenia własne na podstawie danych GUS.



Rys. 5. Wykres PACF liczby urodzeń żywych w Polsce (w tys.)

Źródło: obliczenia własne na podstawie danych GUS.

5.2. Modele urodzeń przed pandemią i po pandemii

W pierwszej kolejności zostały wykonane testy ADF (Dickey-Fuller = $-4,9286$, Lag order = 4, p -value = 0,01) oraz KPSS (KPSS Level = 0,1792, Truncation lag parameter = 3, p -value = 0,1) dla okresu do 9 miesięcy przed pandemią, które to potwierdziły stacjonarność danych. W przypadku drugiego badanego okresu, uwzględniającego pandemię, test na istnienie pierwiastka jednostkowego potwierdza stacjonarność danych (Dickey-Fuller = $-4,3499$, Lag order = 4, p -value = 0,01), jednak test na trendostacjonarność danych odrzuca hipotezę zerową mówiącą o stacjonarności (KPSS Level = 0,44548, Truncation lag parameter = 4, p -value = 0,05755). Ze względu na niskie p -value dla testu KPSS w obu przypadkach dokonano jednokrotnego zróżnicowania danych, aby usunąć niestacjonarność.

Z powodu występowania autokorelacji danych oraz sezonowości w obu przypadkach dopasowano model z autoregresyjnym sezonowym składnikiem różnicowania uwzględniający dwa opóźnienia oraz z autoregresyjnym niesezonowym składnikiem różnicowania uwzględniającym dwa sezony opóźnień, gdzie okres sezonowy wynosi 12 miesięcy i można go zapisać jako $ARIMA(2,1,0)(2,1,0)_{12}$. Współczynniki dla modelu I są nieznacznie mniejsze niż dla modelu II, co oznacza, że ma on mocniejsze autoregresyjne oraz sezonowe składniki niż model I. Estymacja współczynników jest bardziej precyzyjna dla modelu II, na co wskazują niższe błędy standardowe niż w przypadku błędów modelu I. Model II ma natomiast słabsze autoregresyjne i sezonowe składniki niż model I – wynika to prawdopodobnie ze zmiany trendu oraz wzorców, które sprawiają, że model jest mniej skomplikowany (tab. 1).

Tabela 1. Współczynniki oraz błędy standardowe modeli I oraz II

Model	ϕ_1	ϕ_2	Φ_1	Φ_2
Model I	-0,98	-0,65	-0,52	-0,38
	(0,091)	(0,089)	(0,114)	(0,112)
Model II	-0,89	-0,64	-0,56	-0,33
	(0,089)	(0,086)	(0,125)	(0,112)

Źródło: obliczenia własne na podstawie danych GUS.

Model I można przedstawić za pomocą wzoru modelu SARIMA (składniki niesezonowe oraz sezonowe MA są równe 0):

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^{12} - \Phi_2 B^{24})(1 - B)(1 - B^{12})y_t = \varepsilon_t.$$

Dla uproszczenia $(1 - B)(1 - B^{12})y_t$ oznaczone jest jako z_t :

$$(1 - \phi_1 B - \phi_2 B^2)(1 - \Phi_1 B^{12} - \Phi_2 B^{24})z_t = \varepsilon_t.$$

Po przekształceniu z_t wyraża się jako:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \Phi_1 z_{t-12} - \phi_1 \Phi_1 z_{t-13} - \phi_2 \Phi_1 z_{t-14} + \Phi_2 z_{t-24} - \phi_1 \Phi_2 z_{t-25} - \phi_2 \Phi_2 z_{t-26} + \varepsilon_t.$$

Podstawiając współczynniki parametrów, otrzymuje się wyrażenie:

$$z_t = -0,98z_{t-1} - 0,65z_{t-2} - 0,52z_{t-12} - (-0,98)(-0,52)z_{t-13} - (-0,65)(-0,52)z_{t-14} - 0,38z_{t-24} - (-0,98)(-0,38)z_{t-25} - (-0,65)(-0,38)z_{t-26} + \varepsilon_t.$$

Ostateczna postać modelu I ma następującą postać:

$$z_t = -0,98z_{t-1} - 0,65z_{t-2} - 0,52z_{t-12} - 0,51z_{t-13} - 0,34z_{t-14} - 0,38z_{t-24} - 0,37z_{t-25} - 0,25z_{t-26} + \varepsilon_t.$$

Model II również można przedstawić za pomocą wzoru modelu SARIMA (składniki niesezonowe oraz sezonowe MA są równe 0). Podstawiając współczynniki parametrów, otrzymuje się:

$$z_t = -0,89z_{t-1} - 0,64z_{t-2} - 0,56z_{t-12} - (-0,89)(-0,56)z_{t-13} - (-0,64)(-0,56)z_{t-14} - 0,33z_{t-24} - (-0,89)(-0,33)z_{t-25} - (-0,64)(-0,33)z_{t-26} + \varepsilon_t.$$

Ostateczna postać modelu II to:

$$z_t = -0,89z_{t-1} - 0,64z_{t-2} - 0,56z_{t-12} - 0,5z_{t-13} - 0,36z_{t-14} - \\ - 0,33z_{t-24} - 0,3z_{t-25} - 0,21z_{t-26} + \varepsilon_t.$$

Model I dopasowuje się lepiej do danych ze zbioru uczącego (niższa wartość log-likelihood od modelu II) oraz ma mniejszą estymowaną wariancję błędów (mniejsza σ^2). Kryteria informacyjne (AIC, AIC_c oraz BIC) wskazują na lepsze dopasowanie do danych modelu I niż modelu II (tab. 2).

Tabela 2. Dopasowanie modelu oraz kryteria informacyjne modelu I oraz II

Model	log-likelihood	σ^2	AIC	AIC _c	BIC
Model I	-157,23	2,74	324,46	325,26	336,44
Model II	-190,73	3,45	391,45	392,14	404,12

Źródło: obliczenia własne na podstawie danych GUS.

Na podstawie błędów *ex post* zbioru uczącego można stwierdzić, że model I wypada lepiej od modelu II i ma tendencję do przewidywania wartości bliższych rzeczywistym wartościom (tab. 3).

Tabela 3. Błędy prognozy modeli I oraz II na podstawie zbioru uczącego

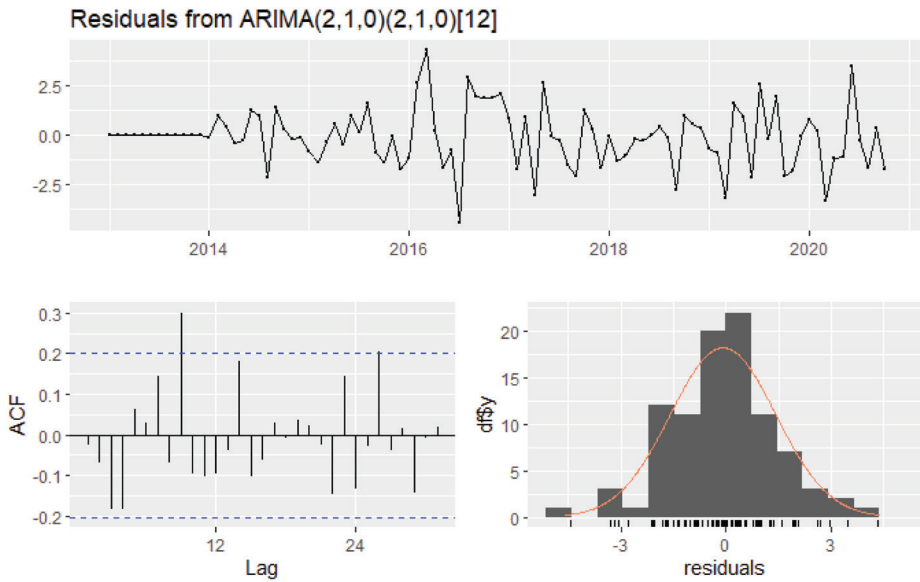
Model	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Model I	-0,08	1,50	1,08	-0,41	3,44	0,63	-0,02
Model II	-0,09	1,70	1,22	-0,45	3,97	0,66	0,01

Źródło: obliczenia własne na podstawie danych GUS.

Rozkład reszt dla modelu II wskazuje na grubsze ogony, niż ma to miejsce w przypadku modelu I (rys. 6 oraz 7). Test Shapiro-Wilka dla modelu I potwierdza normalność rozkładu reszt ($W = 0,9798$, $p\text{-value} = 0,1545$), ale dla modelu II założenie to nie jest spełnione ($W = 0,97497$, $p\text{-value} = 0,04233$), co może wskazywać na słabe dopasowanie modelu do danych. W obu przypadkach nie występuje heteroskedastyczność reszt (rys. 6 oraz 7).

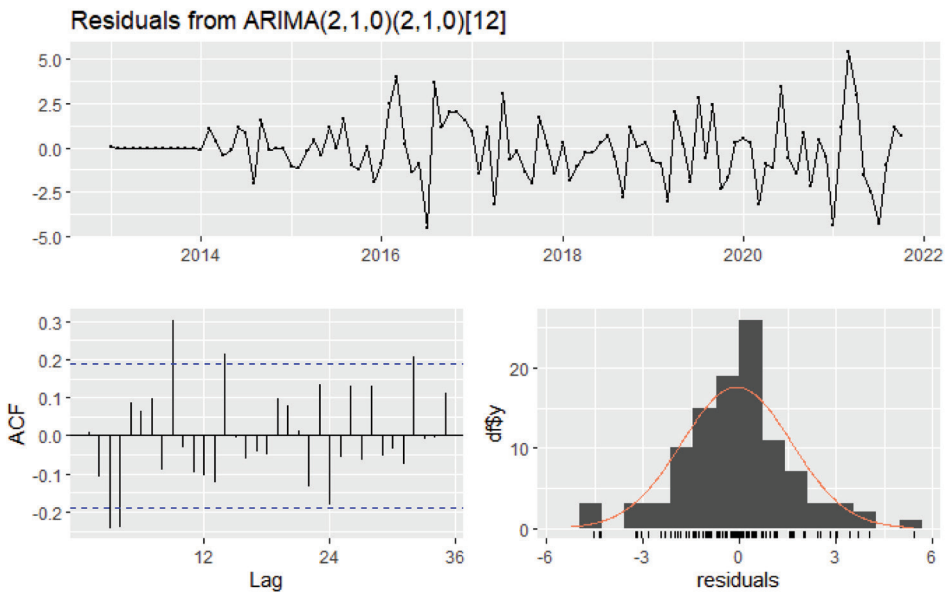
Dla obu modeli ACF wskazuje na występowanie autokorelacji reszt (rys. 6 oraz 7), która może prowadzić do niedokładnych prognoz. Potwierdza to test Ljung-Boxa dla modelu I ($Q^* = 28,741$, $df = 15$, $p\text{-value} = 0,01737$) oraz modelu II ($Q^* = 41,989$, $df = 17$, $p\text{-value} = 0,0006731$).

Na podstawie obu modeli wykonano prognozę dla dwóch okresów: od listopada 2020 roku do października 2021 roku włącznie, na podstawie danych od początku badanego okresu (styczeń 2013 roku) dla modelu I, oraz na kolejne 12 miesięcy (od listopada 2021 roku do października 2022 roku włącznie) na podstawie danych od stycznia 2013 roku do listopada 2021 roku dla modelu II.



Rys. 6. Wykres liniowy, wykres funkcji autokorelacji oraz histogram reszt modelu I

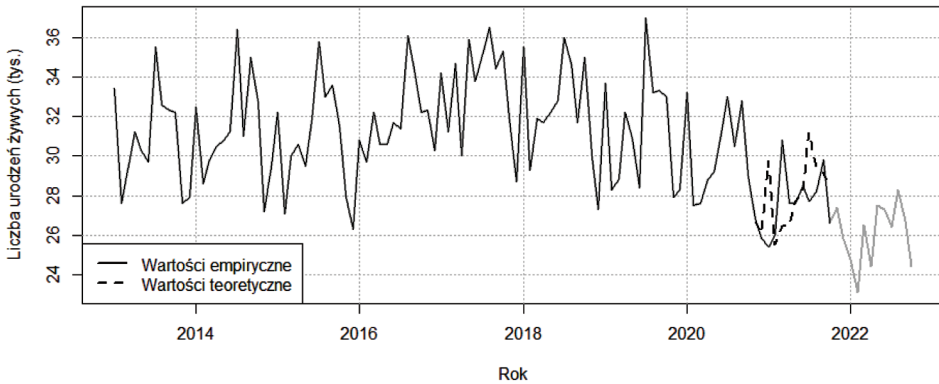
Źródło: obliczenia własne na podstawie danych GUS.



Rys. 7. Wykres liniowy, wykres funkcji autokorelacji oraz histogram reszt modelu II

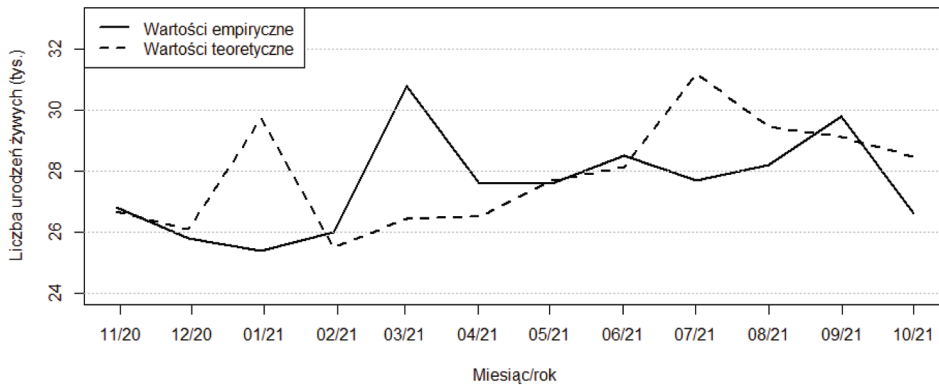
Źródło: obliczenia własne na podstawie danych GUS.

W przypadku modelu I nie ma znaczącej różnicy między wartościami teoretycznymi a wartościami zaobserwowanymi. Dane empiryczne wydają się opóźnione o 3 okresy względem prognozy (rys. 8 oraz 9). Równanie linii trendu dla wartości zaobserwowanych jest równe $y_t = 0,1923x_t + 26,317$. Natomiast dla prognozy jest to $y_t = 0,2796x_t + 26,109$.



Rys. 8. Wykres liniowy dla prognozy modelu I oraz dla danych zaobserwowanych

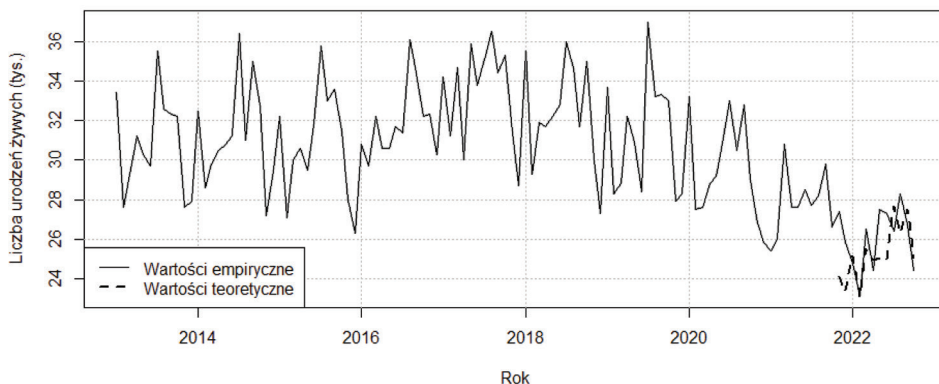
Źródło: obliczenia własne na podstawie danych GUS.



Rys. 9. Wykres liniowy dla prognozy modelu I

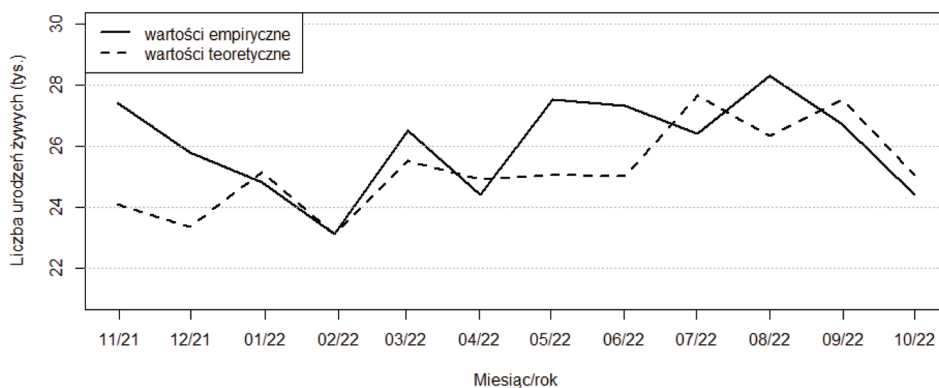
Źródło: obliczenia własne na podstawie danych GUS.

Zdecydowanie gorzej zostały zaprognozowane wartości na podstawie modelu II (rys. 10 oraz 11). Równanie linii trendu dla wartości empirycznych wynosi $y_t = 0,0755x_t + 25,559$. Natomiast dla wartości z prognozy jest to $y_t = 0,2719x_t + 23,455$.



Rys. 10. Wykres liniowy dla prognozy modelu II oraz dla danych zaobserwowanych

Źródło: obliczenia własne na podstawie danych GUS.



Rys. 11. Wykres liniowy dla prognozy modelu II

Źródło: obliczenia własne na podstawie danych GUS.

Błędy *ex post* wskazane na podstawie danych empirycznych dla modelu II są nieznacznie lepsze, niż dla modelu I (tab. 4). Model I przeszacowywał wartości, a model II nie doszacowywał.

Tabela 4. Błędy *ex post* modeli I oraz II

Model	ME	RMSE	MAE	MPE	MAPE
Model I	-0,36	2,18	1,54	-1,58	5,56
Model II	0,83	1,74	1,42	2,98	5,31

Źródło: obliczenia własne na podstawie danych GUS.

6. Zakończenie

Na podstawie analizy można dojść do wniosków wskazujących na to, że modele SARIMA dobrze przewidziały średnią liczbę urodzeń, jednak model I lepiej poradził sobie z prognozowaniem niż model II, który uwzględniał rok pandemii w Polsce. Co ciekawe, w przypadku modelu I dane empiryczne wydają się opóźnione o 3 okresy względem prognozy. Może to oznaczać, że plany prokreacyjne są podejmowane z wyprzedzeniem i pandemia COVID-19 nie zaburzyła planów osób, które wcześniej przygotowały się na posiadanie dzieci. Reakcja społeczeństwa w pierwszych miesiącach pandemii była mniej istotna dla trendów demograficznych, niż można było przypuszczać. Dodatkowo zaobserwowane wartości wyższe niż prognozowane modelem II mogą wskazywać na spełnienie odroczonego planów prokreacyjnych.

Mimo że modele nie uwzględniały wystąpienia wojny w Ukrainie, to wcześniejsze informacje o możliwej agresji Rosji mogły z wyprzedzeniem wpływać na reakcje społeczne. Jest to jedna z wielu trudności analizowania i prognozowania złożonych zjawisk – wpływ społecznych czynników na badane trendy jest trudny do oszacowania.

Dodatkową trudność sprawiło wybranie dokładnego momentu wystąpienia interwencji, co ma wpływ na jakość analizy. Badanie należałoby powtórzyć, wykorzystując bardziej skomplikowane metody, np. transmitancję operatorową (*transfer function*).

Pojawienie się pandemii COVID-19 wywarło istotny wpływ na wiele aspektów życia społecznego i gospodarczego, jednak jej pełne skutki mogą być jeszcze nieznanne. Chaos i niepewność, które wprowadziła pandemia, mogły być czynnikami zapalnymi łańcucha szeregu zdarzeń, które wpływają na urodzenia w Polsce, a które będą wpływać jeszcze przez lata.

Literatura

- Aassve, A., Cavalli, N., Mencarini, L., Plach, S. i Sanders, S. (2021). Early Assessment of the Relationship between the COVID-19 Pandemic and Births in High-income Countries. *Proceedings of the National Academy of Sciences of the United States of America*, 118(36). <https://doi.org/10.1073/pnas.2105709118>
- Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. W: B. N. Petrov i F. Csaki (red.), *Second International Symposium on Information Theory*. Akademiai Kiado.
- Armstrong, J. S. i Green, K. (2015). Simple Versus Complex Forecasting: The Evidence. *Journal of Business Research*, 68(8). <https://doi.org/10.1016/j.jbusres.2015.03.026>
- Beaujouan, E. i Berghammer, C. (2019). The Gap between Lifetime Fertility Intentions and Completed Fertility in Europe and the United States: A Cohort Approach. *Population Research and Policy Review*, (38), 507-535. <https://doi.org/10.1007/s11113-019-09516-3>
- Boef, S. D. i Keele, L. (2008). Taking Time Seriously. *American Journal of Political Science*, (52), 184-200.
- CBOS. (2019). *Preferowane i realizowane modele życia rodzinnego*. Pobrano 4 stycznia 2024 z https://www.cbos.pl/SPISKOM.POL/2019/K_046_19.PDF

- CBOS. (2023). *Bariery zamierzeń prokreacyjnych*. Pobrane 4 stycznia 2024 z https://www.cbos.pl/SPISKOM.POL/2023/K_087_23.PDF
- Committee for the Coordination of Statistical Activities. (2020). *How COVID-19 Is Changing the World: A Statistical Perspective*.
- Degenhardt, L., Day, C., Dietze, P., Pointer, S., Conroy, E., Collins, L. i Hall, W. (2005). Effects of a Sustained Heroin Shortage in three Australian States. *Addiction*, 100(7), 908-920.
- Drozdowski, R., Frąckowiak, M., Krajewski, M., Kubacka, M., Luczys, P., Modrzyk, A., Rogowski, Ł., Rura, P. i Stamm, A. (2020). *Życie codzienne w czasach pandemii. Raport z drugiego etapu badań*. Pobrane 9 stycznia 2024 z https://socjologia.amu.edu.pl/images/pliki/dokumenty/Do_pobrania/Zycie_codzienne_w_czasach_pandemii_Raport_z_drugiego_etapu_badan_wersja_skrcona.pdf
- Gilmour, S., Degenhardt, L., Hall, W. i Day, C. (2006). Using Intervention Time Series Analyses to Assess the Effects of Imperfectly Identifiable Natural Events: a General Method and Example. *BMC Medical Research Methodology*, 6(16) <https://doi.org/10.1186/1471-2288-6-16>
- Hays Poland. (2021). *Kobiety na rynku pracy 2021*. Pobrano 10 lutego 2024 z <https://www.hays.pl/kobiety>
- Hossain, M. (1998). AIC and BIC – The Two Competitive Information Criteria for Model Selection in Economics and Statistics. *The Jahangirnagar Review, Part II: Social Science*, 19, 133-140.
- Hultkrantz, L. i Olsson, C. (1997). Chernobyl Effects on Domestic and Inbound Tourism in Sweden – A Time Series Analysis. *Environmental and Resource Economics*, (9), 239-258.
- Hurvich, C. M. i Tsai, C. L. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, 76, 297-307.
- Kocot-Górecka, K. (2015). Kulturowe role płci i dzietność w Polsce i Norwegii. *Studia Demograficzne*, 2(168), 61-88. <https://doi.org/10.33119/SD.2015.2.4>
- Kotowska, I. E. (2019). Uwagi o urodzeniach i niskiej dzietności w Polsce oraz polityce rodzinnej wspierającej prokreację. *Studia Demograficzne*, 2(176), 11–29. <https://doi.org/10.33119/SD.2019.2.1>
- Kowal, M., Karwowski, M., Coll-Martín, T., Ikizer, G., Rasmussen, J., Lieberoth, A., Eichel, K., Studzińska, A., Koszałkowska, K., Najmussaib, A., Pankowski, D. i Ahmed, O. (2020). Who is the Most Stressed During the COVID-19 Pandemic? Data from 26 Countries and Areas. *Applied Psychology: Health and Well-Being*, 12(4), 946-966, <https://doi.org/10.1111/aphw.12234>
- Krause, E. (2021). Sytuacja kobiet na rynku pracy w czasie pandemii. *Edukacja Ustawiczna Dorosłych*, 1(112), 159-176. <https://doi.org/10.34866/6gaf-2m32>
- Markiewicz, R. (2019). Kryzys w kryzysie. Psychologiczny wymiar pandemii. *Zdrowie Publiczne*, 129, 138-140. <https://doi.org/10.2478/pjph-2019-0031>
- Matysiak, A. (2006). Sharing Professional and Household Duties within the Polish Couples – Preferences and Actual Choices. *Studia Demograficzne*, 1(147), 122-153.
- Ministerstwo Rodziny i Polityki Społecznej. (2021). *Badania warunków do poprawy kondycji demograficznej Polski*. Pobrano 20 stycznia 2024 z <https://www.gov.pl/attachment/b0c2fb5e-6a0c-4928-92e2-f74ee621ed1b>
- Morrison, K. (2012). Searching for Causality in the Wrong Places. *International Journal of Social Research Methodology*, (15), 15-30.
- Okun, B. S. (2016). An Investigation of the Unexpectedly High Fertility of Secular, Native-born Jews in Israel. *Population Studies*, 70(2), 239-257. <https://doi.org/10.1080/00324728.2016.1195913>
- Osiewalska, B., Matysiak, A. i Kurowska, A. (2022) When Are Women Who Work from Home More Likely to Have Children? *Working Papers*, 13(389). Faculty of Economic Sciences, University of Warsaw.
- Rokosz, M., Poprawa, R., Tabiś, K., Barański, M. i Lewandowska, B. (2021). Analiza zmian w zakresie zaspokojenia i frustracji podstawowych potrzeb psychologicznych oraz stresu i radzenia sobie z nim w pierwszych miesiącach pandemii COVID-19 w 2020 roku wśród kobiet. *Polskie Forum Psychologiczne*, 26(3), 330-348. <https://doi.org/10.34767/PFP.2021.03.06>
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6, 461-464.

- Smit, J., Nacer, E., Sikorski, A., Godard, C. i Magdziarz, W. (2023) Consequences for Workers. W: K. Polluveer (red.), *Social and Economic Consequences of COVID-19* (s. 116-136). Policy Department for Economic, Scientific and Quality of Life Policies.
- Smith, J. T., Clarke, R. T. i Saxén, R. (2000). Time-Dependent Behaviour of Radiocaesium: A New Method to Compare the Mobility of Weapons Test and Chernobyl Derived Fallout. *Journal of Environmental Radioactivity*, (49), 65-83. [http://dx.doi.org/10.1016/S0265-931X\(99\)00088-0](http://dx.doi.org/10.1016/S0265-931X(99)00088-0)
- Snopkowski, K. i Kaplan, H. (2015). Fertility Theory: Theory of Intergenerational Wealth Flows. *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (s. 35-39). <https://doi.org/10.1016/B978-0-08-097086-8.31074-1>
- Sobotka, T., Jasilioniene, A., Zeman, K., Nemeth, L., Winkler-Dworak, M., Galarza, A. A., Brzozowska, Z. i Jdanov, D. (2021). Booms, Busts and Trend Reversals? Shifts in Births and Fertility Rates across the Highly Developed Countries during the COVID-19 Pandemic [prezentacja]. *Pandemic Babies? The Covid-19 Pandemic and Its Impact on Fertility and Family Dynamics*. MPIDR.
- Strzelecki, P., Saczuk, K. i Kotowska, I. E. (2013). Godzenie posiadania dzieci i pracy zawodowej – opinie o niektórych rozwiązaniach polityki społecznej. W: J. Czapiński, T. Panek (red.), *Diagnoza Społeczna 2013. Warunki i jakość życia Polaków. Raport*. RMS.
- Styrc, M. (2007). Transfery finansowe i opieki w nieformalnych sieciach wsparcia. W: I.E. Kotowska, U. Sztanderska, I. Wójcicka (red.), *Aktywność zawodowa i edukacyjna a obowiązki rodzinne w Polsce w świetle badań empirycznych*. Scholar.
- UNFPA. (2021). *How Will the COVID-19 Pandemic Affect Births? Technical Brief*. Pobrano 5 stycznia 2024 z <https://www.unfpa.org/sites/default/files/pub-pdf/How%20will%20the%20COVID-19%20pandemic%20affect%20births.pdf>
- Vasileiadou, E. i Vliegenthart, R. (2014). Studying Dynamic Social Processes with ARIMA Modeling. *International Journal of Social Research Methodology*, 17(6), 693-708. <https://doi.org/10.1080/13645579.2013.816257>
- Yanovitzky, I. i VanLear, A. (2008). Time Series Analysis: Traditional and Contemporary Approaches. W: A. F. Hayes, M. D. Slater, L. B. Snyder (red.), *The Sage Sourcebook of Advanced Data Analysis Methods for communication research* (s. 89-124). Sage Publications, Inc. <https://doi.org/10.4135/9781452272054.n4>

Intervention Time Series Analysis Based on the Example of Examining the Impact of the COVID-19 Pandemic on Fertility Using SARIMA Models

Abstract: The pandemic could have affected birth rates differently, depending on many social, economic, and health factors. The research problem is to determine the extent to which the COVID-19 pandemic has affected fertility in Poland. The effects of such an event can be lasting, but they can also be short-term. The work analyses the impact of the COVID-19 pandemic on fertility in Poland, using SARIMA models. Data on the number of births comes from the Central Statistical Office (GUS). Model I predicted the number of births better, suggesting that COVID-19 has not impacted that much previously made reproductive decisions. At the same time, for model II the observed values were higher than predicted. The test should be repeated using more complicated methods, e.g. transfer function. The full effects of the COVID-19 pandemic on fertility may not yet be known, suggesting the need for further analysis.

Keywords: fertility, number of births, pandemic, COVID-19, modeling, SARIMA, forecasting, intervention analysis, time series