

Prognozowanie liczby wybranych szkód ubezpieczeniowych w Polsce

Autor: Jarosław Krajewski

Abstrakt

Artykuł ma na celu zbadanie, czy w zakresie ogólnie dostępnych i oficjalnie publikowanych danych dotyczących liczby szkód mających miejsce na polskim rynku ubezpieczeń można mówić o zjawiskach pozwalających na modelowanie ekonometryczne i w konsekwencji prognozowanie tej liczby.

Poniższy artykuł ma być punktem wyjścia do dalszych analiz rynku ubezpieczeń, a w pierwszej kolejności szkodowości na nim występującej. Badania takie mogą okazać się przydatne w analizach ryzyka rynkowego przeprowadzanego przez brokerów ubezpieczeniowych.

Analiza empiryczna dotyczy zmiennych reprezentujących szkody działu drugiego ubezpieczeń oraz szkody z ubezpieczenia odpowiedzialności cywilnej posiadaczy pojazdów lądowych.

Na podstawie szeregów czasowych o częstotliwości kwartalnej z okresu od pierwszego kwartału 2004 do drugiego kwartału 2014 roku udało się oszacować modele dobrze opisujące liczbę szkód majątkowych występujących na polskim rynku ubezpieczeń oraz wyznaczyć na ich podstawie prognozy.

Najlepszym z prognostycznego punktu widzenia modelem liczby szkód z działu II ubezpieczeń okazał się II+ zawierający trend oraz opóźnienia zmiennej objaśnianej i zmiennej objaśniającej. W zakresie szkód z ubezpieczenia OC posiadaczy pojazdów lądowych był to natomiast sezonowo-trendowy model OC.

Słowa kluczowe: ubezpieczenie, szkoda, szkodowość, prognozowanie

JEL: C01, C5, G22

Wprowadzenie

Polski rynek ubezpieczeniowy w ostatnich latach rozwija się systematycznie, zarówno jeśli chodzi o jego wielkość i otwartość, jak również zróżnicowanie dostępnych i powstawanie nowych produktów. Skutkuje to wzrostem liczby zawieranych umów ubezpieczenia różnego rodzaju. Można zatem przypuszczać, że rośnie w czasie również liczba szkód zgłaszanych z tychże właśnie umów.

Niniejszy artykuł ma na celu zbadanie, czy w zakresie ogólnie dostępnych i oficjalnie publikowanych danych dotyczących liczby szkód mających miejsce na polskim rynku ubezpieczeń można mówić o zjawiskach pozwalających na modelowanie ekonometryczne szkodowości, tj.

tendencji rozwojowych, sezonowości i autoregresji. W dalszej części referatu sprawdzone zostanie również, jak narzędzia ekonometryczne sprawdzają się w prognozowaniu liczby szkód.

Poniższy artykuł ma być punktem wyjścia do dalszych analiz rynku ubezpieczeń, a w pierwszej kolejności szkodowości na nim występującej. W związku z powyższym warto spróbować odpowiedzieć na pytanie, czy w jej zakresie zastosowanie metod ekonometrycznych okaże się skuteczne.

Prognozowanie przedstawione w poniższym artykule może okazać się przydatne w analizach ryzyka rynkowego przeprowadzanego przez brokerów ubezpieczeniowych. Zazwyczaj nie dysponują oni szczegółowymi danymi zbieranymi przez ubezpieczycieli ani działami aktuarialnymi,

w związku z tym taka alternatywa może okazać się atrakcyjna w prowadzonej przez nich działalności.

Teoretyczne podstawy ilościowej analizy szkód ubezpieczeniowych

Szkody są podstawą istnienia rynku ubezpieczeń. To one spowodowały, że umowy ubezpieczenia w ogóle zaczęły funkcjonować. Już w czasach starożytnych znane były przypadki zrzeszania się zainteresowanych podmiotów w celu wspólnego ponoszenia ryzyka na zasadach wzajemności i solidarności. Umowy uczestników karawan w krajach Bliskiego Wschodu znane były już za czasów Hammurabiego (ok. 2 tys. lat p.n.e.). Zobowiązywali się oni do wspólnego pokrywania właśnie ewentualnych szkód, poniesionych przez każdego uczestnika takiej umowy (Sangowski 2001: 11–12).

Szkodą nazwiemy dzisiaj stratę majątkową lub niemajątkową powstałą w wyniku wypadku losowego (Monkiewicz 2000: 64). Jako ciekawy przykład szkody przywołać można urodzenie się dziecka, które może być właśnie w ten sposób traktowane w sferze ubezpieczeń.

Z uwagi na analizy zawarte w niniejszym artykule należy przybliżyć również pojęcie działań ubezpieczeń wynikające bezpośrednio z Ustawy o działalności ubezpieczeniowej (Ustawa z 22 maja 2003 r.). Ustawa ta wyróżnia:

Dział I – Ubezpieczenia na życie, oraz
Dział II – Pozostałe ubezpieczenia osobowe oraz ubezpieczenia majątkowe.

Właśnie na tych drugich w ujęciu ogólnym skupiać się będzie przeprowadzana analiza. Dodatkowo jako szczególny przypadek wybrano wchodzące w skład tego działu ubezpieczenie odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych.

Pojęcie samej szkodowości nie jest natomiast jednoznaczne. Inaczej definiują je przepisy prawa, inaczej rozumieją je klienci, pośrednicy i ubezpieczyciele. Biorąc pod uwagę rozumienie potoczne oraz kontekst, w którym jest używane, zazwyczaj spotykamy dwa główne znaczenia szkodowości: suma odszkodowań i współczynnik odszkodowań. Mając tego świadomość, na potrzeby opracowania autor zdecydował się na wykorzystanie pojęcia szkodowości jako równoważnik dla pojęcia liczby szkód.

Zastosowane modele ekonometryczne

W analizie szkodowości wykorzystane zostały podstawowe dynamiczne modele

ekonometryczne uwzględniające w swej strukturze trendy, sezonowość i autokorelację wybranych zmiennych.

Wykorzystano analizę występowania następujących procesów (Kufel 2011: 81–91): wielomianowych funkcji trendu w postaci:

$$Y_t = \sum_{j=0}^r \alpha_j t^j + \eta_t, \quad (1)$$

gdzie:

t – zmienna czasowa, $t = 1, 2, \dots, n$;

r – stopień wielomianu zmiennej czasowej,

η_t – składnik losowy,

wahań sezonowych:

$$Y_t = a_0 + \sum_{i=1}^{m-1} d_i Q_{it} + \eta_t, \quad (2)$$

gdzie:

Q_{it} – 0-1-kowe zmienne sezonowe,

m – liczba sezonów;

autoregresji:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t, \quad (3)$$

gdzie: p – rząd opóźnienia,

ε_t – składnik losowy.

Ponadto zastosowanie w analizach znalazł model wygładzania wykładniczego Wintersa w postaci multiplikatywnej, którą przedstawia następujący zapis (Cieślak 2001):

$$F_t = \alpha \frac{y_t}{C_{t-r}} + (1 - \alpha)(F_{t-1} + S_{t-1}) \quad (4)$$

$$S_t = \beta (F_t - F_{t-1}) + (1 - \beta) S_{t-1}$$

$$C_t = \gamma \frac{y_t}{F_t} + (1 - \gamma) C_{t-r}$$

W powyższym wzorze

F_t służy do wygładzonej oceny wartości zjawiska, S_t służy do wygładzonej oceny wartości przyrostu trendu,

C_t służy do wygładzonej oceny sezonowości,

$\alpha, \beta, \gamma \in \langle 0; 1 \rangle$ – stałe parametry wygładzania,

r – liczba sezonów w jednym cyklu.

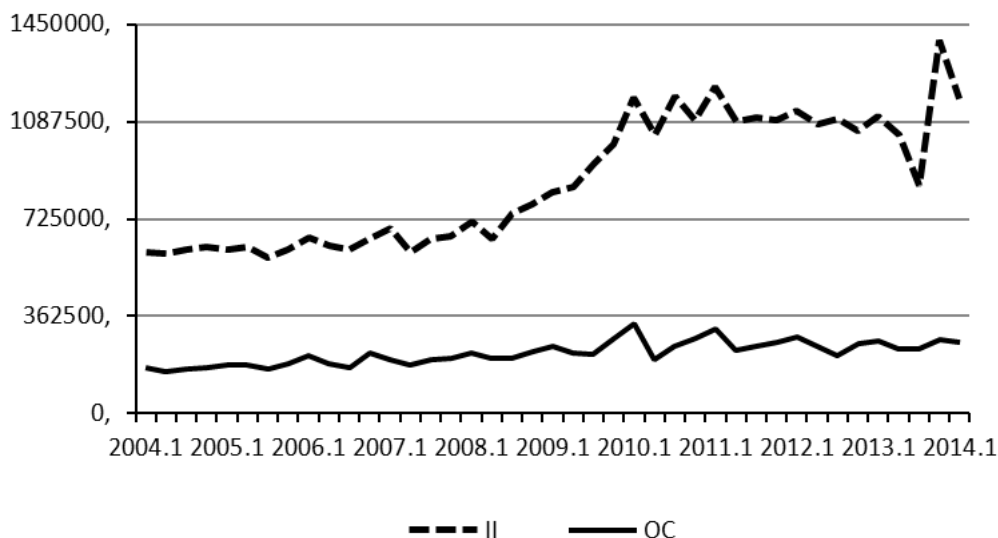
Po oszacowaniu odpowiednich modeli wyznaczone zostały prognozy oraz porównana została ich dokładność za pomocą wybranych mierników.

Dane i analiza wyników empirycznych

W przeprowadzonym badaniu wykorzystane zostały dwie zmienne reprezentujące liczbę szkód:

- liczba wypłat działu II ubezpieczeń ogółem oraz (II),
- liczba wypłat z ubezpieczenia OC pojazdów lądowych (OC).

Wykres 1. Kwartalna liczba szkód w okresie 2004.1–2014.2



Źródło: opracowanie własne na podstawie danych KNF (http://www.knf.gov.pl/opracowania/rynek_ubezpiezen/Dane_o_rynku/Dane_kwartalne/dane_kw.html)

Zbiór danych kwartalnych w postaci szeregów czasowych dotyczy okresu od 1. kwartału 2004 roku do 2. kwartału 2014 roku, a więc szeregi składają się z 42 obserwacji. Wszystkie dane pochodzą z internetowej strony Komisji Nadzoru Finansowego (KNF).

Dane poddane zostały transformacji z uwagi na fakt, że publikowane są w sposób narastający. Sprowadzone zostały zatem do wielkości przypadających na dany okres.

Wartości zmiennych wykorzystanych bezpośrednio w modelowaniu przedstawia wykres 1. Warto zaznaczyć, że liczba wypłat może się różnić od liczby zgłoszonych szkód, bo towarzystwa ubezpieczeniowe nie zawsze mają je zgłoszone i dokonują wypłaty.

Do modelowania szkodowości posłużyła próba zebranego zbioru danych skrócona o ostatnie 2 obserwacje, które w dalszej części badania posłużyły do oceny zdolności prognostycznych prezentowanych narzędzi. Po wstępnym przygotowaniu danych przeprowadzono procedury

wyznaczania stopnia wielomianu trendu za pomocą testu F (Aczel 2000: 375–382). Okazało się, że dla obu rozważanych zmiennych wystarczy przyjąć trend liniowy. Następnie na podstawie analizy autokorelacji wyznaczono odpowiednie rzędy autoregresji. W każdym modelu początkowo uwzględniona została również sezonowość. W kolejnym kroku przeprowadzona została eliminacja nieistotnych statystycznie zmiennych *a posteriori* (Osińska 2007: 25). W wyniku tej procedury uzyskano modele przedstawione w kolejnych tabelach (1 i 2). Odpowiednie zmienne opóźnione oznaczone zostały podkreślnikiem i właściwym rzędem opóźnienia, a zmienne sezonowe symbolem dq .

Zasadna wydaje się również próba znalezienia zmiennej lub zmiennych, które posłużyć mogłyby jako objaśniające w ten sposób zdefiniowaną szkodowość ubezpieczeniową. Dostępną i jednocześnie odpowiednią wydaje się być liczba polis przypisana do całego działu II ubezpieczeń i liczba polis OC posiadaczy pojazdów lądowych.

Tabela 1. Model szkodowości działu II

	Współczynnik	Błąd stand.	t-Studenta	wartość p
const	209 789	64 374,6	3,2589	0,0024
time	14 466,4	1841,55	7,8555	0,0000
II_1	0,1359	0,0705	1,9268	0,0619
II_2	0,1343	0,0716	1,8749	0,0689
Wsp. determ. R-kwadrat		0,8203	Autokorel.reszt - rho1	0,2285

Źródło: opracowanie własne.

Tabela 2. Model szkodowości OC

	<i>Współczynnik</i>	<i>Błąd stand.</i>	<i>t-Studenta</i>	<i>wartość p</i>
const	151 398	11 439,1	13,24	3,47E-15
time	2896,08	302,71	9,567	2,66E-11
dq1	16 011,6	9878,74	1,621	0,114
dq2	-24 695,8	9855,52	-2,506	0,017
dq3	-25 796,1	9841,56	-2,621	0,0129
Wsp. determ. R-kwadrat		0,7666	Autokorel.reszt - rho1	0,2475

Źródło: opracowanie własne.

Modele, które w swej strukturze uwzględniają wyżej wymienione zmienne, prezentują tabele 3 i 4. W tabelach zawierających wyniki estymacji zmienne zostały wyróżnione przedrostkiem w postaci litery *p*.

Współczynniki determinacji R-kwadrat we wszystkich modelach (tabele 1, 2, 3 i 4) wskazują na dobre dopasowanie modeli do danych empirycznych, a współczynniki autokorelacji są

wystarczająco niskie, żeby mówić o jej braku w zaprezentowanych modelach ekonometrycznych. Oszacowane zostały również parametry modelu wygładzania wykładniczego Wintersa (W) w wersji multiplikatywnej. Kryterium ich wyznaczenia była minimalizacja błędu średniego prognoz wygastych. Wartości parametrów dla modeli opisujących rozważane zmienne zawiera kolejna tabela 5.

Tabela 3. Model szkodowości działu II+

	<i>Współczynnik</i>	<i>Błąd stand.</i>	<i>t-Studenta</i>	<i>wartość p</i>
const	430 407	169 824	2,5344	0,016
pII_4	-0,0117	0,00482	-2,4203	0,021
pII_5	0,0097	0,00452	2,1365	0,0399
time	15379,7	2435,94	6,3136	0,0000
II_2	0,1718	0,09058	1,8968	0,0664
Wsp. determ. R-kwadrat		0,8311	Autokorel.reszt - rho1	0,2659

Źródło: opracowanie własne.

Tabela 4. Model szkodowości OC+

	<i>Współczynnik</i>	<i>Błąd stand.</i>	<i>t-Studenta</i>	<i>wartość p</i>
const	138 081	38 915,4	-3,5482	0,0012
pOC_1	0,02418	0,0026	9,3061	0,0000
dq1	13 147,8	10 322,1	1,2738	0,2114
dq2	-11 590	10 314,7	-1,1236	0,269
dq3	-19 404	9732,64	-1,9937	0,0543
OC_1	-0,1411	0,0711	-1,9848	0,0553
Wsp. determ. R-kwadrat		0,7803	Autokorel.reszt - rho1	0,3476

Źródło: opracowanie własne.

Tabela 5. Modele Wintersa

	<i>A</i>	<i>β</i>	<i>Γ</i>
Szkodowość działu II	0,32	0,7	0,18
Szkodowość OC	0,23	0,91	0,31

Źródło: opracowanie własne.

Tabela 6. Autoregresyjny model szkodowości działu II

	współczynnik	błąd standardowy	t-Studenta	wartość p
const	226 844	104 553	2,17	0,0365
II_1	0,3454	0,1061	3,256	0,0024
II_2	0,3817	0,1045	3,652	0,0008

Źródło: opracowanie własne.

Tabela 7. Autoregresyjny model szkodowości działu OC

	współczynnik	błąd standardowy	t-Studenta	wartość p
const	180 159	26 261,4	6,86	3,80E-08
OC_1	0,1955	0,11	1,777	0,0835

Źródło: opracowanie własne.

Wszystkie wymienione wyżej modele posłużyły w dalszym etapie analizy do wyznaczenia prognoz szkodowości. Jako punkt odniesienia i kryterium porównań przyjęte zostały również odpowiednie modele autoregresyjne (AR) analizowanych zmiennych zaprezentowane poniżej. Na podstawie oszacowanych modeli wyznaczono prognozy na dwa pierwsze kwartały roku 2014. Wykorzystując posiadane wartości, wyznaczono wybrane błędy prognoz w celu porównania wyników uzyskanych na podstawie alternatywnych modeli. Do porównań wykorzystano (Welfe 2003):

– średni błąd predykcji:

$$ME = \frac{1}{h} \sum_{T=n+1}^{n+h} (Y_T - Y_{T_p}), \quad (5)$$

– pierwiastek błędu średniokwadratowego:

$$RMSE = \sqrt{\frac{1}{h} \sum_{T=n+1}^{n+h} (Y_T - Y_{T_p})^2}, \quad (6)$$

– średni absolutny błąd procentowy:

$$MAPE = \frac{1}{h} \sum_{T=n+1}^{n+h} \left| \frac{Y_T - Y_{T_p}}{Y_T} \right|, \quad (7)$$

– współczynnik Theila:

$$I^2 = \frac{\sum_{T=n+1}^{n+h} (Y_T - Y_{T_p})^2}{\sum_{T=n+1}^{n+h} Y_T^2}, \quad (8)$$

We wszystkich powyższych wzorach:

Y_T – oznacza realizacje zmiennej prognozowanej,
 Y_{T_p} – oznacza wartości prognoz zmiennej prognozowanej,

h – oznacza horyzont prognozy.

W wyniku przeprowadzonych obliczeń uzyskano wyniki zaprezentowane w tabeli 8.

W przypadku szkodowości działu II jako całości wyniki prognozowania niestety nie wskazują jednoznacznie najlepszego narzędzia. Dwukrotnie najniższe błędy uzyskano dla modelu zawierającego zmienną objaśniającą i jednokrotnie dla modelu bez tej zmiennej i modelu Wintersa. Jednocześnie we wszystkich przypadkach najgorszy okazał się model autoregresyjny, ponieważ błędy mają najwyższą wartość. Bardziej jednoznacznymi okazują się wyniki prognozowania szkodowości OC posiadaczy pojazdów lądowych. Wartości trzech błędów

Tabela 8. Błędy prognoz szkodowości.

Model	RMSE	MAPE	THEIL	ME
II	50 107,00	3,8925	1,1646	9039,6
ARII	198 040,00	14,918	3,7828	189 210,00
WII	91 752,99	7,2597	0,0053	91 721,11
II+	45 534	3,3094	0,7099	17 503
OC	23 077	7,1579	0,385	22 622
AROC	85 106	26,165	1,3554	82 891
WOC	38 923,58	12,427	0,0154	38 855,15
OC+	25 787	8,238	0,519	25 749

Źródło: opracowanie własne.

wskazują na prosty model ekonometryczny bez zmiennej objaśniającej jako najlepsze narzędzie prognostyczne spośród rozważanych, a tylko wartość współczynnika Theila wskazuje na przewagę modelu Wintersa.

Zakończenie

Na podstawie szeregów czasowych o częstotliwości kwartalnej udało się oszacować modele

dobrze opisujące liczbę szkód majątkowych występujących na polskim rynku ubezpieczeń oraz wyznaczyć na ich podstawie prognozy.

Najlepszym z prognostycznego punktu widzenia modelem liczby szkód z działu II ubezpieczeń okazał się II+ zawierający trend oraz opóźnienia zmiennej objaśnianej i zmiennej objaśniającej. W zakresie szkód z ubezpieczenia OC posiadaczy pojazdów lądowych natomiast był to sezonowo-trendowy model OC.

Bibliografia

Aczel A. D. (2000), *Statystyka w zarządzaniu*, Warszawa, Wydawnictwo Naukowe PWN.

Cieślak M. (2001), *Prognozowanie gospodarcze*, Warszawa, Wydawnictwo Naukowe PWN.

Monkiewicz J. (2000), *Podstawy Ubezpieczeń*, tom I: *Mechanizmy i funkcje*, Warszawa, Wydawnictwo Poltext.

KNF. Opracowania. Rynek ubezpieczeń, http://www.knf.gov.pl/opracowania/rynek_ubezpieczen/Dane_o_rynku/Dane_kwartalne/dane_kw.html [19.09.2014].

Kufel T. (2011), *Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu GRETL*, Warszawa, Wydawnictwo Naukowe PWN.

Osińska M. (2007), *Ekonometria współczesna*, TNOiK „Dom Organizatora”, Toruń.

Sangowski T. (2001), *Ubezpieczenia gospodarcze*, Warszawa, Wydawnictwo Poltext.

Ustawa z 22 maja 2003 r. o działalności ubezpieczeniowej. Dz. U. 2003, Nr 124, poz. 1151 z późn. zm.

Welfe A. (2003), *Ekonometria*, Warszawa, PWE.

Forecast of insurance claims number in Poland

Abstract

Presented paper aims to research the possibilities to estimate and forecast claims number in Poland. It is a part of a wider analysis of insurance market in Poland. Empirical analysis concerns claims which relate to the second section of insurance and land vehicles liability, and it was performed on the basis of data from the first quarter of 2004 to the second quarter of 2014.

In the article we compare forecast accuracy of 4 competitive models. The results are encouraging. Model II+ turns out the best model for the second section of insurance whereas for land vehicle liability it was a seasonal-trend model.

Keywords: insurance, claims, forecast