



Kamil Gala

Ubezpieczeniowy Fundusz Gwarancyjny
e-mail: kgala@ufg.pl

LOJALNOŚĆ KLIENTÓW W UBEZPIECZENIACH KOMUNIKACYJNYCH W ŚWIETLE DANYCH UFG

Streszczenie: Artykuł jest poświęcony zagadnieniu statystycznego modelowania lojalności klientów w ubezpieczeniach komunikacyjnych. Głównym celem pracy jest identyfikacja czynników, które wpływają na decyzję klienta o kontynuowaniu umowy ubezpieczenia w tym samym zakładzie ubezpieczeń lub o zmianie ubezpieczyciela. Analizą zostały objęte umowy obowiązkowego ubezpieczenia odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych (OC p.p.m.), a źródłem informacji o umowach ubezpieczenia i klientach była ogólnopolska baza danych Ośrodka Informacji Ubezpieczeniowego Funduszu Gwarancyjnego.

W pracy zostały przedstawione wybrane trendy dotyczące lojalności klientów widoczne w danych UFG, a także wnioski płynące z modelowania badanego zjawiska za pomocą regresji logistycznej. Zostało również omówione to, w jaki sposób informacja dotycząca lojalności klienta może zostać wykorzystana do modelowania liczby szkód z tytułu zawartej umowy ubezpieczenia OC p.p.m.

Słowa kluczowe: lojalność klientów, ubezpieczenia komunikacyjne, ubezpieczenie odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych, uogólnione modele liniowe.

JEL Classification: G22.

Wprowadzenie

W czasach rosnącej konkurencji na rynku i większej świadomości ubezpieczonych dużym wyzwaniem dla zakładu ubezpieczeń jest nie tylko pozyskanie nowych klientów, ale również utrzymanie obecnych. W pewnych przypadkach, na przykład w długoterminowych ubezpieczeniach na życie, przedwczesne zakończenie umowy (*lapse*) może mieć dla zakładu negatywne skutki finansowe. Z tego względu ubezpieczyciele mogą być zainteresowani głębszym zrozumie-

niem zachowania klientów na rynku ubezpieczeń. Zastosowanie znajdują tutaj metody statystyczne, które pomagają dopasować ofertę do potrzeb klienta, a także odpowiednio wcześnie reagować na sygnały świadczące o możliwym odejściu klienta do konkurencji.

Szeroko rozumiana lojalność klienta jest złożonym i wielowymiarowym zagadnieniem, a w literaturze przedmiotu nie istnieje jednolita definicja tego pojęcia [Bennet i Bove, 2002]. Jako dominujące nurty należy wskazać podejście behawioralne [Guest 1944; Brown, 1952], w którym lojalność klienta jest postrzegana przez pryzmat jego rzeczywistych decyzji zakupowych, oraz podejście odnoszące się do nastawienia (*attitudinal approach*), podkreślające rolę emocji oraz nastawienia do konkretnego produktu lub marki [Jacoby i Chestnut, 1978]. Obecnie te nurty silnie łączą się ze sobą, jednak definiowanie lojalności i jej pomiar pozostają złożonym problemem. Przykładowo, Mowen [1995, s. 534] uważa, że lojalność może być związana z konkretnym produktem i nie jest ogólną cechą klienta, z kolei Uncles i Laurent [1997] wskazują, że lojalność wiąże się wyłącznie ze skłonnością do zakupu produktu danej marki, a proces ten cechuje się dużą losowością.

Zagadnienie lojalności klienta było również poruszane w literaturze statystycznej. Ilościowe podejście do zagadnienia zostało przedstawione m.in. w pracy Guillen i in. [2002], gdzie za pomocą regresji logistycznej jest modelowane prawdopodobieństwo wcześniejszego zakończenia umowy ubezpieczenia przez klienta, a także w pracy Schneidera i Bristowa [2007] w odniesieniu do lojalności kibiców względem drużyny baseballowej. Badania polskiego rynku ubezpieczeń były prowadzone m.in. przez Rodzinkę [2010] oraz Wicką i Miedzik [2010].

W niniejszym artykule została podjęta próba odpowiedzi na pytanie, jak kształtuje się lojalność klientów na polskim rynku ubezpieczeń odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych (dalej: OC p.p.m.). W tym celu została przeprowadzona analiza statystyczna danych zgromadzonych w bazie danych Ośrodka Informacji Ubezpieczeniowego Funduszu Gwarancyjnego (dalej: OI UFG). Dane te pozwalają na ustalenie, w jakich zakładach ubezpieczeń była ubezpieczona dana osoba, natomiast nie zawierają informacji o postrzeganiu danego zakładu ubezpieczeń przez klienta, ofertach pozostałych ubezpieczycieli itp. W związku z tym lojalność klienta została zdefiniowana jako skłonność do odnowienia umowy ubezpieczenia w tym samym zakładzie ubezpieczeń. Definicja ta jest zbliżona do definicji przedstawionej w pracy Rodzinki [2010] i wpisuje się w przedstawiony wcześniej nurt behawioralny.

Struktura artykułu jest następująca: w punkcie pierwszym omówiono wykorzystany zbiór danych i przedstawiono wybrane trendy rynkowe. Punkt drugi jest poświęcony modelowaniu prawdopodobieństwa zmiany zakładu ubezpieczeń przez ubezpieczonego, natomiast w punkcie trzecim poruszono kwestię związku między lojalnością klienta a jego szkodowością. Pracę zamykają podsumowanie i wnioski.

1. Eksploracja danych

1.1. O bazie danych Ośrodka Informacji UFG

Ubezpieczeniowy Fundusz Gwarancyjny jest instytucją rynku ubezpieczeniowego, której działalność jest regulowana przepisami Ustawy z dnia 22 maja 2003 r. o ubezpieczeniach obowiązkowych, Ubezpieczeniowym Funduszu Gwarancyjnym i Polskim Biurze Ubezpieczycieli Komunikacyjnych [Dz.U. 2013, poz. 392 tj., dalej: Ustawa]. Podstawowym zadaniem UFG jest zaspokajanie roszczeń z tytułu ubezpieczeń obowiązkowych. W szczególności dotyczy to szkód, które zostały wyrządzone w okolicznościach uzasadniających odpowiedzialność cywilną posiadacza pojazdu mechanicznego lub kierującego pojazdem mechanicznym, a nie ustalono ich tożsamości, a także gdy posiadacz pojazdu mechanicznego, którym wyrządzono szkodę, nie był ubezpieczony obowiązkowym ubezpieczeniem OC p.p.m.¹. Jednocześnie, zgodnie z art. 102 ust. 1 Ustawy Fundusz pełni funkcję ośrodka informacji, zdefiniowaną zgodnie z tzw. dyrektywą komunikacyjną UE [Dyrektywa Parlamentu Europejskiego i Rady 2009/103/WE, art. 23]. W ramach tej funkcji UFG powołał w swojej strukturze organizacyjnej Ośrodek Informacji (OI UFG), który prowadzi, w formie komputerowej bazy danych, rejestr umów ubezpieczenia działu II grupy 3 i 10 załącznika do ustawy o działalności ubezpieczeniowej, z wyłączeniem odpowiedzialności przewoźnika. W bazie tej gromadzone są również dane dotyczące zdarzeń powodujących odpowiedzialność zakładu ubezpieczeń z tytułu zawartej umowy OC p.p.m. lub umowy ubezpieczenia grupy 3 działu II, a także wypłaconych odszkodowań i świadczeń (lub odmów wypłaty) z nimi związanych. Zakres gromadzonych danych jest określony w art. 102 Ustawy i obejmuje wiele informacji o umowie ubezpieczenia oraz ubezpieczonych podmiotach oraz pojazdach. Jednocześnie, zgodnie z art. 102 ust. 7, Fundusz jest uprawniony do przetwarzania gromadzonych danych dla celów niewymienionych w Ustawie, po

¹ Zakres odpowiedzialności UFG jest określony szczegółowo w art. 98 Ustawy.

modyfikacji niepozwalającej na ustalenie tożsamości osoby, której dotyczą dane. Przepis ten otwiera przed UFG szerokie możliwości prowadzenia analiz na wysokim poziomie agregacji, które mogą stanowić cenną informację w zakresie zjawisk zachodzących na rynku ubezpieczeń komunikacyjnych. Rozmiar bazy (ok. 300 mln rekordów w 2015 r.) oraz zakres gromadzonych informacji sprawiają, że jej potencjał analityczny jest bardzo duży, a dane udostępniane przez UFG istotnie wspierają procesy biznesowe zakładów ubezpieczeń – członków Funduszu.

Dane gromadzone w bazie OI UFG są danymi detalicznymi, przekazywanymi drogą teletransmisji przez zakłady ubezpieczeń – członków Funduszu. Ze względu na rozproszony model zasilania stałym wyzwaniem, przed którym stoi OI UFG, jest jakość gromadzonych danych. Działania podejmowane przez Fundusz (m.in. prowadzenie rankingu jakości danych) skutkują ciągłym wzrostem kompletności i jakości gromadzonych informacji, co z kolei przekłada się na wzrost wiarygodności wyników analiz wykonywanych na ich podstawie.

1.2. Opis zbioru danych

Dane wykorzystane do badania zostały przygotowane na podstawie informacji zgromadzonych w opisanym w poprzednim punkcie bazie danych OI UFG. Zbiór danych został przygotowany według następujących założeń:

1. Jednostkami badania są obiekty zdefiniowane jako para podmiot + pojazd. Relacja podmioty-pojazdy jest relacją „wiele do wielu”, więc takie określenie jednostki badania sprawia, że łatwiej zdefiniować lojalność względem zakładu ubezpieczeń.
2. Analiza została ograniczona do podmiotów będących osobą fizyczną; zostały również wykluczone umowy flotowe.
3. Dla każdego obiektu jest wyznaczana historia ubezpieczenia, przy czym brane pod uwagę są umowy ubezpieczenia OC p.p.m. zawarte w latach 2009-2014.
4. Każdy obiekt posiada atrybuty opisujące ubezpieczony podmiot lub pojazd.
5. Dane zostały poddane procesowi czyszczenia pod kątem błędów, a także usunięto obserwacje uznane za nietypowe (anomalie).

Lista zmiennych wykorzystanych w badaniu jest następująca:

- PLEC – płeć ubezpieczonego (M – mężczyzna, K – kobieta);
- WIEK – wiek ubezpieczonego w momencie zawarcia umowy, w podziale na grupy 0-25, 26-35, 36-45, 46-55, 56-65, 65+;
- WOJEWODZTWO/POWIAT – województwo i powiat wyznaczone na podstawie adresu podanego przez ubezpieczonego;

- RODZAJ_POJAZDU – rodzaj ubezpieczonego pojazdu, w podziale na kategorie „Samochód osobowy”, „Samochód ciężarowy”, „Jednoślady”, „Inne”;
- CZY_AC – czy ubezpieczony posiada umowę AC na dany pojazd (1 – tak, 0 – nie);
- CZY_AC_ZU – czy ubezpieczony posiada umowę AC na dany pojazd w tym samym zakładzie ubezpieczeń, w którym ma ubezpieczenie OC (1 – tak, 0 – nie);
- CZY_AC_INNY_ZU – czy ubezpieczony posiada umowę AC na dany pojazd w innym zakładzie ubezpieczeń niż ten, w którym ma ubezpieczenie OC (1 – tak, 0 – nie);
- CZY_WIELE_POJAZDOW – czy ubezpieczony posiada inne pojazdy (1 – tak, 0 – nie);
- CZY_WIELE_PODMIOTOW – czy w badanej umowie występują współubezpieczeni (1 – tak, 0 – nie);
- CZY_SZKODA – czy dla badanej umowy powstała szkoda powodująca odpowiedzialność zakładu ubezpieczeń (1 – tak, 0 – nie);
- LICZBA_SZKOD – liczba szkód powiązanych z badaną umową ubezpieczenia;
- ZMIANA_ZU – czy w kolejnym roku podmiot posiadał ubezpieczenie OC p.p.m. w tym samym zakładzie (wartość = 0), innym zakładzie (wartość = 1) lub czy brak danych o ubezpieczeniu OC w kolejnym roku (wartość = -1);
- HISTORIA – czy w poprzednim roku podmiot był ubezpieczony w tym samym zakładzie (wartość = „bez zmian”), innym zakładzie (wartość = „zmiana”) lub czy brak danych o umowie w poprzednim roku (wartość = „nowy”).

Zbiór danych przygotowany według powyższych założeń zawiera ok. 79 mln rekordów i opisuje 25,8 mln unikatowych obiektów, w tym dane o 13,2 mln różnych osobach.

Warto podkreślić, że ze względu na złożony proces selekcji danych źródłowych przeprowadzone badanie, mimo dużego wolumenu danych wejściowych, nie jest badaniem pełnym. Analiza struktury zbioru wskazuje jednak, że przynajmniej w niektórych wymiarach (np. udziały w rynku poszczególnych zakładów ubezpieczeń) udało się zachować reprezentatywność zbioru. W związku z tym wydaje się, że zebrane dane są odpowiednim materiałem analitycznym do analizy lojalności klientów na rynku ubezpieczeń OC p.p.m.

Należy zwrócić uwagę, że poza zakresem danych OI UFG znajduje się wiele informacji istotnych z punktu widzenia podejmowania decyzji przez konsumenta, np. cena produktu (składka), czy też sytuacja materialna klienta i jego preferencje. Niemniej jednak, wydaje się, że dzięki danym UFG można uzyskać ogólny obraz procesów zachodzących na rynku ubezpieczeń komunikacyjnych, który może stanowić punkt wyjścia do dalszych badań.

1.3. Trendy rynkowe

W niniejszym punkcie zostaną przedstawione wybrane trendy dotyczące lojalności klientów obserwowane w danych OI UFG. Podstawową miarą wykorzystywaną w dalszej analizie będzie wskaźnik mobilności (WM) zdefiniowany jako:

$$WM_i = \frac{Z_i}{U_i}, \quad (1)$$

gdzie:

WM_i – wskaźnik mobilności w roku i ,

U_i – liczba umów ubezpieczenia zawartych w roku i , dla których obiekt był ubezpieczony również w kolejnym roku,

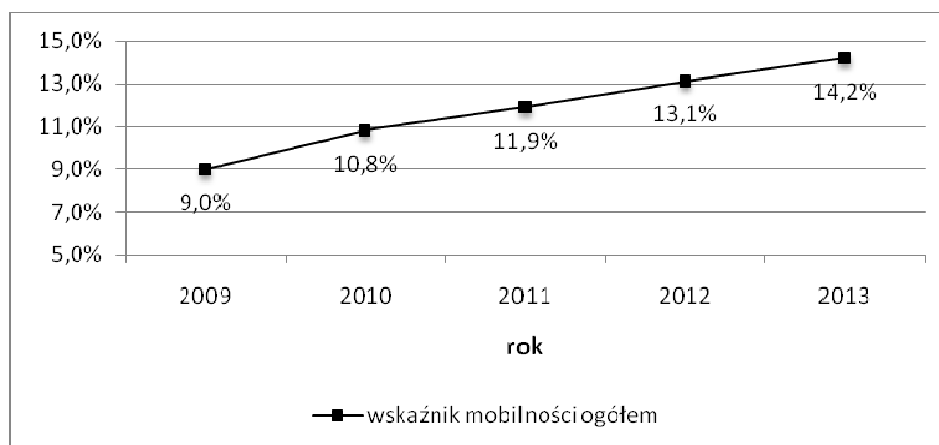
Z_i – liczba umów ubezpieczenia zawartych w roku i , dla których obiekt był ubezpieczony w kolejnym roku w innym zakładzie ubezpieczeń.

Z powyższej definicji wynika, że wartość wskaźnika w roku i odnosi się do umów zawartych w roku i , dla których w roku $i + 1$ nastąpiła zmiana zakładu ubezpieczeń lub odnowienie umowy w tym samym zakładzie. Przykładowo zatem wartość wskaźnika WM_{2009} jest obliczana na podstawie umów zawartych w 2009 r., dla których ewentualna zmiana ubezpieczyciela odbywa się w 2010 r. Według tej definicji, przy dostępnym zakresie danych, jest możliwe wyznaczenie wartości WM_i dla $i \in \{2009, 2010, 2011, 2012, 2013\}$. Warto w tym miejscu zaznaczyć, że podczas obliczania wartości WM_i nie są brane pod uwagę obiekty, dla których historia ubezpieczenia kończy się w roku i . Wynika to z faktu, że koniec historii ubezpieczenia może być wywołany zarówno decyzją klienta (sprzedaż pojazdu, ubezpieczenie pojazdu przez inną osobę z rodziny itp.), jak i jakością danych (brakująca umowa została odfiltrowana na etapie przygotowania danych lub nie została przekazana przez zakład ubezpieczeń do bazy OI UFG).

Na rys. 1 przedstawiono wartości wskaźnika dla całego rynku w latach 2009-2013.

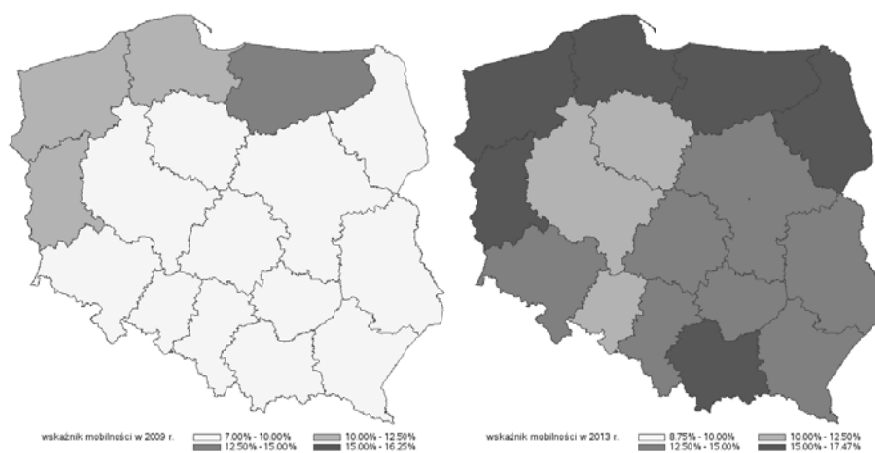
Dane wskazują, że wskaźnik mobilności w latach 2009-2013 wzrósł z 9% do 14,2%, czyli o ponad 50%. Oznacza to, że zakładom ubezpieczeń coraz trudniej utrzymać klienta. Może to wynikać ze wzrostu świadomości ubezpieczeniowej klientów, a także z wojny cenowej trwającej na rynku ubezpieczeń OC p.p.m. Ze względu na krótki okres obserwacji nie można również wykluczyć, że obserwowany trend jest jedynie częścią długookresowej tendencji, np. pewnego cyklu rynku ubezpieczeniowego. Ponadto poza okresem obserwacji znalazł się początek kryzysu finansowego, który mógł istotnie wpłynąć na decyzje podmiotów dotyczące finansów osobistych.

Dostępność informacji o miejscu zamieszkania ubezpieczonego sprawia, że dane można analizować w wymiarze geograficznym. Na rys. 2 przedstawiono porównanie wskaźników mobilności dla lat 2009 i 2013 w podziale na województwa, natomiast rys. 3 przedstawia wskaźniki mobilności w 2013 r. w podziale na powiaty.



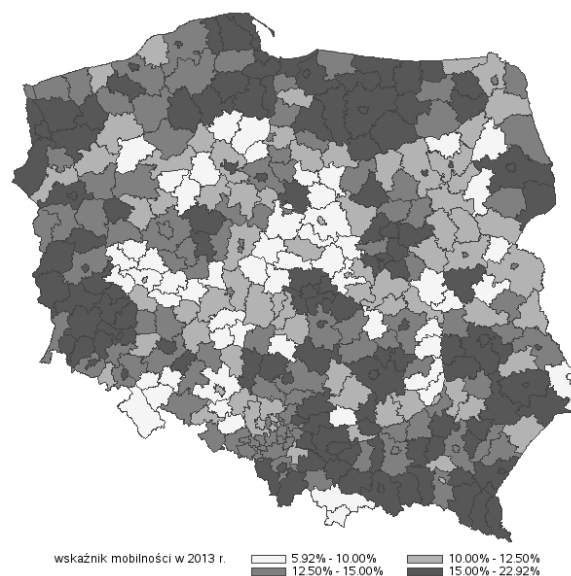
Rys. 1. Wartość wskaźnika mobilności w latach 2009-2013

Źródło: Opracowanie własne.



Rys. 2. Porównanie wskaźników mobilności w podziale na województwa

Źródło: Opracowanie własne.



Rys. 3. Wskaźniki mobilności w 2013 r. w podziale na powiaty

Źródło: Opracowanie własne.

Wskaźnik mobilności wzrósł w latach 2009-2013 dla wszystkich województw, przy czym najwyższy pozostaje w północnej i północno-zachodniej Polsce. Wartość wskaźnika mobilności w podziale na powiaty wskazuje z kolei, że w wielu przypadkach jest on wyższy w mieście na prawach powiatu (powiecie grodzkim), niż w powiecie go otaczającym (powiecie ziemskim). Na uwagę zasługują również skupiska wysokich wartości wskaźnika w zachodniej i północnej Polsce, a także na południowym-wschodzie (Małopolska i Podkarpacie).

Analiza wskaźników mobilności w przekrojach wyznaczonych przez pozostałe zmienne objaśniające potwierdziła, że jest obserwowany istotny spadek lojalności klientów dla większości badanych grup. W tabeli 1 podsumowano wpływ poszczególnych zmiennych na wskaźnik mobilności.

Tabela 1. Wpływ badanych czynników na lojalność klientów

Zmienna	Wpływ na wskaźnik mobilności
1	2
Wiek	spadek
Płeć	wartość wskaźnika większa dla kobiet, różnica nie przekracza 1 p.p.
Wystąpienie szkody	wzrost rzędu 150%-200%
Posiadanie AC w tym samym zakładzie	spadek (różnica mniejsza niż 1 p.p.)
Posiadanie AC w innym zakładzie	wzrost rzędu 200%-300%

cd. tabeli 1

1	2
Występowanie współubezpieczonych	spadek (różnica mniejsza niż 1 p.p.)
Posiadanie wielu pojazdów	spadek (różnica mniejsza niż 1 p.p.)
Odnowienie poprzedniej umowy w tym samym zakładzie	spadek
Zmiana ubezpieczyciela w poprzednim okresie	wzrost
Rodzaj pojazdu	największa wartość – samochody osobowe, najmniejsza wartość – kategoria „inne”

Źródło: Opracowanie własne.

Na zakończenie tej części pracy zostanie poruszona kwestia związku między lojalnością klientów a charakterystykami zakładu ubezpieczeń. Zbadano korelację między wielkością zakładu ubezpieczeń (mierzoną liczbą umów zawartych w danym roku) a wskaźnikiem mobilności dla umów zawartych w tym samym roku. Współczynniki korelacji rang Spearmana między tymi wielkościami wahały się od ok. $-0,2$ dla 2009 r. do ok. $-0,6$ dla 2013 r. Oznacza to, że klienci większych zakładów ubezpieczeń cechują się mniejszą skłonnością do zmiany ubezpieczyciela. Drugim badanym czynnikiem była pozycja zakładu ubezpieczeń w rankingu jakości danych prowadzonym przez UFG. W tym przypadku współczynnik korelacji Spearmana wyniósł ok. $0,3$. Oznacza to, że klienci zakładów zajmujących wyższe pozycje w rankingu cechują się niższą skłonnością do zmiany zakładu ubezpieczeń. Możliwym wyjaśnieniem tego wyniku jest to, że wysoka jakość danych przesyłanych do bazy danych OI UFG jest efektem wysokiego poziomu technologicznego i organizacyjnego, który przekłada się również na jakość pozostałych procesów biznesowych, takich jak likwidacja szkód i obsługa klienta.

2. Statystyczna analiza lojalności klientów w ubezpieczeniach komunikacyjnych

2.1. Metodyka

Przeprowadzona w poprzednim punkcie jednowymiarowa analiza czynników wpływających na skłonność klienta do zmiany zakładu ubezpieczeń nie jest wystarczająca i konieczne jest zastosowanie metod statystyki wielowymiarowej, pozwalającej uwzględnić oddziaływanie wielu zmiennych objaśniających jednocześnie. W związku z tym do identyfikacji determinantów lojalności klienta w ubezpieczeniach komunikacyjnych został zastosowany model regresji logistycznej.

Model ten należy do grupy tzw. uogólnionych modeli liniowych (*generalized linear models* – GLM). GLM są zdefiniowane za pomocą trzech podstawowych elementów [Fox, 2008]:

1. Składnika losowego określającego rozkład warunkowy zmiennej objaśnianej Y , przy ustalonych wartościach zmiennych objaśniających. Zakłada się, że rozkład ten należy do rodziny wykładniczej, tj. jego gęstość lub funkcję prawdopodobieństwa $f_Y(y)$ można zapisać w postaci:

$$f_Y(y; \theta; \psi) = \exp\left(\frac{y\theta - b(\theta)}{\psi} + c(y; \psi)\right), \quad y \in D_\psi, \quad (2)$$

gdzie:

θ i ψ są ustalonymi parametrami,

$b: \mathbb{R} \rightarrow \mathbb{R}$ i $c: \mathbb{R}^2 \rightarrow \mathbb{R}$ to ustalone funkcje,

D_ψ jest nośnikiem rozkładu.

2. Składnika systematycznego postaci:

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}, \quad (3)$$

gdzie:

β_0, \dots, β_k to ustalone parametry,

X_{ij} to wartość zmiennej objaśniającej j ($j = 1, \dots, k$) dla obserwacji i .

3. Funkcji wiążącej (*link function*) g , przekształcającej wartość oczekiwaną zmiennej Y_i , równą μ_i , na predyktor liniowy według wzoru:

$$g(\mu_i) = \eta_i. \quad (4)$$

Funkcja wiążąca określa związek między parametrami rozkładu zmiennej Y a kombinacją liniową zmiennych objaśniających, η_i .

W modelu regresji logistycznej zakłada się, że zmienna objaśniana ma rozkład dwupunktowy, w którym $\mathbb{P}(Y = k) = p^k(1-p)^{1-k}$ dla $k \in \{0,1\}$ oraz $p \in (0,1)$. Rozkład ten otrzymuje się, przyjmując we wzorze (2) $\theta = \ln\left(\frac{p}{1-p}\right)$, $b(\theta) = \ln(1 + \exp(\theta))$, $\psi = 1$ oraz $c(y, \psi) \equiv 0$. Funkcja wiążąca jest postaci $g(p) = \ln\left(\frac{p}{1-p}\right)$, zatem dla obserwacji i zachodzi $p_i = \mathbb{P}(Y_i = 1) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}$.

Zmienną objaśnianą w badaniu była dychotomiczna zmienna ZMIANA_ZU, a zmiennymi objaśniającymi pozostałe zmienne opisane w punkcie 1.2. Próbę podzielono na podpróby ze względu na rok zawarcia umowy – miało to na celu weryfikację stabilności obserwowanych zależności w czasie. W celu zachowania porównywalności modeli analizę ograniczono do lat 2010-2013, ponieważ dla 2009 r. nie była dostępna informacja o poprzedniej umowie ubezpieczenia.

W pierwszym etapie selekcja zmiennych objaśniających do modelu była prowadzona na podstawie kryterium istotności statystycznej. Budowa modelu odbywała się różnymi metodami (m.in. selekcją wsteczną i postępującą), a o uwzględnieniu lub wykluczeniu konkretnego czynnika decydowała jego istotność statystyczna oraz wynik odpowiedniego testu ilorazu wiarygodności. W drugim etapie w procesie doboru zmiennych uwzględniono wartość bayesowskiego kryterium informacyjnego (BIC) i wzięto pod uwagę zdolności predykcyjne modelu.

2.2. Wyniki analizy

Estymacja parametrów modelu została wykonana metodą największej wiarygodności. Wyniki estymacji zostały podsumowane w tabeli 2.

Tabela 2. Oceny parametrów modeli regresji logistycznej

Zmienna	Poziom	Rok			
		2010	2011	2012	2013
1	2	3	4	5	6
wyraz wolny	–	–1,360	–1,270	–1,211	–1,020
czy_ac	–	–0,064	–0,129	–0,128	–0,228
czy_ac_inny_zu	–	0,654	0,534	0,563	0,516
czy_szkoda	–	0,550	0,416	0,391	0,324
czy_wiele_podmiotow	–	–0,361	–0,228	–0,237	–0,157
czy_wiele_pojazdow	–	0,020	0,039	0,054	0,064
historia	BEZ ZMIAN	–0,474	–0,430	–0,324	–0,332
historia	NOWA	–	–	–	–
historia	ZMIANA	0,190	0,220	0,214	0,286
plec	K	–0,063	–0,057	–0,050	–0,056
plec	M	–	–	–	–
rodzaj_pojazdu	CIEŻAROWY	–0,280	–0,247	–0,143	0,000
rodzaj_pojazdu	INNE	–1,624	–1,607	–1,457	–1,647
rodzaj_pojazdu	JEDNOŚLADY	–2,082	–2,142	–2,091	–1,964
rodzaj_pojazdu	OSOBOWY	–	–	–	–
wiek	0-25	–	–	–	–
wiek	26-35	–0,145	–0,151	–0,140	–0,104
wiek	36-45	–0,391	–0,394	–0,354	–0,276
wiek	46-55	–0,467	–0,454	–0,416	–0,351
wiek	56-65	–0,545	–0,541	–0,496	–0,422
wiek	65+	–0,694	–0,679	–0,642	–0,577
województwo	DOLNOŚLĄSKIE	0,332	0,245	0,184	–0,099
województwo	KUJAWSKO-POMORSKIE	0,200	0,107	–0,022	–0,201
województwo	LUBELSKIE	0,231	0,264	0,185	0,032
województwo	LUBUSKIE	0,296	0,298	0,253	0,018

cd. tabeli 2

1	2	3	4	5	6
województwo	MAZOWIECKIE	0,220	0,221	0,120	-0,078
województwo	MAŁOPOLSKIE	0,235	0,187	0,126	-0,002
województwo	OPOLSKIE	-0,009	-0,020	-0,039	-0,227
województwo	PODKARPACKIE	0,029	0,078	0,125	0,025
województwo	PODLASKIE	0,184	0,194	0,162	0,093
województwo	POMORSKIE	0,499	0,380	0,281	0,109
województwo	WARMIŃSKO-MAZURSKIE	0,560	0,364	0,364	0,156
województwo	WIELKOPOLSKIE	-0,035	-0,070	-0,100	-0,263
województwo	ZACHODNIOPOMORSKIE	0,376	0,263	0,198	-0,006
województwo	ŁÓDZKIE	0,168	0,162	0,131	-0,019
województwo	ŚLĄSKIE	0,243	0,168	0,101	-0,112
województwo	ŚWIĘTOKRZYSKIE	-	-	-	-

Źródło: Opracowanie własne.

Znak „-” w powyższej tabeli oznacza poziom odniesienia, względem którego są podawane oceny parametrów dla pozostałych poziomów danej zmiennej. Ujemna ocena parametru oznacza mniejszą skłonność do zmiany zakładu niż dla poziomu odniesienia, natomiast dodatnia ocena parametru – większą.

Należy zwrócić uwagę, że uzyskane wyniki cechują się dużą stabilnością w czasie co do kierunku oddziaływania danej zmiennej, a także – do pewnego stopnia – siły tego efektu. Wyniki można zinterpretować następująco:

- Klienci posiadający AC w tym samym zakładzie ubezpieczeń są mniej skłonni do zmiany ubezpieczyciela OC p.p.m., natomiast klienci posiadający AC w innym zakładzie cechują się większą skłonnością do zmiany zakładu ubezpieczeń.
- Wystąpienie szkody z tytułu umowy OC p.p.m. zwiększa skłonność klienta do zmiany zakładu ubezpieczeń. Można przypuszczać, że ma to związek z obawą przed możliwym wzrostem składki. Interesujące jest to, że siła tego efektu maleje w kolejnych latach. Może to wynikać z tego, że upowszechnia się korzystanie przez ubezpieczycieli z bazy danych UFG, przez co zatajanie historii szkodowej przy zmianie ubezpieczyciela staje się coraz trudniejsze.
- Umowy, w których występują współubezpieczeni częściej zostają przedłużone na kolejny okres.
- Ubezpieczeni posiadający wiele pojazdów są bardziej skłonni do zmiany zakładu ubezpieczeń, choć efekt ten jest niewielki. Może to wynikać z tego, że klient posiadający wiele pojazdów będzie aktywnie poszukiwał najlepszej oferty dla swoich pojazdów.
- Mężczyźni są bardziej skłonni do zmiany ubezpieczyciela niż kobiety, przy czym wpływ tego czynnika jest niewielki.

- Skłonność do zmiany ubezpieczyciela zmniejsza się wraz z wiekiem.
- Największą skłonnością do zmiany ubezpieczyciela cechują się posiadacze samochodów osobowych, najmniejszą – jednośladowców.
- Największą skłonnością do zmiany ubezpieczyciela cechują się klienci, którzy w poprzednim okresie zmienili zakład ubezpieczeń, natomiast najmniejszą skłonnością cechują się klienci, którzy pozostali w tym samym zakładzie ubezpieczeń.
- Największą skłonnością do zmiany zakładu ubezpieczeń cechują się klienci w województwie warmińsko-mazurskim, natomiast najniższą – w województwie wielkopolskim.

Widać, że uzyskane wyniki są w dużej mierze zbieżne z wynikami analizy jednowymiarowej. Godnym uwagi wyjątkiem jest płeć ubezpieczonego – jednocześnie uwzględnienie wielu zmiennych spowodowało, że kierunek szacowanego wpływu odwrócił się. Z powyższych czynników największą siłą oddziaływania charakteryzuje się rodzaj pojazdu, wiek, fakt spowodowania szkody, fakt posiadania AC (w tym samym lub innym zakładzie), a także to, gdzie klient był ubezpieczony w poprzednim okresie (w tym samym zakładzie, w innym zakładzie lub brak umowy w poprzednim roku). Zwłaszcza ostatni z wymienionych czynników jest wart szczególnej uwagi – wyniki analizy wskazują, że klienci, którzy odnowili umowę w tym samym zakładzie ubezpieczeń są bardziej skłonni do zrobienia tego samego w kolejnym roku. W przypadku takich klientów zasadne jest mówienie o lojalności względem ubezpieczyciela. Warto zwrócić uwagę na to, że część przedstawionych wyników jest zbieżna (np. jeśli chodzi o kierunek oddziaływania wieku ubezpieczonego lub faktu spowodowania szkody) z pracą Guillen i in. [2002].

3. Lojalność klienta a szkodowość

3.1. Opis zagadnienia

Wyniki przedstawione w poprzednich punktach dotyczyły opisu lojalności klientów w zależności od cech ich opisujących. Z punktu widzenia praktyki bardzo interesujące wydaje się pytanie o to, w jaki sposób lojalność klientów przekłada się na korzyści ekonomiczne dla zakładu ubezpieczeń. Wśród potencjalnych korzyści można wymienić m.in. lepsze relacje z klientem i wynikającą z tego możliwość zaoferowania mu innych produktów [Bennet i Bove, 2002], mniejszą elastyczność cenową popytu w porównaniu do nowych klientów [Bowen i Shoemaker, 1998], a także niższą szkodowość.

Zakres danych gromadzonych w bazie danych OI UFG pozwala na weryfikację trzeciej hipotezy, tj. zbadanie, czy lojalność klienta jest skorelowana z jego szkodowością.

3.2. Metodyka

Na potrzeby badania jako miarę szkodowości klienta przyjęto częstość szkód, tj. oczekiwaną liczbę szkód powstałych w ciągu roku. Zagadnienie modelowania liczby szkód jest bardzo istotne w tzw. taryfikacji a priori, czyli ustalaniu składki na podstawie obserwowalnych cech ubezpieczonego i jego pojazdu [Ostasiewicz, red., 2000], szeroko stosowanego w ubezpieczeniach komunikacyjnych. W kontekście aktuarialnym do modelowania liczby szkód wykorzystuje się regresję Poissona (i jej odmiany), należącą do omówionej w punkcie 2.1 grupy uogólnionych modeli liniowych. Podstawowy model regresji Poissona otrzymuje się przyjmując logarytmiczną funkcję wiążącą ($g(\mu) = \ln(\mu)$) i zakładając, że składnik losowy ma rozkład Poissona. Dobrze wprowadzenie do zagadnienia stanowią prace Denuit'a i in. [2007] oraz Ohlssona i Johanssona [2010], natomiast opis wykorzystania modeli zmiennych licznikowych (tzn. zmiennych przyjmujących wartości będące nieujemną liczbą całkowitą, wyrażających liczbę wystąpień badanego zjawiska) poza analizami aktuarialnymi można znaleźć w pracy Camerona i Trivediego [2005].

W tej części analizy zmienną objaśnianą była więc liczba szkód obciążających umowę ubezpieczenia OC p.p.m., a zmiennymi objaśnianymi były czynniki opisane w punkcie 1.2, wyznaczone według stanu wiedzy na dzień zawarcia umowy ubezpieczenia.

Oprócz zmiennych objaśniających opisanych w punkcie 1.2, uwzględniono dodatkowo informację o lojalności klienta względem zakładu ubezpieczeń. Na potrzeby tej analizy lojalność klienta określono przez podzielenie klientów ze względu na długość okresu ubezpieczenia w danym zakładzie na następujące grupy (klasy lojalności):

1. Klienci nowi, którzy nie byli ubezpieczeni w poprzednim roku.
2. Klienci, którzy w poprzednim roku byli ubezpieczeni w innym zakładzie ubezpieczeń.
3. Klienci, którzy są ubezpieczeni w danym zakładzie od roku.
4. Klienci, którzy są ubezpieczeni w danym zakładzie od dwóch lat.
5. Klienci, którzy są ubezpieczeni w danym zakładzie od trzech lat.
6. Klienci, którzy są ubezpieczeni w danym zakładzie od czterech lat.

W przypadku obiektów zaobserwowanych po raz pierwszy w 2009 r., ze względu na brak informacji o poprzednich ubezpieczeniach, najwyższą możliwą klasę należy interpretować jako okres ubezpieczenia równy co najmniej rok, co najmniej dwa lata itp. Przykładowo, dla umów zawartych w 2011 r. można wyróżnić klasę „ubezpieczony od roku” oraz „ubezpieczony od co najmniej dwóch lat”.

Selekcja zmiennych objaśniających była prowadzona w sposób analogiczny do opisanego w punkcie 2.1. W pierwszym etapie głównym kryterium była statystyczna istotność poszczególnych czynników, natomiast w drugim uwzględniono również wartość kryteriów informacyjnych oraz zdolności predykcyjne modelu.

3.3. Wyniki analizy

Punktem wyjścia w procesie budowy modelu, zgodnie z przyjętą praktyką, była opisana w poprzednim punkcie regresja Poissona z logarytmiczną funkcją wiążącą. Okazało się jednak, że ze względu na występowanie w danych zjawiskach nadmiernego rozproszenia (*overdispersion*) konieczne było zastosowanie regresji ujemnej dwumianowej, będącej uogólnieniem regresji poissonowskiej [Dionne i Vanasse, 1989]. W tej części analizy uwzględniono umowy zawarte w latach 2011-2013.

Tak jak w przypadku modelowania lojalności, wszystkie uwzględnione czynniki okazały się statystycznie istotne na poziomie istotności 0,01. Poniżej przedstawiono interpretację uzyskanych wyników przy założeniu *ceteris paribus*. Należy wskazać, że cechują się one dużą stabilnością w czasie. Podstawowe wnioski płynące z analizy zostały przedstawione poniżej.

- Oczekiwana liczba szkód maleje wraz z wiekiem, przy czym zależność ta nie jest monotoniczna. Największa szkodowość występuje w grupie wiekowej 0-25, natomiast w pozostałych grupach jest ona niższa o ok. 25%.
- Oczekiwana liczba szkód dla kobiet jest o 4% wyższa niż dla mężczyzn.
- Największa szkodowość występuje w województwie dolnośląskim, najmniejsza – w województwie opolskim.
- Samochody ciężarowe cechują się częstością szkód o ok. 25%-30% wyższą niż samochody osobowe. Jednoślady i pozostałe pojazdy cechują się niższą częstością szkód.
- Ubezpieczeni posiadający więcej niż jeden pojazd powodują ok. 2,5% mniej szkód. Umowy zawarte z wieloma osobami cechują się z kolei częstością szkód o 9% niższą niż umowy z jednym ubezpieczonym.
- Fakt posiadania AC jest związany z częstością szkód wyższą o 14%-16% w porównaniu do sytuacji, gdy ubezpieczony nie posiada AC. Może to być

związane zarówno z wiekiem pojazdu (umowy autocasco są zawierane głównie na nowsze pojazdy), jak i z mniej ostrożną jazdą kierowcy, wynikającą z faktu posiadania ubezpieczenia AC.

- Częstość szkód maleje wraz z klasą lojalności klienta. Największą częstością szkód charakteryzują się nowi ubezpieczeni (pierwsza umowa w dostępnej historii ubezpieczenia). Poniżej podano różnicę w częstości szkód między klientami z określoną długością historii ubezpieczenia w tym samym zakładzie a klientami nowymi (na podstawie modelu dla umów zawartych w 2013 r.):
 - rok – spadek o 13%-18%,
 - dwa lata – spadek o 22%-30%,
 - trzy lata – spadek o 27%-31%,
 - co najmniej cztery lata – spadek o 39%.
- Klient, który w poprzednim okresie był ubezpieczony w innym zakładzie ubezpieczeń, cechuje się częstością szkód porównywalną do klienta, który jest w zakładzie ubezpieczeń od roku. Oznacza to, że związek między lojalnością a oczekiwaną liczbą szkód ujawnia się dopiero po co najmniej dwóch latach ubezpieczenia w tym samym zakładzie ubezpieczeń.

Biorąc pod uwagę wyniki wcześniejszych analiz, można stwierdzić, że o ile klient zmieniający zakład ubezpieczeń cechuje się większą skłonnością do ponownej zmiany zakładu, o tyle jego szkodowość jest porównywalna do szkodowości klienta lojalnego. Nie zmienia to jednak faktu, że oczekiwana częstość szkód maleje wraz z długością okresu ubezpieczenia w danym zakładzie. Można wskazać kilka potencjalnych wyjaśnień tej obserwacji. Po pierwsze, długość okresu ubezpieczenia kierowcy w zakładzie może pośrednio odzwierciedlać nabywanie przez niego doświadczenia. Po drugie, lojalność klienta może wiązać się z jego ogólną skłonnością do ryzyka w życiu codziennym, a ta z kolei może przekładać się na jego zachowanie na drodze. Po trzecie, może to być rezultatem polityki zakładu ubezpieczeń, który dąży do utrzymania mniej szkodowych klientów. Kwestia ta może być przedmiotem dalszych badań.

Podsumowanie

W niniejszym artykule omówiono zagadnienie statystycznego modelowania lojalności klienta w ubezpieczeniach komunikacyjnych. Na podstawie zbioru danych, pochodzącego z bazy danych Ubezpieczeniowego Funduszu Gwarancyjnego, przeprowadzono opisową analizę trendów zachodzących na rynku ubezpieczeń komunikacyjnych oraz zbudowano wiele modeli regresji logistycznej opisujących prawdopodobieństwo zmiany zakładu ubezpieczeń. Wyniki prze-

przebiegu badania wskazują, że możliwa jest identyfikacja cech opisujących klienta (oraz jego pojazd), które pozwalają na skuteczniejsze prognozowanie jego decyzji dotyczącej wyboru ubezpieczyciela OC p.p.m.

W drugiej części analizy zweryfikowano, czy informacja o lojalności klienta może być użyteczna z punktu widzenia oceny ryzyka ubezpieczeniowego. W tym celu zbudowano model regresji ujemnej dwumianowej objaśniający liczbę szkód z tytułu umowy ubezpieczenia. Wyniki wskazują, że nawet po uwzględnieniu innych zmiennych informacja o lojalności klienta jest statystycznie istotna i poprawia dopasowanie modelu do danych. Oczekiwana liczba szkód maleje wraz z długością okresu ubezpieczenia klienta w tym samym zakładzie, przy czym różnica między klientem nowym a klientem ubezpieczonym w tym samym zakładzie od co najmniej czterech lat może sięgać nawet 40%. Wynika stąd, że informacja o lojalności klienta może zostać wykorzystana jako dodatkowa zmienna taryfowa w procesie taryfikacji.

Omawianą analizę można rozszerzyć na kilka sposobów. Po pierwsze, w badaniu nie mogło zostać uwzględnionych wiele czynników dotyczących sytuacji życiowej ubezpieczonego oraz jego preferencji. Informacje te nie są gromadzone przez UFG, a mogą być bardzo ważne w tego rodzaju analizie. Po drugie, istnieje możliwość uwzględnienia dodatkowych czynników na wysokim poziomie agregacji, np. opisujących dany zakład ubezpieczeń lub jednostkę terytorialną. Po trzecie, może zostać rozszerzony zestaw stosowanych metod, np. o modele oparte na łańcuchach Markowa. Zagadnienia te są przedmiotem dalszych badań.

Na zakończenie należy podkreślić, że – ze względu na sposób przygotowania danych – uzyskane wyniki opisują wybraną grupę ubezpieczonych, nie pozwalają natomiast wyciągać jednoznacznych wniosków dotyczących całego rynku. Warto jednak zwrócić uwagę, że w pewnych wymiarach (np. udziały poszczególnych zakładów ubezpieczeń) badana próba dobrze odzwierciedla rynek ubezpieczeń komunikacyjnych. Można się więc spodziewać, że wraz ze wzrostem jakości danych w bazie OI UFG poziom wiarygodności wyników analiz prowadzonych na ich podstawie będzie systematycznie się podnosić.

Literatura

- Bennet R., Bove L. (2002), *Identifying the Key Issues for Measuring Loyalty*, „Australasian Journal of Market Research”, Vol. 9(2).
- Bowen J.T., Shoemaker S. (1998), *Loyalty: A Strategic Commitment*, „Cornell, Hotel and Restaurant Administration Quarterly”, Vol. 39(1).

- Brown G.H. (1952), *Brand Loyalty – Fact or Fiction?*, „Advertising Age”, Vol. 23.
- Cameron A.C., Trivedi P. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.
- Denuit M., Maréchal X., Pitrebois S., Walhin J. (2007), *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*, Wiley, New York.
- Dionne G., Vanasse Ch. (1989), *A Generalization of Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component*, „ASTIN Bulletin”, Vol. 19(2).
- Dyrektywa Parlamentu Europejskiego i Rady 2009/103/WE z dnia 16 września 2009 r. w sprawie ubezpieczenia od odpowiedzialności cywilnej za szkody powstałe w związku z ruchem pojazdów mechanicznych i egzekwowania obowiązku ubezpieczenia od takiej odpowiedzialności. Dz. Urz. UE L 263 z 7.10.2009.
- Fox J. (2008), *Applied Regression Analysis and Generalized Linear Models*, Sage, Los Angeles.
- Guest L.P. (1944), *A Study of Brand Loyalty*, „Journal of Applied Psychology”, Vol. 28.
- Guillen M., Parner J., Densgoe Ch., Perez-Martin A.M. (2002), *Customer Loyalty in the Insurance Industry: A Logistic Regression Approach*, II Conference in Actuarial Science and Finance on Samos, Karlovasi-Samos, Greece.
- Jacoby J., Chestnut R. (1978), *Brand Loyalty: Measurement and Management*, Wiley, New York.
- Mowen J.C. (1995), *Customer Behaviour*, Prentice Hall, Englewood Cliffs.
- Ohlsson E., Johansson B. (2010), *Non-Life Insurance Pricing with Generalized Linear Models*, Springer-Verlag, Berlin, Heidelberg.
- Ostasiewicz W., red. (2004), *Składki i ryzyko ubezpieczeniowe. Modelowanie stochastyczne*, Wydawnictwo AE, Wrocław.
- Rodzinka J. (2010), *Analiza cech determinujących lojalność klientów*, „Wiadomości Ubezpieczeniowe”, nr 4.
- Schneider K.C., Bristow D.N. (2007), *Correlates of Brand Loyalty in the Service Sector: The Case of a “New” Team*, „Marketing Management Journal”, Vol. 17(2).
- Uncles M., Laurent G. (1997), *Editorial*, „International Journal of Research in Marketing”, Vol. 14(5).
- Ustawa z dnia 22 maja 2003 r. o ubezpieczeniach obowiązkowych, Ubezpieczeniowym Funduszu Gwarancyjnym i Polskim Biurze Ubezpieczycieli Komunikacyjnych. Dz.U. 2013, poz. 392 tj.
- Wicka A., Miedzik A. (2010), *Rodzaje ubezpieczeń i czynniki decydujące o wyborze ubezpieczyciela*, „Zeszyty Naukowe SGGW w Warszawie. Polityki Europejskie, Finanse i Marketing”, nr 4(53).

**CUSTOMER LOYALTY IN AUTOMOBILE INSURANCE
– AN EMPIRICAL STUDY**

Summary: This paper addresses the issue of statistical modeling of customer loyalty in motor third party liability (MTPL) insurance. The paper presents the results of the research study conducted on the basis of the data from the Polish Insurance Guarantee Fund. The main aim of the study was to identify key factors influencing the customer's decision on switching to another MTPL insurer. Furthermore, paper discusses market trends visible in the data. Finally, the paper describes how customer loyalty may be beneficial to the insurer and in what way it affects the expected number of claims. The goals of the study were achieved with the use of generalized linear models (GLM) and analysis of correlation.

Keywords: customer loyalty, motor third party liability insurance, generalized linear models.