UDK:368.811:332.88:368.025.3:347.8:368.52:608.4:517.358:368.025.6
DOI: 10.5937/TokOsig2102039P

Branko R. Pavlović[1]

PARCIJALNI INTERNI MODEL U SOLVENTNOSTI II ZA RIZIK OD PREKIDA UGOVORA ŽIVOTNOG OSIGURANJA

ORIGINALNI NAUČNI RAD

Apstrakt

Neizvesnost u pogledu realizacije očekivane stope prekida ugovora životnog osiguranja utiče na rizik od nepreciznog određivanja kapitalnog zahteva za solventnost, iznos minimalnog kapitalnog zahteva i performanse poslovanja osiguravajuće kompanije. Zbog toga precizno projektovanje rizika od prekida ugovora životnog osiguranja ima velik značaj. Brojni faktori utiču na stopu prekida.

Kapitalni zahtev za solventnošću rizika od prekida ugovora životnog osiguranja u režimu Solventnosti II može se odrediti primenom standardne formule ili parcijalnog internog modela. Na primeru podataka sa tržišta osiguranja Srbije, korišćenjem softverskog paketa R, u radu će biti detaljno prikazan izbor faktora za modeliranje zavisnosti stope prekida, postupak formiranja GLM modela prekida ugovora i provera ispunjenosti pretpostavki modela.

Razvijeni parcijalni interni model može biti primenjen za određivanje očekivane stope prekida ugovora osiguravajuće kompanije koja posluje na domaćem tržištu.

Ključne reči: *parcijalni interni model, rizik od prekida ugovora*

I. Uvod

Rizik od prekida je rizik od gubitka ili povećanja obaveza osiguravača koji nastaje zbog odstupanja od očekivanog korišćenja opcija prekida ugovora o životnom

[1] Predsednik Udruženja aktuara Srbije, i-mejl: pavlovic.branko@gmail.com
Rad je primljen: 28.03.2021.
Rad je prihvaćen: 01.04.2021.

osiguranju od strane ugovarača. Ugovor može biti: delimično ili potpuno prekinut; prekinut uz isplatu otkupne vrednosti ili bez nje; prekinut u pogledu obaveza ugovarača da uplaćuje premiju, ali sa zadržavanjem obaveze osiguravača da isplati redukovanu vrednost osigurane sume ili bez kapitalizovane osigurane sume; prekinut sa mogućnošću obnove ugovora u nekom periodu uz plaćanje premija koje bi do obnove dospele ili bez mogućnosti obnove.

1. Kapitalni zahtev za solventnost

Obračun kapitalnog zahteva za solventnost, u režimu Solventnosti II, može se vršiti na sledeće načine: primenom standardne formule, korišćenjem potpunog internog modela i parcijalnog internog modela.

Standardna formula namenjena je prosečnoj evropskoj osiguravajućoj kompaniji i sadrži veliki broj aproksimacija, kao i predloženih dodatnih uprošćavanja za pojedine rizike koji nisu ključni za konkretnu osiguravajuću kompaniju. U svakom modulu i podmodulu rizika, osiguravajuća kompanija može da zameni standardnu formulu sopstvenom metodologijom i da kreira model za određivanje zahtevanog kapitala. Ako se zamena izvrši za sve module i podmodule rizika, onda se takvi interni modeli nazivaju potpuni interni modeli. Ukoliko se zamena izvrši samo za nekoliko modula ili podmodula rizika, dok se preostali rizici obračunavaju standardnom formulom, onda se takvi interni modeli nazivaju parcijalni interni modeli.

Parcijalni interni modeli uvode se da bi se omogućilo preciznije ocenjivanje očekivane stope prekida od onog prikazanog u standardnoj formuli, a samim tim i adekvatnija procena rizika od prekida i odgovarajućeg kapitalnog zahteva.

2. Merenje rizika od prekida u standardnoj formuli Solventnosti II

Kapitalni zahtev za rizik od prekida jednak je najvećem iznosu sledećih kapitalnih zahteva za tri podrizika: za trajno povećanje stopa prekida, trajno smanjenje stopa prekida i za podrizik od masovnog prekida.

$$SCR_{lapse} = \max\,(Lapse_{up}\,;\,Lapse_{down}\,;\,Lapse_{mass}$$

Kapitalni zahtev za svaki od pomenuta tri podrizika računa se kao gubitak osnovnih sopstvenih sredstava osiguravajuće kompanije koji bi nastao zbog trajnog povećanja ili smanjenja stopa iskorišćenja opcija prekida ili trenutnog masovnog prekida ugovora:

$$Lapse_i = \Delta NAV \,|\, lapseshock_i$$

Pomenute tri opcije od kojih se bira ona s najvećim uticajem objašnjene su na primeru u Tabeli 1.

Tabela 1. Primer kretanja stope prekida u skladu sa zahtevima standardne formule

Godina	1	2	3	4	5+
Očekivana stopa prekida	4%	4%	4%	4%	4%
Stalno povećanje prekida 50%	6%	6%	6%	6%	6%
Stalno smanjenje prekida 50%	2%	2%	2%	2%	2%
Masovni prekid 40% svih polisa	40%	4%	4%	4%	4%

Izvor: Obračun autora

3. Merenje rizika od prekida u parcijalnom internom modelu Solventnosti II

U skladu sa pristupom baziranom na principima u režimu Solventnosti II, umesto na pravilima kako je bilo ranije, EIOPA nije propisala formalnu definiciju internog modela, kao ni šta interni model treba da obuhvata. Najznačajnija razlika između internih modela i standardne formule jeste u većem korišćenju stohastičkih tehnika nad sopstvenim podacima u internim modelima.

Primena parcijalnih ili potpunih internih modela za merenje rizika ima za cilj da podstakne osiguravajuće kompanije da preciznije procenjuju i kontrolišu sopstvene rizike. Svrha uvođenja internih modela ne treba da bude smanjenje kapitalnog zahteva za solventnost, nego bolje upravljanje sopstvenim rizicima. Korišćenje internih modela doprinosi adekvatnijem modeliranju sopstvenih rizika osiguravača, što vodi povećanju osetljivosti dobijenih rezultata za kapitalni zahtev za solventnost na rizik specifičan za odgovarajuću osiguravajuću kompaniju. Za uspešno korišćenje internih modela u kompaniji, veoma je važna kontinuirana periodična validacija internih modela.

Mana internih modela je velika složenost, tako da njihovo kreiranje i uvođenje zahteva značajno angažovanje kompanijskih resursa.

Predmet rada je analiza finansijskog uticaja rizika od prekida ugovora životnog osiguranja na solventnost osiguravajućih kompanija za životno osiguranje. Cilj rada je formulisanje parcijalnog internog modela za određivanje zahtevanog kapitala za pokriće rizika od prekida ugovora životnih osiguranja, uz uvažavanje specifičnosti ispoljavanja rizika kako u pojedinačnoj osiguravajućoj kompaniji tako i na tržištu životnog osiguranja u Srbiji.

Istraživanje u ovom radu biće prva empirijska studija stope prekida na tržištu osiguranja u Srbiji. S obzirom na to da je na kraju 2019. godine bilo aktivno oko 900 hiljada polisa životnog osiguranja na domaćem tržištu,[2] i da nema sistematizovanih

[2] *www.nbs.rs*

podatka o prekidima polisa, analiziranje preko 200.000 polisa daće vrlo koristan rezultat. Naročito je zanimljivo ponašanje ugovarača u pogledu prevremenog prekida ugovora u vreme neočekivanih događaja, što će moći da bude analizirano u ovom istraživanju pošto će biti obuhvaćena i 2008. godina, kada je počela svetska ekonomska kriza.

II. Pregled literature o faktorima stope prekida

Brojni autori u svetu proučavali su uticaj različitih faktora na stopu prekida. Prva grupa autora bavila se uticajem karakteristika okruženja, kao što su npr. makroekonomski pokazatelji. U fokusu istraživanja Dara i Doddsa,[3] Outrevillea,[4] grupe autora[5] sa Kouom na čelu i grupe autora sa Russelom na čelu,[6] bili su referentna kamatna stopa, bruto domaći proizvod po stanovniku i stopa nezaposlenosti, dok su se Cox i Lin,[7] Kim,[8] Kiesenbauer[9] bavili uticajem bruto domaćeg proizvoda, razvijenosti tržišta kapitala i veličine osiguravajuće kompanije na stopu prekida. Druga grupa autora istraživala je uticaj podataka iz ugovora o životnom osiguranju na stopu prekida, koristeći generalizovani linearni model. Kagraoka,[10] grupa autora[11] sa Cerchiarom na čelu, grupa autora[12] sa Milhaudom, Eling i Kiesenbauer[13] na čelu proučili su uticaj godine sklapanja ugovora, starosti ugovarača, načina plaćanja premije, kanala prodaje i postojanja dodatnih pokrića. Na kraju, grupa autora[14] sa

[3] Dar, A. & Dodds, C. (1989). Interest Rates, the Emergency Fund Hypothesis and Saving Through Endowment Policies: Some Empirical Evidence for the U.K. *Journal of Risk and Insurance 56*(3), p. 415–433.

[4] Outreville, J. (1990). Whole-life Insurance Lapse Rates and the Emergency Fund Hypothesis. *Insurance: Mathematics and Economics 9*(4), p. 249–255.

[5] Kuo, W., Tsai, C. and Chen, W.-K. (2003). An Empirical Study on the Lapse Rate: The Cointegration Approach. *Journal of Risk and Insurance 70*(3), p. 489-508.

[6] Russel, D., Stephen, J. et al. (2013). An Empirical Analysis of Life Insurance Policy Surrender Activity. *Journal of Insurance Issues 36*(1), p. 35-57.

[7] Cox, S. & Lin, Y. (2006). *Annuity Lapse Rate Modeling: Tobit or Not Tobit?* Society of Actuaries. *http://library.soa.org.*

[8] Kim, C. (2005). Modeling Surrender and Lapse Rates with Economic Variables. *North American Actuarial Journal 9*(4), p. 56–70.

[9] Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal 16*(1), p. 52-73.

[10] Kagraoka, Y. (2005). *Modeling Insurance Surrenders by the Negative Binomial Model.* Working paper.

[11] Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* p. 100–122.

[12] Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2010). *Surrender Triggers in Life Insurance: Classification and Risk Predictions*. Working paper.

[13] Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), p. 241-269.

[14] Cheng, Ch., Hilpert, Ch. et al. (2020). Surender Contagion in Life Insurance. *SSRN Electronic Journal* March 2020.

Chengom na čelu, bavila se analizom rasta stope prekida u zavisnosti od ponašanja drugih učesnika na tržištu koji odlučuju da prekinu svoje ugovore.

Osim Kočović i Jovović,[15] domaći autori se nisu bavili problemom prekida ugovora u životnom osiguranju.

III. Projektovanje rizika od prekida ugovora

1. Značaj projektovanja rizika od prekida ugovora životnog osiguranja

Za osiguravajuću kompaniju, trenutak nastanka obaveze prema osiguraniku je neizvestan. Na primer, po ugovoru o mešovitom osiguranju za slučaj smrti i doživljenja, koji je zaključen na 20 godina, osiguravajuća kompanija već sledećeg dana može da isplati punu osiguranu sumu u slučaju smrti osiguranika, ili otkupnu vrednost bilo kada posle tri godine od početka osiguranja, ili ako ništa od prethodnog ne plati, osiguranu sumu po isteku ugovora koji je osiguranik doživeo.

Neizvesnost u pogledu stope prekida ugovora utiče ne samo na rizik od nepreciznog određivanja kapitalnog zahteva za solventnost i minimalnog kapitalnog zahteva, nego ima i nekoliko drugih važnih efekta na performanse poslovanja osiguravajuće kompanije.

Prvi efekat je vezan za plaćanje otkupne vrednosti. Kada se prekine polisa, osiguravajuća kompanija isplaćuje otkupnu vrednost ugovaraču osiguranja i ukida matematičku rezervu za tu polisu. Na domaćem tržištu, otkupna vrednost uvek je manja od matematičke rezerve ili joj je jednaka, tako da u trenutku prekida osiguravajuća kompanija ima prihod. U slučaju da se realizuje manja stopa prekida, osiguravajuća kompanija će imati manji prihod od očekivanog, što može da donese problem s profitabilnošću ukoliko je očekivana stopa prekida korišćena u formiranju cene usluge.

Drugi efekat je vezan za pokriće akvizicionih troškova. U životnom osiguranju, akvizicioni troškovi su prilično visoki. Akvizicioni troškovi koji nastaju u trenutku sklapanja ugovora obično se priznaju u finansijskim izveštajima kroz redukciju matematičke rezerve, cilmerovanjem. Ukoliko je stopa prekida veća od očekivane, a ne postoji mehanizam povraćaja provizija od agenata za raskinute polise (engl. *clawback*), može da dođe do nemogućnosti pokrića akvizicionih troškova, odnosno do gubitka za kompaniju.

Čak i ako se prosečna stopa prekida ponaša očekivano, može da se javi treći efekat koji utiče na profitabilnost osiguravača. Ukoliko se više nego što je očekivano prekidaju polise zdravijih osiguranika, a manje nego što je očekivano polise osiguranika slabijeg zdravlja, može doći do značajnijeg povećanja stope mortaliteta u portfelju, bez obzira na to što je smrtnost stanovništva nepromenjena.

[15] Kočović, J., Jovović, M. i Kočović, M. (2015). Aktuarski efekti prevremenog raskida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, p. 77–84.

Izostanak očekivanog budućeg profita od raskinutih polisa može biti važan četvrti efekat koji negativno utiče na profitabilnost osiguravajućih kompanija.

Peti efekat je ugrožavanje likvidnosti zbog eventualne neočekivane masovne isplate otkupne vrednosti po raskinutim polisama.

Formiranje matematičke rezerve, iz koje se isplaćuju obaveze prema osiguraniku, zahteva korišćenje različitih pretpostavki u vezi sa smrtnosti, tehničke kamatne stope itd. Navedene pretpostavke najčešće su zasnovane na iskustvu osiguravajuće kompanije, te na aktuarskim principima i propisima.[16] Dovoljnost matematičke rezerve za isplatu preuzetih obaveza direktno je povezana sa solventnošću kompanije. Zato je veoma važno da se ispravno prognoziraju budući novčani tokovi. Šesti efekat povećanja stope prekida jeste uticaj na novčane tokove, a samim tim i na solventnost kompanije.

Ročna neusklađenost imovine i obaveza može da se javi kao sedmi efekat povećane stope prekida ugovora. Budući da je obaveze po standardnim višegodišnjim štednim ugovorima životnog osiguranja potrebno pokriti imovinom sa odgovarajućom duracijom, prekidi takvih ugovora zahtevaju prilagođavanja imovine, što iziskuje određene troškove.

Osmi efekat se odnosi na reputacioni rizik po osiguravajuću kompaniju. Potencijalnim osiguranicima poznato je da se određeni broj ugovora prekida pre isteka, ali su često neobavešteni o svojim pravima prilikom prekida. Pojedini agenti daju zbunjujuće ili čak pogrešne usmene informacije da nema negativnih efekata zbog prekida ugovora i da će se osiguraniku vratiti do tada plaćena premija. Takva praksa u suprotnosti je sa članovima 82. i 83. Zakona o osiguranju,[17] i pored ugrožavanja reputacije osiguravajuće kompanije, može dovesti do kazni koje izriče organ nadzora delatnosti osiguranja.

Na kraju, poslednji negativan efekat povećanja stope prekida ugovora jeste onaj na vrednost portfelja osiguravajuće kompanije (engl. *Embedded Value*), zbog toga što dolazi do smanjenja novčanih tokova koji potiču od budućih premija.[18]

Zbog svih pomenutih efekata, važno je predvideti stopu prekida što preciznije moguće.

2. Faktori koji utiču na stopu prekida

Stopa prekida ugovora zavisi od mnogo racionalnih i iracionalnih razloga koji utiču na ponašanje ugovarača.[19] Primer racionalnog ponašanja jeste reakcija na kretanja na finansijskom tržištu i promenu makroekonomskih varijabli, kao što su

[16] Kočović, J., Jovović, M. i Kočović, M. (2015). Aktuarski efekti prevremenog raskida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, pp. 77-84.

[17] Zakon o osiguranju. *Službeni glasnik RS,* br. 139/14.

[18] Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), p. 241-269.

[19] Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* p. 100-122.

inflacija, valutni kurs, berzanski indeksi, kupovna moć građana, stopa nezaposlenosti i sl. Primer iracionalnog ponašanja može biti prekid ugovora o osiguranju u cilju kupovine automobila od akumuliranog novca.

Velik uticaj na stopu prekida imaju vrsta osiguranja i karakteristike usluge. Ugovarači lakše prekidaju višegodišnje ugovore o osiguranju života za slučaj smrti, jer ne primećuju da na taj način gube novac, dok npr. prekidom ugovora o osiguranju života za slučaj doživljenja gube deo novca, jer je otkupna vrednost po pravilu manja od matematičke rezerve. Sledeće karakteristike usluga osiguranja utiču na eventualnu odluku o prekidu:[20] trajanje ugovora, preostalo vreme do završetka ugovora, visina premije i osigurane sume, frekvencija plaćanja premija, faza rentnog ugovora (uplata premija ili isplata rente), nivo kaznenog umanjenja matematičke rezerve koja će biti isplaćena u slučaju otkupa, način pripisivanja negarantovane dobiti, prinosi fonda u bliskoj prošlosti kod *unit link* ugovora, struktura provizija agenata itd.

Starost i pol ugovarača, njegova lokacija prebivališta kao indikator nivoa prihoda, bračni status, a naročito njegova promena mogu takođe biti korišćeni u analizi stope prekida ugovora.

Pojedini ugovarači ugovor o životnom osiguranju smatraju vidom štednje za slučaj nepredviđenih okolnosti, tako da u slučaju npr. gubitka posla otkupnom vrednošću nadoknađuju nedostatak prihoda u tranzicionom periodu. Investitori koji ulažu novac u životno osiguranje s ciljem da uvećaju svoje bogatstvo, u slučaju rasta kamata na finansijskom tržištu, lako prekidaju ugovor i prelaze u isplativije investicije.

Nepredviđeni događaji, kao što je promena poreske politike ili promene u vlasništvu ili reputaciji osiguravajuće kompanije mogu da dovedu do značajne promene stope prekida.

Na odluku ugovarača o prekidu ugovora utiče često kombinacija više razloga. Pojedini, prethodno pomenuti faktori imaju gotovo savršenu korelaciju, kao npr. starost osiguranika i trajanje ugovora, koji se povećavaju istovremeno. Prilikom modeliranja prekida, ne mogu se koristiti oba takva faktora. Takođe, moguće je da pojedine promenljive zavise od vrednosti druge promenljive.

Svi pomenuti faktori su kandidati za eksplanatorne promenljive koje objašnjavaju ili predviđaju promene stope prekida ugovora životnog osiguranja, kao zavisne promenljive. U zavisnosti od podataka o polisama konkretne osiguravajuće kompanije, biće izabran skup eksplanatornih promenljivih.

Prilikom modeliranja zavisnosti stope prekida od više promenljivih, trebalo bi prvo proučiti podatke osiguravajuće kompanije, i to kroz analizu zavisnosti stope prekida od svake pojedinačne nezavisne promenljive dostupne u skupu podataka, i uz proučavanje korelacije između dostupnih nezavisnih promenljivih.

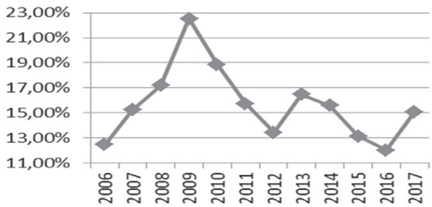
[20] Michorius, C. (2011). *Modeling Lapse Rates – Investigating the Variables that Drive Lapse Rates*. Master Thesis. Enschede: Faculty of Management and Governance, University of Twente.

IV. Podaci

U analizi su korišćeni podaci sa tržišta osiguranja Srbije o polisama koje su izdate i prekinute u periodu od 1. 1. 2006. do 31. 12. 2017. godine.

Slika 1. Stopa prekida

Godina	Stopa prekida	Godina	Stopa prekida
2006	12,46%	2012	13,43%
2007	15,25%	2013	16,48%
2008	17,20%	2014	15,60%
2009	22,53%	2015	13,14%
2010	18,86%	2016	12,04%
2011	15,74%	2017	15,09%



Izvor: Obračun autora

Obradom podataka u programskom paketu Microsoft Access dobijena je vremenska serija stopa prekida po godinama analiziranog perioda, kako je prikazano na Slici 1.

Stacionarnost vremenske serije može se testirati kombinovanjem Kwiatkowski, Phillips, Schmidt and Shin (KPSS) testa i Dickey-Fuller (DF) testa. Pravilo statističkog testiranja kaže da zaključak testa može biti odbacivanje nulte i prihvatanje alternativne hipoteze, ili pak da zaključka nema.

Vremenska serija se prvo testira pomoću KPSS testa. Nulta hipoteza KPSS testa tvrdi da jedinični koren ne postoji. Alternativna hipoteza KPSS testa tvrdi da postoji jedinični koren tj. da vremenska serija nije stacionarna. Nulta hipoteza o stacionarnosti vremenske serije odbacuje se za izabrani nivo značajnosti ako je realizovana vrednost test statistike veća od korespondirajuće kritične vrednosti.[21] Ako se nulta hipoteza KPSS testa odbaci, vremenska serija sadrži jedinični koren. Ako se nulta hipoteza testa KPSS ne može odbaciti, ne može se ni zaključiti da je vremenska serija stacionarna, te se nastavlja testiranje DF testom.

Nulta hipoteza DF testa tvrdi da postoji jedinični koren tj. da vremenska serija nije stacionarna. Alternativna hipoteza DF testa tvrdi da je vremenska serija stacionarna. Nulta hipoteza o postojanju jediničnog korena se odbacuje za dovoljno malu vrednost statistike, tj. kada je izračunata vrednost manja od kritične.[22] Ako DF test odbaci nultu hipotezu, zaključak je da u vremenskoj seriji ne postoji jedinični

[21] Mladenović, Z. (2020). *KPSS test jediničnog korena*, Beograd: Ekonomski fakultet Univerziteta u Beogradu, s. 9.

[22] Mladenović, Z. (2015). *Ekonometrijski metodi i modeli – Dickey-Fuller-ov test jediničnog korena*, Beograd: Ekonomski fakultet Univerziteta u Beogradu, s. 9.

koren, što znači da je vremenska serija stacionarna. Ako test DF ne odbaci nultu hipotezu, i dalje nema zaključka.

Stopa prekida prikazana je na Slici 1. S obzirom na to da vrednost stope prekida fluktuira kroz vremenski period od 11 godina, odnosno nema izraženu tendenciju, pri izboru vrste regresije pretpostavljeno je da vremenska serija ne sadrži komponentu vremenske tendencije.

Tabela 2. Zvanični statistički podaci

God.	Rast BDP	Rast cena	Rast pros. neto zarade u EUR	Stopa nezaposlenosti[23]	Rast indeksa Belex-Line[24]	Ref. kamatna stopa NBS[25]	Rast premije neživotnog osig.	Rast premije životnog osig.	Rast ukupne premije osig.
2006	4,9%	6,6%	23%	20,9%	36,0%	14,00%	7,5%	19,6%	8,6%
2007	6,4%	11,0%	35%	18,1%	44,1%	10,00%	22,4%	28,9%	23,1%
2008	5,7%	8,6%	16%	13,6%	-68,7%	17,75%	13,0%	26,2%	14,4%
2009	-2,7%	6,6%	-16%	16,1%	9,5%	9,50%	-13,5%	8,7%	-10,8%
2010	0,7%	10,3%	-2%	19,2%	-2,2%	11,50%	-5,8%	8,5%	-3,7%
2011	2,0%	7,0%	12%	23,0%	-23,8%	9,75%	1,3%	7,5%	2,3%
2012	-0,7%	12,2%	-2%	23,9%	2,9%	11,25%	-5,5%	7,2%	-3,3%
2013	2,9%	2,2%	6%	22,1%	9,9%	9,50%	0,8%	17,6%	4,1%
2014	-1,6%	1,7%	-2%	19,2%	21,7%	8,00%	3,0%	10,2%	4,6%
2015	1,8%	1,5%	-3%	17,7%	2,7%	4,50%	11,9%	17,8%	13,3%
2016	3,3%	1,6%	2%	15,3%	13,7%	4,00%	5,2%	16,8%	8,0%
2017	2,1%	3,0%	3%	13,5%	5,9%	3,50%	9,1%	0,4%	6,8%

Izvor: Ministarstvo finansija Republike Srbije,[26] Republički zavodi za statistiku,[27] Narodna banka Srbije[28] i Beogradska berza a. d.[29]

Sprovedeno je testiranje stacionarnosti vremenske serije zavisne promenljive stopa prekida ugovora, pri čemu su korišćene kritične vrednosti za petoprocentne intervale pouzdanosti (KPSS: 0,463, DF: -2,862). Testiranje je izvršeno u jeziku R komandama: kpss.test() i adf.test().

Dobijen je sledeći rezultat testiranja stacionarnosti vremenske serije KPSS testom: KPSS Level = 0.40713, što je manje od kritične vrednost 0,463, što nadalje znači da KPSS test nije odbacio nultu hipotezu o stacionarnosti. Dobijen je sledeći

23 Stopa nezaposlenosti je usklađena sa metodologijom ILO od 2004. godine
24 Berzanski indeks BELEXfm, koji je kasnije transformisan u BELEXLine, formiran je u decembru 2004. godine
25 Referentnu stopu Narodna banka Srbije objavljuje od 2006. godine
26 *https://www.mfin.gov.rs*
27 *https://www.stat.gov.rs*
28 *https://www.nbs.rs*
29 *https://www.belex.rs*

rezultat testiranja stacionarnosti vremenske serije ADF testom: Dickey-Fuller = -3.2279. Ta vrednost je manja od kritične vrednost -2,862, što znači da je prošireni DF test odbacilo nultu hipotezu o postojanju jediničnog korena. Zaključak kombinovanog testiranja KPSS i DF testovima jeste da je vremenska serija stacionarna.

Pored podataka sa tržišta osiguranja Srbije, u analizi su korišćeni i zvanični statistički podaci iz Republičkog zavoda za statistiku, Ministarstva finansija Republike Srbije, Narodne banke Srbije i Beogradske berze: BDP, referentna kamatna stopa, prosečne zarade, berzanski indeks BelexLine, inflacija, stopa nezaposlenosti, stope rasta premije životnog i neživotnog osiguranja itd, kao što je prikazano u Tabeli 2. Analiza zavisnosti stope prekida od parametara iz okruženja izvršena je u periodu 2006–2017. godine.

Podaci su obrađeni u programskom paketu R, koji sadrži sav potrebni alat za prediktivnu analizu u programu Microsoft Excel.

V. Izbor faktora za modeliranje zavisnosti stope prekida

U Tabeli 3. prikazani su maksimalni koeficijenti korelacije između devet prediktora iz okruženja i zavisne promenljive stope prekida izračunate u jeziku R, korišćenjem funkcije za korelaciju (engl. *Cross Correlation Function*) *ccf()*.

Tabela 3. Maksimalne korelacije između prediktora iz okruženja i stope prekida

R. br.	Analizirani parametar	Maksimalni koeficijent korelacije	Kašnjenje
1.	Rast BDP	0,6524230	-2
2.	Rast cena	-0,5494015	6
3.	Rast prosečne neto zarade u EUR	0,7072750	-2
4.	Stopa nezaposlenosti	0,6741854	2
5.	Rast berzanskog indeksa BELEXline	-0,4468971	-1
6.	Referentna kamatna stopa NBS	0,6088886	-1
7.	Rast premije neživotnog osiguranja	0,5922836	-2
8.	Rast premije životnog osiguranja	0,5805626	-2
9.	Rast premije osiguranja	0,5887550	-2

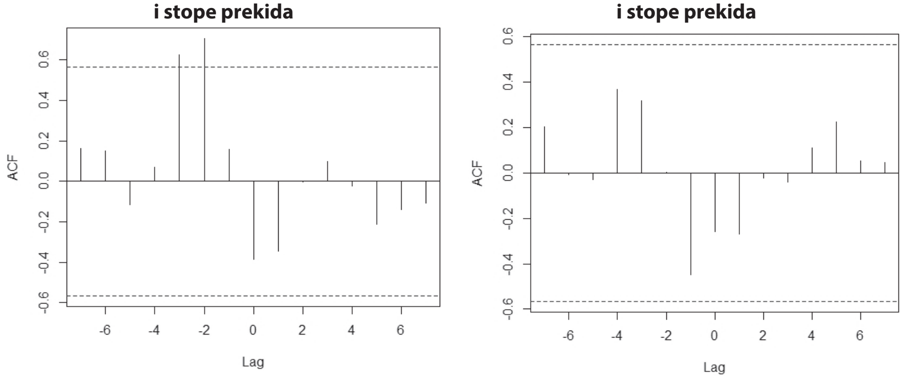
Izvor: Obračun autora

Iz Tabele 3. može se zaključiti da zavisna promenljiva stopa prekida ima pojedinačno najjaču vezu, tj. pozitivnu korelaciju s nezavisnim promenljivim rast prosečne zarade (koeficijent korelacije 0,71), stopa nezaposlenosti (0,67), rast BDP (0,65) i referentna kamatna stopa NBS (0,61), što znači da su navedene četiri nezavisne promenljive kandidati za korišćenje u modeliranju stope prekida.

Na Slici 2. grafički su prikazane korelacije za najveću pojedinačnu korelaciju rasta prosečne zarade i stope prekida i najmanju korelaciju (po apsolutnoj vrednosti) stope rasta indeksa BelexLine i stope prekida.

Slika 2. Korelacije prediktora iz okruženja sa stopom prekidā koji imaju najveći i najmanji maksimalni koeficijent korelacije

a) korelacije prosečne neto zarade i stope prekida **b) korelacije stope rasta indeksa BelexLine i stope prekida**



Izvor: Obračun autora

Istom komandom u jeziku R, *ccf()*, dobijaju se međusobne korelacije svih nezavisnih promenljivih. Rezultat je prikazan u Tabeli 4. Za kvalitetan model treba izabrati dve od četiri nezavisne promenljive koje imaju najjače veze sa stopom prekida, ali su međusobno najslabije povezane. Iz Tabele 4. se vidi da su najbolji kandidati za to nezavisne promenljive rast BDP i referentna stopa NBS, čiji koeficijent korelacije je najmanji i iznosi 0,3777358 uz međusobno kašnjenje parametra od dve vremenske jedinice. Korelacije nezavisnih promenljivih rast BDP i referentna stopa NBS su prikazane na Slici 3.

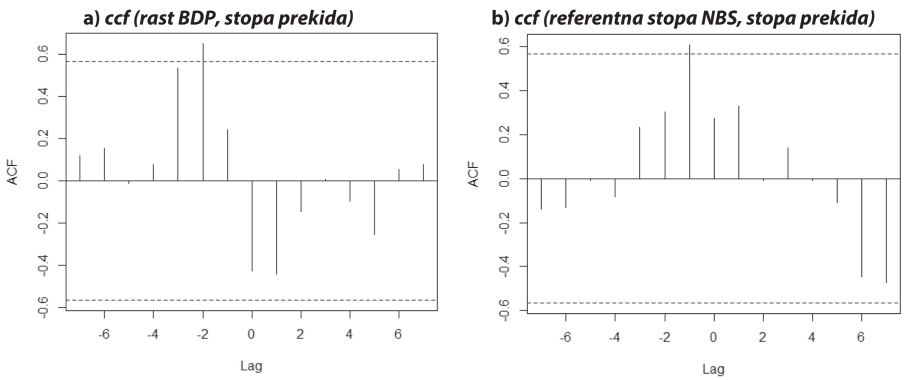
Tabela 4. Maksimalni koeficijenti korelacije između prediktora (u zagradi su data kašnjenja prediktora po kolonama u odnosu na prediktore po redovima)

	Rast BDP	Rast cena	Rast pros. neto zarade u EUR	Stopa ne-zaposlenosti	Rast indeksa Belex-Line	Ref. ka-matna stopa NBS	Rast premije neživot-nog osig.	Rast premije život-nog osig.	Rast ukupne premije osig.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	1	0,59 (4)	0,88 (0)	0,56 (5)	0,75 (-1)	0,38 (2)	0,82 (0)	0,74 (0)	0,83 (0)

	Rast BDP	Rast cena	Rast pros. neto zarade u EUR	Stopa ne-zaposlenosti	Rast indeksa Belex-Line	Ref. ka-matna stopa NBS	Rast premije neživot-nog osig.	Rast premije život-nog osig.	Rast ukupne premije osig.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(2)	0,59 (-4)	1	0,49 (-4)	-0,55 (-3)	0,55 (-5)	0,69 (1)	0,62 (-4)	0,56 (-4)	0,62 (-4)
(3)	0,88 (0)	0,49 (4)	1	0,53 (5)	0,66 (-1)	0,52 (1)	0,77 (0)	0,67 (0)	0,78 (0)
(4)	0,56 (-5)	-0,55 (3)	0,53 (-5)	1	-0,44 (-3)	-0,49 (4)	-0,67 (-1)	-0,66 (-1)	-0,71 (-1)
(5)	0,75 (1)	0,55 (5)	0,66 (1)	-0,44 (3)	1	-0,41 (0)	0,85 (1)	0,54 (1)	0,85 (1)
(6)	0,38 (-2)	0,69 (-1)	0,52 (-1)	-0,49 (-4)	-0,41 (0)	1	0,41 (-5)	0,42 (0)	0,42 (-5)
(7)	0,82 (0)	0,62 (4)	0,77 (0)	-0,67 (1)	0,85 (-1)	0,41 (5)	1	0,65 (0)	0,99 (0)
(8)	0,74 (0)	0,56 (4)	0,67 (0)	-0,66 (1)	0,54 (-1)	0,42 (0)	0,65 (0)	1	0,74 (0)
(9)	0,83 (0)	0,62 (4)	0,78 (0)	-0,71 (1)	0,85 (-1)	0,42 (5)	0,99 (0)	0,74 (0)	1

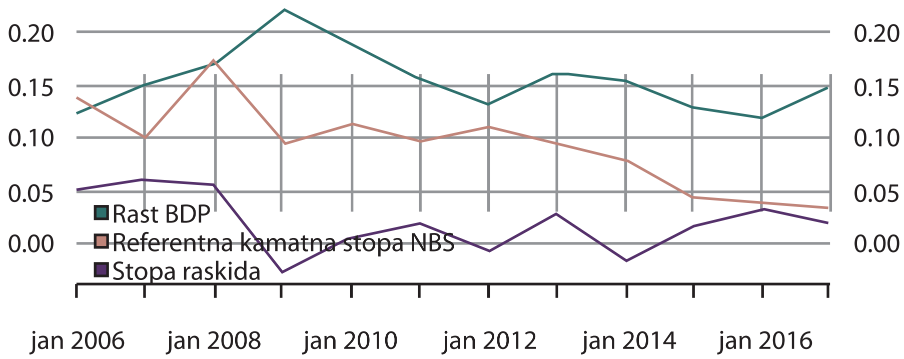
Izvor: Obračun autora

Slika 3. Maksimalne korelacije prediktora iz okruženja koji su izabrani za modeliranje stope prekida



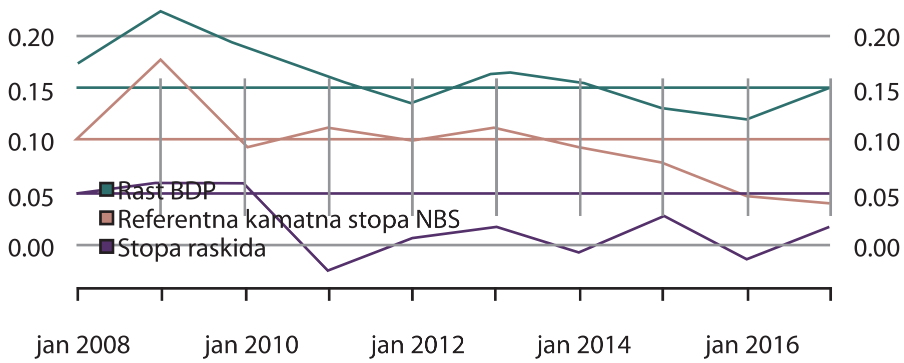
Izvor: Obračun autora

Slika 4. Vremenske serije zavisne i dve nezavisne promenljive bez kašnjenja



Izvor: Obračun autora

Slika 5. Vremenske serije zavisne i dve nezavisne promenljive sa odgovarajućim kašnjenjem nezavisnih promenljivih



Izvor: Obračun autora

Kašnjenje zavisnosti vremenskih serija je važno jer se bez uvođenja kašnjenja ne mogu modelirati odgovarajuće zavisnosti, što je očigledno na Slici 4, gde su prikazane zavisna i dve nezavisne promenljive bez kašnjenja.

Uvođenjem kašnjenja od dve godine za rast BDP i od jedne godine za referentnu stopu NBS, povezanost zavisne i nezavisnih promenljivih postaje vidljiva na Slici 5.

VI. Model

Postoji nekoliko vrsta alata koji mogu da posluže za donošenje odluke o odgovarajućem modelu. Mnogi softverski paketi automatski nude prihvatljive modele, zahvaljujući iterativnom prilagođavanju (fitovanju) modela raspoloživim podacima. Taj način je svakako dobra osnova, ali je dobro proveriti i fino podesiti predložene modele. Jedan od načina je test p-vrednosti kojim se proverava značajnost svake promenljive, tako što se računa značajnost celog modela i značajnost modela bez pojedinačne nezavisne promenljive. Drugi način je da se model oceni na osnovu slučajno izabranih 70% uzorka, a zatim da se proveri na preostalih 30% uzorka i refituje. Treći način je korišćenje nekog od kriterijuma, AIC (engl. *Akaike Information Criteria*) ili BIC (engl. *Bayesian Information Criteria*).

1. Generalizovani linearni model (GLM)

Jedan od poznatijih višefaktorskih regresionih modela, koji se sreće u literaturi koja analizira stopu prekida ugovora u životnom osiguranju, jeste generalizovani linearni model (GLM od engl. *Generalized Linear Model*). Model su formulisali Nelder i Wedderburn 1972. godine.[30] Izabran je zbog toga što se relativno jednostavno može razumeti i vrlo je fleksibilan u pogledu izbora raspodele verovatnoća zavisne promenljive i ulaznih promenljivih. Zahvaljujući link funkciji, može da se koristi za rad sa zavisnim i nezavisnim promenljivim koje mogu biti kontinualne ili binarne, što je naročito bitno kod prekida ugovora, jer promenljiva prekida ima binarnu vrednost na nivou jedne polise (prekinuta ili ne), a kontinualnu vrednost između 0 i 1 za ceo portfelj.

Sedamdesetih godina XX veka razvijen je softver specijalne namene za rad sa GLM modelima pod nazivom GLIM, od engl. *Generalized Linear Interactive Modelling*. Danas su GLM modeli primenjeni u različitim softverskim paketima, kao što je SAS ili SPSS, ali je najpopularnije GLM modeliranje u jeziku R, gde je primenjeno na sledeći način:

glm(formula, family = binomial(link=logit) data, weights, subset, na.action,…)

2. Formiranje modela

Izabran je generalizovani linearni model s normalnom raspodelom i funkcijom identiteta za link funkciju. Na osnovu ulaznih podataka i ispunjenosti navedenih pretpostavki, pozivanjem funkcije *glm()* softverskog alata R dobijen je rezultat, koji je prikazan u Tabeli 5.

[30] Nelder, J. & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society A 135*, p. 370–384.

Tabela 5. Rezultat funkcije *glm()* u jeziku R

```
Deviance Residuals:
    Min        1Q    Median       3Q       Max
-0.025582  -0.003458   0.003365   0.008776   0.017722

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)     0.10624    0.01437   7.394  0.00015 ***
rast_bdp.lag2    0.44828    0.18940   2.367  0.04983 *
ref_ks_NBS.lag1  0.47182    0.15264   3.091  0.01754 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for gaussian family taken to be 0.0002529083)

Null deviance: 0.0084011  on 9  degrees of freedom
Residual deviance: 0.0017704  on 7  degrees of freedom
AIC: -50.013

Number of Fisher Scoring iterations: 2
```

Izvor: Obračun autora

Korišćenjem koeficijenata iz rezultata dobijenog u jeziku R iz Tabele 5. može se prikazati model stope prekida u zavisnosti od rasta BDP i referentne stope NBS:

$$\hat{Y}(t) = 0{,}10624 + 0{,}44828 * Rast\ BDP\ (t-2) + 0{,}47182 * ref.stopa\ NBS\ (t-1)$$

Pokazatelj kvaliteta ocenjenog modela jeste koeficijent determinacije R^2. Koeficijent determinacije za ovaj model iznosi 0,7893, dok je korigovani koeficijent determinacije 0,7291. Navedeni koeficijenti pokazuju koji je deo varijacija zavisne promenljive objašnjen modelom, i imaju prilično visoke vrednosti, što dokazuje da je model adekvatan.

3. Ispunjenost pretpostavki modela

Posle formiranja modela potrebno je proveriti da li podaci ispunjavaju pretpostavke GLM modela.[31]

[31] Janković, D. (2014). *Regresija – Linearni modeli*. Seminarski rad. Beograd: Matematički fakultet Univerziteta u Beogradu.

Pretpostavka br. 1: nema prisustva u opservacijama većeg broja odstupajućih vrednosti, budući da bi to negativno uticalo na model koji se zasniva na fitovanju metodom najmanjih kvadrata. U slučaju da postoji veći broj opservacija čija je standardna devijacija veća od zadatog limita, potrebno ih je izbaciti iz analize. Odstupajuće vrednosti se javljaju, ali nisu dovoljno značajne. Pošto bi izbacivanjem odstupajućih vrednosti bile skraćene vremenske serije, čime bi bila smanjena i tačnost predviđanja modelom, u modelu će ostati sve vrednosti.

Pretpostavka br. 2: zavisna i nezavisne promenljive imaju linearnu vezu, što se može proveriti matematički ili vizuelno, prikazivanjem zavisne i nezavisnih promenljivih na grafikonu. Posmatrajući Sliku 5. lako primećujemo jasnu linearnu vezu između zavisne promenljive i dve nezavisne promenljive.

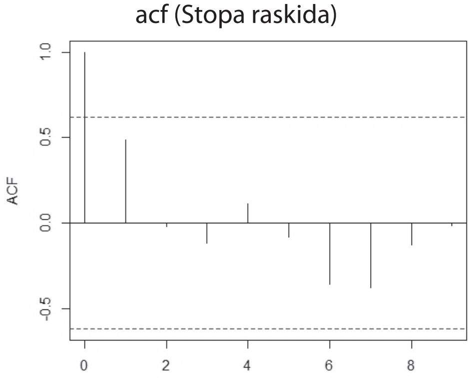
Pretpostavka br. 3: opservacije su međusobno nezavisne, što znači da nema autokorelacije. Ako zavisna promenljiva ima svojstvo autokorelacije, to znači da se vrednosti promenljive iz prošlosti ponavljaju, pa njene istorijske vrednosti mogu biti korišćene kao eksplanatorna promenljiva. Autokorelacija je testirana Ljung-Box statistikom[32] u jeziku R komandom Box.test() i dobijena p-vrednost je 0.07414. P-vrednost veća od 0,05 potvrđuje da nulta hipoteza o odsustvu autokorelacije ne može biti odbačena na nivou značajnosti od 5%. Grafički rezultat Ljung-Box statistike prikazan je na Slici 6, gde se takođe vidi da nema značajnih korelacija između elemenata vremenske serije zavisne promenljive.

Pretpostavka br. 4: odsustvo multikolinearnosti, što u kontekstu GLM modela znači da nema jake korelacije između prediktora. Multikolinearnost se takođe može desiti ukoliko je dostupno malo podataka u poređenju s brojem parametara koje treba proceniti. U slučaju multikolinearnosti, parametarski vektor nema jedinstveno rešenje. Provera prisustva multikolinearnosti može se izvršiti na osnovu VIF[33] faktora (engl. *Variance Inflation Factor*), koji predstavlja skor izračunat na osnovu određenih parametara za izabrani prediktor. VIF faktor pokazuje stepen uvećanja varijance regresione promenljive zbog multikolinearnosti. Najmanji VIF može biti jednak 1, što znači potpuno odsustvo kolinearnosti. VIF veći od 10 zahteva preduzimanje određene akcije u cilju smanjenja multikolinearnosti. U jeziku R VIF vrednosti svih promenljivih izračunavaju se komandom *vif()*. Dobijeni rezultat za obe testirane nezavisne promenljive je isti i iznosi 1.22409, što ukazuje na odsustvo multikolinearnosti.

[32] Ljung, G. & Box, G. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika* Vol. 65, p. 297–-303.

[33] Kassambara, A. (2017). *Machine Learning Essentials: Practical Guide in R*. Marseille: STHDA

Slika 6. Autokorelacija zavisne promenljive stopā prekida

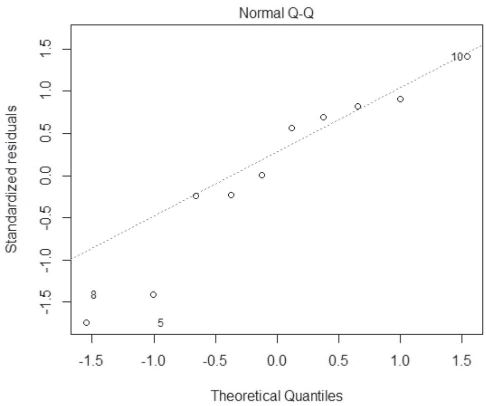


Izvor: Obračun autora

Pretpostavka br. 5: Reprezentativnost uzorka bazira se na slučajnosti selekcije opservacija. U ovom istraživanju će biti uzeti u obzir svi raspoloživi podaci o prekidima polisa u izabranom vremenskom rasponu, tako da se problem reprezentativnosti ne postavlja.

Pretpostavka br. 6: Greške imaju normalnu raspodelu. Pretpostavka se može proveriti vizuelno na grafikonu ili Kolmogorov-Smirnov testom, koji poredi uzorak iz modela s normalnom raspodelom. Na Slici 7. prikazan je Q-Q dijagram, dobijen iz softverskog paketa R, komandom *plot()*, na kome se vidi da odstupanja postoje, ali nisu značajna.

Slika 7. Q-Q dijagram

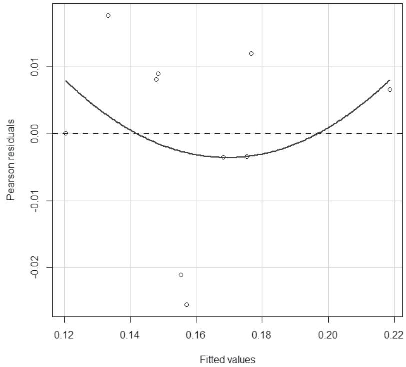


Izvor: Obračun autora

Pirsonovi reziduali ili reziduali modela podeljeni kvadratnim korenom varijance prikazani su na Slici 8. Dobijeni su komandom *residualPlots()*. Reziduali nisu potpuno simetrični, oko vrednosti 0, ali su odstupanja prihvatljiva, jer nije moguće manjim izmenama dobiti efikasniji model.

Pretpostavka br. 7: Homoskedastičnost – homogenost varijanci grešaka, što znači da različite promenljive imaju istu disperziju u svojim greškama, bez obzira na vrednosti početnih promenljivih. Greške su heteroskedastične ukoliko rezultujuće promenljive mogu da variraju u širokom rasponu. Kako bi se odredila heterogena greška disperzije ili kada skup reziduala krši pretpostavke modela o homoskedastičnosti, mudro je potražiti tzv. fening-efekat[34] između rezidualne greške i pretpostavljenih vrednosti. To govori da će biti sistematske promene u apsolutnim ili kvadratnim rezidualima kada se grafički prikažu protiv predvidljivog ishoda. Greška neće čak ni biti raspoređena po regresionoj liniji. U stvari, reziduali se grupišu i prostiru nešto dalje od njihovih predviđenih grafika za veće i manje vrednosti od tačaka sa linearne regresione linije, a srednje-kvadratna greška za model će biti pogrešna. Rezultujuća promenljiva čija je sredina velika tipično će imati veću disperziju od one s malom sredinom. Može se proveriti prikazivanjem na grafikonu standardizovanih grešaka i standardizovanih prediktora. U slučaju da pretpostavka homoskedastičnosti nije ispunjena, model se i dalje može koristiti, ali je kvalitet dobijenih rezultata iz modela smanjen. Homoskedastičnost se može proveriti i Brojš-Paganovim testom, koji se u jeziku R realizuje komandom *bptest()*. Rezultat Brojš-Paganovog testa ima p-vrednost 0.6704. Dobijena p-vrednost testa je veća od 0,05, što ukazuje na to da nulta hipoteza o homoskedastičnosti ne može biti odbačena na datom nivou značajnosti.

Slika 8. Pirsonovi reziduali modela



Izvor: Obračun autora

[34] Janković, D. (2014). *Regresija – Linearni modeli*. Seminarski rad. Beograd: Matematički fakultet Univerziteta u Beogradu.

Pretpostavka br. 8: Odsustvo autokorelacije grešaka podrazumeva da su slučajne greške međusobno nezavisne. Autokorelacija je testirana Ljung-Box statistikom u jeziku R komandom: *Box.test()*. Rezultat dobijen Ljung-Box testom ima p-vrednost 0.6996. Dobijena p-vrednost testa je veća od 0,05, što ukazuje na to da nema autokorelacije grešaka modela.

4. Primena modela

Na osnovu izabranog modela, može se predvideti očekivana stopa prekida ugovora u 2018. godini, i to na sledeći način:

$$\hat{Y}\,(2018) = 0{,}10624 + 0{,}44828 * \mathit{Rast\ BDP}(2016) + 0{,}47182 * \mathit{ref.\ stopa\ NBS}(2017)$$

Predikcioni interval je očekivani opseg vrednosti u kome će se naći vrednost slučajne promenljive sa određenom pouzdanošću. Očekivana stopa prekida u 2018. godini, koja je izračunata u prethodnoj formuli, predstavlja srednju vrednost predikcionog intervala, tako da je predikcioni interval za stopu prekida u 2018. godini $[\hat{Y}(2018) - \mathit{limit}, \hat{Y}(2018) + \mathit{limit}]$ sa pouzdanošću α.

U jeziku R, predikcioni interval sa nivoom poverenja 99,5% dobija se komandom *predict.glm()*. Dobijeni rezultat je prikazan u Tabeli 6.

Tabela 6. Predikcioni interval dobijen u jeziku R

	fit	*lwr*	*upr*
2016-01-10	*0.1375434*	*0.09015011*	*0.1849368*

Izvor: Obračun autora

Predikcioni interval koji daje razvijeni model je [13,75% - 4,73%; 13,75% + 4,74%] ili je [9,01%; 18,49%], Predikcioni interval koji daje standardna formula je [tekuća stopa prekida*0,5; tekuća stopa prekida*1,5] što je [7,55%; 22,64%].

S obzirom na to da je predikcioni interval razvijenog parcijalnog internog modela manji nego predikcioni interval standardne formule, razvijeni model preciznije meri kapitalni zahtev i samim tim rezultuje nižim zahtevanim kapitalom u poređenju sa standardnom formulom.

VII. Zaključak

Analizirano je devet faktora koji utiču na stopu prekida ugovora životnog osiguranja: BDP, referentna kamatna stopa, prosečna zarada, berzanski indeks

BelexLine, inflacija, stopa nezaposlenosti, stopa rasta premije životnog osiguranja, stopa rasta premije neživotnog osiguranja i stopa rasta ukupne premije osiguranja.

Na osnovu istraživanja stranih autora koji su pomenuti u delu ovog rada *Pregled literature o faktorima stope prekida*, stopa prekida ugovora može se adekvatno modelirati na osnovu referentne kamatne stope, bruto domaćeg proizvoda po stanovniku, stope nezaposlenosti, razvijenosti tržišta kapitala i veličine osiguravajuće kompanije na stopu prekida. Istraživanje na domaćim podacima, koje je opisano u ovom radu, kao što bi se i moglo očekivati na osnovu pomenutih istraživanja stranih autora, pokazalo je da su za modeliranje zavisnosti stope prekida konkretne osiguravajuće kompanije najadekvatniji sledeći faktori: rast BDP-a i referentna stopa NBS. Korišćen je generalizovani linearni model s normalnom raspodelom i funkcijom identiteta za link funkciju, koji ispunjava sve zahtevane pretpostavke GLM modela.

Na osnovu sprovedenog istraživanja autor zaključuje da razvijeni parcijalni interni model preciznije meri kapitalni zahtev konkretne osiguravajuće kompanije i ima niži zahtevani kapital u poređenju sa standardnom formulom. Kada se primeni režim Solventnosti II u Srbiji, ovaj zaključak može biti iskorišćen od strane domaćih osiguravajućih kompanija za adekvatnije određivanje kapitalnog zahteva za solventnost formiranjem internog parcijalnog modela koji bi uključio i model za rizik od prekida ugovora životnog osiguranja.

Pravac daljeg razvoja istraživanja može biti analiziranje uticaja internih faktora sa polise na stopu prekida ugovora, kao što su: vrsta osiguranja, broj proteklih godina od zaključenja ugovora, trajanje ugovora, veličina premije, frekvencija plaćanja premije, veličina osigurane sume, kanal prodaje, starost osiguranika, pol osiguranika itd.

Literatura

- Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* p. 100–122.
- Cheng, Ch., Hilpert, Ch. et al. (2020). Surender Contagion in Life Insurance. *SSRN Electronic Journal* March 2020.
- Cox, S. & Lin, Y. (2006). *Annuity Lapse Rate Modeling: Tobit or Not Tobit?* Society of Actuaries. *http://library.soa.org.*
- Dar, A. & Dodds, C. (1989). Interest Rates, the Emergency Fund Hypothesis and Saving Through Endowment Policies: Some Empirical Evidence for the U.K. *Journal of Risk and Insurance 56*(3), p. 415-433.
- Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), p. 241–269.

- Gatzert, N. & Schmeiser, H. (2008). Assessing the Risk Potential of Premium Payment Option in Participating Life Insurance Contracts. *Journal of Risk and Insurance, 75*(3), p. 691–712.
- Janković, D. (2014). *Regresija – Linearni modeli*. Seminarski rad. Beograd: Matematički fakultet Univerziteta u Beogradu
- Kagraoka, Y. (2005). *Modeling Insurance Surrenders by the Negative Binomial Model.* Working paper.
- Kassambara, A. (2017). *Machine Learning Essentials: Practical Guide in R*. Marseille: STHDA
- Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal 16*(1), p. 52-73.
- Kim, C. (2005). Modeling Surrender and Lapse Rates with Economic Variables. *North American Actuarial Journal 9*(4), p. 56-70.
- Kočović, J., Jovović, M. i Kočović, M. (2015). Aktuarski efekti prevremenog prekida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, pp. 77–84.
- Kuo, W., Tsai, C. and Chen, W.-K. (2003). An Empirical Study on the Lapse Rate: The Cointegration Approach. *Journal of Risk and Insurance 70*(3), p. 489-508.
- Ljung, G. & Box, G. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika Vol. 65*, p. 297-303.
- Michorius, C. (2011). *Modeling Lapse Rates – Investigating the Variables that Drive Lapse Rates*. Master Thesis. Enschede: Faculty of Management and Governance, University of Twente.
- Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2010). *Surrender Triggers in Life Insurance: Classification and Risk Predictions*. Working paper.
- Mladenović, Z. (2015). *Ekonometrijski metodi i modeli - Dickey-Fuller-ov test jediničnog korena,* Beograd: Ekonomski fakultet Univerziteta u Beogradu, s. 9.
- Mladenović, Z. (2020). *KPSS test jediničnog korena,* Beograd: Ekonomski fakultet Univerziteta u Beogradu, s. 9.
- Nelder, J. & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society A 135*, p. 370–384.
- Outreville, J. (1990). Whole-life Insurance Lapse Rates and the Emergency Fund Hypothesis. *Insurance: Mathematics and Economics 9*(4), p. 249–255.
- Russel, D., Stephen, J. et al. (2013). An Empirical Analysis of Life Insurance Policy Surrender Activity. *Journal of Insurance Issues 36*(1), p. 35–57.
- Zakon o osiguranju. *Službeni glasnik RS,* br. 139/14.
- *https://www.belex.rs*
- *https://www.mfin.gov.rs*
- *https://www.nbs.rs*
- *https://www.stat.gov.rs*

UDK:368.811:332.88:368.025.3:347.8:368.52:608.4:517.358:368.025.6
DOI: 10.5937/TokOsig2102039P

Branko R. Pavlović[1]

PARTIAL INTERNAL MODEL UNDER THE SOLVENCY II FOR THE LIFE INSURANCE LAPSE RISK

SCIENTIFIC PAPER

Abstract

Uncertainty in terms of the expected lapse rate of life insurance contracts affects the risk of imprecise determination of solvency capital requirement, minimum capital requirement, and performance of an insurance company. Therefore, a precise lapse risk projection of a life insurance contract bears great significance. The lapse rate is influenced by numerous factors.

Solvency capital requirement in terms of the life insurance lapse risk under Solvency II regime may be determined by using a prescribed standard formula or partial internal model. On the example of data obtained from the Serbian insurance market and by using an R software package, this paper will show in detail the selection of lapse rate risk factor modelling. It will also show the formation of the GLM model of lapse rates and verification of assumptions of the GLM.

The developed partial internal model may be applied for determining the lapse rate of an insurance company operating on the domestic market.

Key words: *partial internal model, lapse risk*

I. Introduction

The lapse risk is the risk of loss or increase of insurer's liabilities due to a change in the expected exercise rates of policyholder contractual options in terms of life insurance. A contract may be: partly or fully terminated; terminated with the

[1] President of Serbian Actuarial Association, e-mail: pavlovic.branko@gmail.com
Paper received on: 28 march 2021
Paper accepted on: 1 april 2021

payment of cash value or without such payment; terminated in relation to the liability of the policyholder to pay the premium but with the obligation of the insurer to pay out the reduced value of the sum insured or without a paid-up sum; terminated with the possibility of contract renewal in a particular period with premium payments that would become due until renewal or without the option of renewal.

1. Solvency Capital Requirement

Solvency capital requirement under the Solvency II regime may be calculated as follows: using a prescribed standard formula, using full internal model, and using a partial internal model.

The standard formula is intended for an average European insurance company and contains a large number of approximations and additionally proposed simplifications for particular risks which for a particular insurance company do not represent key risks. In each risk module and sub-module, an insurance company may replace a standard formula with its own methodology, and create a model for determining the capital requirement. If all risk modules and sub-modules are replaced, then such internal models are called full internal models. If replacement is made only for a few risk modules or sub-modules, whereas remaining risks are calculated by standard formula, such internal models are called partial internal models.

Partial internal models are introduced to enable the assessment of the expected lapse rate that would be more precise than that shown in a standard formula and thus, provide a more adequate assessment of the lapse risk and appropriate capital requirement.

2. Measuring Lapse Risk under a Standard Solvency II Formula

Capital requirement for the lapse risk equals the highest amount of the following capital requirements for the three sub-risks: a permanent increase in lapse rates, a permanent decrease in lapse rates, and a mass lapse event.

$$SCR_{lapse} = \max\ (Lapse_{up}\ ; Lapse_{down}\ ; Lapse_{mass}$$

The capital requirement for each of the above three sub-risks is obtained as the loss of net asset value of an insurance company that would occur due to permanent increase or decrease in used lapse rate options or immediate mass lapse event:

$$Lapse_i = \Delta NAV \mid lapseshock_i$$

The mentioned three options among which the one with the highest influence is selected are explained in Table 1.

Table 1 Example of lapse rate trends under the standard formula requirements

Year	1	2	3	4	5+
Expected lapse rate	4%	4%	4%	4%	4%
Permanent increase in lapse 50%	6%	6%	6%	6%	6%
Permanent decrease in lapse 50%	2%	2%	2%	2%	2%
Mass lapse 40% of all policies	40%	4%	4%	4%	4%

Source: Author's calculation

3. Measuring Lapse Risk under the Solvency II Partial Internal Model

In line with the principles-based approach under the Solvency II regime, when replacing the rules-based approach that was previously effective, EIOPA did not prescribe a formal definition of the internal model, nor did it define what internal model should include. The most significant difference between the internal models and the standard formula is in a more extensive use of stochastic techniques over the own data in internal models.

The application of partial or full internal models for risk measuring is aimed at encouraging insurance companies to more accurately assess and control their risks. The purpose of introducing internal models should not be the reduction of solvency capital requirement but better management of own risks. The use of internal models helps insurers to more adequately model their risks, which leads to increased risk-specific sensitivity of obtained results for solvency capital requirement of a given insurance company. Continuous periodical validation of the internal model is crucial for the successful use of internal models in a company.

The disadvantage of internal models lies in their great complexity which requires a considerable deployment of company resources.

The subject of this paper is the analysis of the financial impact of life insurance lapse risk on the solvency of life insurance companies. The aim of the paper is to formulate a partial internal model for determining capital requirements to cover the life insurance lapse risks, observing the specific characteristics of the risk in both individual insurance companies and the life insurance market in Serbia.

The research presented in this paper will be the first empirical study of the lapse rate on the insurance market in Serbia. Since the domestic market accounted for some 900 thousand active life policies at the end of 2019, [2] whereas there are no systematised data on policy lapses, the analysis of more than 200.000 policies will prove very useful. It is particularly interesting to note the behaviour of the policyholders in terms of early termination of the contracts at the time of the unexpected

[2] *www.nbs.rs*

event, which this research will be able to analyse since it also includes the year 2008 when the Global Financial Crisis began.

II. Overview of Literature on the Lapse Rate Factors

Numerous authors around the globe have studied the impact of different factors on the lapse rate. The first group of authors dealt with the impact of environmental characteristics such as macroeconomic indicators. Dar and Dodds,[3] Outreville,[4] group of authors[5] headed by Kuo and group of authors led by Russel,[6] focused on the reference interest rate, gross domestic product per capita, and unemployment rate, whereas Cox and Lin,[7] Kim,[8] and Kiesenbauer[9] analysed the impact on the lapse rate created by gross domestic product, development of the capital market, and size of an insurance company. The other group of authors studied the impact of details contained in the life contract on the lapse rate, using a generalized linear model. Kagraoka,[10] a group of authors[11] led by Cerchiara, a group of authors[12] headed by Milhaud, Eling, and Kiesenbauer[13] studied the impact of the year of the contract, age of policyholders, premium payment method, sales channels, and the existence of additional covers. Eventually, the group of authors[14] headed by Cheng analysed the growth of lapse rate depending on the behaviour of other market players who decide to terminate their contracts.

[3] Dar, A. & Dodds, C. (1989). Interest Rates, the Emergency Fund Hypothesis and Saving Through Endowment Policies: Some Empirical Evidence for the U.K. *Journal of Risk and Insurance 56*(3), pp. 415–433.

[4] Outreville, J. (1990). Whole-life Insurance Lapse Rates and the Emergency Fund Hypothesis. *Insurance: Mathematics and Economics 9*(4), pp. 249–255.

[5] Kuo, W., Tsai, C. and Chen, W.-K. (2003). An Empirical Study on the Lapse Rate: The Cointegration Approach. *Journal of Risk and Insurance 70*(3), pp. 489-508.

[6] Russel, D., Stephen, J. et al. (2013). An Empirical Analysis of Life Insurance Policy Surrender Activity. *Journal of Insurance Issues 36*(1), pp. 35-57.

[7] Cox, S. & Lin, Y. (2006). *Annuity Lapse Rate Modeling: Tobit or Not Tobit?* Society of Actuaries. *http://library.soa.org.*

[8] Kim, C. (2005). Modeling Surrender and Lapse Rates with Economic Variables. *North American Actuarial Journal 9*(4), pp. 56–70.

[9] Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal 16*(1), pp. 52-73.

[10] Kagraoka, Y. (2005). *Modeling Insurance Surrenders by the Negative Binomial Model.* Working paper.

[11] Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* pp. 100–122.

[12] Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2010). *Surrender Triggers in Life Insurance: Classification and Risk Predictions*. Working paper.

[13] Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), pp. 241-269.

[14] Cheng, Ch., Hilpert, Ch. et al. (2020). Surrender Contagion in Life Insurance. *SSRN Electronic Journal* March 2020.

Except for Kočović and Jovović,[15] domestic authors did not deal with the issues of life contract lapses.

III. Projection of the Lapse Risk

1. Importance of Life Lapse Risk Projection

The moment when the liability to the Insured will arise is uncertain for an insurance company. For example, under the endowment insurance contract that covers the death risk and survival, which is concluded for the term of 20 years, in the event of death of the insured person, an insurance company may pay out the full sum insured on the very next day or a surrender value at any time after three years from insurance inception date, or in the event that it does not pay anything of the aforementioned, it may pay out the sum insured upon the insured person's survival of the contract expiry.

Uncertainty relating to the lapse rate does not only influence the risk of imprecise determination of solvency capital requirement and minimum capital requirement but also has several other important effects on the business performance of insurance companies.

The first effect relates to the payment of surrender value. When policy lapses, the insurance company pays out the surrender value to the policyholder and cancels the mathematical reserve for such policy. On the Serbian market, the surrender value is lower than or equal to the mathematical reserve so that at the moment of lapse, the insurance company has revenue. In the event when the lapse rate is lower, the insurance company will have a revenue lower than expected, which may present a profitability problem if the expected lapse rate was used for pricing the service.

Another effect relates to the coverage of acquisition costs. In life insurance, acquisition costs are rather high. In a financial statement, acquisition costs that occur at the moment of contract conclusion are usually recognised through mathematical reserve reduction, using the zillmerization. If the lapse rate is higher than expected and there is no clawback mechanism, the company may not be able to cover the acquisition costs and may face losses.

Even if the average lapse rate behaves expectedly, the third effect may emerge and affect the insurer's profitability. If policies of healthy insured persons lapse more than expected and the policies of insured persons of poor health lapse less than expected, the mortality rate within a portfolio may considerably increase, regardless of the unchanged mortality of population.

[15] Kočović, J., Jovović, M. and Kočović, M. (2015). Aktuarski efekti prevremenog raskida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, pp. 77–84.

The absence of expected future profit from lapsed policies may be an important fourth effect harming the profitability of insurance companies.

The fifth effect is threatened liquidity due to any unexpected mass pay-outs of surrender value for lapsed policies.

Formation of the mathematical reserve from which liabilities to the Insured are paid out, requires the use of different assumptions relating to mortality, technical interest rate, etc. The said assumptions are most often based on the experience of an insurance company and actuarial principles and regulations.[16] The sufficiency of the mathematical reserve for the payment of assumed liabilities is directly related to the company's solvency. Therefore it is very important to properly estimate future cash flows. The sixth effect of the increase in lapse rate is the impact on cash flows and thus on company solvency.

Maturity mismatch between assets and liabilities may occur as the seventh effect of the increased lapse rate. Since liabilities from standard long-term savings life contracts should be covered by assets of appropriate duration, the lapses of such contracts require asset adjustment, which entails particular costs.

The eighth effect relates to the reputational risk of an insurance company. Prospective insureds know that a particular number of contracts may lapse before their expiry but are often uninformed about their rights regarding that matter. Particular agents provide confusing or even wrong oral information that there are no adverse effects due to contract termination and that the premium paid until then will be returned to the insured person. Such practice is contrary to Articles 82 and 83 of the Insurance Law[17] and in addition to undermining the reputation of an insurance company, may lead to penalties imposed by the insurance supervisory authority.

Finally, the last negative effect of the increase in the lapse rate is the effect on the embedded value of an insurance company because of the decrease in cash flows from future premiums.[18]

Because of all the mentioned effects, it is important to foresee the lapse rate as precisely as possible.

2. Factors Influencing Lapse Rate

The lapse rate depends on many rational and irrational reasons influencing the behaviour of policyholders.[19] An example of rational behaviour is the response to

[16] Kočović, J., Jovović, M. and Kočović, M. (2015). Aktuarski efekti prevremenog raskida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, pp. 77-84.

[17] Insurance Law. *Official Gazette of RS,* no. 139/14.

[18] Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), pp. 241-269.

[19] Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* pp. 100-122.

financial market trends and changes in macroeconomic variables such as inflation, foreign exchange rate, stock market indices, purchasing power, unemployment rate, and the like. An example of irrational behaviour may be seen in the termination of the insurance contract to buy a car from the accumulated cash.

The type of insurance and characteristics of provided service considerably influence the lapse rate. Policyholders find it easier to terminate a multi-year term life contract unaware that this is how they lose money. For example, by terminating a life insurance contract with survival benefit, policyholders lose a part of their money because, as a rule, the surrender value is lower than the mathematical reserve. The following characteristics of insurance services impact the possible decision on contract termination:[20] term of the contract, period remaining until the contract expiry, amount of premium and the sum insured, premium payment frequency, phase of the annuity contract (premium payment and annuity pay-outs), level of penalty reduction of the mathematical reserve to be paid in the event of surrender, allocation method of non-guaranteed profit, recent fund yields in unit-linked contract, the structure of agents' commissions, etc.

Age and gender of a policyholder, his or her place of residence as an income level indicator, marital status, and particularly the change of such status may also be used to analyse the lapse rate.

Particular policyholders consider life insurance contracts as a type of savings in case of unforeseen circumstances so that, for example, if they lose their job, they can use the surrender value to make up for the lack of income in the transition period. Investors into life insurance aiming to increase their wealth easily terminate the contract and switch to more lucrative investments when the interest rates start going up.

Unforeseen events such as a change in tax policy and change of ownership or reputation of an insurance company may considerably change the lapse rate.

Policyholders often decide to terminate the contract for more than one reason. Particular, previously mentioned factors, have almost perfect correlation such as, for example, age of the insureds and term of the contract which simultaneously increase. When modelling the lapse, both such factors cannot be used. In addition, particular variables may depend on the value of the other variable.

All the mentioned factors may be used as explanatory variables to explain and foresee the change in life insurance lapse rates, as dependent variables. The set of explanatory variables will be selected depending on the data on policies written by a particular insurance company.

Firstly, one should study the data of insurance companies when modelling the dependency of lapse rate on more than one variable, by analysing the lapse rate dependency on each independent variable available in the data set and by studying the correlation between available independent variables.

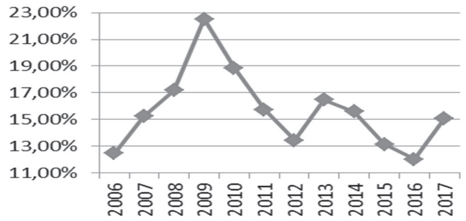
[20] Michorius, C. (2011). *Modeling Lapse Rates – Investigating the Variables that Drive Lapse Rates*. Master Thesis. Enschede: Faculty of Management and Governance, University of Twente.

IV. Data

In the below analysis, the data taken from the Serbian insurance market were used with policies issued or lapsed in the period from 1 January 2006 to 31 December 2017.

Figure 1 Lapse Rate

Year	Lapse rate	Year	Lapse rate
2006	12,46%	2012	13,43%
2007	15,25%	2013	16,48%
2008	17,20%	2014	15,60%
2009	22,53%	2015	13,14%
2010	18,86%	2016	12,04%
2011	15,74%	2017	15,09%



Source: Author's calculation

After the data processing in Microsoft Access, the time series of lapse rates were obtained by years of the analysed period, as shown in Figure 1.

The stationarity of time series may be tested by combining Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test and Dickey-Fuller (DF) test. The rule of statistical testing says that the test conclusion may be the rejection of null hypothesis and acceptance of alternative hypothesis or, yet, that no conclusion can be drawn.

The time series are firstly tested by the KPSS test. According to a null hypothesis of the KPSS test, there is no unit root. According to the alternative hypothesis of the KPSS test, there is a unit root, namely, the time series are not stationary. Null hypothesis on stationary time series is rejected for the selected level of significance if the realised value of test statistics is higher than the corresponding critical value.[21] If the null hypothesis of the KPSS test is rejected, the time series contain a unit root. If the null hypothesis of the KPSS test cannot be rejected, then it cannot be concluded that the time series are stationary, and DF testing continues.

The null hypothesis of the DF test holds that there is a unit root i.e. that time series are not stationary. According to the alternative hypothesis of the DF test, time series are stationary. The null hypothesis that there is a unit root is rejected for a sufficiently low value of statistics i.e. when the calculated value is lower than the critical value.[22] If the DF test rejects the null hypothesis, it can be concluded that there is no unit root in the time series, which means that the time series are stationary. If the DF test fails to reject the null hypothesis, there is still no conclusion.

21 Mladenović, Z. (2020). *KPSS test jediničnog korena*, Beograd: Faculty of Economics of the University of Belgrade, pp. 9.

22 Mladenović, Z. (2015). *Ekonometrijski metodi i modeli – Dickey-Fuller-ov test jediničnog korena*, Beograd: Faculty of Economics of the University of Belgrade, pp. 9.

The lapse rate is shown in Figure 1. Since the lapse rate value fluctuates over the period of 11 years i.e. does not have a pronounced trend, when selecting the type of regression it was assumed that the time series do not contain the time trend component.

Table 2 Official statistics

Year	GDP growth	Price growth	Average net wage growth in EUR	Unem-ploy-ment rate[23]	Belex-Line Index growth[24]	Ref. Inter-est rate of NBS[25]	Non-life pre-mium growth	Life pre-mium growth	Total premium growth
2006	4,9%	6,6%	23%	20,9%	36,0%	14,00%	7,5%	19,6%	8,6%
2007	6,4%	11,0%	35%	18,1%	44,1%	10,00%	22,4%	28,9%	23,1%
2008	5,7%	8,6%	16%	13,6%	-68,7%	17,75%	13,0%	26,2%	14,4%
2009	-2,7%	6,6%	-16%	16,1%	9,5%	9,50%	-13,5%	8,7%	-10,8%
2010	0,7%	10,3%	-2%	19,2%	-2,2%	11,50%	-5,8%	8,5%	-3,7%
2011	2,0%	7,0%	12%	23,0%	-23,8%	9,75%	1,3%	7,5%	2,3%
2012	-0,7%	12,2%	-2%	23,9%	2,9%	11,25%	-5,5%	7,2%	-3,3%
2013	2,9%	2,2%	6%	22,1%	9,9%	9,50%	0,8%	17,6%	4,1%
2014	-1,6%	1,7%	-2%	19,2%	21,7%	8,00%	3,0%	10,2%	4,6%
2015	1,8%	1,5%	-3%	17,7%	2,7%	4,50%	11,9%	17,8%	13,3%
2016	3,3%	1,6%	2%	15,3%	13,7%	4,00%	5,2%	16,8%	8,0%
2017	2,1%	3,0%	3%	13,5%	5,9%	3,50%	9,1%	0,4%	6,8%

Source: Ministry of Finance of the Republic of Serbia,[26] Statistical Office of the Republic of Serbia,[27] National Bank of Serbia[28] and Belgrade Stock Exchange a. d.[29]

The stationarity of the time series for the dependent variable of lapse rate was tested by using critical values for 5 percent confidence intervals (KPSS: 0,463, DF: -2,862). Testing was performed in language R using commands: kpss.test() and adf.test().

In testing the time series stationarity with the KPSS test, the following result was obtained: KPSS Level = 0.40713, which is lower than the critical value 0,463, and which further means that the KPSS test failed to reject the null hypothesis of stationarity. With the ADF test, the following test result of time series stationarity was obtained: Dickey-Fuller = -3.2279. This value is lower than the critical value -2,862, which means that the extended DF test rejected the null hypothesis that there is a unit root. From the combined KPSS and DF tests, it can be concluded that the time series are stationary.

[23] Unemployment rate has been aligned with ILO methodology since 2004.
[24] Stock market index BELEXfm, later transformed into BELEXLine, was formed in December 2004.
[25] Reference rate has been published by the National Bank of Serbia since 2006.
[26] *https://www.mfin.gov.rs*
[27] *https://www.stat.gov.rs*
[28] *https://www.nbs.rs*
[29] *https://www.belex.rs*

In addition to the data obtained from the Serbian insurance market, the analysis used official statistics of the Statistical Office of the Republic of Serbia, Ministry of Finance of the Republic of Serbia, National Bank of Serbia, and Belgrade Stock Exchange: GDP, reference interest rate, average wages, stock market index BelexLine, inflation, unemployment rate, growth rates of life and non-life premium, etc., as shown in Table 2. The analysis of lapse rate dependency from the environmental parameter was performed in the period 2006–2017.

The data were processed in the programme package R which contains all necessary tools for predictive analysis in Microsoft Excel.

V. Selection of Modelling Factors for Lapse Rate Dependency

Table 3 shows maximum correlation coefficients between nine environmental predictors and dependent variable lapse rate calculated in the language R, using Cross-Correlation Function *ccf()*.

Table 3 Maximum correlations between environmental predictors and lapse rate

No.	Analysed parameter	Maximum correlation coefficient	Time lag
1.	GDP growth	0,6524230	-2
2.	Price growth	-0,5494015	6
3.	Average net wage growth in EUR	0,7072750	-2
4.	Unemployment rate	0,6741854	2
5.	Growth of stock market index BELEXline	-0,4468971	-1
6.	Reference interest rate of NBS	0,6088886	-1
7.	Non-life premium growth	0,5922836	-2
8.	Life premium growth	0,5805626	-2
9.	Premium growth	0,5887550	-2

Source: Author's calculation

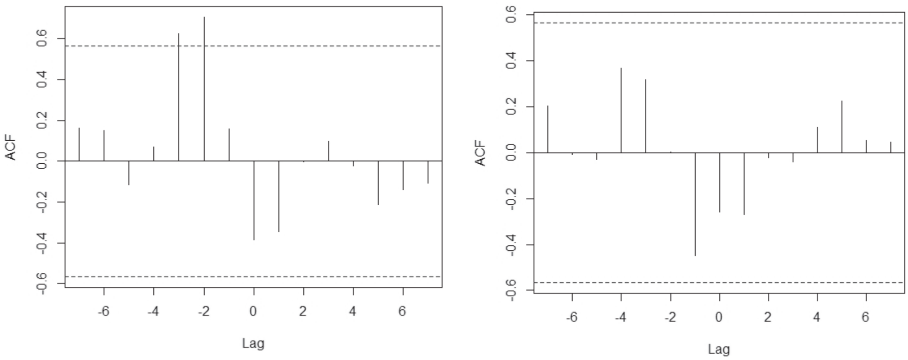
Table 3 shows that dependent variable lapse rate individually has the strongest relationship i.e. positively correlates with the independent variable of average wage growth (correlation coefficient 0.71), unemployment rate (0.67), GDP growth (0.65), and reference interest rate of NBS (0.61), which means that the mentioned four independent variables can be used in lapse rate modelling.

Figure 2 shows correlations for the highest correlation between individual average wage growth and lapse rate and the lowest correlation (in absolute value) between BelexLine index growth and lapse rate.

Figure 2 Correlation between environmental predictors and the lapse rate with the highest and lowest maximum correlation coefficient

a) correlation between average net wage and lapse rate

b) correlation between BelexLine index growth rate and lapse rate



Source: Author's calculation

With the same command in the language R, *ccf()*, mutual correlations of all independent variables are obtained. The result is shown in Table 4. For a model of good quality, one should select two of four independent variables which have the strongest relationships with the lapse rate but the weakest relationship between one another. Table 4 shows that the best choice would be the independent variables of GDP growth and NBS reference rate where the correlation coefficient is the lowest and amounts to 0.3777358, with a mutual delay of parameters of two time units. Correlations between independent variables such as GDP growth and NBS reference rate are shown in Figure 3.

Table 4 Maximum correlation coefficient between predictors (lags of predictors by columns in relation to the predictors by rows are provided in the brackets)

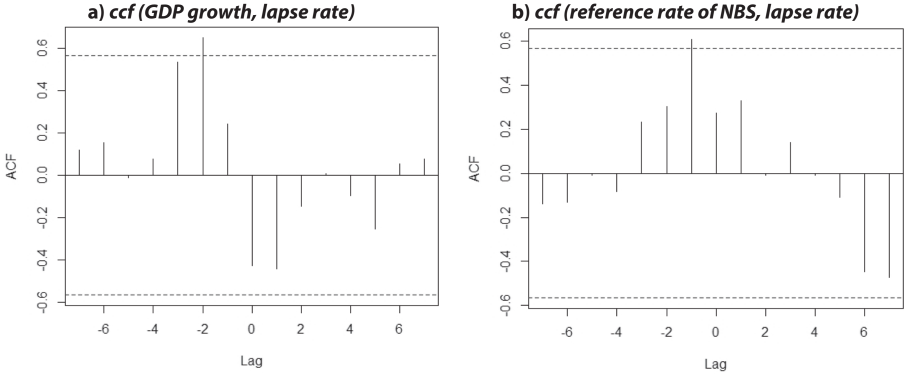
	GDP growth	Price Growth	Growth of average net wage in EUR	Unemployment rate	Belex-Line index growth	Ref. Interest rate of NBS	Non-life premium growth	Life premium growth	Total insurance premium growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1)	1	0,59 (4)	0,88 (0)	0,56 (5)	0,75 (-1)	0,38 (2)	0,82 (0)	0,74 (0)	0,83 (0)
(2)	0,59 (-4)	1	0,49 (-4)	-0,55 (-3)	0,55 (-5)	0,69 (1)	0,62 (-4)	0,56 (-4)	0,62 (-4)

	GDP growth	Price Growth	Growth of average net wage in EUR	Unemployment rate	Belex-Line index growth	Ref. Interest rate of NBS	Non-life premium growth	Life premium growth	Total insurance premium growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(3)	0,88 (0)	0,49 (4)	1	0,53 (5)	0,66 (-1)	0,52 (1)	0,77 (0)	0,67 (0)	0,78 (0)
(4)	0,56 (-5)	-0,55 (3)	0,53 (-5)	1	-0,44 (-3)	-0,49 (4)	-0,67 (-1)	-0,66 (-1)	-0,71 (-1)
(5)	0,75 (1)	0,55 (5)	0,66 (1)	-0,44 (3)	1	-0,41 (0)	0,85 (1)	0,54 (1)	0,85 (1)
(6)	0,38 (-2)	0,69 (-1)	0,52 (-1)	-0,49 (-4)	-0,41 (0)	1	0,41 (-5)	0,42 (0)	0,42 (-5)
(7)	0,82 (0)	0,62 (4)	0,77 (0)	-0,67 (1)	0,85 (-1)	0,41 (5)	1	0,65 (0)	0,99 (0)
(8)	0,74 (0)	0,56 (4)	0,67 (0)	-0,66 (1)	0,54 (-1)	0,42 (0)	0,65 (0)	1	0,74 (0)
(9)	0,83 (0)	0,62 (4)	0,78 (0)	-0,71 (1)	0,85 (-1)	0,42 (5)	0,99 (0)	0,74 (0)	1

Source: Author's calculation

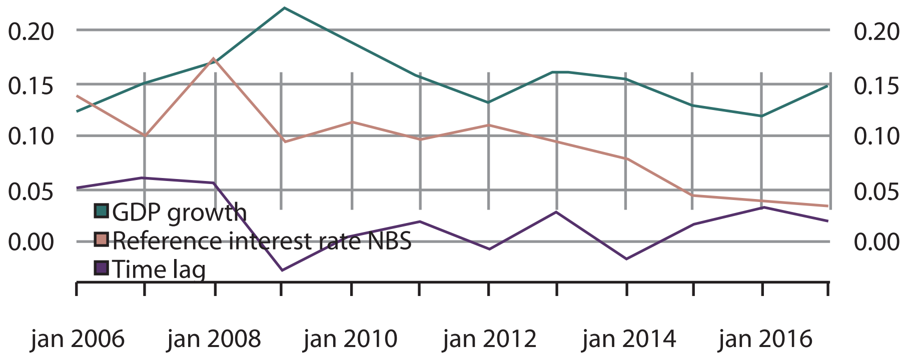
Figure 3 Maximum correlation of environmental predictors selected for lapse rate modelling

a) *ccf (GDP growth, lapse rate)* **b) *ccf (reference rate of NBS, lapse rate)***



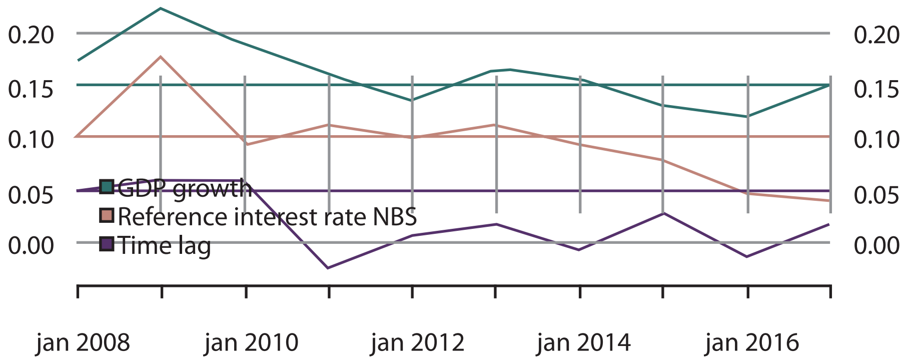
Source: Author's calculation

Figure 4 Time series of a dependent and two independent variables without time lag



Source: Author's calculation

Figure 5 Time series of a dependent and two independent variables with a corresponding time lag of independent variables



Source: Author's calculation

The time lag between time series dependence is important because without the introduction of time lags, appropriate dependences cannot be modelled, as can be seen in Figure 4 which shows one dependent and two independent variables without time lags.

With the introduction of two-year lag for GDP growth and one-year lag for NBS reference rate, the relationship between a dependent variable and independent variables becomes visible in Figure 5.

VI. Model

There are several types of tools that may be used when deciding on the appropriate model. Many software packages offer acceptable models automatically, owing to iterative adjustment (fitting) of the model to the available data. This method is certainly a good start, but it is a good idea to fine-tune proposed models. One of the methods is to test p-values to check the significance of each variable by calculating the significance of the entire model and the significance of the model without individual independent variable. Another method is to assess the model based on randomly selected 70% of samples and then to check and refit the remaining 30% of samples. The third method is to use any of the following criteria: AIC (Akaike Information Criteria) or BIC (Bayesian Information Criteria).

1. Generalized Linear Model (GLM)

Generalized linear model (GLM) deals with the life insurance lapse rate and is one of the well-known multifactor regression models covered in literature. This model was formulated by Nelder and Wedderburn in 1972.[30] It was selected because it is relatively easy to understand and very flexible in terms of the distribution of probabilities of dependent variable and input variables. Owing to link function, it may be used in the work with dependent and independent variables that can be continuous and binary, which is particularly important in lapses, because lapse variable has a binary value at the level of one policy (lapsed or not), and a continuous value between 0 and 1 for the whole portfolio.

In the 1970s, a special-purpose software was developed for work with GLM models and was dubbed GLIM (Generalized Linear Interactive Modelling). Today, GLM models are applied in different software packages such as SAS or SPSS, but the most popular GLM modelling is that in the language R, where GLM is applied as follows:

glm(formula, family = binomial(link=logit) data, weights, subset, na.action,…)

2. Model Formation

A generalized linear model with normal distribution and identity link function was selected. Based on input data and fulfilment of the specified assumptions, by calling the function *glm()* of the software package R, the result was obtained as shown in Table 5.

[30] Nelder, J. & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society A 135*, pp. 370–384.

Table 5 The result of function *glm()* in the language R

Deviance Residuals:
Min 1Q Median 3Q Max
-0.025582 -0.003458 0.003365 0.008776 0.017722

Coefficients:
Estimate Std. Error t value Pr(>|t|)
*(Intercept) 0.10624 0.01437 7.394 0.00015 ****
*rast_bdp.lag2 0.44828 0.18940 2.367 0.04983 **
*ref_ks_NBS.lag1 0.47182 0.15264 3.091 0.01754 **
---
*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

(Dispersion parameter for gaussian family taken to be 0.0002529083)

Null deviance: 0.0084011 on 9 degrees of freedom
Residual deviance: 0.0017704 on 7 degrees of freedom
AIC: -50.013

Number of Fisher Scoring iterations: 2

Source: Author's calculation

By using coefficients from the result obtained in the language R from Table 5, the lapse rate model may be shown depending on GDP growth and NBS reference rate:

The indicator of the quality of the assessed model is the coefficient of determination R^2. The coefficient of determination for this model is 0.7893, whereas the adjusted coefficient of determination is 0.7291. The said coefficients show which part of variations of a dependent variable is explained by the model and have quite high values, which proves that the model is adequate.

3. Fulfilment of Model Assumptions

After the formation of a model, it is necessary to check if the data meet the assumptions of the GLM model.[31]

Assumption no. 1: in the observations, there is no presence of a larger number of deviations since this would have an adverse impact on the model fitted by

[31] Janković, D. (2014). *Regresija – Linearni modeli*. Term paper. Belgrade: Faculty of Mathematics of the University of Belgrade.

the least-squares method. In the event that there is a larger number of observations where the standard deviation is higher than the set limit, such observations should be removed from the analysis. Deviations occur but are not sufficiently significant. Since the removal of deviations would shorten the time series and thus diminish the accuracy of the model forecast, all values will remain in the model.

Assumption no. 2: dependent variable and independent variables have a linear relationship, which can be verified mathematically and visually by showing a dependent variable and independent variables on the graph. In Figure 5, it is easy to notice a clear linear relationship between the dependent variable and two independent variables.

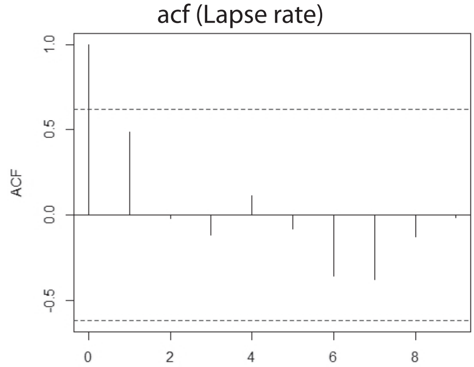
Assumption no. 3: observations are mutually independent, which means that there is no autocorrelation. If a dependent variable has the property of autocorrelation, it means that the past values of the variable repeat, and thus, its historical values may be used as an explanatory variable. Autocorrelation was tested by Ljung-Box statistics[32] in the language R by the command Box.test() where a p-value of 0.07414 was obtained. A P-value greater than 0.05 confirms that the null hypothesis of the absence o autocorrelation may not be rejected for a significance level of 5%. Chart result of Ljung-Box test statistics is shown in Figure 6, where it can also be seen that there are no significant correlations between the time series elements of the dependent variable.

Assumption no. 4: absence of multicollinearity, which in the context of the GLM model means that there is no strong correlation between the predictors. Multicollinearity may also occur if very few data are available in comparison with the number of parameters to be assessed. In the event of multicollinearity, the parameter vector has no unique solution. The presence of multicollinearity may be detected by VIF[33] (Variance Inflation Factor), which represents a score calculated based on particular parameters for the selected predictor. VIF shows the degree of increase in the variance of regression model variable caused by multicollinearity. The smallest VIF value may equal 1, which means a complete absence of collinearity. VIF greater than 10 calls for particular actions toward reducing multicollinearity. In the language R, VIF values of all variables are calculated by the command *vif()*. The result obtained for both tested independent variables is the same and amounts to 1.22409, which indicates the absence of multicollinearity.

[32] Ljung, G. & Box, G. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika* Vol. 65, pp. 297–-303.

[33] Kassambara, A. (2017). *Machine Learning Essentials: Practical Guide in R*. Marseille: STHDA

Figure 6 Autocorrelation of Dependent Variable of Lapse Rates

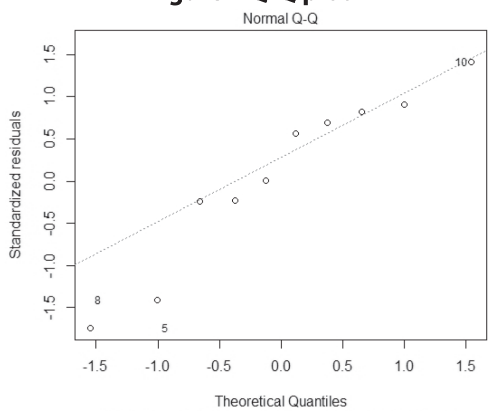


Source: Author's calculation

Assumption no. 5: The representativeness of a sample is based on the randomness of selected observations. This research will take into account all the available data on the policy lapses in the selected time interval so that representativeness is not an issue.

Assumption no. 6: Errors have a normal distribution. The assumption may be verified visually on the graph or by the Kolmogorov-Smirnov test which compares the sample from the model with normal distribution. Figure 7 shows the Q-Q plot obtained from the software package R, by the command *plot()*, where it can be seen that deviations exist, but are not significant.

Figure 7 Q-Q plot

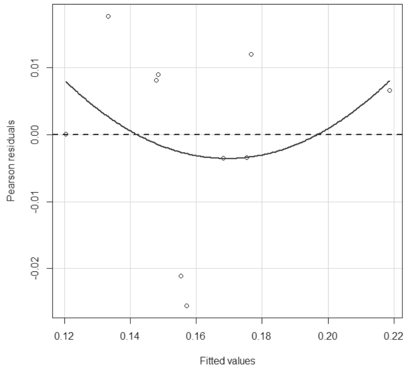


Source: Author's calculation

Pearson residuals or model residuals divided by the square root of the variance function are shown in Figure 8. They are obtained by the command *residualPlots()*. Residuals are not fully symmetric about 0 but deviations are acceptable because with minor changes, it is not possible to obtain a more efficient model.

Assumption no. 7: Homoscedasticity – homogeneity of error variance, which means that different variables have the same dispersion in their errors, regardless of the values of initial variables. Errors are heteroscedastic when the range of response variables is very broad. To check for a heterogeneous dispersion error, or when a pattern of residuals violates model assumptions of homoscedasticity, it is prudent to look for a so-called fanning effect[34] between residual error and predicted values. This is to say there will be a systematic change in the absolute or squared residuals when plotted against the predictive variables. Errors will not be evenly distributed across the regression line. In fact, residuals appear clustered and spread apart on their predicted plots for larger and smaller values along the linear regression line, and the mean squared error for the model will be wrong. A response variable whose mean is large will typically have a greater dispersion than the one whose mean is small. This may be checked by showing standardized errors and standardized predictors on the graph. In the event that the homoscedasticity assumption is not met, the model can still be used, but the quality of the results obtained from the model will be reduced. Homoscedasticity may also be verified by the Breusch-Pagan test which in the language R is realised by the command *bptest()*. The result of Breusch –Pagan test has p-value of 0.6704. The obtained p-value of the test is greater than 0.05, which indicates that the null hypothesis of homoscedasticity may not be rejected at the given level of significance.

Figure 8 Pearson Model Residuals



Source: Author's calculation

[34] Janković, D. (2014). *Regresija – Linearni modeli*. Term paper. Belgrade: Faculty of Mathematics of the University of Belgrade.

Assumption no. 8: Absence of error autocorrelation means that random errors are mutually independent. Autocorrelation was tested with Ljung-Box statistic in the language R by the command *Box.test()*. The result of the Ljung-Box test obtained has a p-value of 0.6996. Obtained tested p-value is greater than 0.05, which indicates that there is no autocorrelation in the errors of the model.

4. Model Application

Based on the selected model, the expected lapse rate in 2018 may be predicted as follows:

$$\hat{Y}\,(2018) = 0{,}10624 + 0{,}44828 * \mathit{GDP\ growth}\,(2016) + 0{,}47182 * \mathit{NBS\ ref.rate}\,(2017)$$

A prediction interval is the expected range of values that will contain the value of a random variable with particular reliability. The expected lapse rate in 2018 calculated according to the above formula represents the mean value of the prediction interval and thus, the prediction interval for the lapse rate in 2018 is with reliability α.

In the language R, the prediction interval with a confidence level of 99.5% is obtained by the command *predict.glm()*. The obtained result is shown in Table 6.

Table 6 Prediction interval obtained in the language R

	fit	lwr	upr
2016-01-10	*0.1375434*	*0.09015011*	*0.1849368*

Source: Author's calculation

Prediction interval which produces the developed model is [13.75% - 4.73%; 13.75% + 4.74%] or [9,01%; 18,49%]. Prediction interval produced by standard formula is [current lapse rate *0.5; current lapse rate *1.5] which is [7.55%; 22.64%].

Since the prediction interval of the developed partial internal model is lower than the prediction interval of the standard formula, the developed model measures the capital requirement more precisely, and results in the capital requirement which is lower than that in the prescribed standard formula.

VII. Conclusion

Nine factors influencing the lapse rate in the life insurance industry were analysed: GDP, reference interest rate, average wage, stock market index BelexLine, inflation, unemployment rate, life premium growth rate, non-life premium growth rate, and the growth rate of total insurance premium.

Based on the research of international authors mentioned in the section of this paper entitled *Overview of Literature on the Lapse Rate Factors*, the lapse rate can be adequately modelled based on the reference interest rate, gross domestic product per capita, unemployment rate, capital market development, and size of an insurance company. The research conducted with the data from the domestic market met the expectations that can be drawn from the research of international authors, showing that for the modelling of lapse rate dependency of a particular insurance company, the most adequate are the following factors: GDP growth and NBS reference rate. Used was the generalized linear model with normal distribution and identity function as the link function, which meets all required GLM assumptions.

Based on the conducted research, the author concludes that the developed partial internal model measures the capital requirement of a particular insurance company more precisely and has a lower capital requirement compared to that of the prescribed standard formula. In the application of the Solvency II regime in Serbia, Serbian insurance companies may use this conclusion for more adequate determination of solvency capital requirement by forming an internal partial model that would include the model for life insurance lapse risk.

Further course that the development of this research may take is to analyse the impact of internal factors of the policy on the lapse rate, such as: type of insurance, number of years passed since the conclusion of the contract, term of the contract, amount of premium, premium payment schedule, level of the sum insured, sales channel, age of policyholders, gender of policyholders, etc.

Literature

- Cerchiara, R.R., Edwards, M. & Gambini, A. (2008). Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II. *Giornale dell'Istituto Italiano degli Attuari 72,* pp. 100–122.
- Cheng, Ch., Hilpert, Ch. et al. (2020). Surrender Contagion in Life Insurance. *SSRN Electronic Journal* March 2020.
- Cox, S. & Lin, Y. (2006). *Annuity Lapse Rate Modelling: Tobit or Not Tobit?* Society of Actuaries. *http://library.soa.org.*
- Dar, A. & Dodds, C. (1989). Interest Rates, the Emergency Fund Hypothesis and Saving Through Endowment Policies: Some Empirical Evidence for the U.K. *Journal of Risk and Insurance 56*(3), pp. 415-433.
- Eling, M. & Kiesenbauer, D. (2013). What Policy Features Determine Life Insurance Lapse? An Analysis of the German Market. *The Journal of Risk and Insurance 81*(2), pp. 241–269.
- Gatzert, N. & Schmeiser, H. (2008). Assessing the Risk Potential of Premium Payment Option in Participating Life Insurance Contracts. *Journal of Risk and Insurance, 75*(3), pp. 691–712.

- Janković, D. (2014). *Regresija – Linearni modeli*. Term paper. Belgrade: Faculty of Mathematics of the University of Belgrade
- Kagraoka, Y. (2005). *Modelling Insurance Surrenders by the Negative Binomial Model.* Working paper.
- Kassambara, A. (2017). *Machine Learning Essentials: Practical Guide in R*. Marseille: STHDA
- Kiesenbauer, D. (2012). Main Determinants of Lapse in the German Life Insurance Industry. *North American Actuarial Journal 16*(1), pp. 52-73.
- Kim, C. (2005). Modelling Surrender and Lapse Rates with Economic Variables. *North American Actuarial Journal 9*(4), pp. 56-70.
- Kočović, J., Jovović, M. and Kočović, M. (2015). Aktuarski efekti prevremenog prekida ugovora o osiguranju života. In: *42nd International Symposium on Operations Research, SYM-OP-IS 2015*, pp. 77–84.
- Kuo, W., Tsai, C. and Chen, W.-K. (2003). An Empirical Study on the Lapse Rate: The Cointegration Approach. *Journal of Risk and Insurance 70*(3), pp. 489-508.
- Ljung, G. & Box, G. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika Vol. 65*, pp. 297-303.
- Michorius, C. (2011). *Modelling Lapse Rates – Investigating the Variables that Drive Lapse Rates*. Master Thesis. Enschede: Faculty of Management and Governance, University of Twente.
- Milhaud, X., Loisel, S. and Maume-Deschamps, V. (2010). *Surrender Triggers in Life Insurance: Classification and Risk Predictions*. Working paper.
- Mladenović, Z. (2015). *Ekonometrijski metodi i modeli - Dickey-Fuller-ov test jediničnog korena* Belgrade: Faculty of Economics of the University of Belgrade, pp. 9.
- Mladenović, Z. (2020). *KPSS test jediničnog korena* Belgrade: Faculty of Economics of the University of Belgrade, pp. 9.
- Nelder, J. & Wedderburn, R. (1972). Generalized linear models. *Journal of the Royal Statistical Society A 135*, pp. 370–384.
- Outreville, J. (1990). Whole-life Insurance Lapse Rates and the Emergency Fund Hypothesis. *Insurance: Mathematics and Economics 9*(4), pp. 249–255.
- Russel, D., Stephen, J. et al. (2013). An Empirical Analysis of Life Insurance Policy Surrender Activity. *Journal of Insurance Issues 36*(1), pp. 35–57.
- Insurance Law. *Official Gazette of RS,* no. 139/14.
- *https://www.belex.rs*
- *https://www.mfin.gov.rs*
- *https://www.nbs.rs*
- *https://www.stat.gov.rs*

Translated by: **Zorica Simović**