8 stycznia 2021 r.

Prof. dr hab. Mirosław Szreder
Kierownik Katedry Statystyki
Uniwersytetu Gdańskiego
ul. Armii Krajowej 101
81-824 Sopot

Recenzja rozprawy doktorskiej mgr. Kamila Gali pt. "Przestrzenne aspekty ryzyka ubezpieczeniowego w ubezpieczeniach komunikacyjnych", napisanej pod kierunkiem dra hab. Wojciecha Bijaka, prof. SGH

Autorem rozprawy doktorskiej pt. „Przestrzenne aspekty ryzyka ubezpieczeniowego w ubezpieczeniach komunikacyjnych" jest mgr. Kamil Gala. Rozprawa dotyczy ważnego w naukach aktuarialnych oraz w praktyce ubezpieczeniowej zagadnienia taryfikacji składek w ubezpieczeniach komunikacyjnych, z uwzględnieniem przestrzennego zróżnicowania czynników ryzyka. Zakres przedmiotowy badań został ograniczony do dwóch rodzajów ubezpieczeń komunikacyjnych: obowiązkowego ubezpieczenia odpowiedzialności cywilnej posiadaczy pojazdów mechanicznych (OC) oraz dobrowolnego ubezpieczenia Autocasco (AC). Słusznie Autor stwierdza (na s. 10), iż stosunkowo niewiele jest prac naukowych omawiających szczegółowo przestrzenne aspekty ryzyka ubezpieczeniowego i sposoby ich uwzględnienia w taryfikacji składek w ubezpieczeniach komunikacyjnych. Mgr Kamil Gala stara się w jakimś stopniu tę lukę wypełnić własnymi badaniami, które zaprojektował, scharakteryzował, i których wyniki omówił w przedłożonej do oceny rozprawie.

Za główny cel rozprawy doktorskiej mgr Kamil Gala przyjął: „ocenę istotności przestrzennych aspektów ryzyka ubezpieczeniowego w ubezpieczeniach komunikacyjnych w procesie ustalania składki za ubezpieczenie oraz wskazanie najlepszej metody wykorzystania przestrzennego charakteru danych ubezpieczeniowych" (s. 12). Autor jest świadom tego, że z jednej strony zróżnicowanie przestrzenne ryzyka ubezpieczeniowego jest ważne w ubezpieczeniach komunikacyjnych, a z drugiej – dostępnych jest coraz więcej danych statystycznych na ten temat, stąd propozycje wartościowych sposobów uwzględnienia tego ryzyka w taryfikacji składek powinny znaleźć uznanie wśród aktuariuszy i ubezpieczycieli. Aby podkreślić trzy wymiary celu swoich badań, Autor sformułował dodatkowo dwa cele poznawcze, dwa cele metodyczne i jeden cel praktyczny (por. s.13). Niezależnie od tego, punktem wyjścia do badań omówionych w rozprawie były dla Autora cztery hipotezy badawcze zapisane na s. 13. Uważam, że problematyka badawcza podjęta przez Doktoranta jest

oryginalna i aktualna, a ponadto ważna z punktu widzenia nauki o ubezpieczeniach, a także praktyki funkcjonowania ubezpieczeń. Poprawnie i jasno sformułowane zostały cele podjętych badań i hipotezy badawcze. Konfrontując ze sobą cel główny badań z brzmieniem hipotezy H4, mam jedynie wątpliwości, czy w obu przypadkach nie byłoby lepszym wyrażeniem „ocena znaczenia", zamiast „ocena istotności", co w przypadku H4 oznaczałoby zastąpienie słów: „jego właściwej oceny", słowami: „jego wpływu" (w obu przypadkach słowo „jego" odnosi się do przestrzennego aspektu ryzyka ubezpieczeniowego).

Rozprawa doktorska mgr. Kamila Gali składa się z czterech rozdziałów, jest napisana w sposób zwięzły, oszczędnym i precyzyjnym językiem, liczy 161 stron tekstu (bez załączników). We Wstępie Autor syntetycznie przedstawił motywy podjęcia badań na wskazany temat oraz scharakteryzował tło ekonomiczne i tło aktuarialne podjętej problematyki, przywołując kilka ważnych pozycji bibliograficznych z rynku krajowego i międzynarodowego.

Treść rozdziału 1 pt. „Taryfikacja a priori w ubezpieczeniach komunikacyjnych" stanowi wprowadzenie do zagadnień oceny ryzyka ubezpieczeniowego, w tym w szczególności do taryfikacji składek. Autor wyjaśnia różnice między czynnikami ryzyka *a priori* oraz czynnikami ryzyka *a posteriori* i deklaruje, że w rozprawie koncentrować się będzie na taryfikacji *a priori*. Konsekwentnie, w następnym punkcie rozdziału (1.2) omawia metody taryfikacji *a priori*, uwzględniając wśród nich te, które szczególnie silnie powiązane są z nowoczesnymi metodami statystycznymi, w tym wykorzystującymi modele regresji, metody *data mining*, sztuczne sieci neuronowe, i inne. W podrozdziale kończącym rozdział 1 scharakteryzowane zostało zagadnienie efektywności taryfikacji a priori. Ważne jest tu zastrzeżenie dokonane przez Autora, iż przez efektywność taryfikacji rozumie on jej skuteczność, czyli stopień, w jakim taryfikacja realizuje cele, które powinna wypełnić. Nie są tu bowiem omówione aspekty ekonomicznej efektywności taryfikacji składek. Moim zdaniem, podrozdział ten (1.3) prezentuje w nieco uproszczony sposób zagadnienie pomiaru skuteczności taryfikacji *a priori*. Autor nie zwraca m.in. uwagi na kwestie pomiaru homogeniczności grup taryfowych. Dość niejednoznacznie pisze również o trafności prognoz (nie jest jasne czego, s. 33), o kryterium ocen (czego?), i mało precyzyjnie o „szeroko rozumianej korelacji" (s. 33). Warto było temu zagadnieniu poświęcić więcej niż 3 stronice tekstu, bo – mimo, że wraca się do niego w rozdz. 3 – jego charakterystyka zawarta w rozdz. 1 uzasadnia w sposób niepełny słuszne stwierdzenie Autora, iż efektywność taryfikacji *a priori* jest zagadnieniem złożonym.

Ważnym elementem dalszych analiz Autora są kwestie wykorzystania danych przestrzennych w ubezpieczeniach, w tym w szczególności w problematyce szacowania ryzyka

ubezpieczeniowego. Treść tego rozdziału wiąże się ściśle z głównym celem badawczym podjętym w rozprawie. W trzech podrozdziałach rozdziału 2, zatytułowanego: „Dane przestrzenne i metody ich analizy", Autor omawia definicje i klasyfikacje danych przestrzennych, charakteryzuje znaczenie aspektu przestrzennego danych dla adekwatnej oceny ryzyka ubezpieczeniowego oraz przedstawia metody analizy i modelowania danych przestrzennych. W zawartych w tym rozdziale prezentacjach i charakterystykach są obecne, chociaż niezbyt licznie, elementy dyskusji i krytycznej oceny metodyki statystycznej analizy danych przestrzennych i przydatności tych danych w modelowaniu zjawisk ubezpieczeniowych. Szczególną uwagę zwraca Autor na kwestie przestrzennej autokorelacji, która może uniemożliwiać zastosowanie niektórych standardowych metod modelowania i wnioskowania statystycznego. Cały ten rozdział stanowi konieczne (w kontekście celu pracy) i dobrze napisane wprowadzenie do problematyki, którą można określić jako: specyfika i modelowanie danych przestrzennych. Autor nie ustrzegł się pewnych uproszczeń, ale nie jest ich dużo. Jednym z nich jest zdanie oznajmiające, iż „w ubezpieczeniach komunikacyjnych obecnie największe znaczenie mają dane obszarowe, przede wszystkim ze względu na stosunkową łatwość ich zbierania i analizowania". Nie uważam, aby był to główny powód wykorzystania danych przestrzennych w ubezpieczeniach komunikacyjnych.

Rozdziały 3 i 4 zawierają treści stanowiące realizację głównego celu badawczego Doktoranta. Są w nich zawarte oryginalne propozycje Autora, odnoszące się do modelowania taryfikacji składek na podstawie danych statystycznych, uwzględniających ich przestrzenne zróżnicowanie.

W rozdziale 3 pt. „Modele taryfikacji *a priori* uwzględniające efekty przestrzenne" została zawarta prezentacja i dyskusja wybranych modeli taryfikacji *a priori* oraz propozycja Doktoranta, odnosząca się do rozszerzenia (uogólnienia) jednej z grup omawianych modeli. Autor scharakteryzował najpierw uogólnione modele liniowe: standardowy (GLM) oraz przestrzenny (SGLM), formułując kilka najważniejszych warunków (postulatów), jakie powinien spełniać model taryfikacji *a priori* uwzględniający przestrzenne zróżnicowanie ryzyka ubezpieczeniowego. Przedyskutowane zostały przez Doktoranta także kwestie estymacyjne parametrów takiego modelu. W dalszej części tego rozdziału mgr Kamil Gala zaproponował i analitycznie uzasadnił rozszerzenie uogólnionego, znanego z literatury modelu Bühlmanna-Strauba. Uogólnienie to polega na uchyleniu warunku niezależności ukrytych czynników ryzyka, co wydaje się konieczne w okolicznościach, gdy jednostkami badania są obszary (regiony) geograficzne. Występowanie autokorelacji przestrzennej jest wówczas bardzo prawdopodobne. Mając to na uwadze, Doktorant sformułował i w sposób analityczny,

nie budzący zastrzeżeń, udowodnił dwa oryginalne twierdzenia dotyczące estymacji parametrów uogólnionego modelu Bühlmanna-Strauba (twierdzenia 4.1 i 4.2 na s. 75-77). Autor słusznie proponuje także rozważenie kalibracji parametrów modelu, w miejsce estymacji, głównie z racji scharakteryzowanych w p. 3.2.2 trudności w oszacowaniu elementów macierzy wariancji-kowariancji. W kolejnym fragmencie rozdziału 3 Autor rozwija koncepcję modelu z czynnikiem wielopoziomowym MLF (ang. *multi-level factor*), za który mogą być uznane np. jednostki geograficzne wynikające z podziału administracyjnego kraju. Rozwinięcie to polega na uwzględnieniu w modelu MLF efektów przestrzennych. W ostatniej części tego rozdziału zostały poprawnie scharakteryzowane metody segmentacji przestrzennej i ich możliwe zastosowania do konstrukcji zmiennych taryfowych o charakterze geograficznym. Obok dobrze znanej metody k-średnich, Doktorant zaproponował własną metodę grafową, której istota sprowadza się do zastosowania algorytmu wyszukiwania społeczności w tzw. grafie nieskierowanym. Spójnym logicznie zakończeniem rozdziału 3 jest szczegółowe omówienie w p. 3.5 sposobów pomiaru efektywności taryfikacji a priori z uwzględnieniem specyfiki scharakteryzowanych wcześniej uogólnionych modeli liniowych. Na podkreślenie zasługuje duży stopień oryginalności treści całego rozdziału 3 i zaproponowanych rozwiązań metodycznych. Wśród propozycji najbardziej wartościowych rozwiązań wymienić należy: zdefiniowanie przestrzennego uogólnionego modelu liniowego (który może być traktowany jako wyjściowy do konstruowania szczegółowych modeli odpowiadających przyjętym warunkom a priori), zaprezentowanie rodziny modeli z czynnikiem wielopoziomowym (pozwalającym na uwzględnienie różnych rodzajów struktury zależności przestrzennej), a ponadto scharakteryzowanie własnej metody segmentacji przestrzennej (grafowej), która może stanowić prostszą w zastosowaniach alternatywę dla modeli z czynnikiem wielopoziomowym.

Rozdział 4 ma charakter empiryczny, w którym do identyfikacji wzorców przestrzennego rozkładu ryzyka ubezpieczeniowego oraz do estymacji omówionych w rozdziale 3 modeli wykorzystane zostały dane statystyczne pochodzące z bazy Ubezpieczeniowego Funduszu Gwarancyjnego (UFG). Treść tego rozdziału stanowi jednocześnie ilustrację i weryfikację zaproponowanych wcześniej przez Doktoranta rozwiązań. Jest także ważnym elementem realizacji celów rozprawy doktorskiej. Mgr Kamil Gala scharakteryzował najpierw zbiór danych UFG oraz szczegółowo przedstawił przyjętą metodykę analizy i etapy badania. Wśród nich wyróżnił: analizę eksploracyjną, estymację parametrów modeli przestrzennych oraz ocenę i porównanie efektywności modeli z efektami przestrzennymi. Walidacji modeli Doktorant dokonał na trzech poziomach: portfela ubezpieczeń, podportfeli określonych przez wybrane zmienne taryfowe oraz indywidualnym.

Opis metodyki badawczej i kolejno realizowanych etapów badania empirycznego uważam za wzorowy. Dla osoby śledzącej tekst rozdziału 4, każdy kolejny krok badania jest wcześniej zapowiedziany i scharakteryzowany, precyzyjnie opisana jest jego realizacja i interesująco – za pomocą zestawień liczb, wykresów i map – przedstawione i zinterpretowane są uzyskane wyniki. Z punktu widzenia poznawczego wyniki te uznać należy za oryginalne i dobrze naukowo uzasadnione. Większość z nich może być także z pożytkiem wykorzystana do weryfikacji przestrzennego zróżnicowania ryzyka w ubezpieczeniach OC i AC przez zakłady ubezpieczeń.

Do głównych osiągnięć badawczych mgr. Kamila Gali omówionych w rozprawie doktorskiej zaliczam:

1) zaproponowanie i scharakteryzowanie właściwości, w tym estymacyjnych, uogólnionego modelu liniowego, uwzględniającego aspekt przestrzenny;
2) zaprezentowanie autorskiej koncepcji uogólnienia modelu Bühlmanna-Strauba poprzez uchylenie założenia o niezależności ukrytych czynników ryzyka;
3) scharakteryzowanie rodziny modeli z czynnikiem wielopoziomowym uwzględniających różne rodzaje struktur zależności przestrzennej;
4) empiryczne uzasadnienie istnienia przestrzennego zróżnicowania ryzyka ubezpieczeniowego w ubezpieczeniach komunikacyjnych w Polsce oraz scharakteryzowanie korelacji między poziomem ryzyka w różnych obszarach geograficznych;
5) wykorzystanie modeli MLF, HMLF i SMLF do wykazania, iż precyzja wnioskowania o ryzyku ubezpieczeniowym na badanym obszarze może być zwiększona poprzez wykorzystanie danych z pozostałych obszarów (idea zbliżona do istoty statystyki małych domen);
6) przekonujące uzasadnienie korzyści wynikających z uwzględnienia wymiaru przestrzennego w modelu taryfikacji składek *a priori* w kontekście jego zdolności predykcyjnych, a także biorąc pod uwagę efektywność procesu taryfikacji.

Uważam, że Autor osiągnął w pełni założone cele badawcze i wniósł opisanymi w rozprawie badaniami oryginalny wkład w rozwój wiedzy na temat aktuarialnych zagadnień modelowania danych przestrzennych i ich praktycznego wykorzystania w procesie taryfikacji składek przez zakłady ubezpieczeń. Mgr Kamil Gala wykazał się dobrą znajomością literatury z zakresu ubezpieczeń oraz zdolnością do krytycznej jej oceny w interesujących go aspektach. Posiadł wiedzę teoretyczną z zakresu ekonomii i finansów, pozwalającą na samodzielne prowadzenie badań naukowych w tej dyscyplinie. Badaniami zawartymi w rozprawie udowodnił, że jest w

stanie zaprojektować i wykonać oryginalne, wartościowe dla nauki i praktyki badanie naukowe. Wysoko oceniam wartość merytoryczną całości badań omówionych w rozprawie. Doktorant udoskonalił i rozwinął metodykę modelowania statystyczno-ekonometrycznego z wykorzystaniem danych przestrzennych, w szczególności (chociaż nie wyłącznie) w odniesieniu do zagadnień ubezpieczeniowych. Swoje propozycje metodyczne zweryfikował i ocenił w oparciu o empiryczne analizy, wykorzystując do tego dane rzeczywiste z polskiego runku ubezpieczeniowego.

Rozprawa charakteryzuje się wysokim poziomem edytorskim i napisana jest poprawnym, precyzyjnym językiem. Nieliczne dostrzeżone przeze mnie błędy są m.in. w słowach „możliwe" (zamiast możliwie) na s. 33, „zależy" (zamiast zależą) na s. 38, „stosowanych" (zamiast stosowanym) na s. 52. W niektórych miejscach rozdziałów 2 i 3 uzupełnienia wymagają objaśnienia części symboli we wzorach.

Merytorycznych zastrzeżeń do treści rozprawy nie zgłaszam. Jedyną sugestię merytoryczną, odnoszącą się do potrzeby szerszego i bardziej precyzyjnego omówienia zagadnienia efektywności taryfikacji składek, zawarłem na s. 2 niniejszej recenzji.

Uważam, że wysoki poziom merytoryczny rozprawy doktorskiej mgr. Kamila Gali, zarówno w warstwie metodycznej, jak i empirycznej, uzasadniają mój wniosek o uznanie tej rozprawy za wyróżniającą. Dla wysunięcia tego wniosku znaczenie ma także fakt, iż podjęty w rozprawie temat jest aktualny i ważny w problematyce aktuarialnej, a wnioski wynikające ze zrealizowanych badań są dobrze uzasadnione naukowo i mogą znaleźć zastosowanie w praktyce ubezpieczeniowej.

Opowiadam się za dopuszczeniem mgr Kamila Gali do publicznej obrony rozprawy doktorskiej pt. „Przestrzenne aspekty ryzyka ubezpieczeniowego w ubezpieczeniach komunikacyjnych", przygotowanej po kierunkiem dr. hab. Wojciecha Bijaka, prof. SGH. Wnioskuję ponadto o uznanie rozprawy za wyróżniającą i rozważenie jej opublikowania (po uwzględnieniu uwag recenzentów).

Mirosław Szreder