

Poznań, 08 sierpnia 2022 r.

dr hab. Grażyna Dehnel, prof. UEP  
Uniwersytet Ekonomiczny w Poznaniu  
Instytut Informatyki i Ekonomii Ilościowej  
Katedra Statystyki  
61-875 Poznań, al. Niepodległości 10

**Recenzja rozprawy doktorskiej mgr Małgorzaty Krzciuk  
pt. „Podejście modelowe w statystyce małych obszarów i jego zastosowania  
w badaniach ekonomicznych”  
napisanej pod kierunkiem naukowym dra hab. Tomasza Żądło, prof. UE,  
Uniwersytet Ekonomiczny w Katowicach**

***Podstawa formalna***

Podstawę formalną przygotowania niniejszej recenzji stanowią:

- Ustawa o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (DZ.U. nr 65, poz.595 z późn. zm.)
- Pismo Przewodniczącego Komitetu Naukowego Dyscypliny Ekonomia i Finanse Uniwersytetu Ekonomicznego w Katowicach prof. dra hab. Grzegorza Kończaka z dnia 13 lipca 2022 roku (wraz z załączoną dokumentacją)

***Tematyka pracy***

Przedstawioną do recenzji pracę doktorską poświęcono nieklasycznym metodom szacunku parametrów podpopulacji ze szczególnym uwzględnieniem obszaru analiz ekonomicznych. Tematyka opracowania wpisuje się w nurt badań naukowych dotyczących metodologii estymacji charakterystyk opartej na małolicznych, czy nawet zerowych próbach.

Obserwowany dynamiczny wzrost popytu na szczegółową i dostarczaną systematycznie informację, charakteryzującą uwarunkowania gospodarcze i społeczne zapoczątkował etap poszukiwań metod estymacji, które sprostałyby tak określonym wymaganiom. Szacunek charakterystyk dla jednostek na niskim poziomie agregacji stanowi często szczególny rodzaj wyzwania ze względu na konieczność spełnienia różnych metodycznych założeń, włączając w

to odpowiednio dużą liczebność próby. Stąd potrzeba poszukiwania nieklasycznych technik estymacji, które na podstawie informacji pochodzących od niewielu jednostek dostarczą wiarygodnych szacunków. Jedną z propozycji z pewnością stanowić mogą metody predykcji zaproponowane w rozprawie doktorskiej. Można je zaliczyć do grupy metod stosowanych w ramach statystyki małych obszarów (SMO), reprezentujących podejście modelowe. Metody SMO w ostatnim czasie nabierają szczególnego znaczenia głównie z tego powodu, że ich stosowanie nie wymaga pozyskiwania dodatkowych informacji. Wykorzystują one bowiem w formie tzw. zmiennych pomocniczych dostępne źródła danych – tak więc prowadzi to do znacznej redukcji kosztów analizy. Ponadto w podejściu modelowym dopuszcza się wykorzystanie informacji pochodzących nie tylko z prób losowych, ale również celowych. Dodatkowo, co stanowi cenny wkład rozprawy, w przypadku proponowanych w pracy metod predykcji uwzględniono wykorzystanie danych wielookresowych. Przedstawiona propozycja metodyczna wychodzi naprzeciw zgłaszanym przez analityków, jednostki badawcze oczekiwaniom. Oceniając zatem wartość pracy z punktu widzenia wagi podjętej problematyki stwierdzam, że wybór tematu rozprawy jest trafny i w pełni uzasadniony.

### ***Cel i struktura pracy***

Cele zostały w rozprawie określone jasno i precyzyjnie. Zdefiniowano jeden cel główny o charakterze teoretyczno-poznawczym. Uzupełnia go pięć bardziej szczegółowo sformułowanych celów o wymiarze praktycznym dotyczących adaptacji, zastosowania i oceny wskazanego w ramach metod statystyki małych obszarów podejścia modelowego. W szczególności cele te koncentrują się na autorskich propozycjach metodycznych zgłoszonych w rozprawie przez Doktorantkę. Stanowią również podstawę do sformułowania czterech pytań badawczych. W mojej ocenie zarówno cele, jak i pytania zostały określone w sposób właściwy. Autorka odwołuje się w nich do kwestii metodycznych zarówno od strony czysto teoretycznej, jak i praktycznego wykorzystania dostępnych źródeł danych w celu „wzmocnienia” procesu estymacji.

Rozprawa doktorska ma właściwą strukturę. Praca liczy 216 stron. Złożona jest ze wstępu, pięciu rozdziałów, zakończenia oraz spisu literatury. W pracy wyróżniłabym trzy części. Pierwszą stanowią dwa pierwsze rozdziały wprowadzające w problematykę badawczą. Dwa

kolejne rozdziały mają charakter metodyczny. W ostatniej części - rozdziale piątym - przedstawiono wyniki badań symulacyjnych.

Kolejne części rozprawy składają się na logiczną całość, która podporządkowana jest realizacji celu głównego. Układ treści poszczególnych rozdziałów nawiązuje do sformułowanych we *Wstępie* problemów badawczych. Autorka w kolejnych częściach pracy podejmuje próbę odpowiedzi na postawione pytania badawcze.

Pierwszy rozdział wprowadza w tematykę statystyki małych obszarów. Przedstawiono w nim podstawy teoretyczne, definicje podstawowych pojęć, wyróżnione w literaturze podejścia. Szczególną uwagę poświęcono podejściu modelowemu oraz charakterystyce liniowych modeli mieszanych co jest w pełni uzasadnione z punktu widzenia celu pracy. W dalszej części rozdziału w syntetyczny sposób przedstawiono rozwój metodyki statystyki małych obszarów, co ważne, podkreślając polski wkład. Opisano wybrane obszary praktycznych zastosowań SMO odwołując się szczególnie do przykładów przeprowadzonych analiz o charakterze ekonomicznym. Na uwagę zasługują zgłoszone w tej części pracy, w obszarze metodyki badań, autorskie propozycje:

- wykorzystania w badaniach wielookresowych liniowych modeli mieszanych ze skorelowanymi wektorami efektów losowych,
- uwzględnienia tej klasy modeli w ramach SMO,
- uogólnienia wybranych predyktorów na przypadek danych przekrojowo–czasowych,
- wykorzystania w weryfikacji istotności parametrów zaproponowanych modeli testów permutacyjnych, włączając w to permutacyjne wersje klasycznych testów,
- testu opartego na parametrycznej metodzie bootstrap, pozwalającego na weryfikację występowania zależności pomiędzy efektami losowymi.

Rozdział drugi poświęcono charakterystyce potencjalnych źródeł danych. Przedstawiono badania statystyczne z uwypukleniem podziału na badania jedno i wielookresowe, przy czym ze względu na analizowaną w pracy metodę predykcji, główny nacisk położono na badania powtarzalne w czasie. Wartościową część rozdziału stanowi podrozdział 2.4, w którym dokonano przeglądu zalet, wad i ograniczeń badań wielookresowych. W tej części pracy wskazano także metody pozwalające na uniknięcie lub złagodzenie problemów, które mogą się pojawić przy prowadzeniu badań powtarzalnych w czasie.

W rozdziale trzecim przedstawiono podstawy teoretyczne metod szacunku w oparciu o predyktory typu BLU oraz EBLU. Szczególną uwagę zwrócono na te klasy predyktorów z punktu

widzenia klasyfikacji liniowych modeli mieszanych na modele typu A oraz B. W przypadku każdego z predyktorów wskazano także na przyjmowane założenia oraz postać błędu średniokwadratowego predykcji. Zaprezentowano również przykłady zastosowania omówionych predyktorów. Jedną z części rozdziału poświęcono estymacji błędu średniokwadratowego predyktorów typu EBLU. Analizą objęto zarówno podejście klasyczne, jak i estymatory oparte na parametrycznej metodzie bootstrap oraz metodzie jackknife. W opisie uwzględnione zostały także własności prezentowanych estymatorów.

Tematem podrozdziału 3.5 jest charakterystyka wybranych modyfikacji predyktora typu EBLU. Wśród nich znalazły się: predyktor odporny REBLUP, predyktor uwzględniający korelacje przestrzenne efektów losowych SEBLUP, predyktor będący połączeniem dwóch wyżej wskazanych modyfikacji – SREBLUP oraz geograficznie ważony predyktor - GWEBLUP, jego odporny wariant - RGWEBLUP, pseudo-EBLUP i nieparametryczna wersja predyktora EBLU - NPEBLUP. Opis predyktorów dopełnia wskazanie zarówno ich zalet, jak i wad oraz przykłady zastosowań.

Podkreślić należy przedstawienie w tej części pracy przez Autorkę kolejnych autorskich propozycji dotyczących:

- wykorzystania predyktorów typu EBLU przy założeniu liniowego modelu mieszanego ze skorelowanymi wektorami efektów losowych w statystyce małych obszarów włącznie z przykładem,
- modyfikacji znanych metod oceny błędu średniokwadratowego predykcji pozwalających na estymację dokładności predyktorów typu EBLU uwzględniających korelacje między wektorami efektów losowych.

W rozdziale czwartym zaprezentowano dwie klasy predyktorów – empiryczne najlepsze predyktory włącznie z odniesieniem się do ich aproksymacji metodą Monte Carlo oraz predyktory typu plug-in. Przedstawiono tu także autorskie propozycje wykorzystania predyktorów należących do tych klas przy założeniu liniowych modeli mieszanych ze skorelowanymi efektami losowymi. W rozważaniach uwzględniono problem oceny błędów średniokwadratowych zaproponowanych predyktorów. Całość dopełniają przykłady zastosowań omawianych predyktorów w badaniach o charakterze ekonomicznym.

Rozdział piąty poświęcono badaniom symulacyjnym. Jest to przykład rzetelnie opisanego analizy począwszy od charakterystyki zbioru danych, poprzez przyjęte założenia, aż po interpretację i ocenę otrzymanych wyników. Przeanalizowane zostały trzy warianty badania. W każdym z

nich z użyciem liniowego modelu mieszanego uwzględniającego korelacje pomiędzy efektami losowymi poddano ocenie dwie charakterystyki domeny studiów: wartość globalną oraz medianę. W pierwszym podejściu wygenerowano dane populacyjne badanej zmiennej w oparciu o rzeczywiste wartości cechy dodatkowej oraz oszacowano parametry modelu uwzględniającego poziom korelacji między efektami losowymi na poziomie  $\rho = -0,83$ . W pozostałych dwóch podejściach modyfikowano siłę zależności między wektorami efektów losowych na etapie generowania danych populacyjnych określając ją odpowiednio na poziomie silniejszym ( $\rho = -0,95$ ) i słabszym ( $\rho = -0,65$ ) w porównaniu z pierwszym wariantem. Wartości współczynników korelacji ustalono na podstawie badania wstępnego. Przeanalizowano własności trzech autorskich propozycji predyktorów należących dla klas BP, EBP oraz plug-in porównując je między innymi z wybranymi predyktorami zakładającymi brak zależności między efektami losowymi oraz wybranymi estymatorami. Wśród wielkości pozwalających na przeprowadzenie analizy własności powyższych metod szacowania charakterystyk domen znalazły się mierniki precyzji predykcji, dokładności predykcji oraz obciążenie predyktora.

Oryginalny wkład stanowią również autorskie kody do wyznaczania predyktorów BP i EBP wartości globalnej oraz mediany w domenie, przy założeniu liniowego modelu mieszanego ze skorelowanymi efektami losowymi oraz do prowadzenia badań symulacyjnych napisane w języku R.

Rozprawę doktorską wieńczy *Zakończenie*, w którym w syntetyczny sposób przedstawiono jej walory i osiągnięcia, odniesiono się do postawionych pytań badawczych oraz wskazano kierunki dalszych badań. Poza teoretycznym, warto podkreślić także praktyczny wymiar rozprawy. Otrzymane wyniki badań mogą być przydatne zarówno jednostkom prowadzącym badania, jak i organom administracji państwowej różnego szczebla.

Rozprawa doktorska oparta jest na bogatej bibliografii, w której znaczącą część stanowi literatura angielskojęzyczna. Są to zarówno książki i artykuły w czasopismach naukowych, jak i różnorodne opracowania i raporty.

Podsumowując można stwierdzić, że struktura całej pracy jest poprawna. Podział na rozdziały logiczny i odpowiadający rozwiązaniu problemu badawczego.

### ***Ocena merytoryczna pracy***

Ogólna ocena rozprawy doktorskiej jest pozytywna. Problem badawczy został określony i rozwiązany poprawnie. Praca stanowi indywidualny i oryginalny wkład do problematyki pomiaru i oceny charakterystyk domen wyróżnionych na niskim poziomie agregacji przestrzenno-rzeczowej. W mojej ocenie jest to przykład rzetelnie zaprojektowanego, przeprowadzonego i opisanego badania. Praca została przygotowana bardzo starannie. Wśród wartościowych osiągnięć, które w moim uznaniu zasługują na podkreślenie znalazły się:

- dobrze postawiony i rozwiązany problem badawczy, właściwe sformułowanie celu pracy oraz pytań badawczych,
- tematyka badawcza – rozprawa wypełnia lukę, jaką stanowi brak prac podejmujących problematykę estymacji parametrów na podstawie danych pochodzących nie tylko z badań częściowych, dla jednostek wyróżnionych na niskim poziomie agregacji,
- znaczny wkład autorski do teorii estymacji charakterystyk podopopulacji w zakresie podejścia modelowego obejmujący zarówno metodykę predykcji parametrów, jak i jej oceny,
- w przypadku proponowanych w pracy metod predykcji uwzględniono wykorzystanie danych wielookresowych,
- w treści rozdziałów liczne powołania na literaturę przedmiotu – świadczy to o bogatych studiach bibliograficznych (ponad 360 pozycji głównie angielskojęzycznych),
- rozważania teoretyczne podparto licznymi przykładami z praktyki badań z obszaru ekonomii,
- wzorcowa struktura rozprawy – kolejne części rozprawy składają się na logiczną całość na końcu każdego rozdziału zawarto podsumowanie, w którym w syntetyczny sposób zebrano dyskutowane w nim problemy badawcze.

W świetle przedstawionej powyżej oceny rezultatów recenzowanej pracy stwierdzam, że rozprawa doktorska mgr Małgorzaty Krzciuk pt. *„Podejście modelowe w statystyce małych obszarów i jego zastosowania w badaniach ekonomicznych”* spełnia wymagania stawiane rozprawom doktorskim określone w ustawie o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki. Tym samym wnoszę o dopuszczenie Pani Małgorzaty Krzciuk

do dalszych etapów przewodu doktorskiego przeprowadzanego w dziedzinie nauk ekonomicznych, w dyscyplinie ekonomia.

Ponadto, biorąc pod uwagę walory naukowe oraz praktyczne recenzowanej rozprawy doktorskiej zamieszczone w recenzji wnioskuję o jej wyróżnienie.

Poznań, 08 sierpnia 2022 r.

*Grzegorz Delu*