

Recenzja pracy doktorskiej
mgr. Wojciecha Starosty
pt „Modeling the Loss Given Default of Retail Contracts”
przygotowanej na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego
pod kierunkiem naukowym
dr. hab. Pawła Baranowskiego, prof. UŁ oraz dr. Mariusza Górajskiego

Niniejsza recenzja została sporządzona na wniosek pana dr hab. Rafała Matery, prof. UŁ, dziekana Wydziału Socjologiczno-Ekonomicznego Uniwersytetu Łódzkiego. Ramy prawne recenzji określa ustawa z dn. 20.07.2018 Prawo o szkolnictwie wyższym i nauce (Dz. U., poz. 1668 z 20.07.2018), zaś jej celem jest w szczególności ocena spełnienia przez Doktoranta wymogów określonych ogólnie w art. 187 Ustawy.

Rozprawa została przygotowana w języku angielskim i składa się z czterech opublikowanych wcześniej artykułów (których jedynym autorem jest Doktorant) oraz poprzedzającego je wprowadzenia. Rozwiązanie to jest jak najbardziej dopuszczone przez ustawę PSWiN pod warunkiem spójności tematycznej artykułów składających się na cykl. Ta nie budzi wątpliwości: zgodnie z samym tytułem rozprawy, wszystkie teksty poświęcone są problematyce modelowania strat z tytułu niewykonania zobowiązania (*loss given default*, LGD) i zawierają cztery oryginalne propozycje rozszerzenia istniejącej metodologii, które omawiam w dalszej części recenzji. Poszczególne eseje nie są więc redundantne (z dokładnością do – w pewnym stopniu – opisów istniejących metod i ich założeń, co nieuniknione zważywszy na fakt, że każdy z nich oryginalnie stanowił odrębny byt), ale jednocześnie ich wspólny mianownik nie wzbudza żadnych wątpliwości.

Pierwszy z tekstów (poza, rzecz jasna, wprowadzeniem) wchodzących w skład rozprawy, [1] „Modelling Recovery Rate for Incomplete Defaults Using Time Varying Predictors”, poświęcony jest problemowi niezakończonych przypadków braku możliwości regulowania zobowiązań. Estymacja LGD powinna być oparta na wszystkich przypadkach braku możliwości regulowania zobowiązań w określonym z góry okresie, także tych, w przypadku których proces odzyskiwania należności nie dobiegł jeszcze końca, w przeciwnym bowiem wypadku uwzględnienie w próbie jedynie przypadków w pełni zakończonych może –

i zapewne będzie – oznaczać autoselekcję próby, a w konsekwencji brak reprezentatywności uzyskanych wyników w skali pełnej populacji. W próbie znajdują się bowiem zapewne wszystkie przypadki w których wartość zabezpieczenia umożliwi szybkie i pełne odzyskanie wierzytelności, zaś pominięte zostaną przypadki, w których z uwagi na niewystarczającą wartość zabezpieczenia proces odzyskiwania wierzytelności będzie długotrwały. W konsekwencji niemal na pewno dojdzie do niebezpiecznego z punktu widzenia systemu zaniżenia LGD. Aby tego uniknąć, w artykule zaproponowano sposób estymacji umożliwiający wykorzystanie także przypadków niezakończonych. Autor dzieli całą próbę uwzględniając czas odzyskiwania wierzytelności na podpróby, po czym dokonuje estymacji RR (stopy odzyskania) – a w konsekwencji LGD – korzystając w każdym przypadku jedynie z odpowiedniej podpróby, co niejako przy okazji dodatkowo umożliwia zaobserwowanie zróżnicowanego wpływu poszczególnych determinant LGD w różnym czasie od wystąpienia braku możliwości regulowania zobowiązań. Należy podkreślić, że w artykule bardzo dużą część stanowi właśnie analiza empiryczna, w której Autor wykorzystuje dużą bazę klientów banku jako podstawę obliczeń.

W artykule [2] „Beyond contract. Client behavior from origination to default as the new set of the Loss Given Default risk drivers” Doktorant koncentruje się na identyfikacji czynników determinujących LGD. Zainteresowanie badaczy LGD dopiero niedawno zaczęło wypierać dominujące analizy poświęcone prawdopodobieństwu niewykonania zobowiązania. W konsekwencji, nie ma konsensusu co do zestawu cech, które w kształtowaniu LGD można uznać za kluczowe (można zresztą spodziewać się, że nie dojdzie do szybkiego ustabilizowania tego zestawu, gdy wraz z rozwojem efektywnych technik textminingowych możliwe stanie się uwzględnianie jako kolejnych czynników choćby nastroju szacowanego na podstawie komunikatów tekstowych zamieszczanych w mediach społecznościowych). Doktorant podejmuje próbę poszerzenia obecnego stanu wiedzy, koncentrując się na samym zestawie determinant LGD bez ingerencji w istniejące rozwiązania ekonometryczne.

Tekst [3] „LGD decomposition using mixture distributions of in-default events” oparty jest na spostrzeżeniu, że przy modelowaniu LGD (czy właściwie stopy odzyskania, $RR=1-LGD$) warto uwzględnić trzy możliwe sytuacje, jakie mogą mieć miejsce w przypadku wystąpienia braku możliwości regulowania zobowiązań: „uzdrowienie”, częściowe odzyskanie oraz spisanie. W przypadku każdej z nich rozkład RR jest inny i powinien podlegać odrębnej estymacji, ponadto zaś częściowe odzyskanie oraz spisanie to zdarzenia, które wzajemnie się dopełniają i wykluczają, a warunkiem ich wystąpienia jest brak możliwości uzdrowienia. Wszystko to razem prowokuje Doktoranta do sformułowania propozycji modelowania RR

(czyli pośrednio – LGD) opierając się na pięciu osobno szacowanych konstrukcjach (prawdopodobieństwach poszczególnych stanów i warunkowych wartościach oczekiwanych RR). Jakkolwiek sam pomysł od strony koncepcyjnej nie jest skomplikowany o wiele bardziej niż wzór na prawdopodobieństwo całkowite, który jest silnikiem w tej maszynie, to w praktyce taka z pozoru nieskomplikowana dekompozycja umożliwia uwzględnienie domniemanej niejednorodności rozkładów LGD w przypadku każdej z rozpatrywanych dalszych ścieżek (uzdrowienie, częściowy odzysk, spisanie), a w szczególności zróżnicowanie jego determinant. Warto zauważyć, że w miejsce klasycznej regresji Doktorant wykorzystuje podejście SVM.

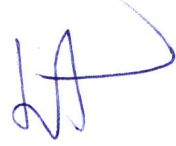
Ostatni z tekstów, [4] „Forecast combination approach in the Loss Given Default estimation”, w głównej mierze koncentruje się na znanym problemie efektywnego łączenia prognoz. W jaki sposób – i czy w ogóle – uśrednić prognozy LGD uzyskane na podstawie odrębnych modeli, z których w każdym uwzględnia się inny typ regresorów? Na to pytanie próbuje odpowiedzieć Doktorant w tym od strony ilościowej najmniej wysublimowanym (to jedyny tekst w którym do estymacji wykorzystano w całości MNK jako metodę podstawową, a nie jako benchmark), ale zgrabnym, krótkim artykule, konkludując, że owszem, łączenie prognoz z różnych modeli prowadzi do poprawy łącznej trafności. Jakkolwiek wniosek ten mógł być przewidziany z góry, wartościowa jest empiria jaką do tego celu przedstawia mgr Starosta.

Temat dysertacji jest niewątpliwie znakomity z co najmniej dwóch powodów. Po pierwsze, cały obszar jakiego dotyczy praca jest zarówno bardzo ciekawy od strony naukowej, jak i istotny od strony praktycznej: znaczenie analizowanej problematyki uwypuklają dodatkowo doświadczenia kryzysu finansowego sprzed ponad dekady. Po drugie, mimo iż mgr Starosta jest jedną z wielu osób analizujących omawianą w pracy tematykę, jest ona jednak na tyle szeroka i stosunkowo jeszcze nowa, że pozostawia przestrzeń na liczne oryginalne rozwiązania.

Praca stanowi bardzo rozsądny kompromis między analizami czysto empirycznymi, których autorzy w niewielkim lub żadnym stopniu poszerzają istniejące metody, koncentrując się na ich replikacji z wykorzystaniem nietypowych zbiorów danych czy też w szczególnie ciekawych kontekstach, a badaniami, w których proponuje się wyłącznie nowe rozwiązania teoretyczne, często na tyle złożone, że ich zastosowanie w praktyce jest bardzo trudne lub w ogóle niemożliwe. Rozszerzenie istniejących podejść, jakie proponuje mgr Starosta, oparte jest na celnych spostrzeżeniach dotyczących samego procesu i zgrabnym połączeniu istniejących rozwiązań metodologicznych w sposób, jaki do tej pory nie był praktykowany – w konsekwencji, efektem pracy nie jest wprowadzenie nowej złożonej metody estymacji lub

rozwiązywania równań, a raczej efektywne połączenie metod istniejących, przy czym rozwiązania klasyczne wykorzystywane są w większości jako benchmark, zaś w większej części obliczeń Autor opiera się na metodach stosunkowo nowych, w szczególności narzędziach big data. Tak dzieje się w artykułach [1] oraz [3] – Doktorant proponuje niesztampowe modyfikacje istniejących podejść, kierując się rozumieniem procesów i uzasadniając ich wprowadzenie faktycznie występującymi mechanizmami, które – co bardzo dobrze daje się odczuć – świetnie rozumie. Z kolei w tekstach [2] i [4] Doktorant w większym stopniu replikuje istniejące metody z wykorzystaniem oryginalnego zbioru danych. Nie stanowi to w żadnym stopniu o słabości artykułów [2] i [4], a jedynie o ich innym charakterze – w szczególności bowiem w tekście [2] dokonuje identyfikacji determinant LGD powiązanych z postępowaniem dłużnika, tak więc wartość dodana tego tekstu ma jedynie inny charakter niż nieco bardziej nakierowanych metodycznie [1] i [3]. Na tym tle stosunkowo najmniej wnosi tekst [4], jest on jednak tekstem najkrótszym i o nieco innym charakterze od pozostałych, w mojej ocenie stanowi jednak ich bardzo udane uzupełnienie.

We wprowadzeniu Doktorant formułuje zarówno cel jak i hipotezy badawcze. Prawdę mówiąc, wprowadzenie jest chyba najsłabszym elementem rozprawy, co niewątpliwie wynika z wybranego sposobu jej przygotowania: każdy z czterech tekstów stanowi odrębną całość, ma swoją logikę i sposób skonstruowania – nie zawsze w pełni kompatybilny z pozostałymi (także dlatego, że całkiem inaczej przygotować trzeba tekst do czasopisma typu *Letters*, inaczej zaś – choćby do CEJEME). Cele przytoczono w każdym z artykułów z osobna. Nie w każdym z nich jasno sformułowano hipotezy badawcze, bo choćby w ostatnim z tekstów nie są one niezbędne, a być może wręcz w kanoniczny sposób podane hipotezy badawcze nie byłyby naturalne, biorąc pod uwagę charakter czasopisma. W konsekwencji, hipotezy we wprowadzeniu sformułowane są nieco na siłę – tak by uczynić zadość kanonowi dysertacji. Nie wszystkie z nich są w kolejnych rozdziałach w jasny sposób zweryfikowane. Oczywiście, z treści poszczególnych rozdziałów taka weryfikacja wynika, Autor jednak do nich nie wraca *per se*. Wydaje się, że biorąc pod uwagę formę dysertacji, pomocne pod tym względem byłoby podsumowanie, w którym Autor mógłby taką weryfikację przedstawić i omówić. Patrząc natomiast na każdy z wchodzących skład rozprawy tekstów osobno, nie mam wątpliwości, że ich konstrukcja od strony formalnej jest prawidłowa. Można by teoretycznie wskazać pewną niespójność: charakter kolejnych artykułów powoduje, że mamy zestawione ze sobą teksty złożone z kilkudziesięciu, a z drugiej strony – kilku stron. Trzeba jednak zauważyć, że głównym celem rozprawy doktorskiej jest prezentacja szeroko pojętego warsztatu przez doktoranta. Z tej



perspektywy różne formy artykułów dają na warsztat ten szersze – i bardzo pozytywne – spojrzenie.

Jak wspomniałem, w literaturze napotkać można sporo badań poświęconych LGD, czy wcześniej – prawdopodobieństwu niewykonania zobowiązania (PD). W swojej rozprawie mgr Starosta udowadnia, że najnowszy stan wiedzy jest mu bardzo dobrze znany. Doktorant opiera się na aktualnej i odpowiednio dobranej literaturze, jednocześnie nie multiplikując jej w sposób sztuczny i nie poszerzając bibliografii przesadnie.

Biorąc pod uwagę aspekty metodologiczne, liczba wykorzystanych narzędzi ekonometrycznych i statystycznych, których znajomość i umiejętność wykorzystania prezentuje Doktorant, imponuje. Bardzo dobre wrażenie sprawia także swoboda, z jaką Doktorant porusza się po tematyce pracy i nie można chyba mieć wątpliwości co do jego głębokiej wiedzy w tym obszarze. W konsekwencji, poszczególne rozdziały opracowane są w mojej opinii na bardzo wysokim poziomie. Doktorant jest przekonujący w swojej argumentacji kiedy wyjaśnia *po co* coś robi. Nieco mniejszy mój entuzjazm wzbudzają opisy *jak* prowadzi analizy oraz *co wynika* z przeprowadzonego badania. W tym miejscu trzeba podkreślić, że solidne są opisy o charakterze ilościowym – w mojej ocenie są one jasne i klarowne. Tu i ówdzie przydałoby się może sformułowanie założeń poszczególnych modeli, brak ten nie wydaje się jednak rażący. Gorzej wypadają opisy samego zjawiska: miejscami tekst bywa hermetyczny i wymaga dość dużego skupienia od czytelnika, który do lektury tekstu musi być odpowiednio przygotowany, bo większość fachowych określeń nie jest wyjaśniona, lub też wyjaśniona jest dość pobieżnie. Można domniemywać, iż z perspektywy Autora, wszystkie wykorzystane w pracy terminy oraz logika wyводу są absolutnie oczywiste, zaś praca doktorska nie jest podręcznikiem – ten punkt widzenia nie jest też z pewnością pozbawiony do końca racji. Mimo wszystko jednak w niektórych miejscach część opisów zyskałaby na większej klarowności. Należy przy tym dodać, że pewnym ułatwieniem dla mniej obeznanego z tą tematyką czytelnika może być wprowadzenie, w którym Doktorant stara się usystematyzować wykorzystywane pojęcia, co jednak nie zmienia tego, że np. szczegółowy opis proponowanego algorytmu od strony koncepcji metody w artykule [1] mógłby być jaśniejszy. Niewątpliwie natomiast odpowiedni jest poziom językowy dysertacji. Wprawdzie tu i ówdzie można znaleźć drobne błędy (przykładowo, w kilku miejscach Doktorant nieprawidłowo odmienia czasowniki, np. dodając „s” na końcu, mimo iż odnosi się do liczby mnogiej), są to jednak szczegóły, które pozostają bez wpływu na ogólny odbiór pracy.

Po lekturze dysertacji nasuwa mi się kilka pytań – częściowo o charakterze polemicznym, pojedyncze z nich to także moje drobne wątpliwości dotyczące użytych w pracy sformułowań lub zastosowanych rozwiązań:

1. W tekście „Modelling Recovery Rate (...)” we wzorze (3) pojawia się – jak mi się wydaje niezdefiniowany – symbol d , oznaczający niewątpliwie stopę dyskontową. Jakiej stopy procentowej – i dlaczego – użyto w badaniu empirycznym? Jaka jest wrażliwość wyników na stopę, w szczególności w kontekście różnych scenariuszy możliwych dla polskiej gospodarki?
2. W tekście „Modelling Recovery Rate (...)” we wzorze (7) Doktorant proponuje wykorzystanie transformacji, która jest konieczna dla zapewnienia, że zmienna zależna (*recovery rate*) mieści się w przedziale $<0;1>$. Podaje przy tym przykład kiedy tak może nie być – jeśli mianowicie przedmiot zabezpieczenia uda się sprzedać za kwotę przekraczającą wartość ekspozycji kredytowej w momencie niewykonania zobowiązania przez kredytobiorcę. Czy jednak faktycznie taka transformacja jest uzasadniona? Intuicyjnie wydaje się, że w przypadku takim jak opisany powyżej adekwatne byłoby raczej ucięcie stopy odzysku do 1, bo jej nadwyżka powinna jak rozumiem i tak zostać zwrócona kredytobiorcy, a w ten sposób unikniemy (niepożądanego?) transformowania całości rozkładu.
3. W tekście „Beyond the contract (...)” Autor na str. 79 wskazuje: „In the estimation process, forward, backward and stepwise methods are used with a p-value of less than 0.05 as the criterion for adding a variable. The maximum value of the correlation between the final set of predictors is fixed at 80%. The variable with the higher loglikelihood is ultimately used, and the second one is removed.” Pierwsza część zdania nie jest dla mnie do końca zrozumiała (czy słowo „and” na pewno jest potrzebne?). Jeśli jednak chodzi o – jak się domyślam – wykorzystanie regresji krokowej postępującej oraz wstecznej, to wykorzystanie zwłaszcza tej pierwszej wzbudza mój sprzeciw. Nieco podobnie moje wątpliwości wzbudza arbitralność progu 80% w dalszej części zdania oraz sama idea postępowania w ten sposób.
4. W tekście „Beyond the contract (...)” w modelu wykorzystano w charakterze regresorów właściwie tylko charakterystyki kredytu i pojedyncze charakterystyki klienta, jednak wyłącznie te, które mają bezpośredni związek z kredytem. Tymczasem nawet wśród cytowanych w pracy opracowań choćby Anolli et al. proponują wykorzystanie stosunkowo szerokiej grupy czynników socjoekonomicznych. Czy taka próba została podjęta, a jeśli tak – to z jakim skutkiem? Ten artykuł w gruncie rzeczy

koncentruje się właśnie na doborze odpowiednich regresorów – czy rezygnacja z części z obecnych w cytowanej literaturze rozwiązań była celowa?

5. W tekście „Loss given default (...)” pojawia się pomysł wykorzystania stosowanej sekwencyjnie regresji logistycznej. Wydaje się, że potencjalnie lepszą alternatywą byłoby w tym kontekście oparcie się na modelu typu *nested logit* – dlaczego Autor tego poniechał?
6. W tekście „Forecast combination (...)” odnoszę wrażenie, że jednym z kluczowych wniosków jest to, że wśród trzech metod uśredniania prognoz najlepsza z punktu widzenia pierwiastka błędu średniokwadratowego jest ta, w której staramy się tak dobrać wagi, by zminimalizować uśredniony błąd średniokwadratowy. Gdzie popełniam błąd, a jeśli go nie popełniam – to czy ten wniosek nie był stosunkowo prosty do przewidzenia (nawet jeśli walidacji dokonano na próbie odłożonej)? Dodatkowo: intuicja stojąca za podejściem Grangera–Ramanathana jest oczywista. Czy można w przystępny sposób pokazać intuicję stojącą za podejściem Mallowsa?
7. Czy alternatywę dla uśredniania prognoz, zaprezentowanego w tekście „Forecast combination (...)” mogłoby stanowić uśrednianie modeli (BMA lub podobne)? Czy byłaby to alternatywa atrakcyjna?
8. Pytanie do wszystkich analiz: które z uzyskanych wyników (np. wniosków o wyższości jednej z metod nad inną albo adekwatności danego czynnika) mogą być uznane za mogące mieć zastosowanie powszechne (nie zaś – w odniesieniu do pojedynczych, czasem bardzo dużych, prób na których Autor przeprowadził badanie empiryczne), które natomiast w opinii Autora wymagałyby pod tym kątem szerszych badań?

Sformułowane powyżej pytania nie stanowią rzecz jasna wątpliwości odnośnie wartości dysertacji – to problemy, na których temat chętnie podyskutowałbym z Doktorantem podczas obrony. Sama bowiem rozprawa niewątpliwie jest po prostu bardzo dobra i trudno w mojej ocenie sformułować pod jej adresem wiele zarzutów, które zaprzeczałyby tej tezie.

Dlatego też jednoznacznie stwierdzam, że w mojej opinii przedłożona rozprawa doktorska spełnia z naddatkiem wymogi stawiane w postępowaniach w sprawie o nadanie stopnia doktora, w szczególności zaś przesłanki określone w art. 187 ustawy PSWiN z 20.07.2018. Wnoszę zatem o dopuszczenie rozprawy do publicznej obrony, a w związku z jej wysokim poziomem naukowym wnoszę ponadto o jej wyróżnienie.

