

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Patricie Mittigová
Advisor:	prof. Ing. Evžen Kočenda, Ph.D.
Title of the thesis:	Consumer Credit Risk Analysis: Evidence from the Czech Republic

OVERALL ASSESSMENT (provided in English, Czech, or Slovak):

Please provide your assessment of each of the following four categories, summary and suggested questions for the discussion. The minimum length of the report is 300 words.

Contribution

Přínos diplomové práce sl. Mittigové spočívá v aplikaci standardních metod kreditního skóringu na ne zcela tradiční data ohledně spotřebitelských úvěrů. Je třeba uvést, že literatura věnující se spotřebitelským úvěrům a jejich determinantám je v české kotlině poměrně řídká, přičemž spotřebitelské úvěry jsou typem úvěrů, u kterého se může projevit eskalace úvěrového rizika poměrně rychle. Zároveň se u nich v poslední době projevují zajímavé a ne zcela intuitivní tendence, jakou je zvyšující se podíl dlouhodobých úvěrů, které v současnosti dynamiku spotřebitelských úvěrů dominují, zvyšující se podíl účelových spotřebitelských úvěrů a úvěrů na refinancování starších úvěrů, vliv zesilující konkurence a uvolněné měnové politiky (viz například Brož, Hlaváček, 2018). V tomto ohledu se tak diplomová práce sl. Mittigové věnuje zajímavému a aktuálnímu tématu.

Jak nicméně uvádím níže, přínos diplomové práce je poněkud erodován nejasnostmi ohledně podkladové databáze, na které je celá analýza postavena. Vzhledem k těmto nejasnostem je tak přidaná hodnota provedené analýzy poněkud omezená.

Methods

Článek aplikuje širokou paletu metod, které se standardně používají při vyhodnocování kreditního rizika v bankách, resp. při skóringu klientů. Tyto metody jsou použity odpovídajícím způsobem a jsou provedeny standardní testy jejich účinnosti, účinnost různých metod je pak korektním způsobem porovnána i mezi sebou. Potud je tedy aplikovaná metodologie odpovídající a poměrně aktuální.

S čím mám nicméně určitý problém je diskuze vhodnosti aplikování těchto metod na zvolená data resp. diskuze kvality podkladových dat. Bohužel čtenář sám nemá příliš mnoho možností kvalitu dat posoudit, neboť u nich není uvedena zdrojová banka. Rovněž není uvedeno, nakolik jsou tato data reprezentativní jak pro český bankovní sektor, tak pro danou banku, zda struktura dat odpovídá struktuře populace podle různých sociodemografických a ekonomických charakteristik dostupných například z agregovaných či subagregovaných dat ČSÚ či ČNB. Bez diskuze reprezentativnosti zdrojových dat jsou nicméně výsledky analýzy pouze určitým cvičením, na kterém autorka ilustruje svoji schopnost s danými metodami pracovat, ale pro posouzení vývoje trhu spotřebitelských úvěrů jsou prakticky bezcenné.

Při 2420 úvěrech za více než 10 let tedy v průměru cca 240 úvěrech za rok přitom reprezentativnost podkladové databáze asi nebude z nejvyšších. Nejisté je rovněž to, zda byla konstrukce databáze provedena korektně. U databáze je uvedeno, že byla vytvořena v březnu 2017 a že pokrývá úvěry z období 2006-2017. Bohužel ale není jasné, zda jsou v databázi zahrnuty i starší úvěry, které byly poskytnuty například v roce 2008, ale byly v minulosti již buď splaceny, nebo naopak došlo již v minulosti k jejich defaultu a byly již odepsány či prodány třetí straně. Z distribuce databáze podle roku poskytnutí úvěru (graf 3.8 na str. 26), v které jednoznačně dominují novější úvěry poskytnuté v roce 2016 (více než polovina vzorku), se spíše domnívám, že databáze pokrývá pouze úvěry, které byly k březnu 2017 stále ještě aktivní. Pro starší úvěry ale tento způsob konstruování databáze znamená zjevný „selection bias“, který ale autorka nikde nekomentuje. Tento selection bias například může znamenat, že jsou pro starší data poskytnutí úvěru v databázi nadreprezentované úvěry

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Patricie Mittigová
Advisor:	prof. Ing. Evžen Kočenda, Ph.D.
Title of the thesis:	Consumer Credit Risk Analysis: Evidence from the Czech Republic

s dlouhou splatností (ty krátkodobé již byly splaceny), nebo že je pro tyto roky naddimenzována propočtená míra defaultu (kvalitní a v minulosti již splacené úvěry již v databázi nefigurují, zůstávají jen ty nesplacené; tomu by naznačoval i vysoký podíl úvěrů s defultem- 22,5 %, který je násobně vyšší než podíl nesplacených úvěrů na agregovaných datech- viz str. 17). Tento selection bias ale může znamenat, že některé z uvedených výsledků jsou špatně (např. výsledky ohledně rizikovosti úvěrů podle roku poskytnutí). Starší údaje tak mohou být distortované samotným způsobem konstrukce databáze, pro nové úvěry (rok 2016) je pak problémem, že se u nich dost dobře nestihlo kreditní riziko projevit.

Možná se zde mýlím, databáze je konstruována korektně a vyšší podíl novějších úvěrů je dán například tím, že diskutovaná banka je „novým hráčem na trhu“, ale popis konstrukce databáze je podle mého názoru neodstatečný.

Krom této obecnější poznámky mám několik konkrétnějších poznámek k vysvětlujícím veličinám:

- 1) Některé z vysvětlujících veličin jsou v diplomové práci definovány poněkud nejednoznačně. Například není přesně zadefinován default, resp. je definován tak, že nabývá hodnoty 1, když úvěr v minulosti zdefaultoval a 0 pokud nikoli. Nebere zde ale v úvahu, že novější úvěry zatím „neměly čas“ zdefaultovat, což opět může ovlivnit výsledky. Standardně se přitom default úvěru vztahuje k nějakému časovému horizontu (např. jeden rok) a úvěry, jejichž trvání je kratší než tento horizont jsou buď z analýzy vyřazeny, nebo se údaje pro ně určitým způsobem upravují.
- 2) V popisných tabulkách (např. tabulky 3.2, 3.5, 3.6) se u mediánu, minima a maxima výše úvěru objevují podivně kulatá čísla zaokrouhlená na tisíce resp. pětitisíce. To by naznačovalo, že jsou podkladová data nějak zaokrouhlená. Pokud tomu tak skutečně je, bylo by vhodné toto uvést.
- 3) Vzhledem k tomu, že databáze pokrývá celkem dlouhou časovou dimenzi (více než 10 let) by stálo za to údaje o příjmu dlužníka buď deflovat, nebo je vztáhnout pro jednotlivé úvěry k pozici příjmu klienta v nějaké standardní příjmové distribuci (ať již vztážené k dané databázi, tak k vhodnější např. k distribuci mezd z ČSÚ). Z grafu o distribuci podle příjmu dlužníka (grafy 3.5a a 3.5b) mám trochu podezření, zda není údaj o měsíčním příjmu příliš hrubý (0-20 tis. měsíčně, 20-40 tis. měsíčně, a dále hranice skupin po 20 tis.). Buď jsou tedy data o příjmech podivná, nebo jsou nevhodně volené hranice pro graf 3.5.
- 4) U grafu uvádějícího strukturu úvěrů podle krajů (fig. 3.4 na str. 22) jednak chybí Středočeský kraj, který je v rámci ČR jeden z největších, jednak chybí popisek osy y (je zde počet úvěrů, jejich celkový objem, či průměrná velikost úvěrů)? Pokud jde o počet úvěrů či jejich objem, měly by být údaje nějakým způsobem standardizovány, např. na počet obyvatel (rozdíly ve velikosti krajů podle počtu obyvatel jsou násobné). Podivně působí také vysoká hondota pro poměrně chudý Moravskoslezský kraj. Každopádně mám pochybnost, zda tento graf ukazuje to, čím autorka argumentuje.
- 5) Jak jsme už uvedl, u demografických čísel by bylo vhodné uvést jejich vztah k agregovaným úvěrům na celé české populaci (tedy například srovnat počet dětí, strukturu vzdělání a pod). Některé z ukazatelů přitom na první pohled vypadají poněkud nereprezentativně (nulový mediánový počet dětí).
- 6) Některé z veličin jsou popsány poněkud podivně a není jasné, co se za nimi skrývá. Například „Loan Duration“ může odrážet splatnost resp. maturitu úvěru (krátkodobé vs. dlouhodobé úvěry), ale také duraci úvěru ve smyslu macaulyho resp. modifikované durace (ta vedle splatnosti bere v úvahu také strukturu splátek úvěrů a může se od ní poměrně výrazně lišit). Hypoteticky by také mohla znamenat délku dostavadního trvání úvěru. Zcela není jasný ani vztah veličiny „Loan Duration“ k veličině „tenure profile“.
- 7) Obdobně je v článku podivně popsána veličina loan type, která odlišuje dva typy úvěrů, jednak „loan for housing“, který zřejmě odpovídá „spotřebitelským úvěrům na bydlení“ dle statistik ČNB, jednak „consolidation loans“, které asi odpovídají „refinancovaným“ úvěrům.

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Patricie Mittigová
Advisor:	prof. Ing. Evžen Kočenda, Ph.D.
Title of the thesis:	Consumer Credit Risk Analysis: Evidence from the Czech Republic

Problémem ale je, že ani jeden z uvedených typů úvěrů není klasickým spotřebitelským úvěrem („spotřebitelské úvěry na bydlení“ dle statistik ČNB mezi spotřebitelských úvěrů vůbec nepatří, parametry refinacovaného úvěru souvisejí spíše s parametry původních úvěrů, které jím byly nahrazeny), takže aplikovatelnost výsledků analýzy je opět diskutabilní.

- 8) U veličin odrážející úroveň vzdělání jsme měl problém porozumět jejich obsahu, neboť neodpovídají standardním číselníkům vzdělávání např. podle metodik ČSÚ.

Literature

Přehled literatury je zpracovaný standardním způsobem, odkazy na literaturu jsou odpovídající. V sekci výsledků jsou tyto výsledky srovnány i s výsledky z literatury, možná se zde až příliš často odkazuje na článek Kočenda a Vojtek (2011). I když tato literatura není příliš obsáhlá, možná bych přivítal trochu podrobnější diskuzi literatury věnované analýze spotřebitelských úvěrů v ČR (např. Seidler, Jakubík, ČNB WP 13/2009). Rovněž by bylo vhodné srovnat vývoj hlavních veličin zahrnutých do analýzy s vývojem celého trhu spotřebitelských úvěrů (pak by například nebyla autorka tak překvapena tím, že velká část spotřebitelských úvěrů má dlouhou splatnost- viz poslední věta kapitoly 3.4.12; k této změně v datech skutečně došlo).

Manuscript form

Srozumitelnosti posuzované diplomové práce příliš nepomáhá poněkud zanedbaný popis vysvětlujících veličin (viz výše), v některých pasážích mi připadala jako poněkud kostrbatá i použitá angličtina. V diplomové práci se objevují určité duplicity, když autorka nejprve v části 3.4 (resp. subakapitolkách 3.4.1 až 3.4.12) poměrně podrobně popisuje data, která vstupují do modelů kreditního rizika, následně v části 5.1 při formulaci hypotéz mnohé z předchozí části opakuje, což činí i v části diskutující výsledky po jednotlivých vysvětlujících veličinách (6.1.1 až 6.1.7). Oproti tomu v diskusi výsledků mi poněkud chybělo jejich kritické posouzení.

Summary and suggested questions for the discussion during the defense

Celkově hodnotím diplomovou práci sl. Mittigové jako solidní v části aplikovaných analytických metod, i přehled literatury je v zásadě odpovídající. Diskutované téma mi také připadne jako velmi aktuální. Hlavní problém analýzy nicméně vidím v ne zcela dobré kvalitě podkladových dat, což by samo o sobě asi nebylo zcela férové autorce vytýkat (pracovala s daty, která byla dostupná). Domnívám se ale, že měla nepříliš dobrou kvalitu podkladových dat otevřeně přiznat, věnovat více energie jejich očištění resp. úpravě a především pečlivě a kriticky diskutovat, jak tyto nedostatky ovlivnily výsledky celé její analýzy. Bez tohoto je přidaná hodnota diplomové práce poněkud omezena. Na diplomové práci mě také zklamalo, že se v podstatě vůbec nevěnovala popisu vývoje trhu spotřebitelských úvěrů na agregovaných datech. Srovnání výsledků analýzy s trendy vývoje na celém trhu spotřebitelských úvěrů by přitom mohlo dodat těmto výsledkům na síle. **Celkově doporučuji přijmout diplomovou práci sl. Mittigové k obhajobě a ohodnotit ji známkou C.**

Při obhajobě by bylo možné diskutovat, jak konstrukce podkladové databáze ovlivnila výsledky analýzy a rovněž srovnat výsledky analýzy s hlavními trendy na trhu spotřebitelských úvěrů v poslední době.

SUMMARY OF POINTS AWARDED (for details, see below):

Report on Bachelor / Master Thesis

Institute of Economic Studies, Faculty of Social Sciences, Charles University in Prague

Student:	Bc. Patricie Mittigová
Advisor:	prof. Ing. Evžen Kočenda, Ph.D.
Title of the thesis:	Consumer Credit Risk Analysis: Evidence from the Czech Republic

CATEGORY	POINTS
<i>Contribution (max. 30 points)</i>	18
<i>Methods (max. 30 points)</i>	22
<i>Literature (max. 20 points)</i>	18
<i>Manuscript Form (max. 20 points)</i>	15
TOTAL POINTS (max. 100 points)	73
GRADE (A – B – C – D – E – F)	C

NAME OF THE REFEREE:
PhDr. Michal Hlaváček, PhD.

DATE OF EVALUATION: 12/6/2018

Referee Signature

EXPLANATION OF CATEGORIES AND SCALE:

CONTRIBUTION: *The author presents original ideas on the topic demonstrating critical thinking and ability to draw conclusions based on the knowledge of relevant theory and empirics. There is a distinct value added of the thesis.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

METHODS: *The tools used are relevant to the research question being investigated, and adequate to the author's level of studies. The thesis topic is comprehensively analyzed.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
30	15	0

LITERATURE REVIEW: *The thesis demonstrates author's full understanding and command of recent literature. The author quotes relevant literature in a proper way.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

MANUSCRIPT FORM: *The thesis is well structured. The student uses appropriate language and style, including academic format for graphs and tables. The text effectively refers to graphs and tables and disposes with a complete bibliography.*

<i>Strong</i>	<i>Average</i>	<i>Weak</i>
20	10	0

Overall grading:

TOTAL	GRADE
91 – 100	A
81 - 90	B
71 - 80	C
61 – 70	D
51 – 60	E
0 – 50	F