

Posudek oponenta k disertační práci

Ing. Adama Čabaly

Doba nezaměstnanosti v České republice pohledem analýzy přežití

Předložená práce o rozsahu 133 stran formátu A4 a CD má všechny předepsané formální náležitosti, jako jsou prohlášení autorství, abstrakty (v češtině a v angličtině), shrnutí výsledků a závěry, seznam literatury a seznam použitých symbolů a zkratk.

Autor ve své disertační práci zvolil přístup, kdy řeší konkrétní věcný problém s použitím známých metod (klasických i nově konstruovaných). Konkrétně se soustředí na délku doby nezaměstnanosti v České republice a na faktory ji ovlivňující. Popis vlivu faktorů je sám o sobě velmi zajímavý a užitečný. Oceňuji, že autor navíc přidává porovnání tří různých období v těsné blízkosti hospodářské krize (před krizí: 2007-2008, během krize: 2010-2011 a po krizi: 2013-2014), což vedlo k významnému nárůstu objemu počítaných modelů. Porovnání je velmi cenným příspěvkem k hodnocení dopadů hospodářské krize na ekonomiku a obyvatelstvo České republiky.

V kapitole 3 považuji navržené metody, testy, modely a software použitý pro jejich implementaci za vhodné a odpovídající cílům práce. Za velmi přínosnou považuji implementaci teorie extrémní hodnoty do problematiky analýzy přežití. V této kapitole autor detailně a srozumitelně popisuje postupy práce (tvorba modelu na str. 41-42 a optimalizační postup odhadu parametrů po částech spojitě distribuční funkce na str. 48). Celá kapitola, ač z velké části teoretická, je psaná čtivým a atraktivním způsobem. Podle mého názoru bude zajímavá i pro čtenáře bez hlubší znalosti problematiky práce s cenzorovanými daty, analýzy přežití a teorie extrémní hodnoty.

Detailní popis jednotlivých modelů je popsán v kapitole číslo 4. Hypotetické faktory ovlivňující délku nezaměstnanosti považuji za relevantní. Pouze u geograficky uspořádané proměnné Velikost obce (*velikostobce*) se domnívám, že bez vzájemných prostorových souvislostí (vzájemná poloha, vzdálenost, sousednost) má jen velmi malou vypovídací schopnost. Příkladem může být malá obec v těsné blízkosti průmyslové zóny.

Je patrné, že autor musel odhadnout obrovské množství modelů, řadu z nich pouze jako mezistupeň k nalezení optimálního výsledného modelu použitého pro konečnou interpretaci. Z toho plyne asi zřejmě největší nedostatek celé práce. V kontroverzi s předchozí kapitolou jsou výsledky hůře čitelné a místy po prvním přečtení jen stěží srozumitelné. A to i přes to, že se autor snažil zjednodušit čtenáři orientaci ve výsledcích přehlednými tabulkami a grafy. Grafy intervalových odhadů parametrů výsledných modelů (konkrétně obrázky 4.19 až 4.23) jsou bohužel kontraproduktivní a působí matoucím dojmem. Jednak v nich není na první pohled patrná pozice referenční hodnoty koeficientu (pouze malý křížek). Zejména jsou v nich však obtížně rozlišitelné jednotlivé modely, kterých je v jednom grafu hned několik těsně pod sebou.

Navíc se jedná o konečné modely se statisticky významnými proměnnými, které i když jsou vedeny v jednom grafu, jsou jen stěží vzájemně porovnatelné. Důvodem je to, že každý model je založen na jiné sadě proměnných. Přestože si uvědomuji složitost některých modelů, doporučoval bych pro porovnatelnost uvádět do grafů intervalové odhady všech koeficientů (i nevýznamných), zejména však těch, spojených s kategoriálními proměnnými. Čtení práce dále znesnadňuje i odkazování na výstupy na přiloženém CD.

Zajímavé jsou výsledky odhadu pravého konce rozdělení doby nezaměstnanosti (kapitola 4.2). Postrádám zde ale porovnání s autorem navrženým postupem v kapitole 3.9.3.

V kapitole závěr jsou stručně a srozumitelně popsány hlavní faktory působící na dobu nezaměstnanosti.

Přes výše uvedené připomínky je závěrem třeba říci, že se celkově jedná o vynikající, technicky náročnou práci, která bezesporu splňuje nároky kladené na doktorandskou práci. Výsledky doporučuji prezentovat v časopise s IF (např. Statistics and Economy Journal). Práci doporučuji k obhajobě před příslušnou komisí a po pozitivním výsledku obhajoby doporučuji udělení titulu Ph.D.

Zjištěné drobné formální nedostatky v práci:

- Str. 20 odst. 2 – příliš složité souvětí o osmi větách
- Str. 25 nadpis kapitoly – vzhledem k tomu, že kapitola je výčtem použitých metod, postupů a pravidel je vhodnější název Metodika
- Str. 32 vzorec 3.20 – nejde o distribuční funkce, ale o funkce přežití
- Str. 40 odstavec 4 – není zřejmé, zda jde o názor autora nebo výsledek podrobnější rešerše
- Str. 56 obrázek 4.3 – graf je příliš malý a nepřehledný, nelze od sebe odlišit jednotlivé křivky
- Str. 50 tabulka 4.8 – tabulka je hůře čitelná kvůli absenci vodorovných čar
- Str. 60 odstavec 2 – autor uvádí výrazný rozdíl v mediánu doby (11 měsíců), avšak v odpovídajících tabulkách 4.10 a 4.11 jsou odhady zcela stejné (10,39)
- Str. 67 tabulka 4.14 – tabulka je hůře čitelná
- Str. 86 obrázek 4.26 – pro snadnější interpretaci by bylo vhodné uvést stupně vzdělání slovně tak, jak je tomu u ostatních kategoriálních proměnných
- Str. 112 zdroj č. 60 není úplný, chybí název článku

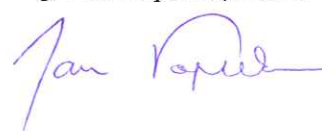
Práci doporučuji k obhajobě před příslušnou komisí.

Připojuji některé dotazy:

1. Jsou intervalové odhady parametrů modelů opravdu symetrické? Neuvažoval jste o aplikaci metody vedoucí k asymetrickým odhadům?
2. Technický dotaz. Automatizoval jste nějakým způsobem navržené postupy optimalizace modelů v programu R nebo v jiném software?

V Ústí nad Labem dne 2. 5. 2017

Ing. Jan Popelka, Ph.D.



Hodnocení disertační práce v oboru Statistika (VŠE v Praze)

(vyplňuje školitel a oponenti - příloha posudku oponenta)

Jméno doktoranda: ING. ADAM ČABALA

Název práce: DOBA NEZAMĚSTNANOSTI V ČESKÉ REPUBLICE
POHLEDEM ANALÝZY PŘEŽITÍ

Jméno hodnotitele: ING. JAN POPELKA, PH.D.

Kritéria hodnocení

1. Odpovídá název práce jeho obsahu (zcela, částečně)? ZCELA
2. Je vymezení cílů vyhovující, je uvedeno v úvodu, abstraktu? ANO
3. Jak jsou vytyčené cíle splněny, je shrnuto v závěru, abstraktu?
CÍLE BYLY SPLNĚNY - SHRNUTO V ZÁVĚRU
4. Jsou zřetelně odlišeny metody převzaté z literatury a vlastní přístupy?
ANO, PŘEVZATÉ METODY JSOU ŘÁDNĚ CITOVÁNY
5. Co je vědeckým přínosem (např. z hlediska teorie statistiky či metod analýzy dat)?
PŘÍNOSEM JE NOVÝ POHLED NA FAKTORŮ PŮSOBÍCÍCH NA DELKU
6. Úroveň rešerše poznatků a použité literatury ve zkoumané oblasti: NEZAM. V ČR
REŠERŠE JE NA VYSOKÉ ODBORNÉ ÚROVNI
7. Úroveň aplikací, hodnocení a porovnání metod (zařazení tabulek a grafů):
APLIKACE METOD JE VHDNA, HODNOCENÍ A VÝSLEDKY JSOU HŮŘE
SPROZUMITELNÉ
8. Způsob vyjadřování, jazyková úroveň:
PRÁCE OBSAHUJE MINIMUM CHYB, MÍSTY ZBYTEČNĚ SLOŽITÁ SOUVĚTÍ
9. Jsou řádně definovány používané pojmy, zkratky a symboly?
ANO
10. Formální úroveň (matematických výrazů, tabulek, grafů):
MAT. VÝRAZY A TABULKY JSOU V PORÁDKU, NĚKTERÉ GRAFY
JSOU HŮŘE POCHOPITELNÉ

Datum: 2.5.2017

..... Jan Popelka
podpis