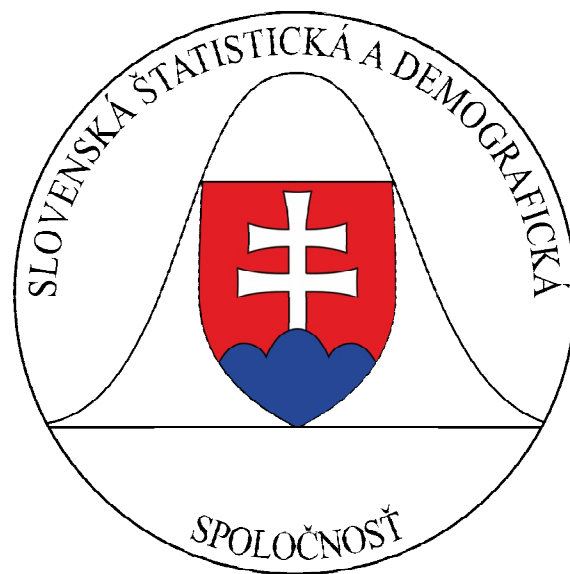
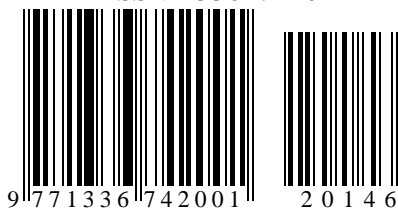


6/2014

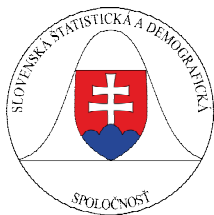
FORUM STATISTICUM SLOVACUM



ISSN 1336-7420



9 771336 742001 2 0 1 4 6



**Slovenská štatistická a demografická
spoločnosť Miletičova 3, 824 67 Bratislava
www.ssds.sk**



Naše najbližšie akcie:

(pozri tiež www.ssds.sk, blok Organizované akcie)

MEDSTAT 2015

1. polrok 2015

Pohl'ady na ekonomiku Slovenska 2015

7. 4. 2015, aula EU v Bratislave

Nitrianske štatistické dni 2015

máj 2015, UKF Nitra

29. medzinárodný seminár EKOMSTAT 2015

24. – 27. 5. 2015, Trenčianske Teplice

15. Slovenská demografická konferencia, Trenčiansky kraj

2015

Regionálne akcie

priebežne

Slávnostná konferencia 50 rokov Slovenskej štatistickej a demografickej spoločnosti

marec 2018, Slovenská republika

Výpočtová štatistika 2015

3. – 4. 12. 2015

FOREWORD

Dear colleagues,

we propose the sixth issue of the tenth volume of the scientific peer-reviewed journal published by the Slovak statistical and demographical society (SSDS). This issue comprises contributions that are content-compatible with the topic „Computational statistics“.

Editors: Iveta Stankovičová, Jozef Chajdiak, Ján Luha, Tomáš Želinský

Reviewers: Jozef Chajdiak, Bohdan Linda, Ján Luha, Iveta Stankovičová, Tomáš Želinský

Assoc. Prof. Ing. Jozef Chajdiak, CSc.

Editor in chief



Continuation of International Year of Statistics 2013 is initiative The World of Statistics. Informations about this initiative you will find in the first issue bulletin *News from the World of Statistics*, which you will find on

<http://www.worldofstatistics.org/files/2014/01/January-22-2014.pdf>.

PREDHOVOR

Vážené kolegyně, vážení kolegovia,

predkladáme šieste číslo desiateho ročníka vedeckého recenzovaného časopisu Slovenskej štatistickej a demografickej spoločnosti (SŠDS). Toto číslo je zostavené z príspevkov, ktoré sú obsahovo orientované v súlade s tematikou „Výpočtová štatistika“.

Editori: doc. Ing. Iveta Stankovičová, PhD., doc. Ing. Jozef Chajdiak, CSc., RNDr. Ján Luha, CSc., doc. Ing. Tomáš Želinský, PhD.

Recenzenti: doc. Ing. Jozef Chajdiak, CSc., doc. RNDr. Bohdan Linda, CSc., RNDr. Ján Luha, CSc., doc. Ing. Iveta Stankovičová, PhD., doc. Ing. Tomáš Želinský, PhD.

Doc. Ing. Jozef Chajdiak, CSc.

Šéfredaktor



Pokračovaním Medzinárodného roku štatistiky 2013 je iniciatíva The World of Statistics. Informácie o tejto iniciatíve nájdete v prvom čísle bulletinu *News from the World of Statistics*, ktorý nájdete na <http://www.worldofstatistics.org/files/2014/01/January-22-2014.pdf>.

Vonkajšia rovnováha po roku 2000 v podmienkach SR a ČR – čistý export a čisté úspory

External balance after 2000 in Slovakia and the Czech Republic - net exports and net savings

Ludmila Bartóková

Abstract: The evolution of the euro area external balance started attracting more attention in the context of the crisis, which highlighted the risks posed by the uneven macroeconomic development of the euro area member countries. This increases the interest in the analysis of external balance, the sustainability of its evolution and evaluation of resemblance between euro area member and candidate countries. This paper aims to verify the similarities in the evolution of the external balance of the Czech Republic and Slovakia through selected variables. Given the broadly comparable macroeconomic development, we can assume certain common features for indicators of external balance.

Abstrakt: Vývoj vonkajšej rovnováhy krajín eurozóny sa dostal do pozornosti v súvislosti s krízou, ktorá poukázala na riziká, ktoré so sebou prináša odlišný makroekonomický vývoj členských krajín eurozóny. Zvýšil sa tak záujem o analýzu v oblasti vonkajšej rovnováhy, udržateľnosť tohto vývoja a hodnotenie podobnosti členských aj kandidátskych krajín eurozóny. Nasledujúci príspevok je zameraný na overenie podobnosti vo vývoji vonkajšej rovnováhy ČR a SR prostredníctvom vybraných premenných. S ohľadom na všeobecne porovnateľný makroekonomický vývoj je predpoklad spoločných znakov aj v oblasti ukazovateľov vonkajšej rovnováhy.

Key words: export, import, savings, investments.

Kľúčové slová: export, import, úspory, investície.

JEL classification: E20, F41

1. Úvod

Nedávne hospodárske problémy eurozóny opäť poukázali na mnohé pretrvávajúce nedostatky a prehlbujúce sa rozdiely medzi krajinami. Jedným z príkladov je vývoj vonkajšej rovnováhy. Obdobie predchádzajúce globálnej finančnej kríze, ktorá sa neskôr rozšírila aj do ďalších oblastí, bolo všeobecne charakterizované zväčšovaním nerovnováh na bežných účtoch platobnej bilancie, nasledované ich čiastočnou korekciou v období po kríze. Vývoj vonkajšej rovnováhy krajín eurozóny, prípadne krajín Európskej únie, sa aj v súvislosti s krízou znovu stal častým predmetom analýz, ktoré sú zvyčajne zamerané na odlišnosti vo vývoji. V prípade krajín, ktoré využívajú spoločnú menu môžu práve tieto odlišnosti negatívne pôsobiť na menovú stabilitu.

Aj napriek tomu, že Slovensko od roku 2009 patrí do eurozóny, je ešte stále, s ohľadom na svoj makroekonomický vývoj, často prirovnávané skôr ku krajinám skupiny V4 (hlavne k Českej republike) ako ostatným členom eurozóny. Makroekonomické analýzy naznačujú podobný vývoj krajín v mnohých oblastiach. Práve z tohto dôvodu sme sa rozhodli analyzovať vývoj SR a ČR v oblasti vonkajšej rovnováhy, úspor a investícií po roku 2000. Cieľom článku je overiť podobnosť tohto vývoja, sledovať trend pre jednotlivé premenné a zistiť, či zmeny vo vývoji v súvislosti s krízou boli porovnateľné aj v podmienkach členskej a nečlenskej krajiny eurozóny.

2. Skúmanie vonkajšej rovnováhy v podmienkach eurozóny

Od vzniku eurozóny došlo k významnému prehĺbeniu nerovnováh na bežných účtoch členských krajín, v dôsledku čoho sa začali objavovať rôzne obavy o udržateľnosť menovej únie. Nerovnováha bežného účtu, či už v podobe deficitu alebo prebytku, odráža predovšetkým vývoj exportu a importu tovarov a služieb, ktoré tradične predstavujú najväčšiu položku v rámci bežného účtu. Deficit bežného účtu znamená, že domáce výdavky presahujú domáce príjmy a naopak prebytok predstavuje zároveň aj prebytok príjmov. Tento vývoj však môžeme dať rovnako do súvisu aj s vývojom domácich úspor a investícií, keďže saldo bežného účtu odráža aj rozdiel medzi domácimi úsporami a investíciami.

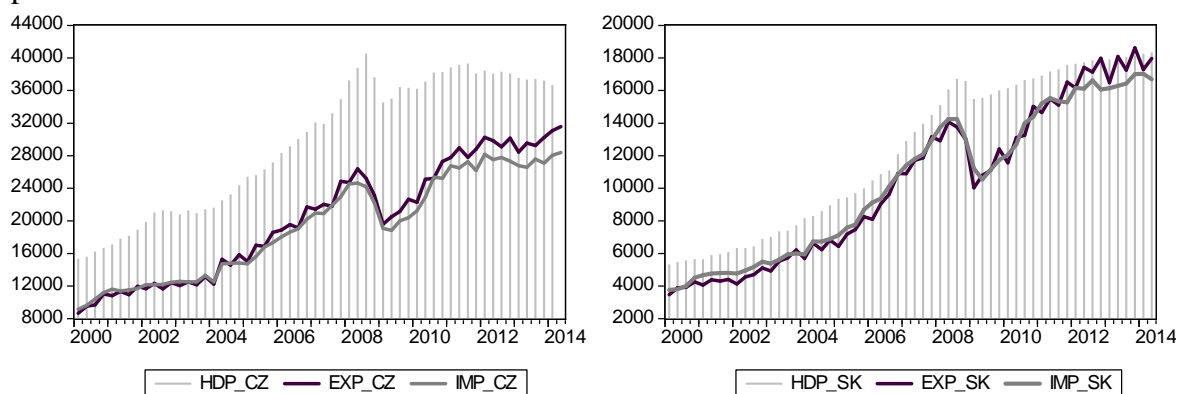
Záujem o problematiku vonkajšej (ne)rovnováhy vzrástol aj v súvislosti s hospodárskou krízou a tiež dlhovou krízou. Často sa možno stretnúť s názorom, že mnohé súčasné problémy eurozóny úzko súvisia práve s dlhodobou nerovnováznym vývojom vonkajšej rovnováhy jednotlivých krajín, kde silné ekonomiky ako napr. Nemecko a Holandsko vykazovali dlhodobé prebytky a krajiny ako Grécko, Taliansko, či Portugalsko, naopak deficity. Na druhej strane však eurozóna ako celok vykazovala takmer vyrovnaný bežný účet voči zvyšku sveta. (Blanchard, Giavazzi, 2003) (Gaulier, Vicard, 2012) Nakoľko sa veľká časť zahraničného obchodu v rámci eurozóny uskutočňuje medzi krajinami navzájom, dochádzalo a dochádza tak neustále k presúvaniu prostriedkov zo skupiny deficitných krajín do skupiny prebytkových. Tieto potom vytvárajú vyššie úrovne úspor, ktoré sú následne presúvané do zahraničia. Väčšina týchto investícií sa kvôli nižšiemu riziku tradične realizovala práve v deficitných krajinách. Ako uvádzajú Schmitz a von Hagen (2011), tiež Campa, Gavilan (2011) podobné presuny kapitálu z „bohatších“ do „chudobnejších“ krajín podporujú do istej miery konvergenciu ekonomík a z toho dôvodu nerovnováhy na bežnom účte nemusia nevyhnutne predstavovať znak nesprávneho fungovania ekonomiky. Na druhej strane však zvyšujú potrebu vhodnej a obozretnej hospodárskej politiky.

Názory na to, či vonkajšie nerovnováhy významne prispievajú k vypuknutiu alebo zhoršeniu kríz, respektíve do akej miery, sa rôznia. Zaujímavý je ale poznatok, že aj v prípade významných nerovnováh dochádza zvyčajne po odznení krízy k ich náprave. Tento vývoj sa potvrdil v prípade viacerých minulých kríz: napr. v podmienkach ázijských krajín, ktoré boli zasiahnuté krízou koncom 90. rokov. (Gruber, Kamin, 2007) Podobné analýzy boli uskutočnené aj pre vývoj vonkajšej rovnováhy v európskych krajinách, predovšetkým pre krajiny vykazujúce deficity. Kang a Shambaugh (2013) skúmali pobaltské krajiny a tzv. periférne ekonomiky eurozóny (Grécko, Španielsko a Írsko), ktorých deficity na bežnom účte mali v období pred krízou skôr rastúci charakter. Aj keď dôvody vonkajšej nerovnováhy týchto krajín boli rôzne, autori poukazujú na to, že po kríze došlo k úprave cien a znižovaniu deficitov, vo väčšine prípadov v dôsledku prepadu importu. Zároveň došlo k zlepšeniu jednotkových nákladov práce, čo prispelo k zvyšovaniu exportu. Blanchard a Giavazzi (2003) skúmali vývoj deficitov bežného účtu v podmienkach malých európskych krajín, členov eurozóny, ktoré vykazovali pred krízou nižšie domáce úspory a vyššie investície. Tento vývoj prispieval k vyšším deficitom na bežnom účte. Zlepšenie vývoja na bežnom účte v súvislosti s krízou skúmali tiež Lane, Milesi-Ferretti (2011). Ich analýza ukázala, že vo väčšine prípadov bola táto náprava dosiahnutá primárne vďaka významnému zníženiu domáceho dopytu a nie zmenami v štruktúre výdavkov. Podobný vývoj možno zaznamenať vo väčšine štátov Európskej únie, vrátane členov, ktorí vstúpili do EÚ v roku 2004, čo by mohlo naznačovať istú podobnosť krajín v tejto oblasti. Takýto pohľad je však značne zjednodušený, umožňuje len hodnotenie na základe všeobecného vývoja premenných, hodnôt skôr smerovanie vývoja, trend a nezaobera sa bližšie napríklad štruktúrou či príčinami deficitu. Presnejšie hodnotenie by si vyžadovalo rozsiahlejšie porovnanie a hlbšiu analýzu.

3. Hodnotenie vývoja vonkajšej nerovnováhy v podmienkach ČR a SR

Pre analýzu boli zvolené dve často porovnávané krajiny – Slovensko a Česká republika. S ohľadom na všeobecne podobný makroekonomický vývoj v minulosti bolo cieľom overiť podobnosť ich vývoja aj v oblasti vonkajšej rovnováhy po roku 2000. Toto obdobie zahŕňa spoločný vstup oboch krajín do EÚ, neskôr vstup SR do eurozóny, ako aj negatívne dopady hospodárskej krízy. Na sledovanie vonkajšej rovnováhy týchto dvoch krajín boli zvolené nasledujúce premenné, získané z databázy Eurostat: export a import (tovary a služby), úspory a investície (hrubá tvorba fixného kapitálu), hrubý domáci produkt. Údaje boli vyjadrené v bežných cenách, v miliónoch eur, čo umožnilo sledovať vývoj týchto premenných v absolútnych hodnotách a tiež sledovať trend. Všetky časové rady boli sezónne očistené.

Nasledujúce grafy (Obr.1) znázorňujú vývoj HDP, exportu aj importu pre obidve krajiny v období 2000Q1-2014Q2. Je zrejmé, že do roku 2009 bol trend vývoja veľmi podobný, so stabilným rastom ekonomiky. Po kríze však už je možné vidieť isté rozdiely, keďže v prípade ČR došlo iba ku krátkodobému obnoveniu rastu, v ďalšom období už bol vývoj HDP skôr stagnujúci respektíve klesajúci. Na druhej strane v prípade SR je zrejmy mierne rastový trend, pričom v posledných kvartáloch sledovaného obdobia došlo tiež k spomaleniu dynamiky HDP. Podobný vývoj (rastový trend, dočasne prerušený hospodárskou krízou) bol typický aj pre export a import tovarov a služieb. Z grafov je tiež zrejmé, že v podmienkach SR dosahovali tieto premenné vyššie podiely na HDP počas celého obdobia. Potvrďuje to významnejšiu a stále sa zvyšujúcu otvorenosť SR, a teda aj vyššiu závislosť našej ekonomiky od vývoja v zahraničí, predovšetkým od hospodárskeho vývoja hlavných obchodných partnerov.

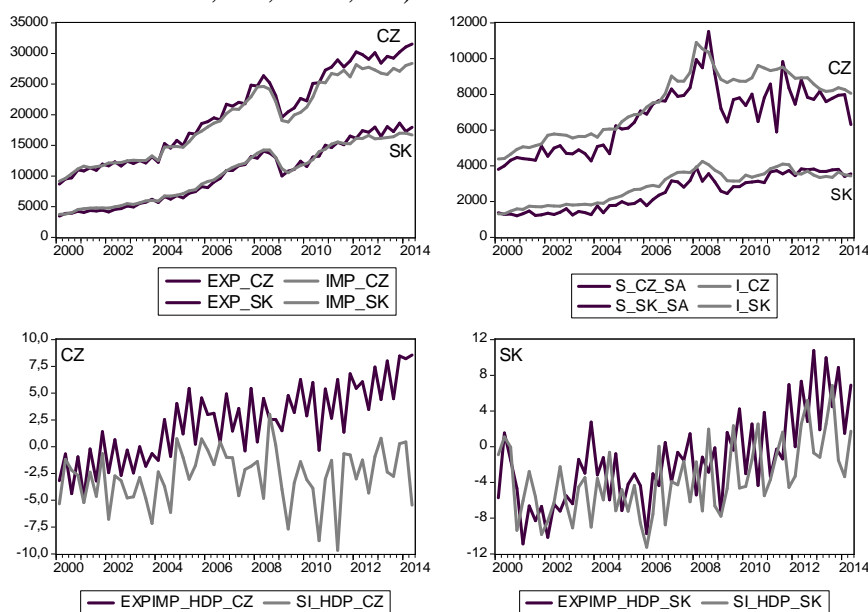


Obr. 1: HDP, export a import (mil. EUR); Zdroj: vlastné spracovanie, Eurostat

Obr.1 zároveň naznačuje aj vývoj čistého exportu oboch krajín. Čistý export, ako rozdiel exportu a importu tovarov a služieb, možno použiť aj ako indikátor vývoja celého bežného účtu platobnej bilancie krajiny, keďže tieto dve položky zvyčajne predstavujú najväčší podiel na bežnom účte. Grafy ukazujú, že kým v prípade ČR bol čistý export záporný len zo začiatku obdobia 2000-2004, v SR nastala významnejšia zmena až v súvislosti s hospodárskou krízou, keď sa dlhodobo záporný čistý export dostal do plusu.

Keďže kladný čistý export (resp. kladné saldo bežného účtu) predstavuje prebytok príjmov nad výdavkami, zvykne sa často porovnávať s rozdielom domácich úspor a investícií (čisté úspory). Nasledujúce grafy na Obr. 2 ukazujú vývoj týchto štyroch premenných pre SR aj ČR. Je zrejmé, že vývoj úspor aj investícií sledoval celkový trend ekonomického vývoja krajín – rast v období pred hospodárskou krízou, prepád po kríze a následné oživenie. Vývoj úspor bol viac volatilný v podmienkach ČR, predovšetkým v období po roku 2009. Trend vývoja by bolo možné označiť skôr ako mierne rastúci. V prípade investícií možno hovoriť skôr klesajúcich hodnotách, predovšetkým v podmienkach ČR. Napriek tomu kvartálne hodnoty investícií v posledných 5 rokoch presahovali hodnoty pre úspory. V SR bol vývoj aj po roku

2009 viac stabilný, nedochádzalo k významným výkyvom ani na strane úspor ani investícií. Posledné 2 grafy na Obr. 2 porovnávajú vývoj salda bežného účtu v %HDP vyjadrený dvoma spôsobmi, tzn. odhadnutý prostredníctvom vývoja čistého exportu a tiež ako rozdiel v hodnote úspor a investícií. Kým v podmienkach SR bol trend vývoja vyjadrený týmito dvoma spôsobmi značne podobný, v ČR to možno povedať len o období do roku 2009. Následne možno pre ČR sledovať nárast hodnôt exportu a importu (tiež salda čistého exportu), ale aj pokles investícií a vyššiu volatilitu úspor (tiež vyššiu volatilitu čistých úspor). Je zrejmé, že vývoj vyjadrený týmito dvoma spôsobmi je viac podobný v prípade SR ako v prípade ČR (korelačný koeficient - ČR: 0,257; SR:0,455).



Obr. 2: Vývoj exportu, importu, úspor a investícií (mil. EUR), čistého exportu a čistých úspor (%HDP); Zdroj: vlastné spracovanie, Eurostat

Tab. 1 uvádza hodnoty pre minimum, maximum, medián a priemer pre čistý export a čisté úspory vyjadrené v mil. EUR a v %HDP pre obidve krajiny. Hodnoty ukazujú vyššiu stabilitu vývoja čistého exportu v podmienkach ČR. V % vyjadrení bol rozptyl hodnôt pre ČR -4,68% HDP až 8,55% HDP, s mediánom na úrovni 2,53% HDP. Naproti tomu v prípade SR sa hodnoty čistého exportu pohybovali v rozmedzí -10,88% HDP až 10,76% HDP. Medián na úrovni -1,51% HDP potvrdzuje prevažne záporný vývoj hodnôt v sledovanom období.

Tab. 1: čistý export a čisté úspory (mil. EUR, %HDP), 2000-2014

	EXIM CZ	SI CZ	EXIM SK	SI SK	EXIMHDP CZ	SIHDP CZ	EXIMHDP SK	SIHDP SK
priemer	892,2121	-781,7707	-30,98448	-347,8276	2,294409	-2,739042	-1,506676	-3,523009
medián	877,5000	-624,4000	-216,4000	-408,6500	2,526838	-2,523299	-1,515282	-4,083450
maximum	3159,700	1230,700	1922,800	1235,200	8,552721	3,037026	10,75963	6,839348
minimum	-800,1000	-3781,600	-1166,200	-1205,900	-4,679577	-9,657064	-10,87557	-11,24832
kvartály	58	58	58	58	58	58	58	58

Zdroj: vlastné spracovanie, Eurostat

4. Porovnanie vývoja vonkajšej nerovnováhy pred a po kríze

Predchádzajúce grafy naznačujú, že v období po roku 2009 došlo k zmene vo vývoji sledovaných ukazovateľov v ČR aj SR, preto bol vývoj porovnaný aj osobitne pre obdobia 2000-2008 a 2009-2014. Pri ukazovateli čistý export v % HDP je v oboch prípadoch zrejmy posun smerom ku kladným hodnotám, teda už spomínaná zmena zo záporného salda na pozitívne. Tento vývoj možno spojiť s viacerými vysvetleniami. Kríza prispela k oslabeniu

domáceho dopytu, čo so sebou prinieslo aj pokles importovanej produkcie. Oživenie po kríze sa prejavilo tiež posilnením domácej priemyselnej produkcie určenej na export a s tým súvisiaci nárast exportu v ČR aj SR. Ako naznačuje Obr. 2, v danom období síce narastal aj import, no jeho prírastky neboli natoľko významné ako v prípade exportu. Ďalším faktorom v podmienkach ČR bolo oslabenie kurzu českej koruny, ktoré podporilo domácich exportérov a rovnako mohlo prispieť k istému príbrzdeniu vo vývoji importu.

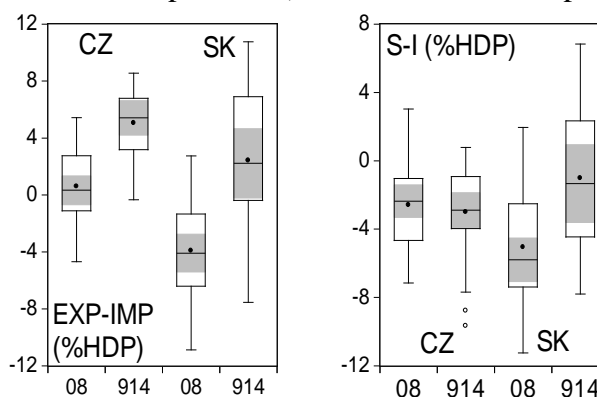
V Tab. 2 sú uvedené hodnoty pre minimum, maximum, medián a priemer pre čistý export a čisté úspory v %HDP pre dve obdobia. Pri porovnaní priemerných hodnôt a mediánov pred a po kríze možno pozorovať, že s výnimkou čistých úspor v %HDP pre ČR, došlo k významnejšej zmene- nárastu hodnôt. Porovnanie minima a maxima však už ukazuje pokles hodnôt čistých úspor v ČR po roku 2009. Na základe toho by sme mohli konštatovať, že vývoj v období pred a po kríze bol odlišný.

Tab. 2: Čistý export a čisté úspory pred a po kríze (2000-2008 a 2009-2014)

	EXIMGDP 08_CZ	EXIMGDP 914_CZ	EXIMGDP 08_SK	EXIMGDP 914_SK	SIGDP 08_CZ	SIGDP 914_CZ	SIGDP 08_SK	SIGDP 914_SK
priemer	0,606247	5,056857	-3,910980	2,427641	-2,578215	-3,002213	-5,057713	-1,011674
medián	0,335906	5,420678	-4,076993	2,229404	-2,356318	-2,883343	-5,789799	-1,321926
maximum	5,437048	8,552721	2,748852	10,75963	3,037026	0,788637	1,957757	6,839348
minimum	-4,679577	-0,324438	-10,87557	-7,534809	-7,151031	-9,657064	-11,24832	-7,791310
kvartály	36	22	36	22	36	22	36	22

Zdroj: vlastné spracovanie, Eurostat

Tieto zistenia potvrdzuje aj Obr. 3, ktorý prostredníctvom boxplotov ilustruje zmenu hodnôt medzi obdobia. Posun smerom nahor ku kladným hodnotám je viditeľný vo všetkých prípadoch, okrem S-I pre ČR. Tam aj napriek vyššej volatilitate hodnôt v peňažnom vyjadrení (viď Obr.2), rozptyl hodnôt v % HDP po roku 2009 mierne poklesol a zároveň došlo k poklesu priemeru a mediánu po kríze (viď sivo označené stĺpce v Tab. 1-2).



Obr. 3: Čistý export a čisté úspory v 2000-2008 a 2009-2014Q2 (%HDP); Zdroj: vlastné spracovanie, Eurostat

Na základe sledovaných premenných v období pred a po kríze možno konštatovať, že došlo k miernej zmene vo vývoji premenných a zmierneniu vonkajšej nerovnováhy týchto krajín, čo potvrdilo istú podobnosť v prípade čistého exportu. Vývoj čistých úspor naznačuje určité odlišnosti, a to minimálnu zmenu čistých úspor v % HDP prípade ČR a nárast a postupné približovanie tohto ukazovateľa k nule v prípade SR. Skúmanie možných príčin rozdielneho vývoja čistých úspor sa s ohľadom na rozsah článku nebolo možné venovať.

5. Záver

Príspevok bol zameraný na porovnanie vývoja vonkajšej rovnováhy v podmienkach SR a ČR, členskej a nečlenskej ekonomiky eurozóny, ktoré vykazujú vo všeobecnosti mnohé

podobné znaky makroekonomického vývoja. Pre analýzu boli zvolené nasledujúce premenné: export a import tovarov a služieb, domáce úspory a investície, čistý export a čisté úspory.

Vývoj exportu aj importu oboch krajín sledoval do krízy podobný trend, aj keď v prípade SR boli hodnoty vo vzťahu k HDP významne vyššie. Po kríze je možné sledovať v oboch prípadoch posilnenie exportu a isté spomalenie dynamiky importu, následkom čoho došlo k preklopeniu salda čistého exportu do plusových hodnôt. Toto bolo vo väčšej miere pozorovateľné v SR. Pri porovnaní vývoja úspor, investícií a čistých úspor vidíme výraznejšie odlišnosti. Kým v SR boli významnejšie výkyvy len v období krízy, v ČR sa volatilita úspor a investícií zvýšila práve po kríze. Vyjadrenie v % HDP rovnako potvrdzuje vyššiu volatilitu týchto ukazovateľov a naznačuje isté zhoršenie a posun smerom k záporným hodnotám. Tento posun hodnôt bol však pri porovnaní dvoch období len minimálny. Naproti tomu boxploty pre všetky ostatné ukazovatele naznačujú posun smerom ku kladným hodnotám a „zlepšovanie“ v oblasti vonkajšej rovnováhy. Možno teda konštatovať, že vývoj vonkajšej rovnováhy v SR a ČR bol vo všeobecnosti ovplyvnený krízou skôr pozitívne a spôsobom, ktorý zodpovedá vývoju vo väčšine členských krajín eurozóny. Vyššia podobnosť sa preukázala pri čistom exporte ako v prípade čistých úspor.

Literatúra

BLANCHARD, O. – GIAVAZZI, F. 2002. Current Account Deficits in the Euro Area. The End of the Feldstein Horioka Puzzle? SSRN Working Paper Series 2003-05.

CAMPA, J. M. - GAVILAN, A. 2011. Current accounts in the euro area: An intertemporal approach. In: *Journal of International Money and Finance*, roč. 30, č. 1, s. 205–228. ISSN 0261-5606.

GAULIER, G.- VICARD, V. 2012. Current account imbalances in the euro area: competitiveness or demand shock? Banque de France Quarterly Selection of Articles, č. 27

GRUBER, J. W. - KAMIN, S. B. 2007. Explaining the global pattern of current account imbalances. In: *Journal of International Money and Finance*, roč. 26, č. 4, s. 500–522. ISSN: 0261-5606.

EUROSTAT, 2014. kvartálne údaje o vývoji hrubého domáceho produktu, exportu a importu tovarov a služieb, úspor a investícií ČR a SR za obdobie 1999-2014 [online]. Dostupné na: <http://epp.eurostat.ec.europa.eu>

KANG, J. S. - SHAMBAUGH, J. C. 2013. The Evolution of Current Account Deficits in the Euro Area Periphery and the Baltics: Many Paths to the Same Endpoint. IMF Working Paper WP/13/169. [online]. Dostupné na: <http://www.imf.org/external/pubs/ft/wp/2013/wp13169.pdf>

LANE, P. – MILESI-FERRETTI, J.M. 2012. External adjustment and the global crisis. In: *Journal of international economics*, roč. 88, č. 2, s. 252-265. ISSN: 0022-1996

SCHMITZ, B. - VON HAGEN, J. 2011. Current account imbalances and financial integration in the euro area. In: *Journal of International Money and Finance*, roč. 30, č. 8, s. 1676–1695. ISSN: 0261-5606

Príspevok vznikol v rámci riešenia projektu VEGA č. 1/0892/13

Adresa autorky:

Ludmila Bartóková, Ing. PhD.
Ekonomická fakulta TUKE
Němcovej 32, 040 01 Košice
ludmila.bartokova@tuke.sk

Vývoj prostorové distribuce příjmové nerovnosti v Evropě v letech 2007-2012

Development of spatial distribution of income inequality in Europe in the years 2007-2012

Jitka Bartošová, Jakub Stejskal

Abstract: This paper focuses on the analysis of the development of the spatial distribution of income inequality in European countries in the years 2007 – 2012. The level of income inequality are evaluated by the Gini index and the coefficient of income inequality s_{80}/s_{20} . Values for the period 2007 – 2012, when the European economy was hit by the global economic crisis and affected the trend and stability of its development, were used to short-term trend analysis of indicators of inequality. The crisis worsened the social situation of the most vulnerable groups of the population, especially young people integrating into the workforce, but also for people approaching retirement age who are at the peak of their career. The main cause is the reduction of jobs, accompanied by the rising level of unemployment. It is expected that this Europe-wide phenomenon will be reflected in the level of income inequality in each country.

Abstrakt: Předložený příspěvek je zaměřen na analýzu vývoje prostorové distribuce příjmové nerovnosti v zemích Evropy v letech 2007 – 2012. Míra příjmové nerovnosti je zde hodnocena prostřednictvím Giniho indexu a koeficientu příjmové nerovnosti s_{80}/s_{20} . K analýzám krátkodobého vývoje ukazatelů nerovnosti, byly použity hodnoty za období 2007 – 2012, kdy evropskou ekonomiku zasáhla celosvětová hospodářská krize a ovlivnila trend i stabilitu jejího vývoje. V důsledku krize došlo ke zhoršení sociální situace těch nejslabších skupin obyvatelstva, především mladých lidí, začleňujících se do pracovního procesu, ale i lidí v předdůchodovém věku, kteří jsou na vrcholu své pracovní kariéry. Hlavní příčinou je redukce pracovních míst doprovázená růstem nezaměstnanosti. Lze očekávat, že tento celoevropský fenomén se odrazí i v úrovni příjmové nerovnosti v jednotlivých zemích.

Key words: personsequivalised income, EU-SILC, the Gini index, quintile coefficient of income inequality.

Klíčové slová: ekvivalizovaný příjem osoby, EU-SILC, Giniho index, kvantilový koeficient příjmové nerovnosti.

JEL classification: C14, D31, D63

1. Introduction

The global economic crisis, accompanied by high unemployment rate and business uncertainty, together with the growing indebtedness of public and private sector, has resulted in the overall economic and political destabilization and adversely affects cooperation of states within the EU. In this difficult situation where governments are forced to reduce their deficit budgets by all possible means, there is also more global phenomenon of aging population coupled with other mandatory increases in state spending. The poorest people were most affected by the current economic crisis, i.e. households at risk of monetary poverty and social exclusion. European countries are currently faced with the task to design and implement individual measures, which – in line with the global strategy of 2020 – would lead to mitigation of the negative impact of the economic crisis on the population.

Awkwardness and topicality of this issue is demonstrated by a number of articles devoted to the analysis of income inequality, the risk of monetary poverty and material deprivation, unemployment and demographic developments that were recently published in professional

and scientific journals in the Czech Republic and abroad. From the Czech and Slovak publications we can mention works of Bartošová and Želinský (2013), Bílková (2012, b), Fiala and Langhamrová (2013 and 2014), Löster and Langhamrová (2012), Michálek and Veselovská (2012), Pauhofová (2012), Stankovičová, Vlačuha and Ivančíková (2013), Tartal'ová and Želinský (2012), Večerník (2013), Želinský (2012), Želinský and Pauhofová (2013). Numerous other publications closely related to this issue deal with modelling of revenue, respectively expenditure distribution and diagnostic factors that affect it. For example, in recent years were published works by Malá (2013), Marek (2013), Mark and Vrabec (2013), Pacáková, Linda and Sipková (2012), Řezanková and Löster (2013), Řezanková and Želinský (2014), Sipková and Sipka (2012), Šimpach (2012 and 2013), Želinský and Stankovičová (2012) and others.

2. Indicators of income inequality

Income inequality is usually represented in terms of the *Lorenz curve* and the basic characteristics (the *Gini index*, *Robin Hood index*), which were derived by Gini, Pareto, Pietro and other statisticians. These measures, however, do not allow to quantify the contribution of inequality of individual subgroups (subpopulations) to total inequality across heterogeneous population. Other less well-known statistics designed by Thiel – *Theil L index* or *Theil T index* – can be used for this purpose. A well-known and very simple indicator is the *quintile coefficient of income inequality* s_{80}/s_{20} which is robust in comparison with measures mentioned before.

2.1. Lorenz curve and its derivative indicators of inequality

Lorenz curve ($L(p)$) serves for graphical display of inequality in income distribution. It displays the cumulative frequency of total revenue or expenditure (y -axis) in dependence on the cumulative number of individuals (i.e. households) on the x -axis, sorted in ascending order according to the size of their income.

The *Gini index* (G) which indicates the extent to which the distribution of income among individuals (i.e. households) deviates from a perfectly uniform distribution is the best known indicator of inequality derived from the Lorenz curve. Its value is calculated by dividing the area between the line of absolute equality (diagonal $y = p$) and the Lorenz curve $L(p)$ to the total area below the diagonal.

Another measure of inequality based on the Lorenz curve is the *Robin Hood Index* (RHI), which determines the percentage of income would need to be redistributed in order to achieve absolute uniformity. Its size is equivalent to the maximum vertical distance between the line of perfect equality and the Lorenz curve.

2.2. Other indicators of inequality

Atkinson index of inequality (A) is constructed on the basis of on the so-called *fair average income* x_e ¹. Its value is determined by the relative difference between the real and fair value of the average income of the group.

Less known are the measures of inequality based on generalized entropy – *Theil L index*, respectively *Theil T index*. Decomposition of the overall rate of inequality on the individual components is the main advantage of these indicators compared to the *Gini index* (see Labudová, Vojtková and Linda, 2010 or Labudová, 2012).

¹Fair average income is defined as income that in case of evenly distributed among the receivers creates the same level of social welfare as the present distribution.

The *quintile coefficient of income inequality* s_{80}/s_{20} is very simple and often published indicator which represents the ratio of eighty and twentieth quintile values equivalised income. This indicator is robust, unlike previous measurements, which are not.

3. The data base

The source of information from which these quantitative analyses are based, is a large sample survey EU-SILC (*European Union Statistics on Income and Living Conditions*). This is an annual survey, which is mandatory for all Member States of the European Union and allows to generalize the results of the analyses. Its purpose is to obtain representative data on household income, quality and financial cost of housing, availability of objects of consumer consumption as well as working, physical and health conditions of adults living in the household. Data sets EU-SILC are the main source of information about living standards and poverty level in the Member States of the European Union and some other European countries.

3.1. Consumer unit and equivalised income

Spatial benchmarking, differentiation and inequality of income distribution within the European Union requires the representation of income in the single currency (euro) and conversion of value due to purchasing power parity. The comparability of the financial situation of households with different number of members of a different age structures is ensured by recalculation of total disposable income to so-called *consumer unit*. *Equivalised value income* is acquired by the recalculation. Construction of *consumer unit* is generally based on the assumption that the social situation of households depends on the total annual household income and, secondly:

- the joint expenses serving for running the household (housing, water and energy, objects of consumer consumption, etc.)
- on individual spending, which serves to satisfy the needs of individuals living in the same household (food and non-alcoholic beverages, alcohol and tobacco, clothing and footwear etc.).

For this reason, the size of consumer units, representing the equivalent household size, fully (weight 1) includes only one person – i.e. head of the household. Weights of the other members of the household are lower. All other members of the household are categorized into two groups according to age. The first group consists of children aged 0-13 years who is assigned a weight of 0.5, the second group consists of other household members who are in the size of the consumer unit counted with a weight of 0.3. Thus we get so-called *OECD¹ modified unit*, which represents the equivalent size of the household and is used by Eurostat to calculate the official statistics within the EU.

Equivalised income is obtained by dividing the total disposable household income by the equivalent size (consumer unit). In assessing the situation of individuals (persons in the household) equivalised value of income is assigned to each member of the household.

4. Development of income inequality in European countries in the years 2007 - 2012

Development of income inequality of people in European countries since the start of the global economic crisis – particularly between 2007 and 2012 – is shown in Table 1. Non-robust *Gini index* and robust *quintile coefficient of income inequality* is used to the assessment of income inequality, which is the proportion of quintile s_{80}/s_{20} .

¹Organization for Economic Cooperation and Development

Tab.2: Income inequality of people in European countries in the years 2007 – 2012 as measured by the Gini index (G) and the coefficient of income inequality

Country code ¹	2007		2008		2009		2010		2011		2012		Differences 2012–2007	
	G	S _{80/S20}	G	S _{80/S20}	G	S _{80/S20}	G	S _{80/S20}	G	S _{80/S20}	G	S _{80/S20}	G	S _{80/S20}
EU27	30,6	5,0	30,9	5,0	30,5	5,0	30,5	5,0	30,8	5,1	30,6	5,1	0,0	0,1
AT	26,2	3,8	26,2	3,7	25,7	3,7	26,1	3,7	26,3	3,8	27,6	4,2	0,6	0,4
BE	26,3	3,9	27,5	4,1	26,4	3,9	26,6	3,9	26,3	3,9	26,6	3,9	0,3	0,0
BG	35,3	7,0	35,9	6,5	33,4	5,9	33,2	5,9	35,0	6,5	33,6	6,1	-1,7	-0,9
CY	29,8	4,4	29,0	4,3	29,5	4,4	30,1	4,5	29,2	4,3	31,0	4,7	0,2	0,3
CZ	25,3	3,5	24,7	3,4	25,1	3,5	24,9	3,5	25,2	3,5	24,9	3,5	-0,4	0,0
DE	30,4	4,9	30,2	4,8	29,1	4,5	29,3	4,5	29,0	4,5	28,3	4,3	-2,1	-0,6
DK	25,2	3,7	25,1	3,6	26,9	4,6	26,9	4,4	27,8	4,4	28,1	4,5	2,9	0,8
EE	33,4	5,5	30,9	5,0	31,4	5,0	31,3	5,0	31,9	5,3	32,5	5,4	-0,9	-0,1
ES	31,9	5,5	31,9	5,7	33,0	6,4	34,4	7,2	34,5	7,1	35,0	7,2	3,1	1,7
FI	26,2	3,7	26,3	3,8	25,9	3,7	25,4	3,6	25,8	3,7	25,9	3,7	-0,3	0,0
FR	26,6	3,9	29,8	4,4	29,9	4,4	29,8	4,4	30,8	4,6	30,5	4,5	3,9	0,6
GR	34,3	6,0	33,4	5,9	33,1	5,8	32,9	5,6	33,5	6,0	34,3	6,6	0,0	0,6
HU	25,6	3,7	25,2	3,6	24,7	3,5	24,1	3,4	26,8	3,9	26,9	4,0	1,3	0,3
CH	30,3	4,7	31,1	4,9	30,7	4,8	29,6	4,5	29,7	4,5	28,8	4,4	-1,5	-0,3
IC	28,0	3,9	27,3	3,8	29,6	4,2	25,7	3,6	23,6	3,3	24,0	3,4	-4,0	-0,5
IE	31,3	4,8	29,9	4,4	28,8	4,2	30,7	4,7	29,8	4,6	29,9	4,7	-1,4	-0,1
IT	32,2	5,5	31,0	5,1	31,5	5,2	31,2	5,2	31,9	5,6	31,9	5,5	-0,3	0,0
LT	33,8	5,9	34,0	5,9	35,9	6,4	37,0	7,3	33,0	5,8	32,0	5,3	-1,8	-0,6
LU	27,4	4,0	27,7	4,1	29,2	4,3	27,9	4,1	27,2	4,0	28,0	4,1	-0,6	0,1
LV	35,4	6,4	37,5	7,3	37,5	7,4	35,9	6,8	35,1	6,5	35,7	6,5	-0,3	0,1
MT	26,3	3,9	28,1	4,3	27,4	4,0	28,6	4,3	27,2	4,0	27,1	3,9	-0,7	0,0
NL	27,6	4,0	27,6	4,0	27,2	4,0	25,5	3,7	25,8	3,8	25,4	3,6	-2,2	-0,4
NO	23,7	3,5	25,1	3,7	24,1	3,5	23,6	3,4	22,9	3,3	22,5	3,2	-1,2	-0,3
PL	32,2	5,3	32,0	5,1	31,4	5,0	31,1	5,0	31,1	5,0	30,9	4,9	-1,3	-0,4
PT	36,8	6,5	35,8	6,1	35,4	6,0	33,7	5,6	34,2	5,7	34,5	5,8	-2,3	-0,7
RO	37,8	7,8	36,0	7,0	34,9	6,7	33,3	6,0	33,2	6,2	33,2	6,3	-4,6	-1,5
SE	23,4	3,3	24,0	3,5	24,8	3,7	24,1	3,5	24,4	3,6	24,8	3,7	1,4	0,4
SK	24,5	3,5	23,7	3,4	24,8	3,6	25,9	3,8	25,7	3,8	25,3	3,7	0,8	0,2
SL	23,2	3,3	23,4	3,4	22,7	3,2	23,8	3,4	23,8	3,5	23,7	3,4	0,5	0,1
UK	32,6	5,3	33,9	5,6	32,4	5,3	32,9	5,4	33,0	5,3	32,8	5,4	0,2	0,1

From the values given in Table 1 it is clear that the order of states with the lowest inequality in 2012 compared to 2007 has changed somewhat. In 2007, the lowest inequality

¹EU27 – 27 Member States of the EU, AT - Austria BE - Belgium BG - Bulgaria CY - Cyprus CZ - Czech Republic DE - Germany. DK - Denmark, EE - Estonia, ES - Spain FI - Finland FR - France, GR - Greece, HU - Hungary, CH - Switzerland, IC - Island IE - Ireland IT - Italy, LT - Latvia, LU - Luxembourg , LV - Lithuania, MT - Malta, NL - Netherlands NO - Norway PL - Poland PT - Portugal, RO - Romania SE - Sweden, SL - Slovenia SK - Slovakia UK - United Kingdom

measured by the *Gini index* was in Slovenia (23.2), followed by Sweden (23.4), then by Norway (23.7), Slovakia (24.5), Denmark (25.2), Czech Republic (25.3) and other countries. When using the *quintile coefficient of income inequality* s_{80}/s_{20} the order has been slightly altered – from Slovenia and Sweden (3.3) over Norway, Czech Republic and Slovakia (3.5), Denmark, Finland and Hungary (3.7) and Austria (3.8) to Belgium, France, Iceland and Malta (3.9) and others.

In 2012, the lowest inequality measured by the *Gini index* was in Norway (22.5), followed by Slovenia (23.7), followed by Iceland (24.0), Sweden (24.8), Czech Republic (24.9), Slovakia (25.3) and others. When using the *quintile coefficient of income inequality*, situation was quite similar – Norway (3.2), Iceland and Slovenia (3.4), Czech Republic (3.5), the Netherlands (3.6) and Finland, Sweden, Slovakia (3.7) and others.

Conversely, the highest values of the *Gini index* in 2007 were in Romania (37.8), followed by Portugal (36.8), Lithuania (35.4), Bulgaria (35.3), Greece (34.3), Latvia (33.8), Estonia (33.4), and other countries. *Quintile coefficient of income inequality* order of the states with the highest inequality partly altered, and arrayed them from Romania (7.8) over Bulgaria (7.0), Portugal (6.5), Lithuania (6.4), Greece (6.0) and Latvia (5.9) to Spain (5.5) and others.

In 2012, the *Gini index* reached the highest levels in Spain (35), followed by Portugal (34.5), Greece (34.3), Bulgaria (33.6), Romania (33.2), United Kingdom (32, 8) and other countries. *Quintile coefficient of income inequality* order of the states with the highest inequality changed somewhat, and arrayed them from Spain (7.2) via Greece (6.6), Lithuania (6.5), Romania (6.3) and Bulgaria (6.1) to Portugal (5.8) and others.

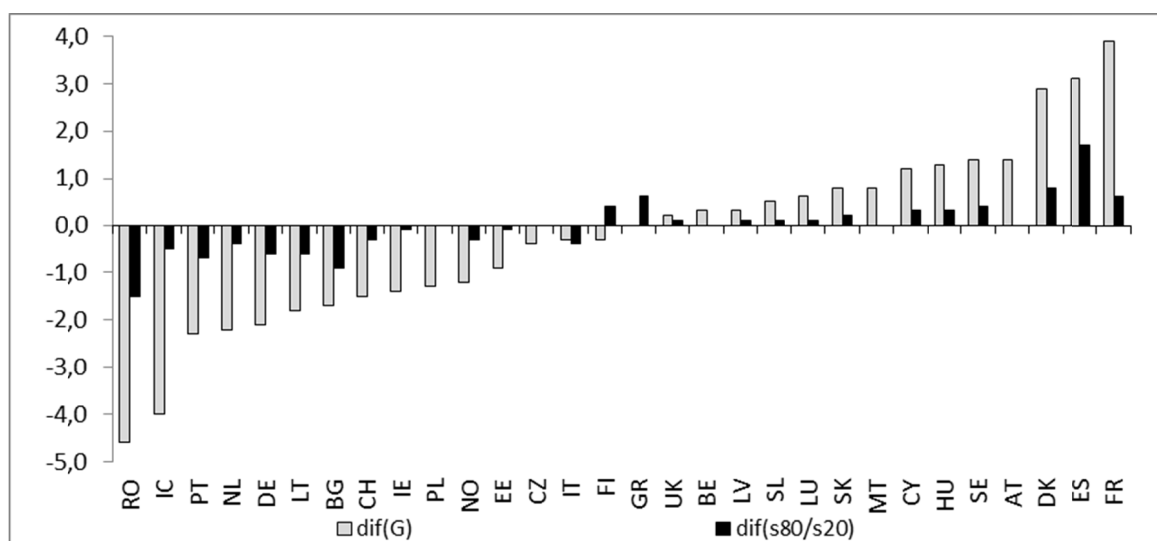


Fig.2: Average differences of the Gini index ($dif(G)$) and quintile coefficient of income inequality ($dif(s_{80}/s_{20})$) in the period 2007 – 2012

Table 1 and Figure 1 also shows that in 2012 – compared to 2007 – the level of inequality increased in 11 European countries – Austria (0.6), Belgium (0.2), Cyprus (0.3), Denmark (-2.9), Spain (-3.1), France (3.9), Hungary (1.3), Sweden (1.4), Slovakia (0.8), Slovenia (0.5) and the UK (0.2) according to the *Gini index*. The maximum increase in the value of the *Gini index* was therefore in France, second and third place belongs to Spain and Denmark respectively. We got similar results, even when using the robust *quintile coefficient of income inequality* s_{80}/s_{20} which indicated an increase in inequality in 13 European countries – Austria, Cyprus, Denmark, Spain, France, Greece, Hungary, Luxembourg, Lithuania,

Sweden, Slovakia, Slovenia and the United Kingdom. The maximum increase was recorded in Spain (1.7), second place went to Denmark (0.8) and third were France and Greece (0.6).

Next figures (Fig. 2 – Fig. 4) demonstrate development of the *Gini index* in European countries in the period 2007 – 2012. Figure 2 includes development in 9 countries (including the Czech Republic and Slovakia) with the lowest values of the Gini index – countries located deep below the EU27 average. Figure 3 shows the development of the Gini index in 8 other countries that are below the EU27 average. The last figure (Fig. 4) shows the development of the Gini index in countries situated above the EU27 average.

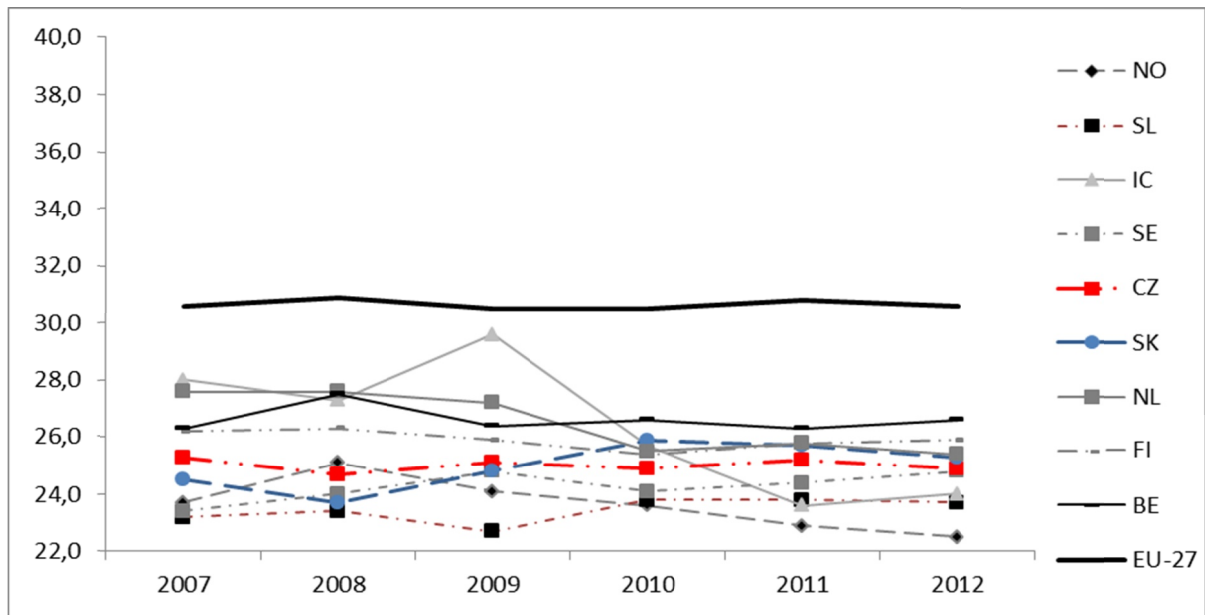


Fig.2: Development of the Gini index in the years 2007 – 2012 in countries located deep below the EU27 average

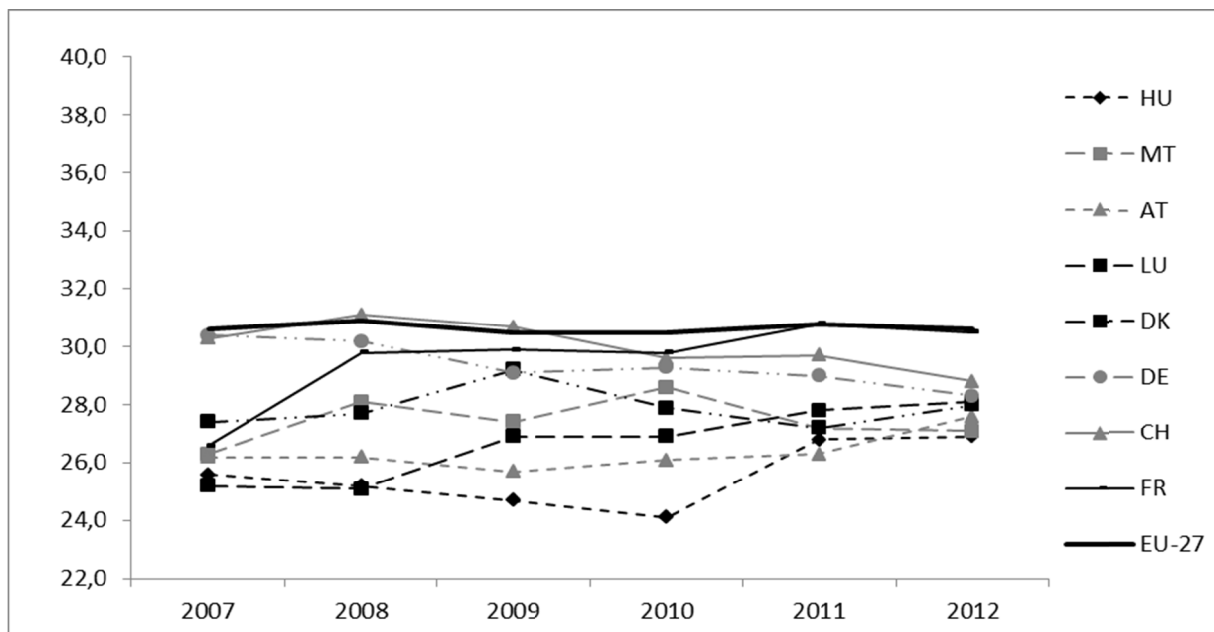


Fig.3: Development of the Gini index in the years 2007 – 2012 in countries situated just below the EU27 average

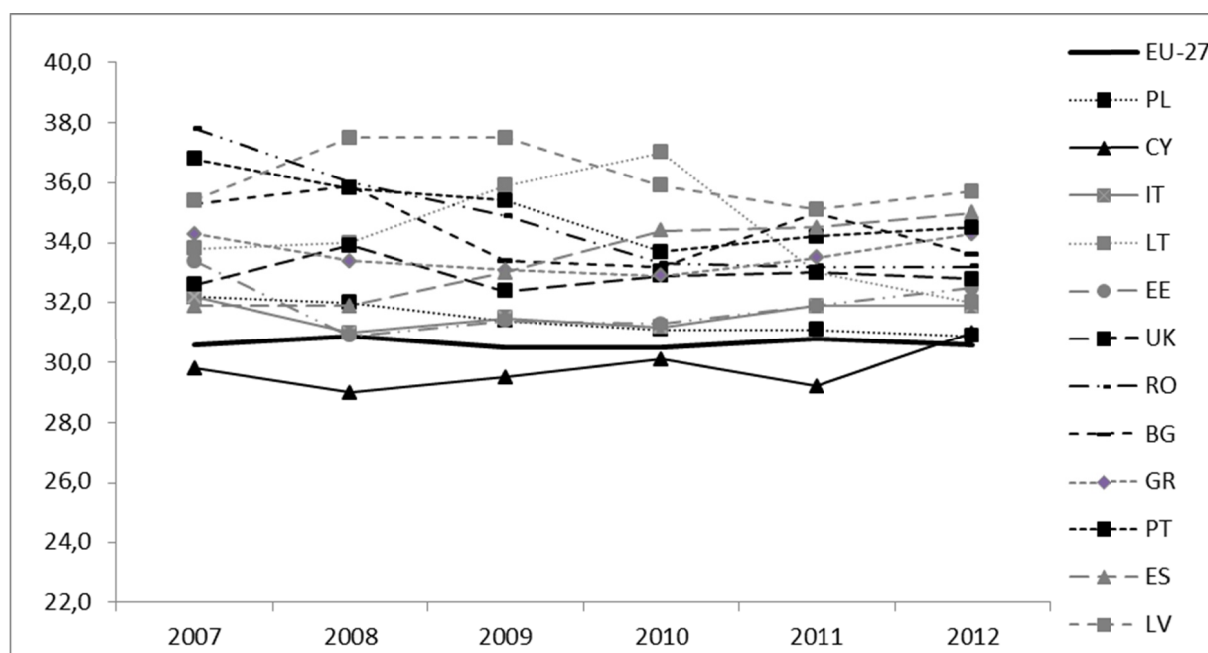


Fig.4: Development of the Gini index in the years 2007 - 2012 in countries situated above the EU27 average

5. Conclusion

An analysis of the current development of spatial distribution of income inequality in European countries between 2007 and 2012 (from the beginning of the global economic crisis) was performed in this paper. Changes in income inequality population of individual states and their position in the European ranking and the changes that have occurred due to the economic crisis since 2007 were measured through two well-known and often officially published indicators – non-robust *Gini index* and robust *quantile coefficient of inequality*. Countries were divided into three categories according to their position towards the EU27 average to increase the clarity of development of the *Gini index*.

The Czech Republic and Slovakia belongs to the group of countries with the lowest income inequality – they are both located deep below the EU27 average. The value of the *Gini index* ranked Czech Republic and Slovakia (in that order) on the 6th and 4th place in 2007 and the 5th and 6th position in 2012. According to the values of *quantile coefficient of inequality* in 2007 Czech Republic and Slovakia jointly shared the 2nd, 3rd and 4th place in European rankings with Norway. In 2012, the Czech Republic was fourth and Slovakia shared 6th, 7th and 8th place with Finland and Sweden. Extreme positions were estimated by both indicators consistently – countries with the lowest inequality were Slovenia in 2007 (according to robust indicator along with Sweden) and Norway in 2012. Romania was the country with the highest inequality in both years. It is clear that the choice of indicator of income inequality affects the results obtained in both absolute (value) and relative way (order of ranking). Indicator selection plays an important part when assessing the spatial distribution of income inequality in Europe and ranking states in relevant rankings.

By the calculation of the average absolute changes (average difference), it was found that in most cases (15 countries) a decrease of the *Gini index* was recorded. The *Gini index* increased in 11 remaining countries and did not change in one case (Greece). It can therefore be concluded that the impact of the economic crisis on income inequality in Europe was moderate and a positive development of inequality reduction continued in most cases.

Acknowledgment

This research was supported by internal project of the University of Economics in Prague (no. F6/30/2014).

References

- BARTOŠOVÁ, J. 2013. Finanční potenciál domácností – kvantitativní metody a analýzy. Praha: Professional Publishing.
- BARTOŠOVÁ, J. – ŽELINSKÝ, T. 2013. Extent of poverty in the Czech and Slovak Republics fifteen years after split. In: *Post-Communist Economies*, Vol. 25, No. 1. pp. 119–131.
- BÍLKOVÁ, D. 2012a. Recent Development of the Wage and Income Distribution in the Czech Republic. In: *Prague Economic Papers*. Vol. 21, No. 2. pp. 233–250.
- BÍLKOVÁ, D. 2012b. Development of Wage Distribution of the Czech Republic in Recent Years by Highest Education Attainment and Forecasts for 2011 nad 2012. In: Löster, T., Pavelka, T. (eds.): 6th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 162–182.
- FIALA, T. – LANGHAMROVÁ, J. 2013. Vývoj ekonomického a sociálního zatížení a stárnutí populace. In: *Politická ekonomie*, Vol. 61, No. 3. pp. 338–355.
- FIALA, T. – LANGHAMROVÁ, J. 2014. Modelling of the Future Development of the Total Amount of Premium Paid and Total Amount of Old-Age Pensions in the Czech Republic. In: *Politická ekonomie*, Vol. 62, No. 2. pp. 232–248.
- LABUDOVÁ, V. – VOJTKOVÁ, M. – LINDA, B. 2010. Application of multidimensional methods to measure poverty. In: *E+M Ekonomie a management*. Vol. 13, No. 1, pp. 6–21.
- LÖSTER, T. – LANGHAMROVÁ, J. 2012. Disparities between Regions of the Czech Republic for Non-business Aspects of Labour Market. In: Löster, T., Pavelka, T. (eds.): 6th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 689–702.
- MALÁ, I. 2013. Použití konečných směsí logaritmicko-normálních rozdělení pro modelování příjmů českých domácností. In: *Politická ekonomie*, Vol. 61, No. 3. pp. 356–372.
- MAREK, L. 2013. Some Aspects of Average Wage Evolution in the Czech Republic. In: Löster, T., Pavelka, T. (eds.): 7th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 947–958.
- MAREK, L. – VRABEC, M. 2013. Probability Models of Wage Distribution. In: Vojáčková, H. (ed.): *Proceedings of 31th International Conference Mathematical Methods in Economics 2013*. Jihlava: Vysoká škola polytechnická. pp. 575–581.
- MICHÁLEK, A. – VESELOVSKÁ, Z. 2012. Vývoj a komparácia vybraných charakteristík nerovnosti a chudoby v krajinách EU. In: Pauhofová, I., Želinský, T. (eds.): *Nerovnosť a chudoba v Európskej únii a na Slovensku*. Košice: TU Košice. s. 11–22.
- PACÁKOVÁ, V. – LINDA, B. – SIPKOVÁ, L'. 2012. Distribution and Factors of the Highest Wages in the Slovak Republic. In: *Ekonomický časopis*, Vol. 60, No. 9. pp. 918–934..
- PAUHOFOVÁ, I. 2012. Generovanie chudoby vo vidieckych regiónoch Slovenska v krízovom období. In: Pauhofová, I., Želinský, T. (eds.): *Nerovnosť a chudoba v Európskej únii a na Slovensku*. Košice: TU Košice. s. 49–56.
- ŘEZANKOVÁ, H. – LÖSTER, T. 2013. Shluková analýza domácností charakterizovaných kategoriálními ukazateli. In: *E+M. Ekonomie a Management*, Vol. 16, No. 3. pp. 139–147.

- ŘEZANKOVÁ, H. – ŽELINSKÝ, T. 2014. Factors of Material Deprivation Rate in the Czech Republic by Household Type. In: Ekonomický časopis, Vol. 62, No. 4. pp. 394–410.
- SIPKOVÁ, L'. – SIPKO, J. 2012. Analysis of Income Inequality of Employees in the Slovak Republic. In: Löster, T., Pavelka, T. (eds.): 6th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 1032–1042.
- STANKOVIČOVÁ, I. – VLAČUHA, R. – IVANČÍKOVÁ, L. 2013. Trend Analysis of Monetary Poverty Measures in the Slovak and Czech Republic. In: Löster, T., Pavelka, T. (eds.): 7th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 1334–1343.
- ŠIMPACH, O. 2012. Logit and Probit Models in the Probability Analysis: Change in the Probability of Death of Celiac Disease Patients. In: Statistika, Vol. 49, No. 4. pp. 67–80.
- ŠIMPACH, O. 2013. Application of Cluster Analysis on Demographic Development of Municipalities in the Districts of Liberecký Region. In: Löster, T., Pavelka, T. (eds.): 7th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 1390–1399.
- VEČERNÍK, J. 2013. The changing role of education in the distribution of earnings and household income: the Czech Republic in 1988–2009. In: Economics of Transition, Vol. 21, No.1. pp. 111–133.
- ŽELINSKÝ, T. 2013. Estimation of Subjective (Income) Poverty Lines :Sensitivity-Specificity Curves Approach. In: Hu, J. (ed.): Social Sciences and Society. Book Series: Advances in Education Research, Vol. 6. pp. 42–47.
- ŽELINSKÝ, T. – PAUHOFŮVÁ, I. 2013. High Income Households in Eastern EU Countries. In: Hu, J. (ed.): 3rd Conference on Applied Social Science (ICASS 2013), Vol. 1. pp. 73–78.
- ŽELINSKÝ, T. – STANKOVIČOVÁ, I. 2012. Spatial Aspects of Poverty in Slovakia. In: Löster, T., Pavelka, T. (eds.): 6th International Days of Statistics and Economics, Conference Proceedings. Slaný: Melandrium. pp. 1228–1235.

Authors' addresses:

Jitka Bartošová, doc. RNDr., Ph.D.
Fakulta managementu, VŠE v Praze
Jarošovská 1117/II, 37701 Jindřichův Hradec
bartosov@fm.vse.cz

Jakub Stejskal, Ing.
Fakulta managementu, VŠE v Praze
Jarošovská 1117/II, 37701 Jindřichův Hradec
jakubstejskal@seznam.cz

Poznámka k váženému dynamickému rozkladu finančného ukazovateľa na báze multiplikatívnej väzby

A note on weighted dynamic decomposition of a financial indicator based on a multiplicative relationship

Martin Boďa

Abstract: Following a recent suggestion as to how to incorporate weights of subjective significance into the dynamic decomposition of a financial indicator under a multiplicative relationship, an improved procedure is posited and demonstrated in the paper.

Abstrakt: Vychádzajúc z nedávneho návrhu, ako zahrnúť do dynamického multiplikatívneho rozkladu finančného ukazovateľa váhy vyjadrujúce subjektívnu dôležitosť, článok predkladá a demonštruje vylepšenie tejto procedúry.

Key words: decomposition of financial indicator, multiplicative relationship, weights, logarithmic method.

Kľúčové slová: rozklad finančného indikátora, multiplikatívna väzba, váhy, logaritmická metóda.

JEL classification: C65, M21.

1. Úvod

Článok je motivovaný nedávnym príspevkom Úradníčka (2014) k problematike zahrnutia exogénne určených váh do dynamického pyramidálneho rozkladu založenom na multiplikatívnych väzbách. V oblasti finančnej a ekonomickej analýzy podnikovej činnosti sú pomerne obľúbené tzv. pyramidálne rozklady niektorých ukazovateľov na parciálne činitele a v rámci ex post hodnotenia bývalého vývoja sa sleduje, ako tieto parciálne činitele svojou zmenou prispeli k vlastnej zmene cieľového ukazovateľa. Tieto rozklady sa vykonávajú najmä nad ukazovateľmi výkonnosti, akými sú napr. ukazovateľ rentability vlastného imania či ukazovateľ rentability aktív, a potom následne výsledky analýzy poskytujú dozadný pohľad na zdroje podnikovej výkonnosti. Ilustrujúc súvislosti na pyramidálnom rozklade ukazovateľa rentability vlastného imania, citovaný autor uvažuje multiplikatívny rozklad syntetického ukazovateľa na parciálne determinujúce činitele a pri kvantifikácii miery vplyvu zmeny jednotlivých determinujúcich činiteľov na zmenu syntetického ukazovateľa medzi dvoma časovými obdobiami (tzn. základným obdobím a bežným obdobím) používa koncepčne zrozumiteľnú logaritmickú metódu. Do problematiky prispieva dvojakým spôsobom:

- na jednej strane navrhuje metodiku pre zohľadnenie subjektívnej dôleživosti prisudzovanej jednotlivým determinujúcim činiteľom formou váh odrážajúcimi preferencie hodnotiaceho analytika,
- na druhej strane prihliada na aj na možnosť časovej zmeny subjektívnej dôleživosti determinujúcich činiteľov pri podieľaní sa na syntetickom činiteľi, čo následne rezultuje v rôzne váhy v jednotlivých obdobiach.

Váhy citovaný autor v príspevku odvodzuje od tzv. Saatyho metódy, ktorá patrí medzi štandardné spôsoby stanovenia váh (Ficzová et al., 2002, s. 132-134; Balajty a Boďa, 2011). Citovaný autor pritom váhy vyjadruje tak, že – hoci sa môžu rôzniť v jednotlivých obdobiach – musí byť medzi nimi multiplikatívny vzťah a musia sa multiplikovať do jednotky, čomu prispôsobuje aj návrh metodiky. Spôsob, akým sú váhy zakomponované do dynamického pyramidálneho rozkladu, je síce matematicky korektný a názorný, nie je však univerzálne použiteľný a nie je prirodzený. Prax finančnej a ekonomickej analýzy totiž uprednostňuje

váhy v aditívnom vzťahu, ktoré sa sumujú do jednotky (teda sú normalizované). Vychádzajúc z originálnej myšlienky Úradníčka (2014), tento článok ukazuje, ako je možné do pyramidálneho rozkladu na báze multiplikatívnych väzieb prirodzenejšie zahrnúť váhy vyjadrujúce rôznu dôležitosť determinujúcich činiteľov v jednotlivých obdobiach v prípade, že sa použije logaritmická metóda. Medzi týmito váhami môže byť aditívny vzťah (ak sú napr. určené vhodnou expertnou metódou alebo ako konečný výsledok použitia Saatyho metódy) alebo multiplikatívny vzťah (ak sú napr. určené ako medziprodukt použitia Saatyho metódy). S intenciou zabezpečiť porovnateľnosť je aplikácia navrhutej metodiky ilustrovaná na rovnakej prípadovej štúdií ako v príspevku Úradníčka (2014), teda na pyramidálnom rozklade ukazovateľa rentability vlastného imania slovenského podniku PANACEA, s. r. o., za roky 2012 a 2013.

Článok je organizovaný do 5 častí. Po tomto uvedení problematiky (časť 1) nasleduje nevyhnutné a stručné ozrejmienie logaritmickej metódy kvantifikácie vplyvu zmeny determinujúceho činiteľa na zmenu syntetického ukazovateľa a Saatyho metódy pre stanovovanie váh vyjadrujúcich relatívnu dôležitosť (časť 2). Ďalšie časti obsahujú opis metodiky pre zohľadnenie váh do dynamického pyramidálneho rozkladu založeného na použití logaritmickej metódy navrhutej Úradníčkom (2014) a novoformulovanej metodiky (časť 3) a empirickú aplikáciu (časť 4). Záverečná časť sumarizuje prínos článku a má diskusný charakter (časť 5).

2. Logaritmická metóda a Saatyho metóda

Pri ozrejmieni logaritmickej metódy sa bude bez ujmy na obecnosti predpokladať, že syntetický ukazovateľ X je určený súčinom troch parciálnych činiteľov a , b a c , medzi ktorými je multiplikatívna väzba, v dôsledku čoho platí $X = a \cdot b \cdot c$. Hodnoty ukazovateľa a jeho parciálnych činiteľov v základnom období budú vyznačované dolným pravým indexom 0 a v bežnom období dolným pravým indexom 1. V jednotlivých obdobiach bude teda tento rozklad nadobúdať podobu $X_0 = a_0 \cdot b_0 \cdot c_0$ v základnom období a $X_1 = a_1 \cdot b_1 \cdot c_1$ v bežnom období. Použitie tejto metódy si vynucuje, aby v každom období parciálne činitele nadobúdali nenulové hodnoty a nemenili medzi sledovaným základným a bežným obdobím znamienko (vo sfére finančno-ekonomickej praxe je primerané požadovať, aby ich hodnoty boli v oboch obdobiach kladné). Symbolom i bude denotovaný index (koeficient) zmeny ukazovateľa, ktorého označenie bude indikované v dolnom pravom indexe, napr. i_a bude indikovať index zmeny činiteľa a a bude platiť $i_a = a_1/a_0$. Porovnaním podielom uvažovaný rozklad v bežnom období a základnom období sa získava vzťah $i_X = i_a \cdot i_b \cdot i_c$. Logaritmicizáciou oboch strán rovnosti (napr. pri prirodzenom základe) sa prevedie multiplikatívna väzba prevedie na aditívny vzťah a relatívny (tzn. percentuálny) vplyv zmeny jednotlivých parciálnych činiteľov a , b a c na zmenu syntetického ukazovateľa X je potom určený vzťahmi $\Delta X_{Da} = \ln i_a / \ln i_X$, $\Delta X_{Db} = \ln i_b / \ln i_X$ a $\Delta X_{Dc} = \ln i_c / \ln i_X$. Kontribúcia každého z činiteľov k zmene syntetického činiteľa pritom nereflektuje ich prípadne odlišnú dôležitosť. Absolútny vplyv je pre činitele a , b a c stanovený potom formulami $\Delta X_{Da}^{\text{abs}} = \Delta X \cdot \Delta X_{Da}$, $\Delta X_{Db}^{\text{abs}} = \Delta X \cdot \Delta X_{Db}$ a $\Delta X_{Dc}^{\text{abs}} = \Delta X \cdot \Delta X_{Dc}$, v ktorých je $\Delta X = X_1 - X_0$.

Pokiaľ ide o Saatyho metódu stanovovania váh odrážajúcich dôležitosť jednotlivých činiteľov, táto spočíva v ich postupnom párovom porovnávaní a vyjadrovaní ich vzájomného násobného významu. Párové preferencie sa kódujú do nezáporných hodnôt s_{mn} a zapisujú do matice $\mathbf{S} = (s_{mn})_{k \times k}$, pričom k označuje počet činiteľov, ktorých relatívna dôležitosť sa ohodnocuje, a s_{mn} je typický element tejto matice uložený na m -tom riadku a n -tom stĺpci (kde $m, n \in \{1, \dots, k\}$). Hodnota s_{mn} vyjadruje, koľkokrát je m -tý činiteľ (uložený v riadku) významnejší ako n -tý činiteľ (uložený v stĺpci). Dbá sa pritom na tieto dva pravidlá: (1.) Pri porovnávaní platí recipročná symetria; ak je činiteľ m h -krát dôležitejší ako činiteľ n , je

činiteľ n nutne $1/h$ -krát dôležitejší ako činiteľ m , v dôsledku čoho je $s_{mn} = 1/s_{nm}$ pre všetky činitele m a n . (2.) Ak sú porovnávané činitele m a n rovnako dôležité, je $s_{mn} = 1$. Samozrejme, komparácia násobkov významnosti sa obmedzuje výlučne na kladné hodnoty, a teda $s_{mn} > 0$ pre všetky činitele m a n . Od matice \mathbf{S} sa váhy odvodzujú dvojako: buď sa použije tzv. exaktný Saatyho postup, alebo sa využije tzv. aproximatívny Saatyho postup. Potrebu názornosti tu najlepšie zodpovedá práve aproximatívny Saatyho postup, a preto sa pozornosť sústreďuje výlučne naň (síce bez dôsledkov na všeobecnosť). Pre každý ukazovateľ sa určí priemerná násobná dôležitosť za pomoci geometrického priemeru. Pre m -tý ukazovateľ je jeho takto priemerná násobná dôležitosť určená vzťahom

$$v_m = \left(\prod_{n=1}^{n=k} s_{mn} \right)^{1/k}. \quad (1)$$

Je veľmi jednoduché ukázať, že platí $v_1 \cdot \dots \cdot v_k = 1$, čo má význam vtedy, ak je potrebné stanoviť váhy, medzi ktorými je multiplikatívny vzťah a sú normalizované (tzn. ich súčin je jedna). Z hodnôt v_1, \dots, v_k sa ale obvykle určia normalizované aditívne váhy u_1, \dots, u_k (normalizované v tom zmysle, že ich súčet je jedna). Pre m -tý ukazovateľ je aditívna normalizovaná váha potom

$$u_m = v_m \left(\sum_{n=1}^{n=k} v_n \right)^{-1}. \quad (2)$$

3. Zohľadnenie váh v logaritmickej metóde

Úvahy zase budú pre jednoduchosť obmedzené na prípad multiplikatívneho vzťahu s tromi činiteľmi $X = a \cdot b \cdot c$ a váhami sa budú rozumieť vhodne reštrikované nezáporné konštanty. Pre normalizované váhy v základnom období bude akceptovaná notácia v_{a0}, v_{b0}, v_{c0} a v bežnom období v_{a1}, v_{b1}, v_{c1} . Normalizácia znamená, že $v_{a0} \cdot v_{b0} \cdot v_{c0} = 1$ a tiež $v_{a1} \cdot v_{b1} \cdot v_{c1} = 1$. Preto platí $X_0 = (v_{a0}a_0)(v_{b0}b_0)(v_{c0}c_0)$ a $X_1 = (v_{a1}a_1)(v_{b1}b_1)(v_{c1}c_1)$ postupne pre základné obdobie a bežné obdobie. Rovnaké úvahy ako v časti 2 článku aplikoval Úradníček (2014) pre takúto situáciu a získal rozklad $i_X = i_{va} \cdot i_{vb} \cdot i_{vc}$ mapujúci zmenu hodnôt činiteľov a, b a c a aj zmenu ich váh medzi základným a bežným obdobím. Pritom teraz indexy zahrňujú v sebe aj zmenu váhy, napr. $i_{va} = (v_{a1}a_1)/(v_{a0}a_0)$. Aj teraz je relatívny vplyv zmeny jednotlivých parciálnych činiteľov a, b a c na zmenu syntetického ukazovateľa X analogicky určený vzťahmi $\Delta X_{Da} = \ln i_{va} / \ln i_X$, $\Delta X_{Db} = \ln i_{vb} / \ln i_X$ a $\Delta X_{Dc} = \ln i_{vc} / \ln i_X$. V skutočnosti sa takto meria relatívny vplyv zmeny jednotlivých parciálnych činiteľov a ich váh (!) na zmenu príslušného syntetického ukazovateľa.

V tomto článku sa však navrhuje zohľadňovať váhy vo forme exponentov tak, ako je to bežné pri produkčnej analýze založenej na produkčnej funkcii cobb-douglasovského typu alebo pri váženom geometrickom priemere. Najprv sa ukáže, ako zohľadniť váhy a ich zmeny v prípade normalizovaných váh s multiplikatívnym vzťahom v_{a0}, v_{b0}, v_{c0} ($v_{a0} \cdot v_{b0} \cdot v_{c0} = 1$) v základnom období a v_{a1}, v_{b1}, v_{c1} ($v_{a1} \cdot v_{b1} \cdot v_{c1} = 1$) v bežnom období. Takto konštruované váhy informujú, koľkonásobne je príslušný činiteľ v porovnaní s ostatnými uvažovanými činiteľmi priemerne dôležitý v danom období (hoci v skutočnosti pre váhy konštruované podľa vzťahu (1) sa ale priemeruje násobná dôležitosť nielen s ostatnými činiteľmi, ale aj ním samotným).

Nech v základnom období prislúchajú váham v_{a0}, v_{b0}, v_{c0} technické koeficienty w_{a0}, w_{b0}, w_{c0} , ktoré v sebe nesú rovnakú informáciu o relatívnej dôležitosti ako váhy samotné. Tieto koeficienty musia byť spoločne úmerné pôvodným váham, od ktorých sa odvodzujú, t. j. musí platiť pre nejakú konštantu úmernosti z_0 ($z_0 > 0$), že $w_{a0} = z_0 \cdot v_{a0}$, $w_{b0} = z_0 \cdot v_{b0}$ a $w_{c0} = z_0 \cdot v_{c0}$. Kládne sa požiadavka rovnosti $X_0 = a_0 \cdot b_0 \cdot c_0 = a_0^{w_{a0}} \cdot b_0^{w_{b0}} \cdot c_0^{w_{c0}}$. Dosadiac vzťahy úmernosti do tejto rovnosti, získava sa $X_0 = a_0^{z_0 v_{a0}} \cdot b_0^{z_0 v_{b0}} \cdot c_0^{z_0 v_{c0}}$. Z tejto rovnosti potom ihneď vyplýva hodnota konštanty z_0 , a síce $z_0 = \ln X_0 / (v_{a0} \ln a_0 + v_{b0} \ln b_0 + v_{c0} \ln c_0)$. Analogicky nech

v bežnom období existujú technické koeficienty w_{a1}, w_{b1}, w_{c1} , že platí $X_1 = a_1 \cdot b_1 \cdot c_1 = a_1^{w_{a1}} \cdot b_1^{w_{b1}} \cdot c_1^{w_{c1}}$ a pre nejakú konštantu úmernosti z_1 ($z_1 > 0$) je $w_{a1} = z_1 \cdot v_{a1}$, $w_{b1} = z_1 \cdot v_{b1}$ a $w_{c1} = z_1 \cdot v_{c1}$. Potom ďalej $z_1 = \ln X_1 / (v_{a1} \ln a_1 + v_{b1} \ln b_1 + v_{c1} \ln c_1)$ plynie z replikácie úvahy zo základného obdobia. Z pomerovania vzťahov pre X_1 a X_0 sa získava podielový multiplikatívny vzťah $i_X = i_{a,w} \cdot i_{b,w} \cdot i_{c,w}$, v ktorom je postupne $i_{a,w} = a_1^{w_{a1}} / a_0^{w_{a0}}$, $i_{b,w} = b_1^{w_{b1}} / b_0^{w_{b0}}$ a $i_{c,w} = c_1^{w_{c1}} / c_0^{w_{c0}}$. Takto konštruované indexy $i_{a,w}$, $i_{b,w}$, $i_{c,w}$ odrážajú jednak samotnú zmenu činiteľa a jednak zmenu jeho relatívnej dôležitosti. Aplikujúc rovnakú ideovú muštru ako v ostatných dvoch prípadoch, je relatívna zmena syntetického ukazovateľa X atribuovateľná kombinovanej zmene činiteľa a jeho relatívnej dôležitosti určená pre činiteľ a vzťahom $\Delta X_{Da} = \ln i_{a,w} / \ln i_X = (w_{a1} \ln a_1 - w_{b1} \ln b_1) / (\ln X_1 - \ln X_0)$ a analogicky pre zvyšné dva činitele b a c .

Ak by ale boli v základnom období k dispozícii normalizované váhy s aditívnym vzťahom u_{a0}, u_{b0}, u_{c0} (zatiaľ čo $u_{a0} + u_{b0} + u_{c0} = 1$) a v bežnom období u_{a1}, u_{b1}, u_{c1} (splňujúc $u_{a1} + u_{b1} + u_{c1} = 1$), mal by postup spočívať v ich konverzii na normalizované váhy s multiplikatívnym vzťahom. Stačí však iba zohľadniť, že podľa vzťahu (2) sú normalizované váhy s aditívnym vzťahom spoločne úmerné (tzn. s identickou konštantou úmernosti) normalizovaným váham s multiplikatívnym vzťahom. V dôsledku toho možno túto procedúru uplatniť bezo zmeny aj v takomto prípade, pričom v priebehu jej aplikácie sa váhy i tak prenormalizujú, aby bola zaistená zhoda medzi hodnotou syntetického činiteľa a súčinom exponenciálne vážených parciálnych činiteľov.

Bez ohľadu na modifikáciu logaritmickej metódy je absolútny vplyv zmeny parciálnych činiteľov a , b a c na zmenu $\Delta X = X_1 - X_0$ syntetického ukazovateľa X určený vzťahmi $\Delta X_{Da}^{\text{abs}} = \Delta X \cdot \Delta X_{Da}$, $\Delta X_{Db}^{\text{abs}} = \Delta X \cdot \Delta X_{Db}$ a $\Delta X_{Dc}^{\text{abs}} = \Delta X \cdot \Delta X_{Dc}$.

Tieto úvahy sú ukázané na empirickom príklade v nasledujúcej časti.

4. Demonštrácia navrhutej procedúry pre zohľadnenie váh v logaritmickej metóde

Pre porovnateľnosť je navrhnuté rozšírenie logaritmickej metódy rešpektujúce diferencovanú dôležitosť jednotlivých parciálnych činiteľov na generovaní syntetického činiteľa formou váh a ich potenciálnej zmeny porovnané s tradičnou logaritmickou metódou a modifikovanou metódou navrhnutou Úradníčkom (2014). Je pritom zvolený tá istá prípadová štúdia, ktorú zvolil pôvodný autor: ide o pyramidálny rozklad ukazovateľa rentability vlastného imania a o jeho aplikáciu na žilinský podnik PANACEA, s. r. o., pre účtovné obdobia rokov 2012 a 2013. Interpretáciu rozkladu možno nájsť napr. u Ficzovej et al. (2002, s. 92, 17-32) alebo u Úradníčka (2014, s. 184-186).

Pre určenie rozkladu boli z účtovnej závierky podniku PANACEA, s. r. o., za roky 2012 a 2013 prevzaté položky výsledku hospodárenia za účtovné obdobie po zdanení (EAT: earnings after taxes), tržieb (S: sales), aktív (A: assets), vlastného imania (E: equity), následne boli spočítané ukazovatele rentability tržieb (ROS: return on assets), obrat aktív v tržbách (AT: asset turnover) a finančná páka (EM: equity multiplier), rešpektujúc pritom definíčné vzťahy $ROS = EAT/S$, $AT = S/A$, $EM = A/E$. Ukazovateľ rentability vlastného imania (ROE: return on equity) je definovaný vzťahom $ROE = EAT/E$ a splňuje rozklad na multiplikatívnu väzbu v podobe $ROE = ROS \times AT \times EM$. Vstupné podklady a hodnoty vypočítaných ukazovateľov sú uvedené v tabuľke 1, ktoré obsahuje okrem iného váhy použité v článku od Úradníčka (2014). Hoci autor navrhuje metodiku pre normalizované váhy s multiplikatívnou väzbou, ním uvedené váhy sú normalizované vzhľadom k aditívnej väzbe a takéto váhy používa aj priamo pri výpočte, ako možno jednoducho overiť. Váhy uvedené v tabuľke ako „aditívne váhy“ sú prevzaté z tabuľky „5“ od Úradníčka (2014, s. 187) a zodpovedajú im normalizované váhy s multiplikatívnym vzťahom prezentované v tabuľke 1 pod označením „multiplikatívne váhy“.

Tab. 3: Vstupné podklady, hodnoty ukazovateľov rozkladu ukazovateľa ROE a váhy¹

Vstup (€)	Účtovné obdobie		Ukazovateľ	Účtovné obdobie		Aditívna váha	Účtovné obdobie		Multiplikatívna váha	Účtovné obdobie	
	2012	2013		2012	2013		2012	2013		2012	2013
EAT	376 061	114 652	ROS	0.1427	0.0506	ROS	0.566	0.582	ROS	2.083	2.268
S	2 634 528	2 268 018	AT	0.7886	0.8588	AT	0.325	0.330	AT	1.196	1.286
A	3 340 587	2 640 863	EM	1.5408	1.4199	EM	0.109	0.088	EM	0.401	0.343
E	2 168 045	1 859 862	ROE	0.1735	0.0616	Spolu	1.000	1.000	Súčin	1.000	1.000

ΔROE -0.1118

Ukazovateľ rentability vlastného imania sa v roku 2013 klesol na hodnotu 6.16 % v porovnaní s rokom 2012, keď dosahoval hodnotu 17.35 %. Motiváciou je rozkládať medziročný pokles ukazovateľa rentability vlastného imania (ROE) o 11.18 percentuálnych bodov medzi ukazovateľa rentability tržieb (ROS), obrat aktív v tržbách (AT) a finančná páka (EM), ktorých zmena túto celkovú zmenu vyvolala. Na tento účel je aplikovaná (1.) tradičná logaritmická metóda nezohľadňujúca diferencovanú dôležitosť jednotlivých parciálnych činiteľov, (2.) metóda navrhnutá Úradníčkom (2014), v ktorej sa nie celkom korektné použijú priamym dosadením váhy s aditívnou väzbou v tabuľke 1, (3.) metóda navrhnutá Úradníčkom (2014), v ktorej sa aditívne váhy v tabuľke 1 konvertujú zodpovedajúcim spôsobom na váhy s multiplikatívnou väzbou, (4.) novoformulovaná metodika použiteľná v prípade váh aj s aditívnym vzťahom, aj s multiplikatívnym vzťahom. Výsledky aplikácie týchto štyroch metód sú uvedené v tabuľke 2, pričom ľavá časť tabuľky sa zhoduje s výsledkami reportovanými v tabuľke 2 u Úradníčka (2014, s. 188).

Tab. 4: Výsledky aplikácie jednotlivých metód

Tradičná logaritmická metóda bez zahrnutia váh			Logaritmická metóda à la Úradníček (2014) s aditívnou väzbou váh			Logaritmická metóda à la Úradníček (2014) s multiplikatívnou väzbou váh			Navrhnutá logaritmická metóda zohľadňujúca váhy		
Vplyv	Relatívny	Absolútny	Vplyv	Relatívny	Absolútny	Vplyv	Relatívny	Absolútny	Vplyv	Relatívny	Absolútny
ROS	1.003	-0.112	ROS	0.838	-0.094	ROS	0.921	-0.103	ROS	1.015	-0.114
AT	-0.082	0.009	AT	-0.083	0.009	AT	-0.152	0.017	AT	-0.038	0.004
EM	0.079	-0.009	EM	0.245	-0.027	EM	0.231	-0.026	EM	0.023	-0.003
Spolu	1.000	-0.112	Spolu	1.000	-0.112	Spolu	1.000	-0.112	Spolu	1.000	-0.112

Možno si ešte v tabuľke 1 všimnúť, že nie je veľký rozdiel medzi váhami (či už aditívnymi, alebo multiplikatívnymi), a teda pri aplikácii metód zohľadňujúcich aj zmenu váh medzi hodnotiacimi obdobiami do výslednej dekompozície ukazovateľa rentability vlastného imania vstupuje predovšetkým relatívna dôležitosť troch parciálnych činiteľov, a nie jej zmena.

Keď sa porovnávajú výsledky dekompozície podľa jednotlivých metód, nieto z praktického hľadiska veľkého rozdielu v identifikovanom vplyve parciálnych činiteľov rentability tržieb (ROS), obrat aktív v tržbách (AT) a finančná páka (EM) na zmenu syntetického ukazovateľa rentability vlastného imania ROE. Podrobná interpretácia výsledkov je u Úradníčka (2014, s. 187-188), ale jednoznačným záverom je, že smerodajným pre negatívny vývoj (pokles) ukazovateľa rentability vlastného imania v rokoch 2012 a 2013 podniku PANACEA, s. r. o., bolo zhoršenie ukazovateľa rentability vlastného imania (ROS), a teda zníženie schopnosti uplatňovať v tržbách vyššie (čisté) obchodné marže. Inými slovami, oslabenie kvalitatívnej stránky dosahovania tržieb podniku PANACEA, s. r. o., z roku 2012 na rok 2013 rezultovalo do zníženia rentability vlastného imania a celkovej výkonnosti podniku meranej na báze tohto klasického ukazovateľa výkonnosti.

¹ Podklady pre výpočet (EAT, S, A, E) sú prevzaté z účtovnej závierky podniku PANACEA, s. r. o., dostupnej v online Registri účtovných závierok Ministerstva financií Slovenskej republiky (<http://www.registeruz.sk>), aditívne váhy sú prevzaté z článku Úradníčka (2014, s. 87).

5. Záver

V článku bol navrhnutý metodický postup, ako do logaritmickej metódy kvantifikácie vplyvu zmien parciálnych činiteľov na zmenu syntetického činiteľa v pyramidálnom rozklade s multiplikatívnymi väzbami zahrnúť váhy vyjadrujúce relatívnu dôležitosť jednotlivých parciálnych činiteľov. Bez ohľadu na to, či sú vstupné váhy normalizované vzhľadom k aditívnej väzbe, alebo normalizované vzhľadom k multiplikatívnej väzbe, zohľadňujú sa v pyramidálnom rozklade v exponente jednotlivých parciálnych činiteľov. Týmto použitie logaritmickej metódy približujú postupom obvyklým v indexnej analýze produkčného procesu, v ktorej sa nezriedka priemery definujú geometrickým priemerom a jednotlivé činitele sa vážia pomocou normalizovaných váh v aditívnej väzbe v ich exponente. Jediným technickým prvkom celej metódy predstavujúcu potenciálne obmedzenie alebo prekážku pri jej implementácii je potreba renormalizovať váhy tak, aby bola v analyzovanom bežnom aj základnom období zaručená zhoda medzi hodnotou syntetického ukazovateľa a (cez exponenty) váženom súčine parciálnych činiteľov. V článku je odporúčané váhy stanovovať za pomoci aproximatívnej Saatyho metódy, ktorá umožňuje v jednotlivých krokoch určiť logicky konzistentné váhy plne reflektujúcu subjektívnu dôležitosť parciálnych činiteľov aj v multiplikatívne normalizovanej podobe, aj v aditívne normalizovanom tvare. Impulz pre zohľadňovanie relatívnej dôležitosti jednotlivých parciálnych činiteľov v pyramidálnom rozklade (vo všeobecnosti, nielen v prípade multiplikatívnej väzby) možno nájsť v nedávnom príspevku u Úradníčka (2014) a bezpochyby si zasluhuje ďalšie skúmanie. Metodika proponovaná v článku bola pre porovnanie aplikovaná na tej istej prípadovej štúdiu ako u citovaného autora, tzn. na pyramidálnom rozklade ukazovateľa rentability vlastného imania slovenského podniku PANACEA, s. r. o., pre účtovné obdobia rokov 2012 a 2013. Koncepcijnej ucelenosti navrhutej metodiky a jej logickej správnosti nasvedčujú aj výsledky tejto prípadovej štúdie. Napriek tomuto povzbudeniu námetmi ďalšej práce zostávajú výzvy

- na stochastické zahrnutie neistoty do pyramidálneho rozkladu a kvantifikácie vplyvu zmien determinujúcich činiteľov na zmenu syntetického ukazovateľa,
- na skúmanie podmienok, za ktorých dochádza k relevantným odchylnostiam medzi výsledkami jednotlivých metód,
- na oddelenie efektu zmeny relatívnej dôležitosti a váh medzi obdobia od vplyvu zmeny prisudzovanej jednotlivým činiteľom na zmenu syntetického činiteľa.

Tieto úlohy zostáva adresovať v ďalšej vedeckej práci autora.

Literatúra

BALAJTY, M., BOĎA, M. 2011. Aplikácia metód viackriteriálneho hodnotenia pri skúmaní vplyvu ekonomicko-právnych aspektov na zakladanie nadnárodných obchodných spoločností v Európskej únii. In *Forum Statisticum Slovacum*. 2011, č. 2, roč. 7. S. 14-18.

FICZOVÁ, I., SEDLÁČEK, J., ÚRADNÍČEK, V. 2002. *Finančno-ekonomická analýza podniku: praktikum. Časť I*. Banská Bystrica: Občianske združenie Financ, 2002. 208 s. ISBN 80-968702-1-1.

ÚRADNÍČEK, V. 2014. Dynamický pyramidálny rozklad finančného ukazovateľa. In *Forum Statisticum Slovacum*, roč. X, č. 4, 2014, s. 184 – 189.

Adresa autora:

Martin Boďa, Mgr. Ing., PhD.
Univerzita Mateja Bela v Banskej Bystrici
Ekonomická fakulta
Tajovského 10, 975 90 Banská Bystrica
martin.boda@umb.sk

Civilizačné choroby a stredná dĺžka života

Civilization diseases and life expectancy

Kornélia Cséfalvaiová, Jana Langhamrová, Jitka Langhamrová

Abstract: Added years of life and years spent in good health are undoubtedly due to enhanced level of health care and advances in medicine. However, civilization diseases may be a problem for advanced societies of the 21st century. Civilization diseases are highly influenced by lifestyle. The aim of presented paper is to characterize the relationship between life expectancy, healthy life expectancy and selected civilization diseases. The article mainly follows the relationship between overweight, obesity and life expectancy at the age 65 in selected European countries. It also examines the time evolution of selected characteristics.

Abstrakt: Za pridané roky života a roky strávené v dobrom zdraví vďačíme nepochybne zvýšenej úrovni zdravotníckej starostlivosti a pokrokom v oblasti medicíny. Pre vyspelé spoločnosti 21. storočia však výzvu predstavujú civilizačné choroby, ktoré sú vo vysokej miere ovplyvnené životosprávou. Cieľom predloženého príspevku je charakterizovať závislosť medzi strednou dĺžkou života, zdravou dĺžkou života a vybranými civilizačnými chorobami. Príspevok sleduje predovšetkým vzťah medzi nadváhou, obezitou a strednou dĺžkou života vo veku 65 rokov vo vybraných európskych krajinách. Taktiež skúma vývoj vybraných charakteristík v čase.

Key words: Health, Life expectancy, Healthy life expectancy, Obesity, Civilization diseases.

Kľúčové slová: zdravie, stredná dĺžka života, zdravá dĺžka života, obezita, civilizačné choroby.

JEL classification: I130, I140.

1. Úvod

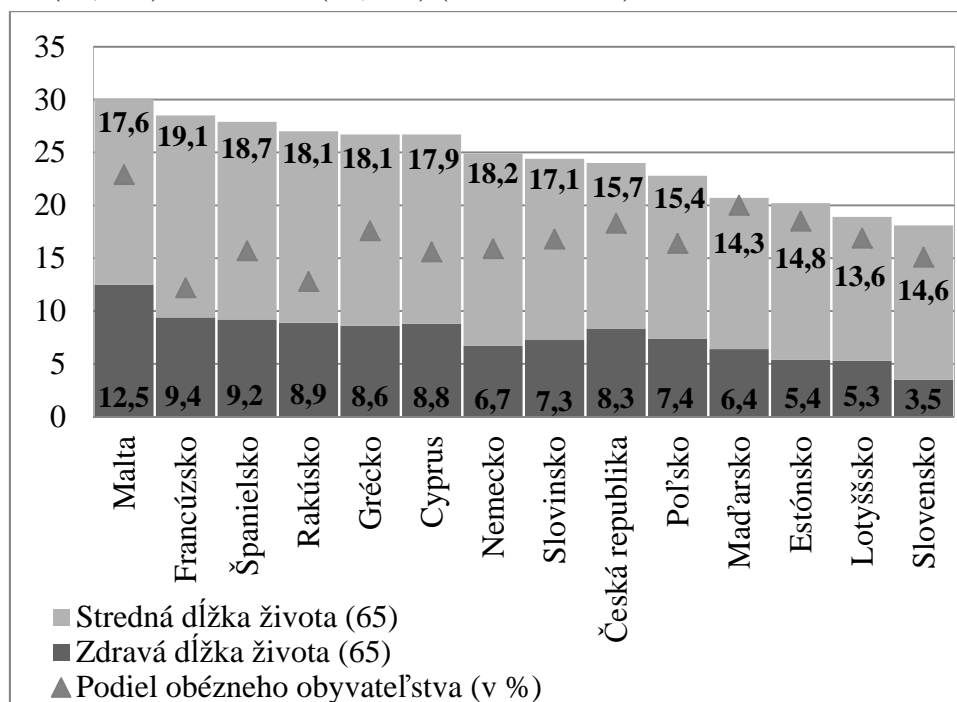
Zdravie je stav úplného telesného, psychického a sociálneho blaha. Na základe údajov zverejnených Svetovou zdravotníckou organizáciou (WHO) 15–30 % populácie sveta trpí civilizačnými chorobami. Výskyt civilizačných chorôb je najviac prítomný vo vyspelých krajinách, kde výskyt týchto chorôb predstavuje až 30–40 %. Počet ochorení na civilizačné choroby každoročne stúpa a preto sa v súčasnosti stávajú prvoradým záujmom národnej stratégie ochrany a podpory zdravia a prevencie chorôb. Ďalej v tomto príspevku bude hlavná pozornosť venovaná problematike nadváhy a obezity, ktoré meriame pomocou indexu BMI (index telesnej hmotnosti), v súvislosti so strednou dĺžkou života a zdravou dĺžkou života. V Českej republike je podiel mužov trpiacich obezitou 18,4 % a podiel žien tvorí 18,3 %. Na Slovensku je podiel obéznych mužov 14,5 % a podiel obéznych žien tvorí 15,7 % (Eurostat, 2008). Z výsledkov niekoľkých štúdií vyplýva, že obezita skracuje strednú dĺžku života. Nasledujúca časť príspevku je venovaná vzťahu medzi týmito charakteristikami.

2. Stredná dĺžka života, zdravá dĺžka života, nadváha a obezita

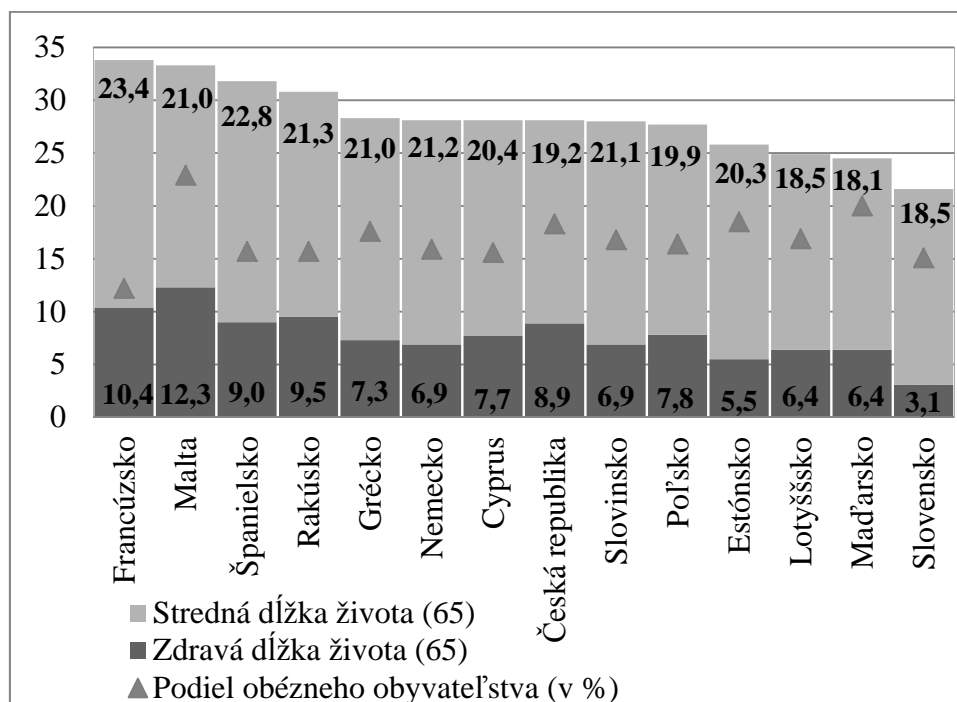
Stredná dĺžka života (očakávaná dĺžka života) vyjadruje priemerný počet rokov, ktorého by sa dožil jedinec vo veku x rokov pri zachovaní súčasnej úmrtnosti (Langhamrová, 2013). Index telesnej hmotnosti (BMI) patrí medzi najbežnejšie metódy merania podvýživy, nadváhy a obezity a počíta sa ako podiel telesnej hmotnosti v kilogramoch delená druhou mocninou telesnej výšky v metroch. Za nadváhu považujeme stav, kedy je hodnota indexu BMI medzi 25–30 a o obezite hovoríme v prípade, že je hodnota indexu BMI ≥ 30 (WHO, 2014).

Stredná dĺžka života a zdravá dĺžka života vo veku 65 rokov je obecné vyššia v krajinách západnej a južnej Európy v porovnaní s východnou a strednou Európou.

Krajiny ako Estónsko, Lotyšsko a Slovensko patria medzi krajiny s nízkou nádejou dožitia. Česká republika sa nachádza na pomyselných hranici medzi krajinami západnej a východnej Európy. Stredná dĺžka života 65-ročných žien je vo všetkých sledovaných krajinách vyššie pre ženy ako pre mužov. Podiel obézneho obyvateľstva je najvyšší na Malte (22,9 %), v Maďarsku (20,0 %) a Estónsku (18,5 %) (viď obr. 1 a 2).

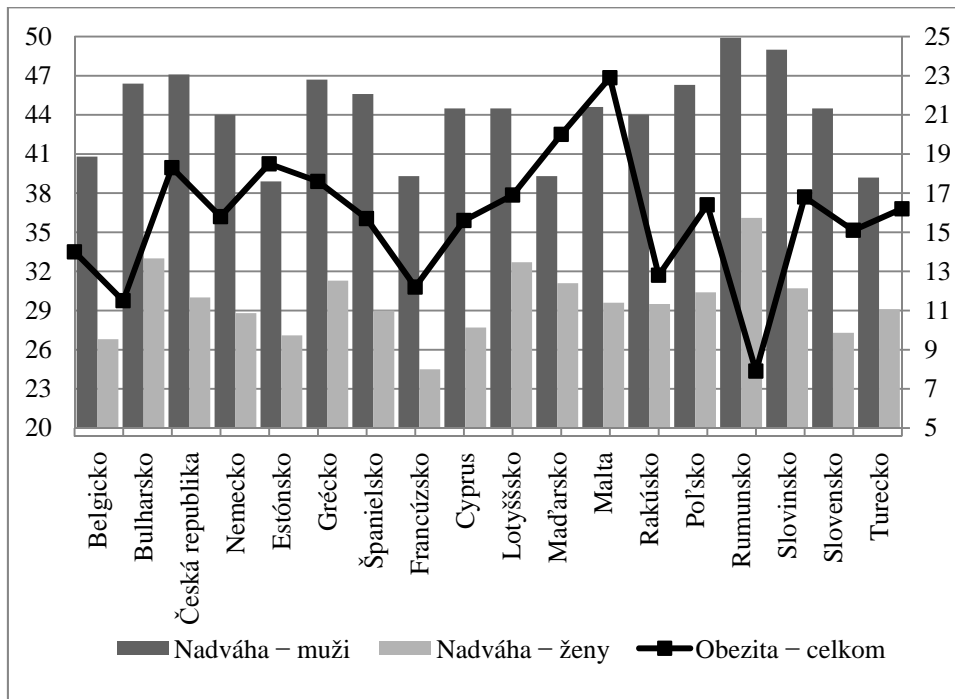


Obr. 5: Zdravá dĺžka života mužov vo veku 65 rokov, stredná dĺžka života mužov vo veku 65 rokov a podiel obézneho obyvateľstva (v %) v roku 2012



Obr. 2: Zdravá dĺžka života žien vo veku 65 rokov, stredná dĺžka života žien vo veku 65 rokov a podiel obézneho obyvateľstva (v %) v roku 2012

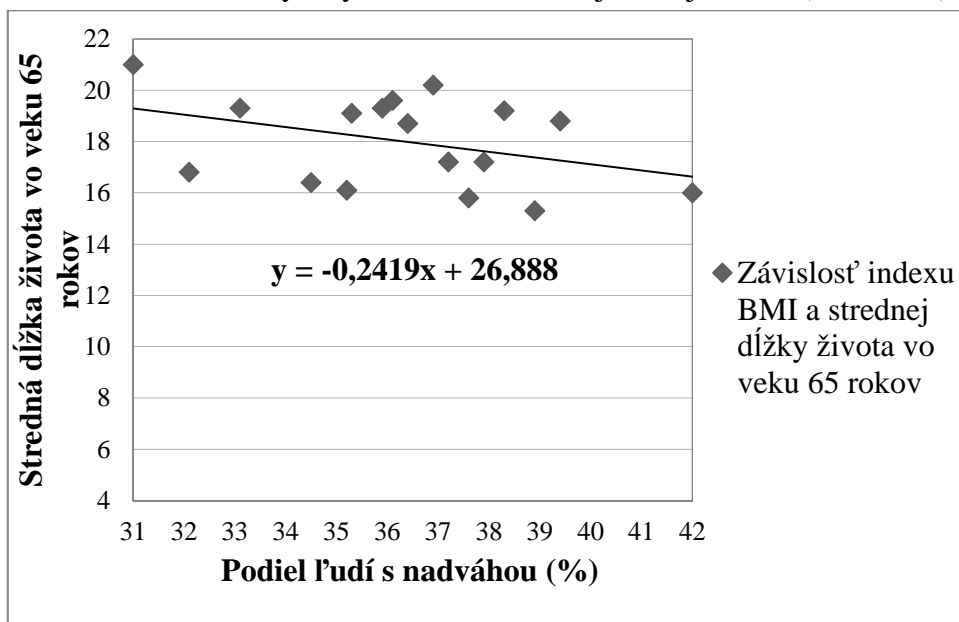
Zastúpenie mužov a žien s nadváhou je najvyššie v Rumunsku (muži = 49,9 %; ženy = 36,1 %), Slovinsku (muži = 49,0 %; ženy = 30,7 %) a Českej republike (muži = 47,1 % a ženy = 30,0 %). Podiel obéznych ľudí (BMI \geq 30) je percentuálne najvyšší v Maďarsku (20,0 %), na Malte (22,9 %) a v Estónsku (18,5 %).



Obr. 3: Nadváha mužov a žien a obezita, vyjadrené indexom BMI (v %) v roku 2008

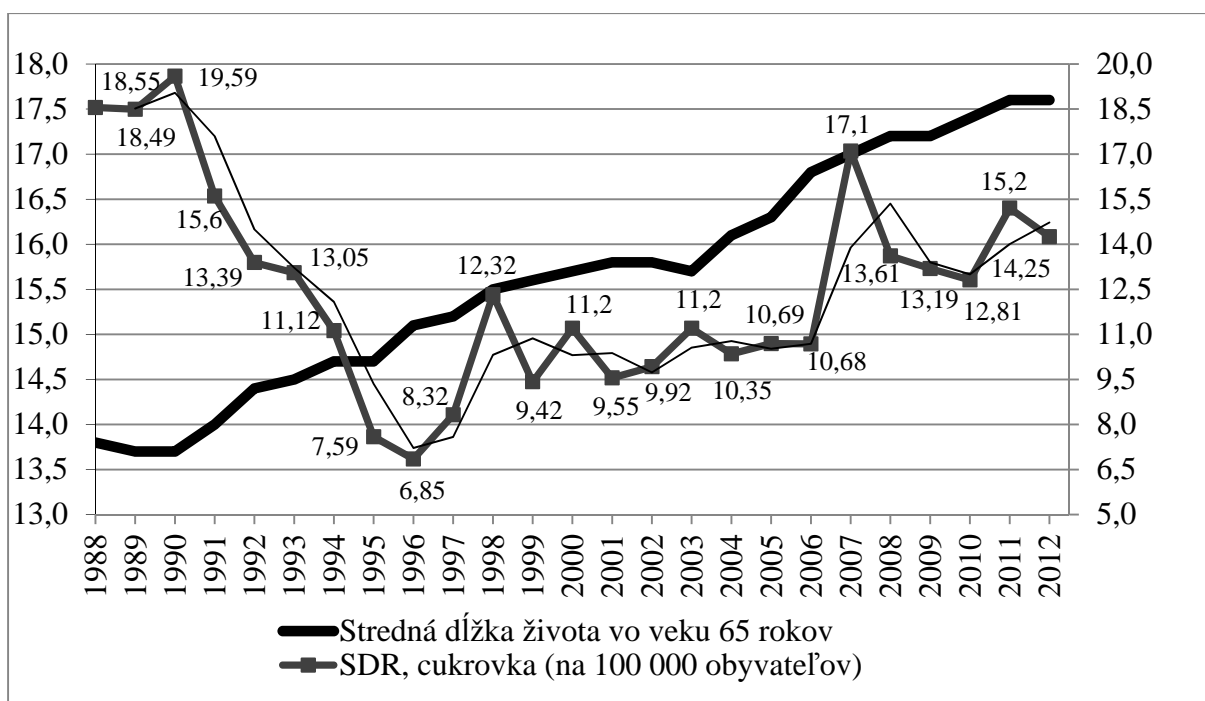
3. Závislosť medzi strednou dĺžkou života vo veku 65 rokov a vybranými charakteristikami

Pri skúmaní závislosti medzi strednou dĺžkou života vo veku 65 rokov a nadváhou (BMI 25–30) bola preukázaná nepriama závislosť ($r = -0,38$), teda predpokladáme, že nadváha má vplyv na strednú dĺžku života. Vysoký index BMI skracaie nádej dožitia (viď obr. 4).



Obr. 4: Závislosť indexu BMI a strednej dĺžky života vo veku 65 rokov vo vybraných európskych krajinách, 2008

U ľudí obéznych sledujeme vyššiu úmrtnosť, ktorá je ovplyvnená napríklad prítomnosťou cukrovky. Najmä pri vyšších stupňoch obezity sa efekt na dĺžku života stáva štatisticky významným (Hainer, 2011). Civilizačné choroby predstavujú častú príčinu úmrtia, obrázok 5 predstavuje štandardizovanú mieru úmrtnosti na cukrovku na 100 000 obyvateľov v Českej republike z časového pohľadu v porovnaní s vývojom nádeje dožitia vo veku 65 rokov. Obecne sa uvádza, že nadváha má negatívny vplyv na mortalitu a na strednú dĺžku života. Z obrázku 5 ďalej vidíme, že vývoj strednej dĺžky života vo veku 65 rokov má hladký priebeh, kým štandardizovaná miera úmrtnosti je charakterizovaná značnými výkyvmi medzi sledovanými obdobiami.



Obr. 5: Vývoj strednej dĺžky života vo veku 65 rokov a štandardizovanej miery úmrtnosti na cukrovku v Českej republike v období 1988–2012

4. Záver

Napriek nízkej dostupnosti dát civilizačné choroby a najmä nadváha a obezita budú v budúcnosti predstavovať dôležitý objekt záujmu štúdií a výskumov, pretože vplyvom súčasnej životosprávy sa podiel osôb trpiacich nadváhou, obezitou, cukrovkou, apod. zvyšuje. V príspevku sme skúmali závislosť medzi strednou dĺžkou života vo veku 65 rokov a nadváhou, pri ktorej sme preukázali nepriamu lineárnu závislosť. Kladením väčšieho dôrazu na telesnú hmotnosť by sme predĺžili roky prežité v dobrom zdraví a taktiež pozitívne ovplyvnili dlhovekosť.

PodĎakovanie

Článok bol pripravený v spolupráci s Internou grantovou agentúrou Vysoké školy ekonomickej v Prahe, číslo 68/2014 pod názvom „Ekonomické a zdravotní súvislosti stárnutí populace“.

Literatúra

ČSÚ. Český statistický úřad. Dostupné na: http://www.czso.cz/csu/2012edicniplan.nsf/publ/1417-12-n_2012.

EUROSTAT. European Statistical Office Database. Dostupné na: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database

HAINER, J. *Základy klinické obezitologie*. Praha: Grada Publishing a.s., 2011. ISBN 978-80-247-3252-7.

LANGHAMROVÁ, J. *Základy demografie (materiály ke cvičením)*. Praha: Oeconomica, 2013. ISBN 978-80-245-1956-2.

NAGAI, M., KURIYAMA, S., KAKIZAKI M, *et al. Impact of obesity, overweight and underweight on life expectancy and lifetime medical expenditures: the Ohsaki Cohort Study. BMJ Open. 2012. Dostupné na: <http://bmjopen.bmj.com/content/2/3/e000940>*

WHO. World Health Organization Database. Dostupné na: <http://data.euro.who.int/hfad/>

Adresa autorov:

Ing. Kornélia Cséfalvaiová
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67, Praha 3
xcsek00@vse.cz

Ing. Jana Langhamrová
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67, Praha 3
jana.langhamrova@vse.cz

doc. Ing. Jitka Langhamrová, CSc.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67, Praha 3
jitka.langhamrova@vse.cz

Zvyšování důchodového věku – jedna z možných cest k finanční udržitelnosti důchodových systémů v ČR, SR i Evropě

Rising of Retirement Age – A Possible Way to Sustainability of Pension Systems in Czech Republic, Slovakia and Europe

Tomáš Fiala, Zdeněk Pavlík

Abstract: The article concerns the influence of increasing the retirement age on the period of receiving old-age pensions and the value of the old-age dependency ratio. The estimate of the period of receiving the pension is based on the life expectancy from generation life tables being computed according to the latest population projection for the period 2013–2080 (Eurostat, 2014), the computation of the old-age dependency ratio follows from the same projection. For each country the increasing retirement age following in stabilization of the average period of receiving pension at the level of 20 years or following in stabilization the old-age dependency ratio at the level of 33 %. The development of the period of receiving pension as well as the old-age dependency ratio for constant retirement age 65 years is presented,

Abstrakt: Článek se zabývá vlivem zvyšování důchodového věku na dobu pobírání starobních důchodů a hodnotu indexu závislosti seniorů. Odhad doby pobírání důchodu je založen na střední délce života z generačních úmrtnostních tabulek vypočtených na základě poslední populační projekce evropských zemí na období 2013–2080 (Eurostat, 2014), výpočet indexu závislosti seniorů vychází z téže projekce. Pro každou zemi byly provedeny výpočty zvyšování důchodového věku vedoucího ke stabilizaci průměrné doby pobírání důchodu na úrovni 20 let nebo ke stabilizaci hodnoty indexu závislosti na úrovni 33 %. Uveden je rovněž vývoj doby pobírání důchodu a indexu závislosti při konstantním důchodovém věku 65 let.

Key words: old-age pension, retirement age, length of period of receiving a pension, old-age dependency ratio, population projection, Czech Republic, Slovakia, European countries.

Klíčová slova: starobní důchod, důchodový věk, doba pobírání důchodu, index závislosti seniorů, populační projekce, Česká republika, Slovensko, evropské země.

JEL classification: H55, J11, J26.

1. Úvod

Jedním z velmi často diskutovaných důsledků stárnutí populace je otázka růstu finančního zatížení důchodových systémů. Protože se v nejbližších desetiletích předpokládá další snižování úmrtnosti i zlepšování zdraví populace, objevují se stále častěji názory, aby úměrně předpokládanému zvyšování délky života rostl rovněž důchodový věk s cílem omezit dobu pobírání starobního důchodu. V jednom z posledních doporučení (Doporučení rady, 2014) doporučuje Evropská komise jasnější provázání tohoto věku se změnami ve střední délce života. Navrhuje se například, že důchodový věk by měl být nastaven tak, aby se průměrná doba pobírání starobního důchodu pohybovala kolem 20 let. Jiným možným kritériem by bylo zvolit důchodový věk tak, aby vedl k zachování přibližně stabilní hodnoty indexu závislosti (poměru počtu osob v důchodovém věku a počtu osob v produktivním věku).

Článek ukazuje, jaká by byla průměrná doba pobírání důchodu a hodnoty indexu závislosti v ČR při důchodovém věku podle současné právní úpravy (resp. při jeho navrhovaném zastropování na úrovni 65 let) a jak vysoký by měl být důchodový věk, aby průměrná doba pobírání důchodu činila přesně 20 let, resp. aby se hodnota indexu závislosti pohybovala kolem 33 %. Jsou zde rovněž uvedeny některé analogické údaje pro ostatní země Evropy ve

vybraných letech. Všechny výpočty vycházejí z hlavní varianty poslední demografické projekce Eurostatu z roku 2013 (Eurostat, 2014a).

2. Metodologické poznámky

Střední délkou života generace daného ročníku narození budeme rozumět délku života vypočtenou na základě generačních úmrtnostních tabulek. Výchozími daty pro výpočet těchto tabulek byly průřezové míry úmrtnosti za období 2013–2080 z hlavní varianty demografické projekce Eurostatu (Eurostat, 2014b). Po roce 2080 se předpokládalo, že specifické míry budou nadále geometricky klesat s koeficientem poklesu rovným geometrickému průměru za posledních 10 let období projekce

$$m_{t+1,x} = \sqrt[10]{\frac{m_{2080,x}}{m_{2070,x}}} \cdot m_{t,x} \quad (1)$$

Generační specifické míry úmrtnosti pro ročník narození g pak byly vypočteny podle vzorce:

$$m_x^{(g)} = \sqrt{m_{r+x,x} \cdot m_{r+x+1,x}} \quad (2)$$

neboť osoby narozené v roce g prožijí část dokončeného věku x v roce $g+x$ a druhou část v roce $g+x+1$. Na základě těchto měr úmrtností pak byly vypočteny generační úmrtnostní tabulky obvyklým postupem.

Důchodem budeme rozumět vždy starobní důchod pobíraný od okamžiku dosažení důchodového věku. Důchodový věk předpokládáme stejný pro muže i ženy (pokud nebude výslovně uvedeno jinak). Indexem závislosti budeme v tomto článku vždy rozumět poměr počtu osob v důchodovém věku a počtu osob od 20 let do důchodového věku.

Za průměrnou dobu pobírání důchodu osob dané generace pro dané pohlaví byla považována generační střední délka života v okamžiku dosažení důchodového věku. Pro neceločíselný důchodový věk byla provedena lineární interpolace mezi nejbližšími celočíselnými hodnotami. Průměrná doba pobírání důchodu pak byla vypočtena jako aritmetický průměr doby pobírání důchodu mužů a žen. Rovněž pro výpočet důchodového věku, při němž by doba pobírání důchodu činila přesně 20 let či věku, při němž by hodnota indexu závislosti byla přesně 33 %, byla použita lineární interpolace mezi nejbližšími známými hodnotami.

3. Důchodový věk v ČR a SR

Podle aktuální právní úpravy (Zákon 155/1995 Sb.) se má důchodový věk v ČR po celé toto století plynule zvyšovat, v roce 2100 by měl dosáhnout 75 let. Pokud by však mělo být splněno doporučení Evropské komise, musel by být po celé období důchodový věk v ČR ještě o něco vyšší, zejména důchodový věk žen (Obr. 2). Odborná komise pro důchodovou reformu v ČR proto zatím nenavrhuje změnu důchodového věku, pouze se uvažuje o pravidelných revizích na základě aktualizovaných prognóz demografického vývoje.

I při současném růstu důchodového věku by došlo v 50. letech, kdy budou do důchodu odcházet početně nejsilnější generace narozených v 70. letech minulého století, ke zvýšení indexu závislosti seniorů. Důchodový věk zajišťující stabilizaci tohoto indexu na úrovni 33 % by proto musel dosáhnout 73 let již kolem roku 2065. Doba pobírání důchodu by přitom klesla zhruba na 18 let.

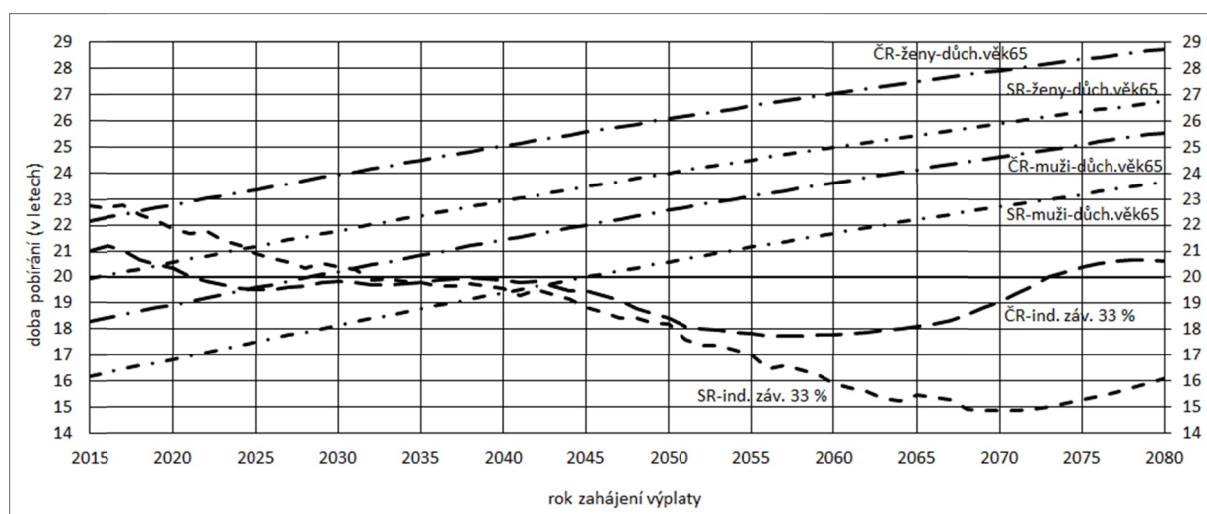
Naopak při „zastropování“ důchodového věku v ČR na hranici 65 let po roce 2030 by doba pobírání důchodu postupně rostla, v roce 2080 by dosahovala téměř 26 let pro muže a 29 let

pro ženy (Obr. 1). Index závislosti seniorů by v takovém případě dosahoval ve 2. polovině tohoto století hodnot vyšších než 50 % (Obr. 3).

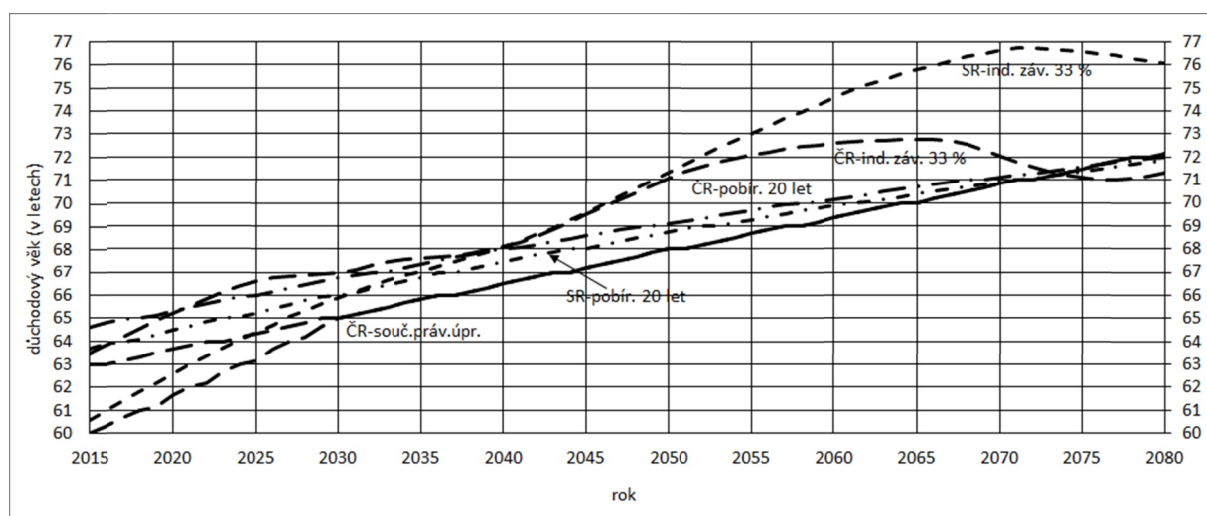
Důchodový věk na Slovensku vedoucí ke snížení průměrné doby pobírání důchodu na 20 let by vzhledem k předpokládané nižší střední délce života mohl být o něco nižší než důchodový věk podle tohoto kritéria pro ČR, ale vyšší než důchodový věk v ČR podle současné právní úpravy. V roce 2080 by měl dosahovat, stejně jako v ČR, zhruba 72 let.

Vzhledem k vyšší plodnosti v minulých letech je současná hodnota indexu závislosti seniorů na Slovensku jedna z nejnižších v EU, jen o málo vyšší než 20 %. Projekce Eurostatu však předpokládá pro Slovensko v tomto století poměrně nízkou plodnost i poměrně malý migrační přírůstek, což by mělo za následek výrazné stárnutí obyvatelstva. Pro udržení indexu závislosti na úrovni 33 % by bylo nutné do roku 2070 postupně zvýšit důchodový věk až na 77 let. Doba pobírání důchodu by se v takovém případě pohybovala kolem 15 let.

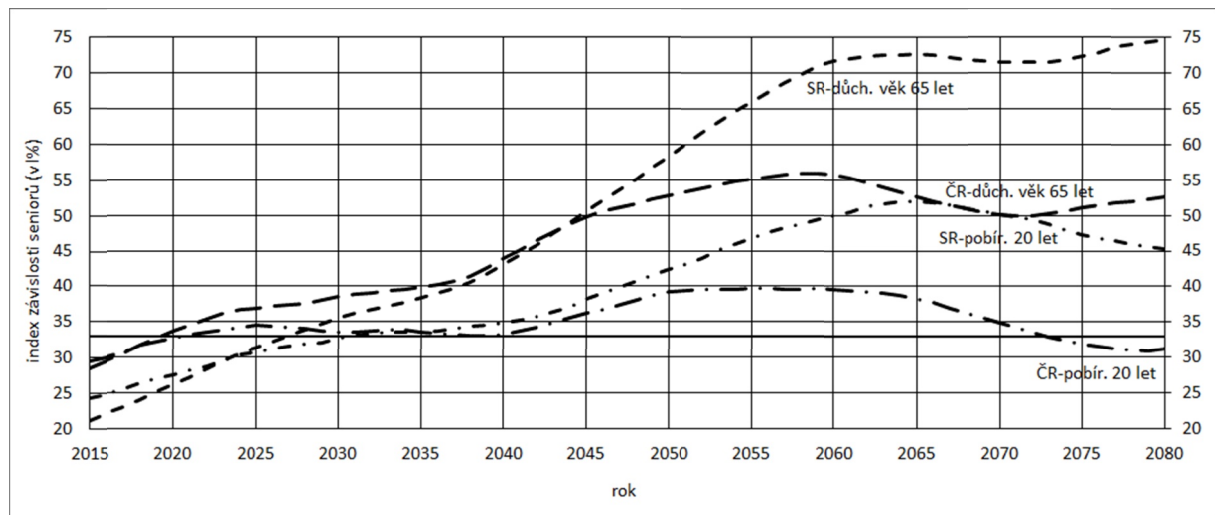
Při zvýšení pouze na 65 let by po roce 2060 byly hodnoty indexu závislosti vyšší než 70. Protože podíl zaměstnaných osob ze všech osob produktivního věku se pohybuje pouze kolem 70–80 %, znamenalo by to, že na 100 zaměstnaných osob by připadalo téměř 100 starobních důchodců.



Obr. 6: Průměrná doba pobírání důchodu při různém stanovení důchodového věku



Obr. 2: Důchodový věk podle různých kritérií jeho stanovení



Obr. 3: Index závislosti seniorů při různém stanovení důchodového věku

4. Důchodový věk v zemích EU28, Islandu, Norsku a Švýcarsku

Současná situace i předpokládaný budoucí vývoj v ostatních evropských zemích je podobný. Vzhledem k vyšší střední délce života by důchodový věk vedoucí k udržení doby pobírání důchodu na úrovni 20 let musel být v řadě zemí vyšší než v ČR (Tab. 1). Již v roce 2015 by musel být důchodový věk na Islandu vyšší než 68 let a i v řadě dalších zemí vyšší než 67 let. Naopak v Bulharsku by (vzhledem k poměrně vysoké úmrtnosti) „stačil“ důchodový věk pouze necelých 63 let. Vzhledem k předpokládanému růstu délky života a snižování rozdílů mezi jednotlivými zeměmi by se roku 2080 musel důchodový věk ve všech zemích zvýšit na více než 71 let.

Pokud by měl být důchodový věk stanoven tak, aby se hodnota indexu závislosti pohybovala kolem 33 %, byly by rozdíly vývoje v jednotlivých zemích daleko vyšší, neboť kromě úmrtnosti zde hraje významnou roli i vývoj plodnosti a migrace. Zatímco v roce 2015 by se důchodový věk ve všech zemích pohyboval zhruba mezi 60 (Irsko) a 66 roky (Itálie), v roce 2080 by v Irsku stačil důchodový věk necelých 68 let, zatímco na Slovensku a v Portugalsku více než 76 let (s dobou pobírání důchodu jen o málo vyšší než 16 let), v Německu více než 75 let.

Zachování důchodového věku na úrovni 65 let (což je současná hodnota v řadě evropských zemí) by znamenalo postupné zvýšení doby pobírání důchodu ve všech zemích na 26–28 let. Došlo by ovšem k výraznému nárůstu indexu závislosti seniorů. Zatímco v roce 2015 by neměla být v žádné zemi hodnota tohoto indexu vyšší než 36 % (Itálie), v roce 2080 by se nejnižší hodnoty pohybovaly kolem 40 %, ve většině zemí by však byl index vyšší než 50 % a v Portugalsku a na Slovensku by se blížil 75 %. Při stanovení důchodového věku na takové úrovni, aby průměrná doba pobírání důchodu činila zhruba 20 let, by ve většině zemí hodnota indexu neměla překročit 40 %, na Slovensku by dosahovala nejvýše 45 %.

Tab. 1: Vývoj důchodového věku, doby pobírání důchodu a indexu závislosti seniorů v zemích EU28, Islandu, Norsku a Švýcarsku

Země	Důchodový věk				Doba pobírání důchodu				Index závislosti seniorů			
	Pobírání důchodu 20 let		Index závislosti 33 %		Důchodový věk 65 let		Index závislosti 33 %		Důchodový věk 65 let		Pobírání důchodu 20 let	
	2015	2080	2015	2080	2015	2080	2015	2080	2015	2080	2015	2080
Belgie	66,6	72,5	63,9	71,1	21,3	27,2	23,0	21,3	30,2	49,9	26,5	29,7
Bulharsko	62,7	71,5	64,8	72,8	17,9	26,2	19,0	18,8	32,2	58,7	39,8	36,6
ČR	64,6	72,0	63,5	71,3	19,6	26,7	21,0	20,6	28,5	52,6	29,6	31,2
Dánsko	66,1	72,4	64,6	72,2	21,0	27,2	21,4	20,3	31,9	52,3	29,1	32,3
Estonsko	64,6	72,1	64,2	72,0	19,6	26,8	20,3	20,1	30,9	56,0	31,9	32,8
Finsko	66,8	72,6	65,4	72,6	21,6	27,3	21,2	20,0	34,4	54,9	28,8	33,0
Francie	68,1	73,0	64,8	71,8	22,7	27,7	24,0	21,1	32,3	51,6	24,2	30,3
Chorvatsko	64,0	71,8	64,2	73,5	19,1	26,6	19,8	18,4	30,8	59,8	33,6	37,5
Irsko	66,9	72,6	60,1	67,9	21,5	27,4	26,2	24,6	22,0	40,9	18,6	21,6
Island	68,3	73,0	60,8	70,0	23,0	27,8	28,0	22,8	22,9	45,8	17,0	26,6
Itálie	67,4	72,7	66,1	74,0	22,1	27,5	21,1	18,8	36,0	62,1	29,7	36,4
Kypr	66,7	72,4	60,2	70,0	21,5	27,2	26,2	22,3	22,2	46,6	18,8	27,8
Litva	63,7	72,0	64,2	66,7	18,9	26,6	19,5	26,0	30,9	37,7	34,3	21,6
Lotyšsko	63,0	71,8	64,4	69,1	18,3	26,4	18,9	22,5	31,4	44,5	36,7	26,6
Lucembursko	67,5	72,8	60,2	70,8	22,2	27,5	26,7	21,9	22,6	48,6	18,3	28,5
Maďarsko	63,2	71,9	63,5	73,3	18,5	26,5	19,8	18,6	28,6	58,1	33,7	36,9
Malta	66,8	72,6	64,1	72,7	21,7	27,3	22,6	19,8	30,2	53,9	24,9	33,4
Německo	66,8	72,6	65,9	75,4	21,5	27,3	21,0	17,4	35,3	65,5	30,9	40,3
Nizozemsko	66,7	72,5	63,8	73,6	21,4	27,3	23,0	19,0	29,9	57,7	25,6	35,6
Norsko	66,9	72,6	62,6	71,5	21,7	27,3	24,1	21,0	27,1	50,8	22,9	30,4
Polsko	64,7	72,1	61,7	74,6	19,6	26,8	23,0	17,7	23,6	65,2	24,4	39,6
Portugalsko	66,6	72,5	65,2	76,4	21,4	27,2	21,2	16,4	33,6	74,4	29,5	45,2
Rakousko	66,9	72,6	63,7	73,8	21,5	27,3	22,8	18,9	30,0	59,4	25,7	36,1
Rumunsko	63,1	71,7	62,9	72,5	18,4	26,4	21,0	19,2	27,3	54,7	32,7	35,1
Řecko	66,7	72,6	65,6	73,9	21,5	27,3	21,0	18,7	34,6	62,3	30,4	36,7
Slovensko	63,7	71,9	60,6	76,0	18,8	26,6	22,7	16,1	21,3	74,6	24,3	45,3
Slovinsko	66,2	72,4	63,4	71,8	21,1	27,2	22,6	20,6	28,5	54,5	25,6	31,5
Španělsko	67,8	72,8	63,6	71,4	22,4	27,6	23,8	21,3	29,9	52,8	23,9	29,3
Švédsko	67,0	72,6	65,4	70,9	21,8	27,4	21,5	21,6	34,1	49,3	28,6	29,2
Švýcarsko	67,8	72,8	63,2	72,0	22,6	27,6	24,2	20,9	28,8	52,6	22,9	31,0
UK	66,7	72,6	63,9	70,9	21,6	27,4	23,0	21,6	30,2	49,5	26,0	29,1

Zdroj: vlastní výpočet na základě projekce Eurostatu

5. Závěr

Trvalé zvyšování důchodového věku je jednou z možností, jak bránit růstu průměrné doby pobírání starobních důchodů růstu indexu závislosti seniorů. To by vedlo ke stabilizaci finančního zatížení důchodových systémů. Růst důchodového věku v ČR podle současné právní úpravy (předpokládající do roku 2030 sjednocení důchodového věku mužů a žen a do roku 2080 zvýšení důchodového věku na 72 let) by (za předpokladu, že se skutečný demografický vývoj nebude výrazně lišit od projekce Eurostatu) vedl k postupné stabilizaci průměrné doby pobírání důchodu na úrovni o málo vyšší než 20 let i ke stabilizaci hodnoty

indexu závislosti seniorů na úrovni mezi 30 a 40 %. Na Slovensku, kde projekce předpokládá nižší plodnost i migraci, by pro udržení indexu závislosti bylo nutno zvyšovat důchodový věk o něco více.

Situace v ostatních evropských zemích je podobná. Je však zřejmé, že v zemích s vyšší plodností a imigrací není nutno pro zachování indexu závislosti tak výrazně zvyšovat důchodový věk.

V každém případě bude nezbytné, aby současně s růstem důchodového věku docházelo k vytváření podmínek a příležitostí k zaměstnávání osob starších 60let. V opačném případě by zvyšování důchodového věku mohlo mít za následek sice snížení počtu starobních důchodců, ale za cenu zvyšování počtu předčasných důchodů, invalidních důchodů, žadatelů o sociální dávky a především zvyšování rizika chudoby seniorů.

Poděkování

Článek vznikl za podpory Interní grantové agentury Vysoké školy ekonomické v Praze F4/24/2013 Úmrtnost a stárnutí obyvatelstva ČR.

Literatura

DOPORUČENÍ RADY (2014) ze dne 8. července 2014 k národnímu programu reforem České republiky na rok 2014 a stanovisko Rady ke konvergenčnímu programu České republiky z roku 2014. *Úřední věstník Evropské unie*. 2014/C 247/03

http://ec.europa.eu/europe2020/pdf/csr2014/csr2014_council_czech_cs.pdf

EUROSTAT (2014). Statistics Database. Database by themes. Population and social conditions. Population (populat). Population projections (proj). EUROPOP2013 - *Population projections at national level (proj_13n)*. Projected population (proj_13np). Main scenario - Population on 1st January by sex and single year age (proj_13npms). [Cit. 2014-10-25].

http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=proj_13naasfr&lang=en.

EUROSTAT (2014). Statistics Database. Database by themes. Population and social conditions. Population (populat). Population projections (proj). EUROPOP2013 - *Population projections at national level (proj_13n)*. Assumptions (proj_13na). Age specific mortality rates by sex (proj_13naasmr). [Cit. 2014-10-25].

http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=proj_13naasmr&lang=en

KUČERA, Tomáš a PT2. (2014). *Návrh revizního systému nastavení hranice důchodového věku (včetně návrhu případného způsobu posunu této hranice)*. Předběžné výsledky dosavadního jednání pracovního týmu 2 Odborné komise pro důchodovou reformu.

<http://www.duchodova-komise.cz/wp-content/uploads/2014/10/N%C3%A1vrh-revizn%C3%ADho-syst%C3%A9mu-nastaven%C3%AD-hranice-d%C5%AFchodov%C3%A9ho-v%C4%9Bku-23.-%C5%99%C3%ADjna-2014.pdf>.

[Cit. 2014-11-17]

ZÁKON 155/1995 Sb. *O důchodovém pojištění*, aktuální znění, Příloha

Adresa autorů:

Tomáš Fiala, RNDr. CSc.
Katedra demografie
Fakulta informatiky a statistiky
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67 Praha 3
fiala@vse.cz

Zdeněk Pavlík, prof. Ing. DrSc.
Katedra demografie
Fakulta informatiky a statistiky
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67 Praha 3
fiala@vse.cz

Režim vývoja burzového indexu z hľadiska klastrovej analýzy

Stock Exchange Index Development Regime as an Application of the Cluster Analysis

Lukáš Friga, Vladimír Gazda

Abstract: Cluster analysis is performed on the non-overlapping 20 day windows of the stock exchange index A1FIN returns. The expected value, standard deviation, skewness and kurtosis were computed for each window. Then, the data were re-normalized and clustered via means of the cluster analysis methods. We show that the stock exchange development regimes have tendency to persist in the same cluster. In addition, a seemingly equivalent evolution of the index can be represented by different clusters.

Abstrakt: V predloženom príspevku je urobená klastrová analýza na neprekrývajúcich sa kľzavých oknách výnosností burzového indexu A1FIN. Okná majú dĺžku 20 obchodných dní a za každé okno sú počítané ukazovatele strednej hodnoty, štandardnej odchýlky, šikmosti a špicatosti. Takto získané údaje sú následne prenormované a je na nich urobená zhluková analýza. Je ukázané, že režimy vývoja burzového indexu majú tendenciu pretrvávajúť v tom istom zhľuku a zdanlivo rovnaký vývoj burzového indexu je reprezentovaný rozličnými zhľukmi.

Key words: Cluster analysis, Canberra distance, Ward method, return, A1FIN.

Kľúčové slová: zhluková analýza, vzdialenosť Canberra, Wardova metóda, výnos, A1FIN.

JEL classification: G12, C58

1. Úvod

Informačná explózia a dostupnosť obrovského množstva údajov vytvára priestor pre uplatňovanie nových štatistických metód zachytávajúcích komplexné vzťahy medzi jednotlivými súbormi údajov. V tejto oblasti sa do praxe dostávajú viacrozmerne štatistické techniky, medzi ktorými má nezastupiteľné miesto aj zhluková analýza. Tu je určitým prelomom uplatňovanie metód tejto analýzy práve na časových radoch, pričom sú tu (na rozdiel od statických údajov) údaje usporiadané v čase. Údaje majú tiež premenlivý - stochastický charakter. Zmysluplný výskum je ale možné robiť na stacionárnych časových radoch, ktoré v čase nemenia svoje základné stochastické charakteristiky. Pokiaľ uvažované časové rady nemajú vlastnosť stacionarity, je potrebné ju nejakým spôsobom dosiahnuť (odpočítaním deterministického trendu, diferenciáciou a pod.). Prehľad používaných metód klastrovania na časových radoch poskytol Liao (2005).

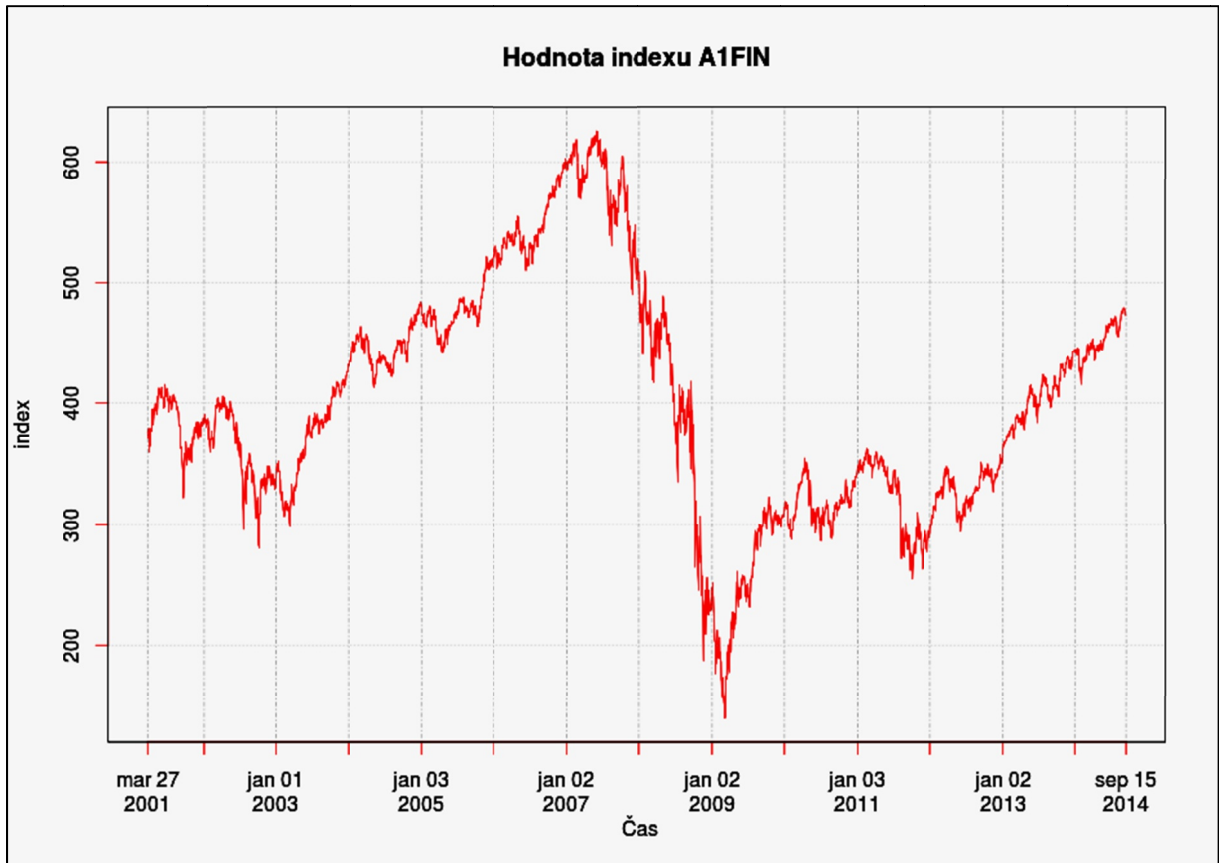
Predmetom našej analýzy je zhlukovanie vybraných charakteristík časového radu burzového indexu a ich konfrontácia s tradičným pohľadom.

2. Údajová základňa a metódy

V našom výskume sme sa zamerali na sektorový index A1FIN (Dow Jones Americas Financial Stock Index). Index bol vybraný z toho dôvodu, že na rozdiel od základného indexu Dow Jones Industrial prudšie reagoval na krízové výkyvy primárne spojené so zlyhaním finančného systému. Za takéto krízové poklesy považujeme pokles vyvolaný hypotekárnou krízou ako aj krízu financovania dlhov národných ekonomík Európskej menovej únie.

Do výskumu sme zahrnuli obdobie Marec 2001 - September 2014, teda obdobie, ktoré zahŕňalo niekoľko krízových poklesov. Krízový pokles trvajúci od marca po november v r. 2001 zásadne neovplyvňoval hodnoty burzových indexov v USA, pričom tieto si zachovali

relatívne stabilný vývoj neovplyvnený poklesom Internetového sektora (tzv. prasknutím Dot-com bubliny). Na vývoj burzových indexov Spojených štátov vplývali teroristické útoky na USA z 11 septembra 2001. Americké burzové indexy potom klesali až do januára 2003, kedy sa obnovil ich stabilný rast. V roku 2006 sa už prejavovali dôsledky hypotekárnej krízy, ktorá priniesla prudký pokles cien nehnuteľností a následné nesplácanie hypotekárnych úverov. To poškodilo najmä finančné inštitúcie. To sa potom prejavilo v poklese cien ich akcií, čo sa vyvolalo prepady burzových indexov trvajúce od konca roka r. 2007 až do začiatku r. 2009. Potom sa začalo sa obdobie stabilného rastu, s obdobiami čiastočnej stagnácie v r. 2010 a 2011 spôsobených zrejme krízou v Európskej menovej únii a ich neschopnosťou splácať dlhy národných ekonomík.



Obr. 7: Vývoj indexu A1FIN - Dow Jones Americas Financials Index

Hodnoty indexu A1FIN - Dow Jones Americas Financials Index sú uvedené na Obr. 1. Hodnoty indexu boli transformované na logaritmické výnosy

$$r_t = \ln(I_t) - \ln(I_{t-1}) \quad (1)$$

kde I_t je hodnota burzového indexu v čase t .

Potom v rámci jednotlivých neprekrývajúcich sa posuvných okien (Košmelj a Bagatelj, 1990; Kumar, 2002) o dĺžke 20 pozorovaní, kde $t = 20, 40, \dots, 174$ sme kvantifikovali nasledovné charakteristiky:

a. Očakávaný výnos

$$E(r_t) = 1/20 \sum_{i=0}^{19} r_{t-i} \quad (2)$$

b. Štandardná odchýlka

$$s(r_t) = \sqrt{1/20 \sum_{i=0}^{19} (r_{t-i} - \bar{r}_t)^2} \quad (3)$$

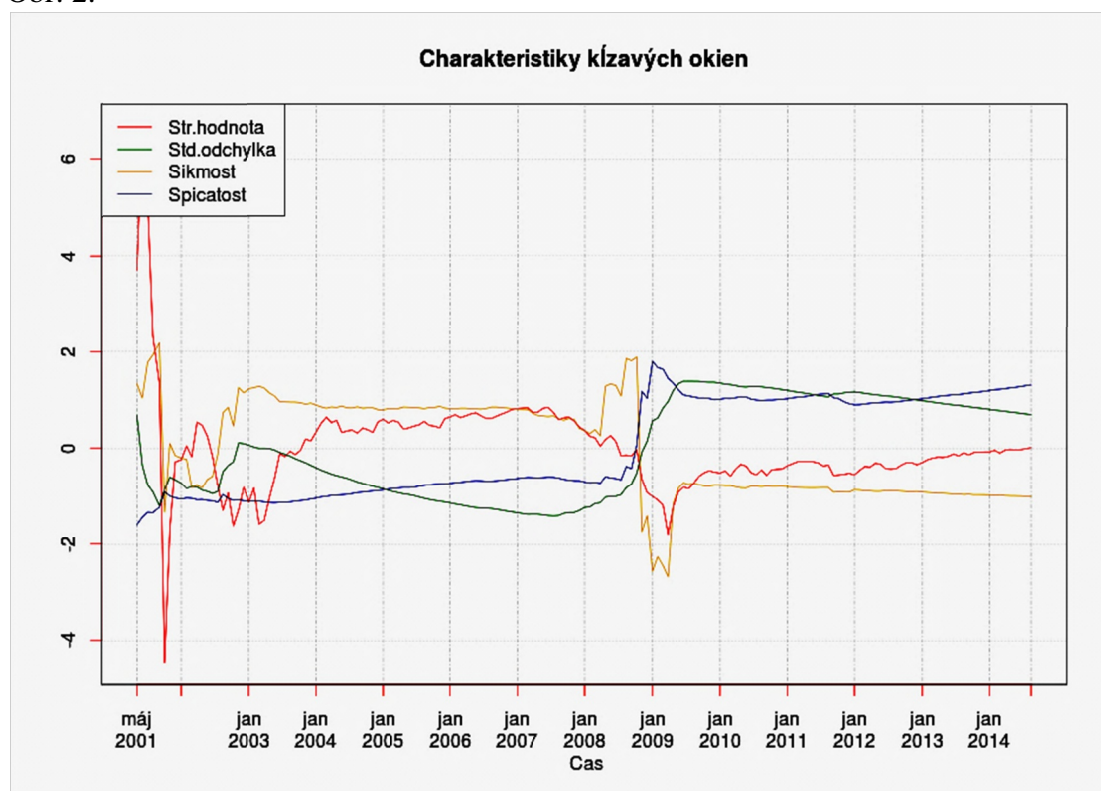
c. Šikmost'

$$skew(r_t) = \frac{1/20 \sum_{i=0}^{19} (r_{t-i} - \bar{r}_t)^3}{\left(1/20 \sum_{i=0}^{19} (r_{t-i} - \bar{r}_t)^2\right)^{3/2}} \quad (4)$$

d. Špicatost'

$$kurt(r_t) = \frac{1/20 \sum_{i=0}^{19} (r_{t-i} - \bar{r}_t)^4}{\left(1/20 \sum_{i=0}^{19} (r_{t-i} - \bar{r}_t)^2\right)^2} - 3 \quad (5)$$

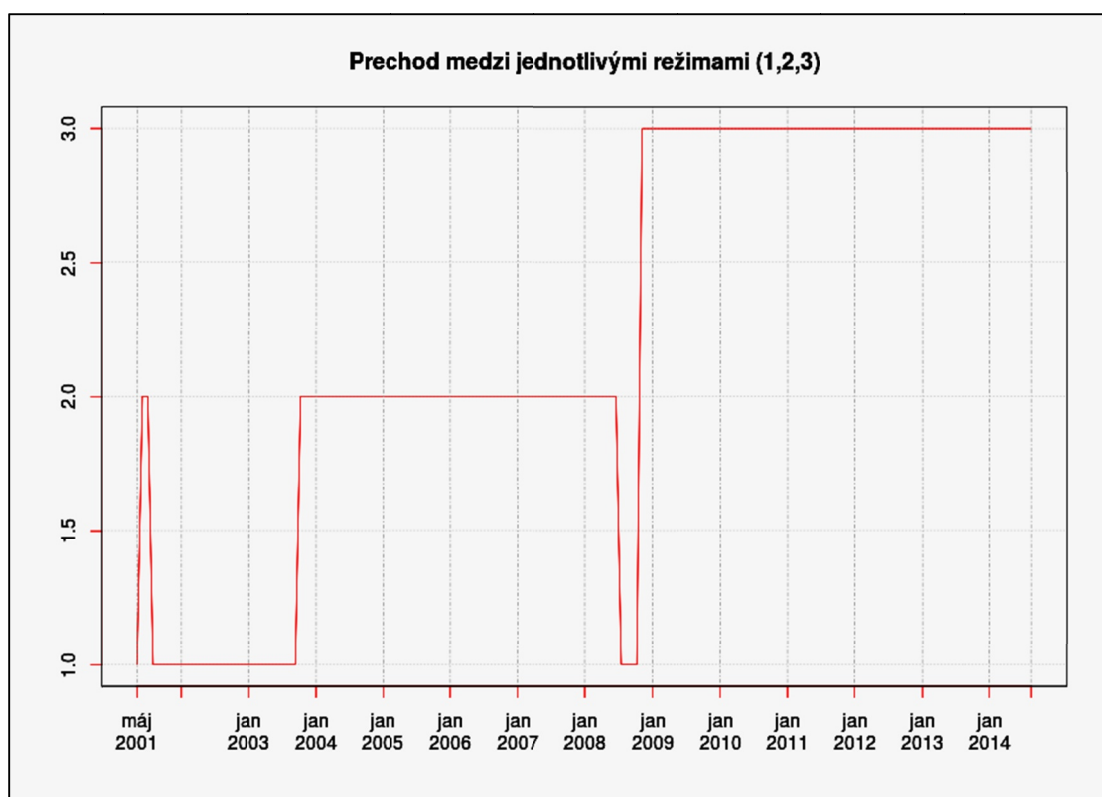
kde \bar{r}_t je priemerný výnos v okne t . Vyššie popísaným spôsobom sme získali 174 okien, pričom každé z nich bolo charakterizované štyrmi vyššie popísanými charakteristikami. Jednotlivé charakteristiky sme normalizovali, pričom po normalizácii každá z nich mala priemernú hodnotu 0 a rozptyl 1. Graf vývoja normalizovaných charakteristík je uvedený na Obr. 2.



Obr. 8: Očakávaná (stredná) hodnota, štandardná odchýlka, šikmost' a špicatost' počítané na kľzavých oknách (20 obchodných dní)

Z grafu je zrejmých niekoľko vlastností:

- Obdobie poklesov je sprevádzané nestabilitou jednotlivých charakteristík
- V prvom období stabilného rastu (2003-2007) vysoké hodnoty očakávaných výnosností dominujú nad ich štandardnými odchýlkami. Naopak, v druhom období rastu (roky 2009-2014) sa pozícia štandardnej odchýlky a očakávaného výnosu vymení.
- Pozitívna šikmosť v prvom období rastu je vystriedaná s negatívnou šikmosťou v druhom období rastu. Pozitívnu šikmosť tu interpretujeme ako dominanciu pozitívnych extrémnych výnosov (neočakávaných extrémnych prírastkov hodnôt indexu) nad negatívnymi extrémnymi výnosmi (neočakávanými extrémnymi poklesmi). U negatívnej šikmosti použijeme opačnú interpretáciu.
- V prvom období rastu je ukazovateľ špicatosti menší, ako v druhom období rastu. Znamená to, že neočakávané zmeny burzového indexu sú častejšie v druhom období rastu, ak to porovnáme s obdobím prvým.



Obr. 9: Očakávaná (stredná) hodnota, štandardná odchýlka, šikmosť a špicatosť počítané na kĺzavých oknách (20 obchodných dní)

V každom časovom okne teda odhadujeme prvky vektora

$$\mathbf{v}_t = [E(r_t), S(r_t), skew(r_t), kurt(r_t)] \quad (5)$$

pričom jednotlivé prvky vektora sú normalizované na nulovú strednú hodnotu a jednotkový rozptyl. Postupnosť vektorov je potom východiskom pre uplatnenie zhlukovej analýzy, ktorej cieľom je identifikovať zhluky – režimy trhu spolu s ich individuálnymi vlastnosťami. Pre výpočet vzdialeností medzi jednotlivými vektormi sme použili definíciu vzdialenosti typu Canberra, kde

$$d(v_t, v_{t'}) = \frac{\sum_{i=1}^4 |v_t^i - v_{t'}^i|}{\sum_{i=1}^4 |v_t^i| + |v_{t'}^i|}, t \neq t' \quad (6)$$

a následne Wardovu zhlukovaciú metódu. Na základe tvaru dendogramu sme usúdili, že optimálny počet zhlukov je 3. Základné charakteristiky jednotlivých zhlukov sú uvedené v Tab. 1 a prechody medzi jednotlivými zhlukmi v čase sú znázornené na Obr. 3. Z uvedeného vyplýva, že obdobiam krízových poklesov zodpovedá zhluk 1, prvému obdobiu rastu (2003-2007) zhluk 2 a druhému obdobiu rastu zhluk 3. Okrem toho je zaujímavé, že napriek využitiu neprekrývajúcich sa okien je demonštrovaná značná zotrvačnosť (najmä zhluky 2 a 3) a stabilita jednotlivých zhlukov v čase. Okrem toho zhluk 1 je charakterizovaný nízkou špicatost'ou (malý výskyt extrémnych výnosov), čo znamená, že v zhluku 1 dochádza k systematickému poklesu burzového indexu. Naopak, zhluk 2 vykazuje veľmi nízku štandardnú odchýlku, čo môže demonštrovať nízke riziká a stabilný rast burzového indexu v čase. Zhluk 3 podobne ako zhluk 2 vykazuje rovnako rastúci trend, ten je však na rozdiel od zhluku 2 dosahovaný za prítomnosti vysokých štandardných odchýlok výnosností a tiež vysokom koeficiente špicatosti čo svedčí o relatívne vysokom počte extrémnych výnosností.

Tab. 3: Základné charakteristiky jednotlivých zhlukov (aritmetické priemery v rámci okna)

	Stredná hodnota	Štand. odchýlka	Šikmost'	Špicatost'
Zhluk 1	-0,38	-0,45	0,67	-1,01
Zhluk 2	0,68	-1,01	0,81	-0,78
Zhluk 3	-0,41	1,04	-0,98	1,1

3. Záver

V predloženom príspevku bola vykonaná zhluková analýza na časovom rade výnosností burzového indexu A1FIN. Z prezentovaného výskumu vyplýva, že režimy vývoja burzového indexu majú tendenciu zotrvať v tom istom zhluku a aj zdanlivo rovnaký vývoj burzového indexu môže byť na základe niektorých charakteristík reprezentovaný rozličnými zhlukmi.

Literatúra

KOŠMELJ, K., BATAGELJ, V. 1990. Cross-sectional approach for clustering time varying data. In: *Journal of Classification*, roč. 7, č.1, s. 99-109.

KUMAR, M., PATEL, M., WOO, J. 2002. Clustering seasonality patterns in the presence of errors. In: *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM.

LIAO, W. 2004. Clustering of time series data – a survey. In: *Pattern Recognition*, roč. 38, č. 11, s. 1857 – 1874.

Adresa autora (-ov):

Lukáš Friga, Ing.
 Ekonomická fakulta TUKE
 Němcovej 32, 040 01 Košice
lukas.friga@student.tuke.sk

Vladimír Gazda, prof. Ing. PhD.
 Ekonomická fakulta TUKE
 Němcovej 32, 040 01 Košice
vladimir.gazda@tuke.sk

Zdraví a lidský kapitál v České republice a na Slovensku Health and human capital in the Czech Republic and the Slovakia

Simona Fučíková, Michaela Antovová

Abstract: Education and health are often together closely connected. The influence of education on the health status of the population has been already demonstrated in many studies. The aim of this paper is find out, how the public health and education in the Czech Republic and Slovakia was developing in the last years. The health status of the population is described in this paper through healthy life years. The area of education is presented through the expected years of study. The development of these two indicators is compared and found, how is in these countries fulfilled the condition, that education affects health status of the population of the Czech republic and Slovakia.

Abstrakt: Vzdelání a zdraví obyvatelstva bývají velmi často úzce propojeny. Vliv vzdělání na zdravotní stav obyvatelstva byl prokázán již v řadě výzkumů. Cílem tohoto příspěvku je zjistit, jakým způsobem se v posledních letech vyvíjel zdravotní stav obyvatelstva a oblast vzdělávání v České republice a na Slovensku. Zdravotní stav obyvatelstva je v tomto příspěvku popsán prostřednictvím délky života ve zdraví. Oblast vzdělávání je prezentována prostřednictvím předpokládané doby studia. Vývoj těchto dvou ukazatelů je následně porovnán a zjištěno, zda je v těchto zemích splněn předpoklad, že vzdělání ovlivňuje zdravotní stav populace České republiky a Slovenska.

Key words: healthy life years, human capital, Czech republic, Slovakia

Klíčové slová: zdravá délka života, lidský kapitál, Česká republika, Slovensko

JEL classification: I10, I21

1. Úvod

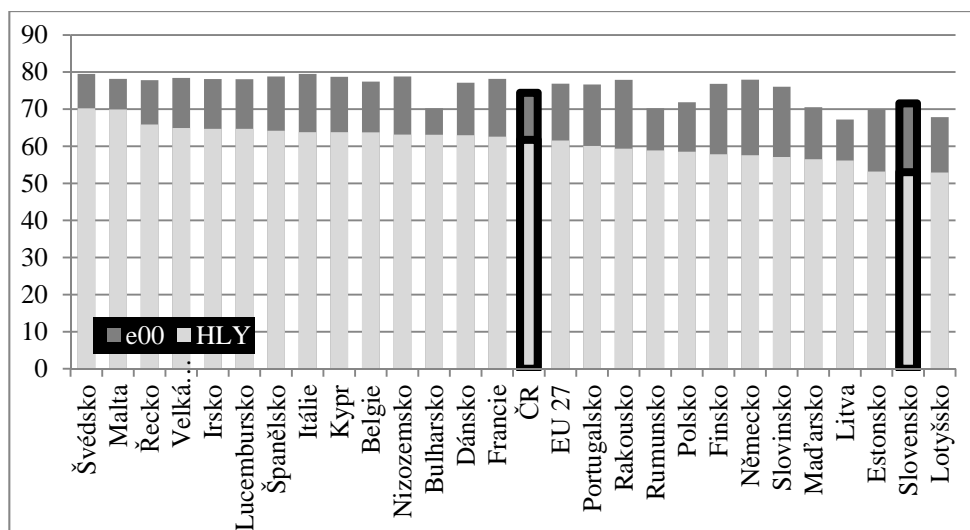
Vzdělání a zdraví obyvatelstva bývají velmi často úzce propojeny. Vliv vzdělání na zdravotní stav populace byl prokázán již v řadě výzkumů. (Mertl, Vychová, 2007) Rychtaříková například využila multinomiální logistickou regresi, pomocí které zjistila, že s rostoucím vzděláním roste i subjektivně vnímaná kvalita života. (Rychtaříková, 2006)

Cílem tohoto příspěvku je zachytit, jakým způsobem se v posledních letech vyvíjelo zdraví populace a oblast vzdělávání v České republice a na Slovensku. Tyto vývoje pak následně porovnat a zjistit, zda je v těchto zemích splněn předpoklad, že vzdělání ovlivňuje zdravotní stav populace České republiky a Slovenska.

2. Zdravá délka života

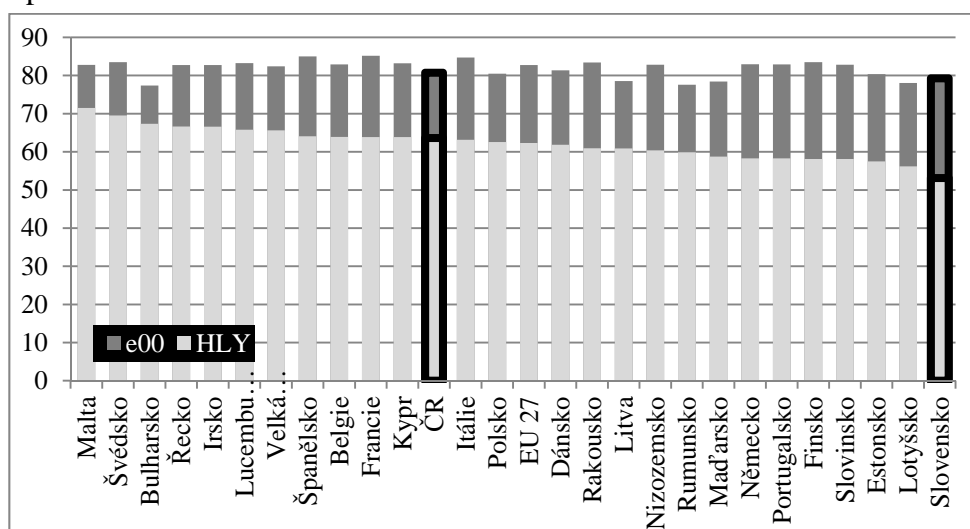
Délka lidského života se sleduje již řadu let z mnoha různých pohledů. V posledních letech však začala být pozorována nejen samotná délka lidského života, ale také kvalita tohoto života. Tím se především má na mysli, kolik let jedinec prožije v úplném zdraví a kolik v nemoci, popř. s nějakým zdravotním omezením. V roce 2000 byla oblast Veřejného zdraví začleněna do Lisabonské strategie. Na základě toho byl vytvořen ukazatel HLY (Healthy Life Years). HLY je kombinací střední délky života a subjektivního hodnocení zdravotního stavu, který je zjišťován pomocí výběrových šetření. (Daňková, Hrkal, 2005)

Vývoj délky života ve zdraví je velmi proměnlivý, proto bude vhodnější provést srovnání pomocí průměru za sledované období 2007 – 2012. (Langhamrová, 2013) Počátek sledovaného období je vymezen vstupem Rumunska a Bulharska do EU a konec sledovaného období je umístěn do roku 2012, tedy před vstupem Chorvatska do EU. I z toho důvodu je průměr EU uváděn za 27 členských států, nikoliv za současných 28.



Obr. 1: Střední délka života a délka života ve zdraví mužů při narození (průměr za roky 2007 – 2012)

Ve sledovaném období byla nejvyšší průměrná délka života ve zdraví mužů při narození ve Švédsku, a to 70,2 let. Česká republika se umístila v rámci EU 27 na 15. místě s délkou života ve zdraví 61,7 let. Muži se pohybují nad průměrem EU 27, ale pouze o 0,2 let. Slovensko se v evropském srovnání umístilo na předposledním místě pouze před Lotyšskem. Délka života ve zdraví byla na Slovensko průměrně v tomto období 53 let, což je o téměř 9 let méně než v České republice.



Obr. 2: Střední délka života a délka života ve zdraví žen při narození (průměr za roky 2007 – 2012)

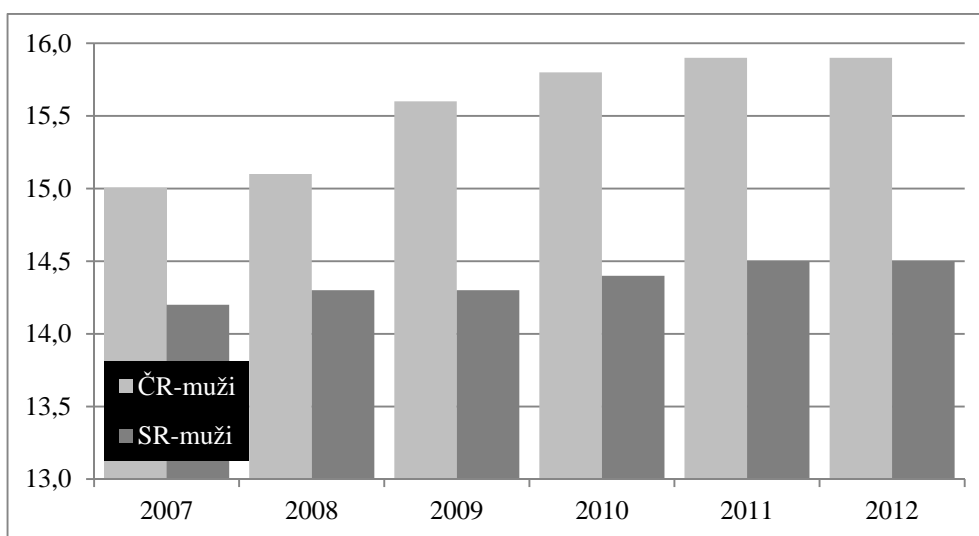
V evropském srovnání je ve sledovaném období nejvyšší průměrná délka života ve zdraví ženy při narození na Maltě (71,5). Česká republika se umístila s délkou života ve zdraví 63,6 let na 12. místě v rámci EU 27 a dokonce se pohybuje nad průměrem těchto zemí (62,3). Poslední místo pak patří druhé sledované zemi – Slovensku. Slovenské ženy měly za období let 2007 – 2012 průměrnou délku života ve zdraví při narození 53,1 let, což je o 10,5 roku méně než Česká republika a pod průměrem EU 27 je o 9,2 let.

3. Předpokládaná doba studia

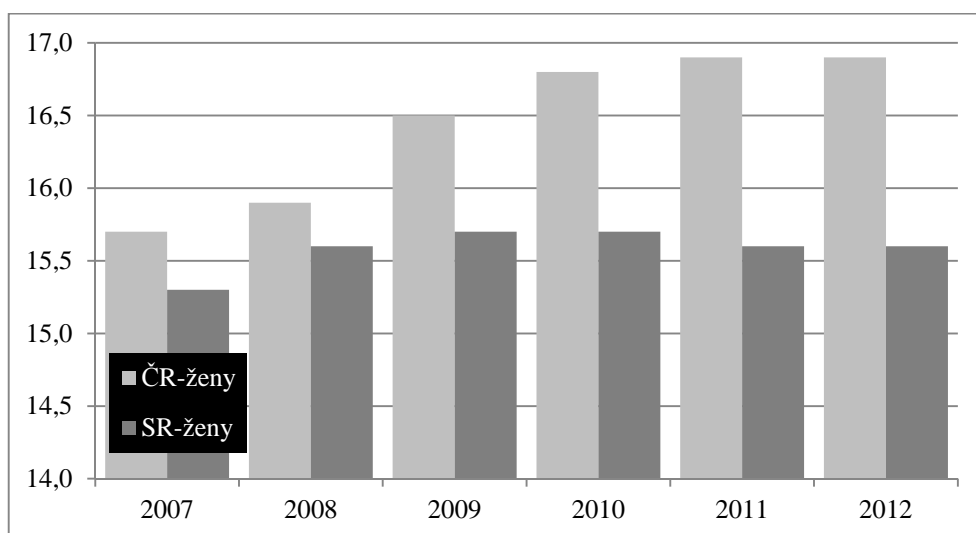
V rámci tohoto příspěvku je proces vzdělávání popsán prostřednictvím tzv. předpokládané doby studia. Lze ji interpretovat jako dobu, kterou bude přibližně studovat školák, který právě

nastoupil povinnou školní docházku. (UNDP, 2014) Zvolená byla z toho důvodu, že zachycuje nárůst osob v terciárním vzdělání.

Z obr. 3 je patrné, že předpokládaná doba studia u mužů v obou zemích ve sledovaném období rostla. V České republice vzrostla předpokládaná doba studia ve sledovaném období téměř o rok. Na Slovensku pak předpokládaná doba studia vzrostla pouze o 0,3 let. Z grafu je taktéž viditelné, že v České republice je předpokládaná doba studia po celé období vyšší, než je tomu na Slovensku. Muži v České republice měli v roce 2007 předpokládanou dobu studia o 0,8 let delší, než muži na Slovensku. V roce 2012 pak tento rozdíl činil dokonce už 1,4 let. Je tedy patrné, že český muž měl v roce 2012 v 6 letech předpokládanou dobu studia 15,9 let, tak slovenský muž bude mít předpokládanou dobu studia 14,5 let.



Obr. 3: Předpokládaná doba studia v České republice a na Slovensku v letech 2007 – 2012, muži



Obr. 4: Předpokládaná doba studia v České republice a na Slovensku v letech 2007 – 2012, ženy

Vývoj předpokládané doby studia u žen vykazuje podobný vývoj, jako tomu bylo u mužů. I zde je patrný nárůst v průběhu období 2007 – 2012. V České republice byla v roce 2007 předpokládaná doba studia 15,7 let a do roku 2012 vzrostla o 1,2 let na 16,9 let. Na Slovensku

v roce 2007 dosahovala hodnoty 15,3 let a stejně jako u mužů vzrostla pouze o 0,3 let. Podobně jako u mužů i zde je předpokládaná doba studia žen v České republice vyšší, než tomu je na Slovensku. V roce 2007 byla předpokládaná doba studia žen v České republice vyšší pouze o 0,4 let, než na Slovensku. Zatímco v roce 2012 byl rozdíl již markantnější, a to o 1,3 let.

4. Srovnání délky života ve zdraví a předpokládané doby studia

Závislost délky života ve zdraví na předpokládané době studia byla zkoumána pomocí analýzy kointegrace časových řad. Testovalo se zvláště pro muže, ženy, ČR a SR. Test jednotkového koeficientu určil, zda jsou časové řady a rezidua stacionární či ne. Podle toho bylo rozhodnuto, zda by mohl být mezi časovými řadami vztah.

U mužů ČR byla prokázána nestacionarita časových řad a reziduí. Jde o případ zdánlivé regrese, kde by mohl být krátkodobý vztah. Po diferenciaci (stacionarizaci) časových řad nebyl zjištěn žádný vztah (parametr ve zkoumaném modelu nebyl statisticky významný).

Časové řady žen ČR vykazovaly stejné výsledky jako muži ČR. Jednalo se o případ zdánlivé regrese a nebyl prokázán krátkodobý vztah.

V případě mužů SR byla zjištěna nestacionarita časových řad a stacionarita reziduí a kointegrační regrese. V tomto případě se ověřovalo, zda mezi časovými řadami není krátkodobý i dlouhodobý vztah. Analýza krátkodobých vztahů byla provedena modelem ADL. Z korelogramu bylo patrné, že zde není autokorelace (nebylo nutné dynamizovat model, tedy přidávat zpožděné proměnné). Diagnostická kontrola (test autokorelace, test normality, test podmíněné heteroskedasticity) ověřila správnost modelu.

Dependent Variable: ZDRAVA_DELKA_ZIVOTA
Method: Least Squares
Date: 11/18/14 Time: 23:15
Sample: 2007 2012
Included observations: 6

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DELKA_VZDELANI	3.688538	0.046443	79.42143	0.0000
R-squared	-0.442384	Mean dependent var		53.00000
Adjusted R-squared	-0.442384	S.D. dependent var		1.360882
S.E. of regression	1.634410	Akaike info criterion		3.971453
Sum squared resid	13.35648	Schwarz criterion		3.936746
Log likelihood	-10.91436	Hannan-Quinn criter.		3.832519
Durbin-Watson stat	1.297606			

Obr. 5: Model krátkodobého vztahu časových řad mužů SR

Model lze přepsat rovnicí: $\text{délka_života_ve_zdraví} = 3,689 \text{ předpokládaná_doba_studia}_t + a_t$. Byla zjištěna přímá závislost, tzn. pokud roste doba studia, zvyšuje se délka života ve zdraví. Kointegrační test prokázal, že mezi časovými řadami není dlouhodobý vztah.

U žen SR byla zjištěna stacionarita u časových řad i reziduí. V tomto případě lze zkoumat pouze krátkodobý vztah (postup je stejný jako při analýze mužů SR).

Dependent Variable: ZDRAVA_DELKA_ZIVOTA
Method: Least Squares
Date: 11/18/14 Time: 23:32
Sample: 2007 2012
Included observations: 6

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	203.6923	23.92004	8.515549	0.0010
DELKA_VZDELANI	-9.661538	1.534919	-6.294494	0.0033
R-squared	0.908300	Mean dependent var		53.13333
Adjusted R-squared	0.885375	S.D. dependent var		1.492202
S.E. of regression	0.505204	Akaike info criterion		1.733491
Sum squared resid	1.020923	Schwarz criterion		1.664078
Log likelihood	-3.200474	Hannan-Quinn criter.		1.455624
F-statistic	39.62066	Durbin-Watson stat		2.715477
Prob(F-statistic)	0.003255			

Obr. 6: Model krátkodobého vztahu časových řad žen SR

Model rovnice: $\text{délka_života_ve_zdraví}_c = 203,692 - 9,662 \text{ předpokládaná_doba_studia}_t + a_t$.
Pokud se sníží doba studia, zvýší se délka života ve zdraví.

5. Závěr

U mužů a žen ČR byla zjištěna zdánlivá regrese a po diferenciaci časových řad se nepotvrdil krátkodobý vztah (důvodem mohou být krátké časové řady - nestihla se projevit závislost). Kointegrační regrese byla zaznamenána u mužů SR. Zde byl prokázán pouze krátkodobý vztah a nepřímá závislost mezi časovými řadami. Stacionarita časových řad a reziduí u žen SR umožnila zkoumat pouze krátkodobý vztah, jež byl ověřen a jednalo se o nepřímou závislost.

Poděkování

Tento příspěvek byl zpracován za podpory grantu IGA VŠE 68/2014 „Ekonomické a zdravotní souvislosti stárnutí populace“.

Literatura

- DAŇKOVÁ, Š. – Hrkal, J. 2005. ANALÝZA: Zdravá délka života u obyvatel EU. In: *Demografický informační portál* [online]. [cit. 2014-10-27]. Dostupné z: http://www.demografie.info/?cz_detail_clanku=&artclID=107
- EUROSTAT. 2014. Search Database [online]. [cit. 2014-10-26]. Dostupné z: http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database
- LANGHAMROVÁ, J. 2013. *Střední délka života a délka života ve zdraví v evropských zemích*. Plakát. Konference České demografické společnosti.
- MERTL, J. – VYCHOVÁ, H. 2007. *Úloha vzdělání a zdraví v ekonomickém rozvoji: teoretická analýza a její aplikace v podmínkách ČR*. Praha: Národohospodářský ústav Josefa Hlávky, 171 s., 1/2007. ISBN 80-867-2932-X.
- UNITED NATIONS DEVELOPMENT PROGRAMME. UNDP [online]. 2014 [cit. 2014-10-26]. Dostupné z: <http://hdr.undp.org/en>
- RYCHTAŘÍKOVÁ J. 2006. Zdravá délka života v současné české populaci. In: *Demografie*, roč. 48, č. 3, s. 166 – 178.

Adresa autorů:

Simona Fučíková, Ing.
Vysoká škola ekonomická v Praze
Náměstí Winstona Churchilla 4
130 67 Praha 3
simona.fucikova@vse.cz

Michaela Antovová, Ing.
Vysoká škola ekonomická v Praze
Náměstí Winstona Churchilla 4
130 67 Praha 3
michaela.antovova@vse.cz

Ukazovatele centrality pri hodnotení polohy predajní Centrality measures in retail unit location evaluation

Marek Gróf

Abstract: In the presented article we propose a graph representation of selected urban areas to provide the base for studying customer and retailer behaviour, taking into account travel (search) costs of customers when shopping. Vertices of the graph represent individual customers and retailers, while edges represent paths traversable by customers when shopping. Graph representation of a real selected urban area is also provided, with current retailer locations described using basic centrality measures.

Abstrakt: V práci predstavujeme alternatívu reprezentácie skúmanej oblasti pomocou teórie grafov, kde vo výslednom grafe jednotlivé vrcholy predstavujú jednotlivých zákazníkov alebo predajcov, pričom hrany grafu predstavujú cesty využiteľné zákazníkmi pri nakupovaní. Daná reprezentácia by mala slúžiť ako podklad pre analýzu správania sa zákazníkov a predajcov ak berieme do úvahy aj náklady vyhľadávania ktoré zákazníci pri nakupovaní podstupujú. Reprezentácia reálnej zvolenej oblasti je v práci taktiež predstavená. Súčasná poloha predajní je následne popísaná základnými kritériami centrality.

Key words: Graph, Customer, Retailer, Centrality measures, Topology

Kľúčové slová: graf, zákazník, predajca, kritériá centrality, topológia

JEL classification: D0, O18

1. Úvod

Klasický pohľad na cenotvorbu je založený na predpoklade dokonalej konkurencie, ktorý predpokladá existenciu jedinej ceny akceptovanej všetkými predajcami. Rovnako sa typicky predpokladá, že kupujúci ignorujú náklady vyhľadávania vhodnej možnosti nákupu. V skutočnosti však empirické výskumy ukazujú, že zákazníci tieto náklady berú do úvahy a minimalizujú prejdenú vzdialenosť pri nakupovaní (pozri štúdiu o nakupovaní chodcov od autorov Gärling E. a Gärling T., 1988).

Rôzne modely boli navrhnuté za účelom zohľadnenia nákladov vzdialenosti pri nakupovaní zákazníkov. Hlavným rozdielom medzi uvádzanými prácami je však forma reprezentácie zvolenej oblasti za účelom modelovania nákladov vzdialenosti. Iba niekoľko z nich je však založených na princípe grafov. Klasickým prístupom (príklad Drezner, 1982) je umiestňovanie predajní na Karteziánskej rovine, predstavujú tak dvojrozmerné rozšírenie pôvodného modelu z práce Hotellinga (1929), kde dvaja predajcovia umiestňujú svoje predajne v zvolenej oblasti reprezentovanej úsečkou. Posunom od týchto spojených modelov je využitie diskretných umiestnení na kruhovej reprezentácii oblasti v práci Huang a Levinson (2011). Prístup teórie grafov bol rovnako použitý v prácach Granot et al. (2010), kde boli v neorientovanom váženom grafe predajcovia reprezentovaní vrcholmi a vážené hrany predstaovali zákazníkov, a Hakimi (1983) kde boli zákazníci reprezentovaní ohodnotenými vrcholmi a predajne boli umiestňované na hranách grafu.

V našej práci pokračujeme v prístupe teórie grafov a pokúšame sa zvolenú oblasť reprezentovať grafom, kde jednotlivé vrcholy predstavujú tak zákazníkov ako aj predajcov, pričom topológia grafu by mala odrážať skutočné umiestnenie zákazníkov a predajcov. Následne hodnotíme umiestnenie predajní pomocou zvolených kritérií centrality.

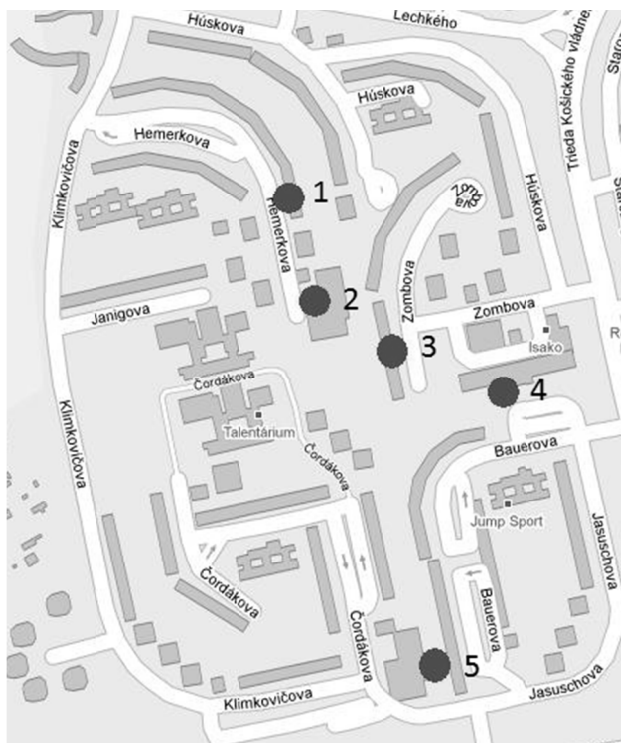
2. Popis zvolenej oblasti a transformácia do grafu

Pre potreby štúdie sme si zvolili sídlisko Košického Vládneho Programu (KVP) v Košiciach. Mapa oblasti je zobrazená na obrázku 1. Vo zvolenej oblasti sa spolu nachádza 5 predajní, ktorých označenie, názvy, číslo vrcholu vo výslednom grafe ako aj predajná plocha sú vedené v tabuľke 1.

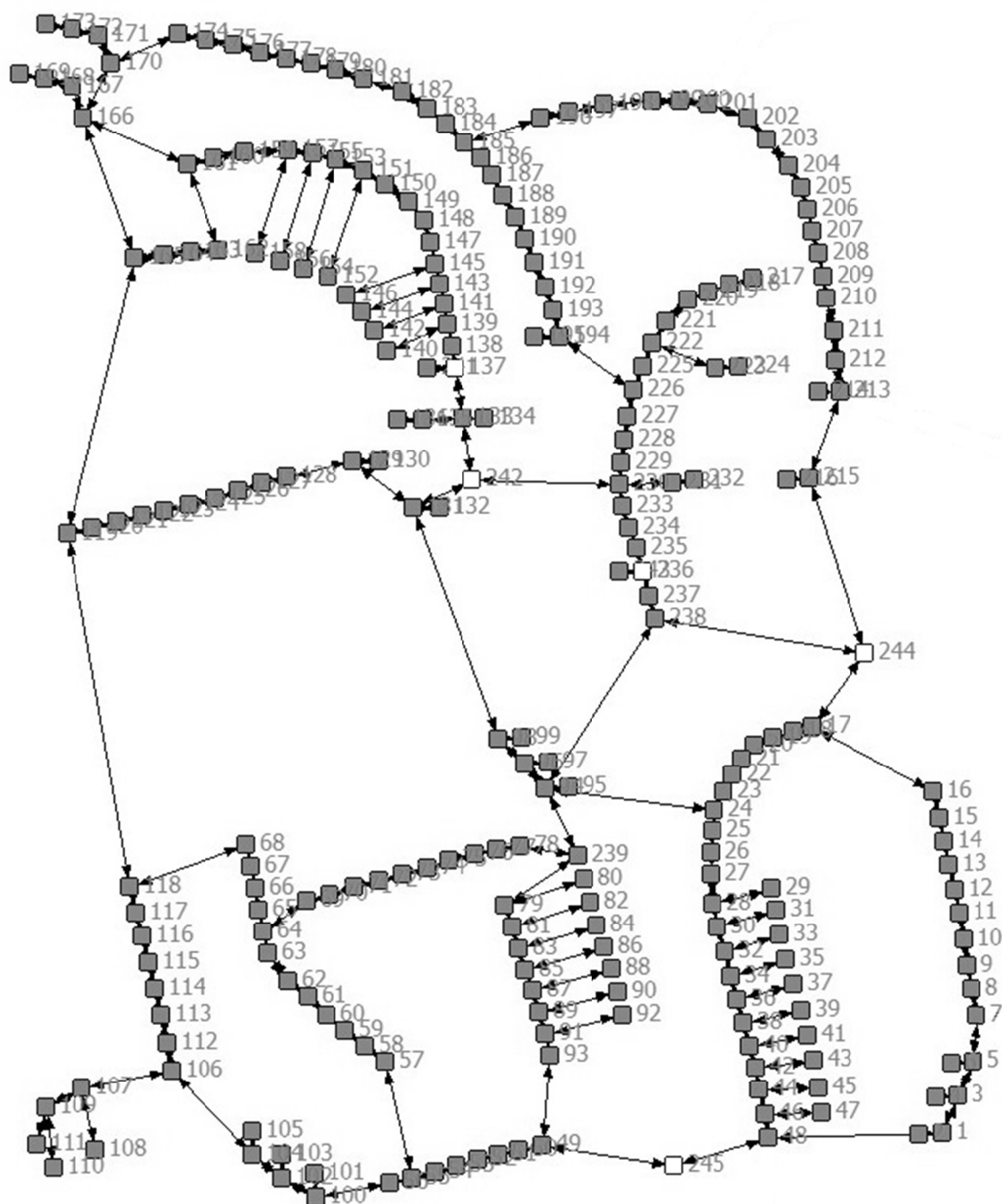
Tab.4: Označenie, názvy, čísla vrcholov v grafe a rozloha predajní v sledovanej oblasti

	Predajňa	Vrchol	Rozloha v m ²
1	Potraviny Hemerková	137	15
2	Bala	242	300
3	Milk Agro	236	40
4	Billa	244	450
5	Fresh	245	220

Pri transformácii oblasti do grafu sme postupovali nasledujúcim spôsobom. Každý vchod bytového bloku považujeme za jedného zákazníka. Keďže vo zvolenej oblasti sa nachádzajú dva rôzne typy obytných blokov, z ktorých má jeden typ približne dvojnásobný počet bytov pripadajúcich na jeden vchod, sú tieto vchody vo výslednom grafe následne reprezentované jedným alebo dvoma vrcholmi. V grafe sú následne rovnako zahrnuté predajne nachádzajúce sa vo zvolenej oblasti, z ktorých je každá reprezentovaná jedným vrcholom. Graf tak obsahuje vrcholy reprezentujúce tak zákazníkov ako aj predajcov. Hrany grafu predstavujú cesty a chodníky využiteľné zákazníkmi pri nakupovaní. V tejto fáze budeme zatiaľ všetky vzdialenosti v danom grafe, teda dĺžky jednotlivých hrán, považovať za ekvidistantné. Výsledná reprezentácia zvolenej oblasti v podobe grafu je predstavená na obrázku 2. Vrcholy reprezentujúce zákazníkov sú označené ako plné, vrcholy predstavujúce predajne sú označené ako prázdne.



Obr.10: Mapa zvolenej oblasti a umiestnenie jednotlivých predajní



Obr.2: Transformácia zvolenej oblasti do grafu, predajne sú označené ako biele

3. Kritériá centrality

Reprezentácia zvolenej oblasti predstavená v druhej časti by mala slúžiť ako podklad pre podrobnú analýzu lokalizácie predajní a následne rozhodovania a tvorby cenovej stratégie predajcami. V tejto práci však popíšeme umiestnenie predajní iba za pomoci základných kritérií centrality z teórie grafov, porovnávajúc tak na jednej strane samotné umiestnenie predajní ako ja vhodnosť využitia daných kritérií centrality. Za týmto účelom budeme používať nasledujúce kritériá centrality.

Betweenness kritérium centrality predstavuje hodnotenie vrcholu na základe počtu najkratších ciest medzi všetkými dvojicami vrcholov (bez zahrnutia samotného vrcholu) v grafe ktoré prechádzajú daným vrcholom, ku počtu všetkých najkratších ciest medzi danou dvojicou vrcholov. Matematicky môžeme toto kritérium vyjadriť ako:

$$C_B(v) = \sum_{s \neq t \neq v} \frac{q_{st}(v)}{q_{st}}, \quad (1)$$

kde q_{st} je najkratšia cesta medzi vrcholmi s a t , a $q_{st}(v)$ je najkratšia cesta medzi týmito dvoma vrcholmi, ktorá navyše však prechádza vrcholom v , pre ktorý dané kritérium počítame.

Closeness centralita predstavuje prevrátenú hodnotu kumulatívnej vzdialenosti sledovaného vrcholu od všetkých ostatných vrcholov grafu, pričom táto vzdialenosť je vyjadrená ako dĺžka najkratšej cesty medzi danou dvojicou vrcholov. Matematicky tento vzťah môžeme vyjadriť ako:

$$C_C(v) = \sum_{s \neq v} \frac{1}{|q_{sv}|}, \quad (2)$$

Tretím kritériom centrality ktoré budeme v práci sledovať je Eigenvektor centralita. Toto kritérium zachytáva vplyv zvoleného vrcholu v grafe. Jednotlivým vrcholom sú priradzované hodnotenia, pričom ak vrchol susedí s vrcholom ktorý ma vysoké hodnotenie, bude to výrazne prispievať k jeho vlastnému hodnoteniu, na rozdiel od susedstva s vrcholom s nízkym hodnotením. Eigenvektor centralita tak zachytáva relatívnu dôležitosť jednotlivých vrcholov grafu. V nasledujúcom texte ju uvádzame ako G_E .

4. Hodnotenie jednotlivých predajní

Pre všetky vrcholy grafu sme vypočítali jednotlivé kritériá centrality uvedené v tretej časti práce. Následne sme sledovali na ktorom mieste sa jednotlivé predajne umiestnili v hodnotení podľa jednotlivých kritérií. Výsledky sú prezentované v tabuľke 2.

Tab.2: Umiestnenie predajní v hodnotení podľa jednotlivých kritérií centrality

Predajňa	C_B	C_C	C_E
Potraviny Hemerková	37	42	27
Bala	6	5	10
Milk Agro	26	165	24
Billa	17	9	8
Fresh	124	54	119

Ako vidíme z výsledkov, predajne číslo 2 a 4 sa umiestnili v najlepšej 20tke vrcholov podľa všetkých zvolených kritérií. Ako je uvedené v tabuľke 1, ide o predajne s najväčšou rozlohou. Môžeme predpokladať, že vzhľadom na snahu o využitie danej predajnej plochy si tieto predajne zámerne volili na svoje umiestnenie vrcholy s výrazne centrálnou polohou. V prípade predajní číslo 1 a 3 ide o predajne s výrazne obmedzenou predajnou plochou. Môžeme predpokladať, že táto obmedzená predajná plocha mala za následok nižšiu prioritu kladenú na samotné umiestnenie predajne, nakoľko obmedzenú kapacitu by tieto predajne nemali mať problém naplniť. Aj napriek tomu, sa podľa jednotlivých kritérií umiestnili dané predajne pomerne vysoko, vzhľadom na celkový počet vrcholov ktorý bol rovný 245. V prípade predajne číslo 5 sme zaznamenali umiestnenie približne v polovici poľa podľa kritérií Betweenness a Eigenvektor centrality, no oveľa vyššie umiestnenie pokiaľ ide o Closeness centralitu. Tento fakt môže byť spôsobený umiestnením predajne na okraji zvoleného územia,

pričom v skutočnosti môže spádová oblasť predajne zasahovať aj mimo zvolenej oblasti. Nepomer medzi jednotlivými kritériami v tomto prípade je spôsobený práve okrajovým umiestnením vrcholu, ktoré je pomerne prísne posudzované Betweenness a Eigenvektor centralitami, no zároveň Cloeseness centralita naznačuje, že daný vrchol má, v danej oblasti, strategicky výhodné postavenie.

5. Záver

V práci sme predstavili alternatívny spôsob reprezentácie reálnej topológie oblasti prostredníctvom grafu, v ktorom jednotlivé vrcholy predstavovali tak zákazníkov ako predajcov, pričom hrany grafu predstavovali schodné cesty a chodníky využiteľné zákazníkmi pri nakupovaní. Prostredníctvom daného prístupu sa snažíme poskytnúť podklad pre štúdium rozhodovania tak zákazníkov ako aj predajcov, v prípade ak sa do klasického modelu rozhodovania a cenotvorby rozhodneme zahrnúť aj náklady vyhľadávania. Ako príklad je uvedená reprezentácia zvolenej mestskej časti Košického Vládneho Programu v Košiciach prostredníctvom daného grafu. Následne sme uviedli umiestnenie jednotlivých predajní v hodnotení všetkých vrcholov použitím troch základných kritérií centrality. Zo sledovaných predajní, dve predajne s najväčšou predajnou plochou sa umiestnili v daných hodnoteniach relatívne vysoko. Naopak predajne s výrazne nižšou predajnou kapacitou nemajú v grafe rovnako výrazné centrálné postavenie.

Literatúra

- HOTELLING, H. 1929. Stability in competition. In: *Economic Journal*, 39, 41 – 57.
- GÄRLING, T. – GÄRLING, E. 1988. Distance minimization in downtown pedestrian shopping. In: *Environment and Planning A*, 20(4), 547 – 554.
- GRANOT, D. – GRANOT, F. – RAVIV, T. 2010. On competitive sequential location in a network with a decreasing demand. In: *European Journal of Operational Research*, 205, 301 – 312.
- DREZNER, Z. 1982. Competitive location strategies for two facilities. In: *Regional Science and Urban Economics*, 12, 485 – 493.
- HAKIMI, S.L. 1983. On locating new facilities in a competitive environment. In: *European Journal of Operational Research*, 12(1), 29 – 35.
- HUANG, A. – LEVINSON, D. 2011. Why retailers cluster: an agent model of location choice on supply chains. In: *Environment and Planning B: Planning and Design*, 38(1), 82 – 94.

Adresa autora:

Marek Gróf, Ing. PhD.
Technická Univerzita v Košiciach,
Ekonomická fakulta
Němcovej 32, 040 01 Košice
marek.grof@tuke.sk

Miery inovácie – prehľad príkladov

Degree of innovation - summary of examples

Jozef Chajdiak

Abstract: Article constitutes a continuation of the previous number. It contains an ordered set of indicators summarizing the rate of innovation. Indicators reflecting the rate of innovation are classified according to the complexity of calculation, namely the extent innovation expressing the rate of innovation simple to complex rates of innovation.

Abstrakt: Článok predstavuje pokračovanie z predchádzajúceho čísla. Obsahuje usporiadanú množinu ukazovateľov s prehľadom mier inovácie. Ukazovatele vyjadrujúce mieru inovácie sú usporiadané podľa zložitosti výpočtu, konkrétne od mier inovácie vyjadrujúce jednoduchú mieru inovácie k zložitým mieram inovácie.

Key words: Degree of innovation, Overview rate of innovation, calculation

Kľúčové slová: miery inovácie, prehľad mier inovácie, výpočet

JEL classification: A10, Z00

1. Úvod

Miery inovácie v tomto príspevku predstavujú pomer parametra/ov inovácie na jednotku parametra/ov ekonomiky.

2. Parametre ekonomiky

Parametre ekonomiky predstavujú buď parametre výstupu (tržby, výnosy), buď parametre vstupu (osobné náklady, odpisy, výrobná spotreba alebo jej časti) a buď parametre efektívnosti (zisk, rentabilita vlastného imania, ziskovosť, , pridaná hodnota podiel pridanej hodnoty v tržbách).

3. Parametre vlastnosti inovácie

Parametre vlastnosti inovácie majú bohatšie zastúpenie, tak ako máme rozličné zastúpenie inovácií. Môžeme ich členiť na naturálne, na užitočnosťné, na efektívnosťné, ekonomické a možno menovať aj ďalšie. Vhodné je si tu uvedomiť žiaduci smer vývoja. Môže byť rast a môže byť aj pokles.

Každý z okruhov má svoju množinu konkrétnych hodnôt s konkrétnym smerom vývoja. Napríklad naturálne môžu mať parameter inovácie pokles hmotnosti. Parameter hmotnosti má svoje vymedzenie na celý výrobok, ale môže byť aj hmotnosť inovatívne nosnej časti výrobku alebo kombinácie poklesov a rastu nosných a nenosných častí výrobku. A to je len hmotnosť v rámci naturálnych aspektov inovovaného výrobku/ služby.

4. Miery inovácie

3.1 Jednoduchá miera inovácie

Miery inovácie predstavuje hodnota parametra inovácie na jednotku ekonomického výsledku v čase. Má zmysel ich merať v dvoch časových situáciách. Môžeme tiež použiť priestorové porovnanie dvoch priestorových celkov pri rovnakom výrobku. Nech tieto situácie označíme indexmi 0 a 1. Nech parameter ekonomie má hodnoty e_0 a e_1 a parameter inovácie predstavuje hmotnosť s hodnotami h_{mot0} a h_{mot1} .

Žiaduci smer inovačného vývoja, pri žiaducom vývoji hmotnosti vo forme poklesu, je pokles (chceme ľahšie výrobky) a v prípade, že žiaduci smer vývoja je rast (budeme chcieť vyššiu kilogramovú cenu).

$$\frac{H_{mot1}}{e1} \quad ? \quad \frac{h_{mot0}}{e0}$$

resp.

$$\frac{e1}{H_{mot1}} \quad ? \quad \frac{e0}{H_{mot0}}$$

$$\{H_{mot1}/e1\} > \{H_{mot0}/e0\}$$

Keď namiesto otáznika (?) dáme znak „väčšie“ ide o rastúci inovačný proces a v prípade, že dáme „menšie“ ide o neinovačný proces či presnejšie zhoršujúci sa inovačný proces. Posúdenie vývoja inovácie predstavuje vyhodnotenie relácie a žiaducich smerov vývoja jednotlivých ukazovateľov.

Príklad 1

Nech v roku 2010 a v roku 2012 hodnota tržieb dosiahla 240 € a potom sa znížila na 220 € na jeden výrobok. Inovatívnu vlastnosť sme hodnotili hmotnosťou výrobku, ktorá v roku 2010 bola 12 kg a v roku 2012 sa znížila na 10 kg. Nech v roku 2012 dosiahol iný výrobca 250 € tržieb na výrobok pri hmotnosti 11 kg. Žiaduci smer vývoja hmotnosti je pokles. Posúďte vývoj inovácie u prvého výrobcu a porovnajte inováciu v roku 2012.

$$\frac{10}{220} > \frac{12}{240} \quad \frac{10}{220} > \frac{11}{250}$$

V roku 2010 dosiahol výrobca 0,05 kg výrobku na 1 euro tržieb (12/240). V roku 2012 sa objem kg znížil na 0,045 kg výrobku na 1 euro tržieb (žaduci je pokles kilogramovej ceny, t.j. inovačný vývoj je pozitívny; v čase rade $(h_{mot1}/e1) < (h_{mot0}/e0)$

V roku 2012 dosiahol prvý výrobca 0,045 kg výrobku na 1 euro tržieb (10/220) a druhý výrobca o 1 gram nižšiu hodnotu rovnú 0,044 kg výrobku na 1 euro tržieb (11/250). U druhého výrobcu je proces inovácie o niečo lepší ako u prvého výrobcu $(h_{mot1}/e1) < (h_{mot2}/e2)$.

4.2 Jednoduchá miera inovácie ku komplexu ekonomických parametrov

Máme parameter, ktorého hodnoty vyjadrujú mieru inovácie a komplex parametrov vyjadrujúci ekonomickú podstatu tvorby príslušného výrobku. Úlohou je zostrojiť príslušnú skupinu mier inovácie. Riešením je zlomok, ktorý má v čitateli parameter inovácie a v menovateli môže mať:

- Pracovnú silu vyjadrenú v osobných nákladoch, alebo
- Agregovanú spotrebu výrobných zdrojov potrebných k výrobe parametra inovácie.

Pracovná sila je vyjadrená v osobách. V podmienkach trhovej ekonomiky sa zdá byť presnejšie pracovnú silu v osobách nahradiť nákladmi na pracovnú silu. Tak vo vzorci miery inovácie dostaneme **mieru inovácie meranú na spotrebu osobných nákladov**. Pracovná sila meraná počtom pracovníkov, počtom odpracovaných časových jednotiek (rokov) či osobnými nákladmi (ročnými) vyjadruje aktívnu zložku tvorby výstupu ekonomického procesu – produktu resp. služby a tak miera inovácie meraná na spotrebu osobných nákladov vyjadruje aktívnu zložku inovačného procesu. V prípade, že použijeme všetky zložky spotreby vo výrobnom procese plus zisk/stratu navyše dostávame mieru inovácie meranú celkovou spotrebou výrobných faktorov (MZD, ODP, MAT a ZISK).

Príklad 2

K skutočnostiam v príklade 1 vieme, že u prvého výrobcu sa spotreba výrobných faktorov skladala v roku 2010 z (MZD = 50 eur, ODP = 30 eur, MAT = 120 eur a ZISK = 40 eur) a v roku 2012 z (MZD = 55 eur, ODP = 35 eur, MAT = 110 eur a ZISK = 50 eur) a u druhého výrobcu v roku 2012 z (MZD = 55 eur, ODP = 30 eur, MAT = 125 eur a ZISK = 40 eur) na jeden výrobok. Vypočítajme mieru inovácie ku osobným nákladom a k tržbám.

Uvedené údaje sú zoradené v priloženej tabuľka . Stĺpce 2010 a 2012 sú venované výsledkom prvého výrobcu a posledný stĺpec je venovaný druhému výrobcovi. V riadkoch tabuľky máme: hmot – hmotnosť jedného výrobku, MZD – osobné náklady na jeden výrobok, ODP – odpisy na jeden výrobok, MAT – ostatné náklady na jeden výrobok, ZISK – zisk/strata na jeden výrobok,

	2010	2012	2012
hmot	12	10	11
MZD	50	45	55
ODP	30	30	35
MAT	120	115	110
ZISK	40	30	50
TRZBY	240	220	250
miMZD	0,240	0,222	0,200
miTRZBY	0,050	0,045	0,044

TRŽBY – objem tržieb na 1 výrobok, miMZD – miera inovácie z pohľadu osobných nákladov a miTRŽBY – miera inovácie z pohľadu celkových tržieb.

Pri podrobnejšom čítaní tabuľky môžeme konštatovať zhodu výsledkov prvého príkladu (riadok miTRŽBY je zhodný z hodnotami miery inovácie v Príklade 1. Novým je porovnanie hmotnosti k osobným nákladom na jednotku výroby. Vidíme zlepšenie miery inovácie v roku 2012 oproti roku 2010 (pokles hodnoty miMZD v roku 2012 oproti roku 2010 pri prvom výrobcovi. Porovnanie roku 2012 prvého a druhého výrobcu ide, podobne ako pri tržbách, je mierne lepšie u druhého výrobcu (hmotnosť u druhého výrobcu je menšia).

Keďže čitateľ miery inovácie je v prvej aj druhej skupine rovnaký a menovateľ pri ukazovateľoch druhej skupiny predstavuje rovnicu $TRZBY = MZD + ODP + MAT + ZISK$ je rozdiel medzi prvou skupinou a druhou skupinou mier inovácie prakticky len formálny.

3.3 Zložená miera inovácie ku komplexu faktorov inovácie

Tretiu verziu predstavuje jav, keď inováciu vyjadrujeme viacerým parametrami. Nech syntetickú vlastnosť „inovácia“ vyjadruje m parametrov p_1, p_2, \dots, p_m . Nech žiaduci smer vývoja všetkých parametrov je zhodný zo žiaducim smerom vývoja použitej miery inovácie.

Menovateľ miery inflácie nech predstavujú tržby za daný inováčný produkt.

Problémom môže byť agregácia obsahu čitateľa do jedného čísla. V prípade, že čiastkové ukazovatele inovácie sú v meraciach jednotkách prakticky z jedného rozdelenia hodnôt situácia je jednoduchá – syntetický ukazovateľ, ktorý dosadíme do čitateľa je súčtom resp. priemerom potenciálnych čiastkových ukazovateľov vyjadrujúcich jednotlivé aspekty inovácie.

V prípade, že ukazovatele sú rôzne, môžeme použiť ich transformovať na rovnaké expertným odhadom, transformáciou. Jej podstata spočíva v transformácii pôvodných hodnôt na hodnoty z intervalu od -2 po +2 zo stredom v nule. Potenciálna najhoršia hodnota v pôvodnej stupnici dostane hodnotu -2, potenciálne najlepšia hodnota +2 a stredná hodnota hodnotu 0. Hodnoty medzi -2 a 0 resp. 0 až +2 sú priradené podľa hodnôt pôvodnej premennej expertom.

Výsledný agregát dostaneme spočítaním resp. priemerovaním týchto transformovaných hodnôt.

Vlastné hodnotenie miery inovácie je potom jednoduché. Máme experta. Máme pôvodné hodnoty premenných p_1, p_2, \dots, p_m . Máme hodnotu menovateľa miery inovácie. Expert na základe pôvodných hodnôt určí pre každý ukazovateľ p_i čitateľa transformovanú hodnotu t_i tohto ukazovateľa ($p_i \rightarrow t_i$). Transformované hodnoty potom agregujeme úhrnom alebo priemerom a vypočítame transformovanú hodnotu miery inovácie.

Záver

Príspevok poskytuje návod ako merať inovácie. V reálnych aplikáciách je problém v ochote použiť resp. poskytovať potrebné údaje.

Literatúra:

CHAJDIK, J.: Viackriteriálne hodnotenie – meranie inovácie. In. Forum Statisticum Slovacum. 5/2014, str.

Adresa autora:

doc. Ing. Jozef Chajdiak, CSc.
Ústav manažmentu STU
Vazovova 5, Bratislava
chajdiak@statis.biz

Príspevok bol vypracovaný v rámci riešenia úlohy VEGA č. 1/1164/12: "Možnosti uplatnenia informačných a komunikačných technológií na zvyšovanie efektívnosti medzinárodnej spolupráce malých a stredných podnikov SR v oblasti inovácií" a úlohy VEGA č. 1/0335/13: "Štatistická analýza vybraných ukazovateľov konkurencieschopnosti na súbore podvojne účtujúcich podnikov SR.

Rozbor súboru vinárskych podnikov pomocou všeobecného lineárneho regresného modelu

Analysis of a wine companies file using GLM

Jozef Chajdiak, Lucia Coskun

Abstract: The publication describes an analysis of wine companies using general linear model method. GLM represents a basic application of regression models with the widest application. Estimation method of model data is the method of least squares. Reliability of data estimation can be determined by R^2 value. The method has been described using excel. Processing has been exemplified on the wine companies file from 2013.

Abstrakt: Príspevok obsahuje opis postupu analýzy metódou všeobecného lineárneho regresného modelu na príklade rozboru súboru vinárskych podnikov. Všeobecný lineárny model predstavuje základnú aplikáciu regresných modelov, ktorá má najširšie použitie. Metódou odhadu parametrov modelu je metóda najmenších štvorcov. Spoľahlivosť odhadu parametrov môžeme určiť hodnotou R^2 . Je popísaný spôsob použitia GLM v exceli. Spracovanie je demonštrované na súbore vinárskych podnikov z roku 2013.

Key words: regression model, R^2 , wine companies, sales, assets

Kľúčové slová: regresný model, R^2 , vinárske podniky, tržby, majetok

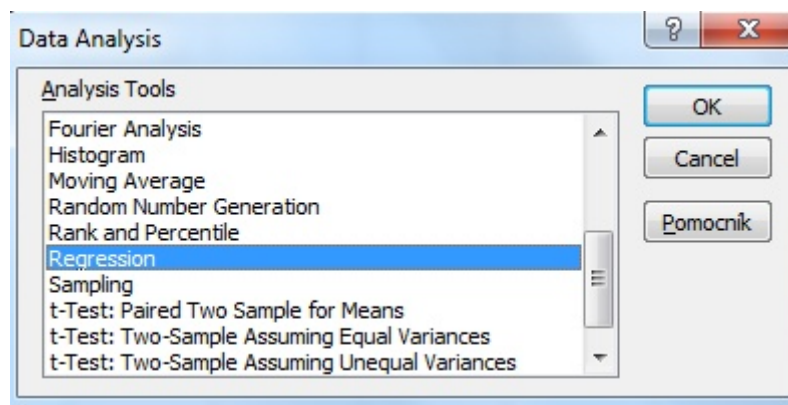
JEL classification: C1, C4

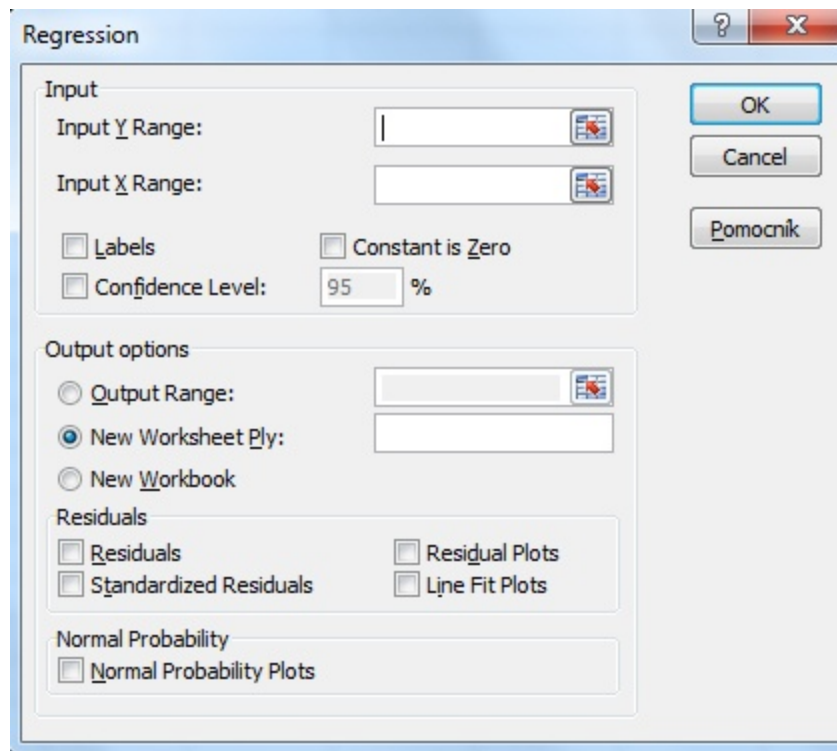
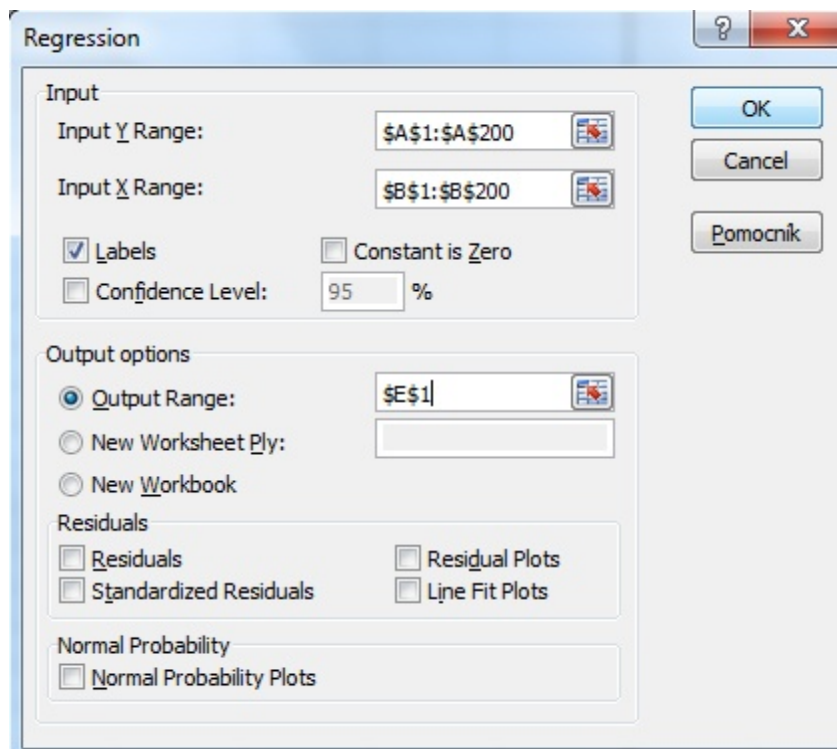
1. Úvod

Všeobecný lineárny regresný model má tvar $y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \dots + a_m \cdot x_m + e$, kde x_1, x_2, \dots, x_m sú hodnoty nezávislých premenných, y je hodnota závisle premennej, a_1, a_2, \dots, a_m sú parametre lineárneho regresného modelu a e je chyba odhadu pomocou tohto modelu. Zvyčajne sa k odhadu parametrov používa metóda najmenších štvorcov. Jej odhadom hodnôt závisle premennej je $\hat{y} = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \dots + a_m \cdot x_m$.

2. Postup riešenia v exceli

Skontrolujeme nastavenie použitia nástrojov v exceli. Po prvom nastavení (musíme ho urobiť) sú nástroje v exceli trvale prítomné. V prvom slpci sú uložené hodnoty pozorovaní závisle premennej y a v ďalších slpciach hodnoty nezávisle premenných. V ústrednej ponuke aktivujeme tlačítko údaje a v jeho rámci zvolíme tlačítko Data Analysis. V ponuke Data Analysis (Obr. 1) zvolíme Regression. Objaví sa okno Regression (Obr. 2).



Obr. 11: Ponuka okna nástrojov*Obr.2: Ponuka okna Regression**Obr. 3: Vyplnené okno Regression (závislosť obratu od majetku)*

Všeobecný lineárny regresný model umožňuje riešiť úlohy typu priamka alebo aj úlohy v ktorých pridávame ďalšie premenné (priamka, parabola, 3- rozmerný priestor...m-rozmerný priestor). Pri priamke ($\hat{y} = a_0 + a_1x_1$) v okne Regression zvolíme nasledovné: v políčku Input Y Range zadáme pole s hodnotami závisle premennej v prvom riadku s názvom stĺpca a v druhom až m+1 riadku s hodnotami jednotlivých podnikov. Hodnoty je žiaduce usporiadať. Riadky s chýbajúcimi hodnotami (prázdne políčka) treba vynechať. Riadky s políčkami s nulovými hodnotami je podľa charakteru úloh možné aj ponechať. V našom prípade sú to hodnoty A1 až A194. V políčkach B1 až B194 sú hodnoty premennej MAJ (majetok), t.j. v prvom riadku názov premennej MAJ a v 2. Až 194. riadku hodnoty premennej majetok (vynechali sme hodnoty v riadkoch 195 až 200 pri ktorých chýba hodnota osobných nákladov). Políčko Labels odľakujeme (hodnoty premenných začínajú názvami premenných). V druhej polovici vyberieme kde chceme umietniť výstup buď Output Range (na tento hárok) do špecifikovaného poľa – v našom prípade začína ľavým horným rohom v políčku E1; alebo New Worksheet Ply (do nového hárku). Výstup pokračuje uložením reziduí – možnosti môžeme odľakovať: Residuals – reziduá, Standardized residuals – štandardizované reziduá a grafy: Residual Plots - graf reziduí; Line Fit Plots – grafické znázornenie vypočítaných hodnôt závisle premennej a Normal Probability Plots – normálny pravdepodobnostný graf.

E	F	G	H	I	J
SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0,8012				
R Square	0,6419				
Adjusted R Square	0,6400				
Standard Error	3527589				
Observations	193				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>gnificance</i>
Regression	1	4,26029E+15	4,26E+15	342,3605	1,84E-44
Residual	191	2,37678E+15	1,24E+13		
Total	192	6,63708E+15			
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>
Intercept	-55720,9468	265511,9131	-0,20986	0,833999	-579433
MAJ	0,75069813	0,040571729	18,50299	1,84E-44	0,670672

Obr. 4: Výstup výpočtov regresnej priamky (obrat od majetku)

V konkrétnej realizácii sme ako závislú premennú použili obrat a ako nezávisle premenné v prípade priamky použili majetok. Usporiadané hodnoty podľa obratu a majetku sú v políčkach A1 až C200. Posledných 6 riadkov obsahuje chýbajúce hodnoty a preto ich vynecháme. Zo zvyšných hodnôt po stlačení tlačidla OK dostaneme výstup (obr. 4). Na obr. 4 nás zaujíma bodový odhad regresnej funkcie a podiel vysvetlenej variability R^2 . Bodový odhad sa nachádza v ľavom spodnom rohu výstupnej zostavy. V našom prípade je to:

$$\text{ObratHAT} = -55721 + 0,75 * \text{MAJ} \quad (1)$$

Parameter a_0 je rovný -55721 (t.j. nulovému majetku zodpovedá 55721 obratu) a parameter a_1 je rovný 0,75 (t.j. zväčšenie majetku o 1 znamená zvýšenie obratu o 0,75 jednotiek majetku).

Podiel vysvetlenej variability predstavuje $R^2 = 0,64$ (t.j. modelom je vysvetlených 64% variability obratu).

Poznámka: Za zamyslenie stojí fakt, že pri F teste časti analýzy rozptylu napríklad hodnote $R^2 = 0,06$ môže byť konštatovaná štatisticky významná závislosť hoci 6% - ný podiel vysvetlenej variability je prakticky bezcenný.

Priamka predstavuje najjednoduchšiu rovnicu všeobecného regresného modelu. V prípade, že pridáme priamke ďalšiu premennú (t.j. dostaneme rovinu) sa situácia len trochu zmení. K regresnej priamke $\text{Obrat} = f(\text{Maj})$ pridáme osobné náklady ($\text{Obrat} = f(\text{Maj}, \text{ON})$), konkrétne

$$\hat{y} = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 \quad (2)$$

$$\text{Obrat} = a_0 + a_1 \cdot \text{Maj} + a_2 \cdot \text{ON} \quad (3)$$

Vo vstupných údajoch okna Regression zmeníme hodnoty Input Y Range: na A1:A122 a hodnoty Input X Range na B1:C122. V poli Output Range zadáme E21. Po stlačení OK dostávame nasledujúci výstup:

SUMMARY OUTPUT					
<i>Regression Statistics</i>					
Multiple R	0,966033455				
R Square	0,933220637				
Adjusted R Square	0,932088783				
Standard Error	1907144,705				
Observations	121				
<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>gnificance</i>
Regression	2	5,99779E+15	3E+15	824,5065	4,51E-70
Residual	118	4,2919E+14	3,64E+12		
Total	120	6,42698E+15			
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>
Intercept	-393375,805	185904,4173	-2,11601	0,036446	-761517
MAJ	0,247636425	0,031379363	7,891697	1,7E-12	0,185497
ON	7,575220758	0,329298582	23,00411	1,99E-45	6,92312

Obr. 5: Výstup výpočtov regresnej paraboly (obrat od majetku a osobých nákladov)

Bodový odhad regresnej roviny $\text{Obrat} = a_0 + a_1 \cdot \text{Maj} + a_2 \cdot \text{ON}$ je:

$$\text{Obrat} = -393376 + 0,24676 \cdot \text{Maj} + 7,575 \cdot \text{ON} \quad (4)$$

Podiel vysvetlenej variability touto rovnicou je $R^2 = 0,93$. Parametre regresnej rovnice znamenajú, že ak sa majetok zmení o 1, tak obrat zmení o 0, 24676 jednotiek majetku a ak sa objem osobných nákladov zmení o 1, tak sa obrat zmení o 7, 575 jednotiek majetku.

Veľmi vysoká hodnota R^2 závisí jednak od toho faktu, že sme rozšírili priamku na parabolu (máme viac premenných, t.j. aj vyšší podiel vysvetlenej variability) a druhým podstatným faktorom je aj fakt zníženia počtu pozorovaní o chýbajúce hodnoty (kým pri priamke sme mali zníženie z 200 na 154 pri regresnej rovine je to z 200 na 122 pozorovaní).

3. Záver

Výsledky výpočtov ukazujú, že použitie regresnej roviny (obrat závisí od majetku a osobných nákladov) je presnejšie ako použitie regresnej priamky (obrat závisí od majetku). Kým pri regresnej priamke je $R^2=0,64$, tak pri regresnej rovine je až $R^2=0,93$. Čo hovorí v prospech použitia modelu regresnej paraboly.

Literatúra

CHAJDIK, J. 2013. *Štatistika jednoducho v Exceli*. Bratislava Statis. ISBN 978-80-85659-74-0.

CHAJDIK, J. 2010. *Štatistika jednoducho*. 3. Vydanie. Bratislava Statis. ISBN 978-80-85659-60-3.

COSKUN, L., 2013, *Vypracovanie štatistickej charakteristiky súboru vinárskych podnikov v SR v roku 2010 v exceli*. FORUM STATISTICUM SLOVACUM. Roč. IX., 07/2013, s.32-37, ISSN 1336-7420.

COSKUN, L., 2014, *Analýza a prognózovanie vývoja vinárskych podnikov*. FORUM STATISTICUM SLOVACUM. Roč. X., 03/2014, s.18-22, ISSN 1336-7420.

Príspevok bol napísaný v rámci riešenia úlohy VEGA č. 1/ 0335/13 Štatistická analýza vybraných ukazovateľov konkurencieschopnosti na súbore podvojne účtujúcich podnikov SR.

Adresy autorov:

Doc. Ing. Jozef Chajdiak, CSc.
Ústav manažmentu
Vazovova 5, 812 43 Bratislava
chajdiak@statis.biz

Ing. Lucia Coskun
Ústav manažmentu, externá doktorandka
Vazovova 5, 812 43 Bratislava
luciacoskun@stuba.sk

Vzt'ah investícií a hrubej pridanej hodnoty v IKT sektore

Relation between investment and gross value added in ICT sector.

Štefan Kováč

Abstract: Information and communication technologies (ICT) have become the important drivers for socioeconomic development in recent years. This article presents an empirical analysis of gross value added (GVA) levels determinants in ICT sector in Slovakia. The analysis is based on panel data covering the period 2001–2010 and using regional data at NUTS3 level. Our results suggest that the level of regional GVA in the ICT sector is positively affected by the level of investment (which are proxied by the amount of loans) in the ICT sector. A 1% increase in the level of investment into ICT leads to increase in the level of regional GVA in the ICT sector by 0,42%.

Abstrakt: Informačno-komunikačné technológie (IKT) sa stali dôležitou súčasťou socioekonomického rozvoja. Príspevok sa venuje analýze determinantov hrubej pridanej hodnoty (HPH) tvorenej v IKT sektore na Slovensku. Analýza je postavená na panelových údajoch z obdobia rokov 2001-2010 na úroveň krajov (NUTS3). Výsledky analýzy ukazujú, že investície v sektore IKT majú pozitívny dopad na celkovú hrubú pridanú hodnotu v odvetví – 1% zvýšenie úrovne investícií v podnikoch podnikajúcich v sektore IKT vedie k zvýšeniu HPH tohto sektora o 0,42%.

Key words: gross value added, investments, loans, ICT, panel data

Kľúčové slová: hrubá pridaná hodnota, investície, úvery, IKT, panelové údaje, informačno-komunikačné technológie

JEL classification: C51, L20

1. Úvod

Ekonomika krajiny je obrovský „organizmus“, ktorý pozostáva z množstva menších súčastí, ktoré môžu byť navzájom funkčne alebo inak prepojené. Pre pochopenie a skúmanie ekonomiky ako celku je potrebné pochopiť a skúmať aj jej jednotlivé súčasti. Existujú rôzne determinanty, ktoré ovplyvňujú jej fungovanie, rast či spôsobujú recesiu. V poslednej dobe sa hovorí o dôležitej úlohe informačno-komunikačných technológií v štruktúre hospodárstva krajiny ako aj o vplyve úrovne zavedenia IKT v krajine a hospodárstve na ekonomický rast (Kim, Kang, Sanders, Lee, 2008; Welfens, Perret, 2013; Khuong, 2013; Miygawa, Yukito, Nobuyuki, 2004). Viacerí autori preukázali pozitívnu závislosť, resp. súvis medzi ekonomickým rastom krajiny a úrovňou zavedenia informačno-komunikačných technológií. IKT sektor funguje v hospodárstve ako samostatné odvetvie, zároveň však IKT tvorí výraznú podporu a súčasť takmer všetkých ostatných odvetví a verejného sektora. Rastom inovácií a odvetvia IKT, rastie aj jeho prepojenosť a podpora v ostatných odvetviach, ktoré by sa už dnes nezaobíšli bez podporných informačných technológií. Jedným z komplexných indexov pre porovnanie sektorov IKT v jednotlivých krajinách je *The Networked Readiness Index* publikovaný inštitúciou *The World Economic Forum*. Slovensko sa v rebríčku krajín podľa tohto indexu nachádza na 59. mieste, pričom ostatné krajiny V4 - Česká republika, Maďarsko, Poľsko sa nachádzajú na 42. 47. a 49. mieste. V rámci tohto indexu existuje niekoľko subindexov. V subindexe *Government Usage* sme až na 106. mieste – pre porovnanie Juhoafrická republiky je na 103. mieste (*The Networked Readiness Index Report 2014*). Zatiaľ čo autori v uvedených štúdiách skúmali vplyv úrovne IKT na ekonomický rast, my sa v našom príspevku zameriame na samotné odvetvie hospodárstva, reprezentujúce IKT, teda odvetvie J podľa štatistickej klasifikácie ekonomických činností SK-NACE. Ide o podniky,

ktoré „tvoria IKT“ v chápaní vyššie spomenutých autorov a ich výskumov (teda tvoria konkrétne informačno technologické výstupy). IKT odvetvie sa na Slovensku veľmi rýchlo rozvíja, o čom svedčí neustály rast firiem podnikajúcich v tomto odvetví (v roku 2001 to bolo 1091 firiem, v roku 2010 to je už 5845 firiem).

2. Hypotéza

V príspevku sa zameriavame na súvis medzi investíciami vo firmách podnikajúcich v odvetví IKT a rastom tohto odvetvia prostredníctvom ukazovateľa HPH na regionálnej úrovni. Základnú hypotézu si môžeme stanoviť nasledovne:

H: Existuje preukázateľný súvis medzi úrovňou regionálnej hrubej pridanej hodnoty v odvetví IKT a investíciami v podnikoch IKT odvetvia

Skúmaním stanovenej hypotézy uskutočňujeme ďalší krok v reťazci skúmania závislosti medzi úrovňou využívania informačno-komunikačných technológií v hospodárstve a ekonomickým rastom krajiny. Nakoľko podľa uvedených štúdií bol preukázaný súvis medzi rastom úrovne IKT v hospodárstve a ekonomickým rastom, my predpokladáme, že množstvo IKT v hospodárstve súvisí s množstvom tvorby pridanej hodnoty v IKT podnikoch. Ak tieto podniky produkujú a vyrábajú, tak musia vyrábať pre niekoho. Istá časť týchto služieb ide na export, ale časť ostáva v domácom hospodárstve, v ktorom sa musí prejaviť. Takže rast produkcie, resp. hrubej pridanej hodnoty v odvetví IKT sa sekundárne prejavuje aj v raste úrovne IKT v slovenskom hospodárstve a následne podľa zistených záverov aj v ekonomickom raste krajiny.

V príspevku nás bude zaujímať vplyv investícií v podnikoch na celkový output odvetvia resp. na celkovú hrubú pridanú hodnotu odvetvia. Investície sú vo všeobecnosti považované za katalyzátor rastu a rozvoja. Investície budú reprezentované rôznymi pomerovými ukazovateľmi odvetvia, ktoré s investíciami úzko súvisia a sú dobre interpretovateľné.

3. Model

V príspevku budeme uvažovať sa nasledovne špecifikovaným modeloml:

$$\log(GVA) = \alpha + \beta_1 \log(LOA) + \beta_2 RAT1 + \beta_3 RAT2 + \beta_4 RAT3 + \beta_5 RAT4 + u_t$$

- *HPH – hrubá pridaná hodnota*
- *LOA – úroveň úverov*
- *RAT1 – pomer mzdových nákladov a tržieb*
- *RAT2 – pomer odpisov a celkových aktív*
- *RAT3 – pomer úverov a neobežného majetku*
- *RAT4 – pomer úverov a pridanej hodnoty*

Na odhad modelu je použitá štandardná metodika odhadu regresných modelov pre panelové údaje. Na odhady modelov je použitý štandardný balík plm (Croissant, Millo, 2008) v softvéri R (R Core Team, 2014).

Odhadnutý model musí byť podrobený testom, či spĺňa apriórne teoretické predpoklady o modeli, teda či sa jedná o model fixných efektov, náhodných efektov s významnými časovými alebo individuálnymi efektmi. Zároveň je nutné analyzovať či vyhovuje všetkým predpokladom a štandardom, ktoré sú kladené na takýto typ modelu. Jedná sa o testovanie tzv. „poolability“, testovanie významnosti časových, individuálnych, poprípade oboch typov efektov, čo sa uskutočňuje prostredníctvom testov Breuscha a Pagana, Hondu. Testovanie prierezovej závislosti Pesaranovým testom prierezovej závislosti v panelových dátach, poradovej korelácie Breusch-Godfreydov, Wooldridgeov test. Test stacionarity Maddala-Wu testom jednotkového koreňa pre panelové dáta a heteroskedasticity Breusch-Paganovým

testom.(Baltagi, 2005). Test multikolinearity zahŕňa použitie VIF – Variance Inflation Factor, ktorá vyjadruje nárast variability regresných koeficientov vplyvom multikolinearity (Fox, Weisberg, 2011).

V mnohých prípadoch je poradová korelácia, heteroskedasticita ako aj prierezová závislosť prítomná v ekonometrických modeloch, čo spôsobuje skreslený odhad štandardných odchýlok odhadnutých koeficientov, a to nás môže do viesť k nesprávnym výsledkom a záverom týkajúcich sa významnosti regresných koeficientov. Z týchto dôvodov sa odporúča využiť robustnú variančno-kovariančnú maticu (Croissant, Millo, 2008)

4. Údaje a premenné v modeli

Údaje sú čerpané z dvoch zdrojov. Prvým je Eurostat, odkiaľ sme získali údaje za HPH pre jednotlivé regióny a odvetvie J podľa štatistickej klasifikácie ekonomických činností SK-NACE (Informácie a komunikácia).

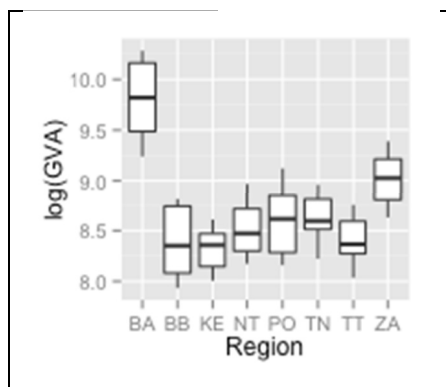
Druhým zdrojom sú anonymizované údaje z finančných výkazov podnikov podnikajúcich v tomto odvetví poskytnuté Finančnou správou SR na výskumné účely.

Používané dáta zahŕňajú sledované obdobie 2001-2010 a sú klasifikované regionálne na úroveň NUTS3.

Hrubá pridaná hodnota (vysvetľovaná premenná)

Ako ukazovateľ rastu ekonomiky sme si zvolili HPH. Hrubú pridanú hodnotu môžeme definovať ako „tú časť produkcie, ktorú novo vytvorila jednotka svojou činnosťou. Ukazovateľ zahŕňa hodnotu hrubej produkcie (hrubého obratu) zmenšenú o medzispotrebu za predpokladu, že oba ukazovatele sú ocenené v rovnakých cenách.

Obrázok č.1 ukazuje rozdiely v regionálnej tvorbe HPH za posledné uvažované obdobie (2010).



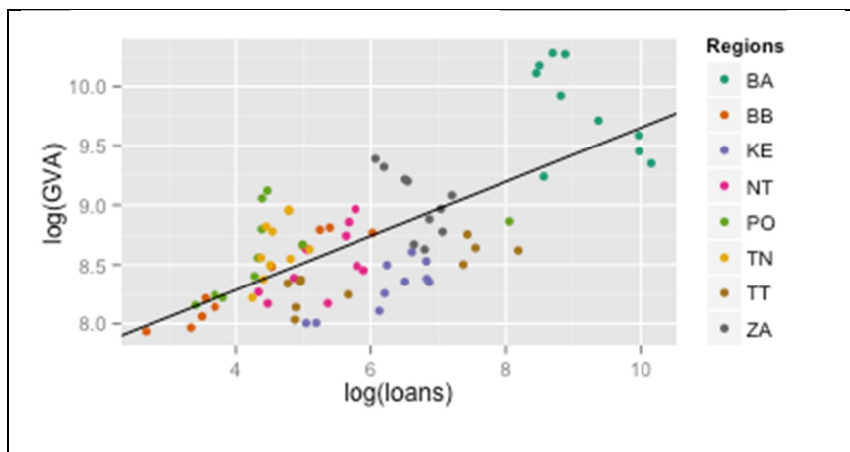
Obr. č.1: Rozdiely v regionálnej tvorbe HPH v roku 2010

S ohľadom na vysokú variabilitu a vysoké hodnoty tejto premennej budeme v ďalšej analýze používať logaritmované hodnoty premennej.

Úroveň úverov

Prvou vysvetľujúcou premennou v našom modeli je agregovaná hodnota úverov (krátkodobé + dlhodobé + fin. výpomoci), nakoľko úvery úzko súvisia s investíciami, resp. s financovaním investícií v podnikoch a dajú sa používať ako miera cudzieho kapitálu v podnikaní.

Na obrázku č.2 je zobrazená závislosť medzi vysvetľovanou premennou – HPH a premennou úroveň úverov - v jednotlivých rokoch pozorovania.



Obr. 2: Závislosť medzi vysvetľovanou premennou a premennou "úvery" v sledovanom období

Pomer odpisov a aktív celkom

Odpisy súvisia s úrovňou investícií v podnikoch, nakoľko do odpisov vstupuje každý hmotný a nehmotný majetok, ktorý má hodnotu vyššiu ako 1700€ (v sledovanom období sa táto hodnota menila len raz – v roku 2004 sa zvýšila z vtedajších 996€ na 1700€) a zároveň dobu použiteľnosti dlhšiu ako 1 rok. Aktíva celkom predstavujú hodnotu nárastu majetku, čo svedčí o „investíciách“ – ide o majetok hmotný aj nehmotný vo forme softvérov licencií, ale aj hnutelný a nehmotný – autá, budovy a podobne. Zmeny tohto pomeru predstavujú zmeny v úrovni investícií v podnikoch.

Pomer osobných nákladov a tržieb

Tento pomer predstavuje mzdovú náročnosť prislúchajúcu hrubým tržbám. Vo všeobecnosti platí, že ak hodnota tohto pomeru rastie, tak pripadajú do úvahy dve možnosti:

- 1.) zamestnávanie expertov, ktorých hodnota na pracovnom trhu je vysoká (vysoká mzda),
- 2.) zvyšovanie počtu zamestnancov – vyššia úroveň a kvalita poskytovaných služieb

Pomer úverov a neobežného majetku

Pomerový ukazovateľ dáva do pomeru hodnotu úverov s hodnotou neobežného majetku (ktorý predstavuje hmotné a nehmotné investície v podniku) a hovorí o miere financovania tohto majetku prostredníctvom úverov, resp. o miere využitia úverov na iné výdavky (napr. hradenie prevádzkových nákladov).

Pomer úverov a pridanej hodnoty

Hodnota tohto pomeru predstavuje rentabilitu úverov, resp. ich návratnosť v podobe pridanej hodnoty.

5. Výsledky a diskusia

Odhad modelu použitím robustnej variančno-kovariančnej matice z dôvodu porušenia predpokladov je uvedený v tabuľke 1.

Tab.1: Odhad modelu

Premenné	Odhad	Std. Errr	t-value	Pr(> t)	
log(LOA)	0,4108	0,0369	11,1254	< 0,00001	***
RAT1	1,2787	0,4107	3,1136	0,0027	**
RAT2	-5,4496	1,1912	-4,5749	< 0,00001	***
RAT3	-0,9560	0,1299	-7,3612	< 0,00001	***
RAT4	-0,4000	0,0718	-5,5698	< 0,00001	***

Tab. 2: Diagnostické testy modelu

	statistic	p-value
Individuálne a časové efekty		
F-test	2,0774	1.351e-15
LM Test	normal = 9,4916	<2,2e-16
Poradová korelácia		
Wooldridgeov test	$z = 1.4287$	0,1531
Breusch-Godfrey/Wooldridge test	$chisq = 45.6263$	1.676e-06
Wooldridgeov test	$chisq = 24.2432$	8.491e-07
Prierezová korelácia		
Pesaran CD test	$z = 2.951$	0.003167
Stacionarita		
Maddala-Wu Test		< 2.2e-16
Heteroscedasticita		
Breusch-Pagan test	20.4299	0.001038

F-statistic: F=26,1068, **p-hodnota** =1,57e-14 **R²** = 0,66082

Úroveň multikolinearity je v norme – hodnoty VIF pre jednotlivé premenné sa nachádzajú v intervale (1,38-1,65). (Belsley et al., 2004).

Výsledky odhadnutého regresného modelu podporujú predpoklad o existencii preukázateľného súvisu medzi úrovňou regionálnej hrubej pridanej hodnoty v odvetví IKT a investíciami v podnikoch IKT odvetvia. K verifikácii stanovenej hypotézy tiež prispieva koeficient determinácie na úrovni 0,66, ktorý vypovedá o veľmi kvalitnom modeli nakoľko sa jedná o panelové údaje, za dostatočujúce sa v tomto prípade považujú hodnoty >0,30. Hodnota regresného koeficientu pri premennej log(LOA) 0,41 svedčí o pozitívnom vplyve tejto vysvetľujúcej premennej na hodnotu HPH, čo je možné interpretovať, že 1% zmena v úrovni úverov v odvetví IKT spôsobí pozitívnu zmenu hodnoty HPH a to konkrétne o 0,41%. Investície resp. výška úverov v tomto odvetví má pozitívny dopad na hrubú pridanú hodnotu v odvetví.

Ďalším zaujímavým poznatkom sú zistenia súvisiace s pomerovým ukazovateľom RAT1, ktorý predstavuje pomer osobných nákladov k tržbám. Ako bolo už bolo uvedené – zvyšovanie hodnôt tohto pomeru predstavuje zvyšovanie miezd a zamestnanosti v tomto sektore. Toto zvyšovanie ma na základe nášho skúmania pozitívny dopad na tvorbu HPH, a to konkrétne zvýšenie pomeru o 1 % zvýši tvorbu HPH o 1,27%. Rast tohto pomeru môže byť však dosiahnutý nie len zvyšovaním mzdovej náročnosti, resp. zvyšovaním miezd alebo rozširovaním pracovných pozícií. Tento pomer môže rásť aj v prípade, ak sa znižujú tržby a úroveň osobných nákladov ostáva zachovaná. Pozitívnu informáciou takejto zmeny môže byť stabilita a garancia pracovných miest aj počas recesie.

Pomer RAT3 hovorí o zaúverovanosti firiem, resp. o pomerovom ukazovateli, ktorý predstavuje časť úverov investovaných do neobežného majetku (hmotné a nehmotné aktíva – reálne investície napr. licencie, softvéry, nehnuteľnosti a pod.). Ak tento pomer vzrastie o 1%, HPH sa zníži o 0,95%. Tento pomer rastie, ak rastie úroveň úverov a zároveň sa hodnota neobežného majetku nemení. Teda firma čerpá úver, nevyužíva ho na investície, ale na iné účely (napr. prevádzkové náklady, platenie daní a pod.). To je zlý signál – tiež sa to negatívne prejavuje na hodnote vysvetľovanej premennej HPH.

6. Záver

IKT odvetvie je postavené na kvalitnej a kvalifikovanej pracovnej sile. Zvyšovanie mzdových nákladov na zamestnancov má tiež pozitívny dopad na tvorbu hrubej pridanej hodnoty. Je to jeden z indikátorov, ktorý svedčí o dobrom hospodárení podnikov.

Zistenia okrem verifikovania stanovenej hypotézy preukázali pozitívny dopad investícií na tvorbu HPH, pričom zvýšenie úrovne úverov vo firmách o 1% spôsobí 0,46% zvýšenie hodnoty HPH. Avšak je potrebné pozerat' sa na model ako na celok – tu sa okrem iného dá vyčítať, že tieto úvery musia byť použité na investície do firmy a nesmú byť použité na hradenie prevádzkových nákladov spoločnosti, ináč je efekt opačný.

Literatúra

- BALTAGI, B. H. 2005: *Econometric Analysis of Panel Data*. 3rd ed. Chichester: John Wiley and Sons,.
- BELSLEY D.A., KUH E., WELSCH R.E. 2004: *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, John Wiley and Sons.
- CROISSANT Y., MILLO, G. 2008: Panel Data Econometrics in R: The plm Package. *Journal of Statistical Software*, roč. 27, č. 2, s. 1-51.
- FOX, J. and WEISBERG, S. 2011: *An R Companion to Applied Regression, Second Edition*, Sage.
- KHUONG M. V. 2013: Information and Communication Technology (ICT) and Singapore's economic growth. In: *Information Economics and Policy*, roč. 25, s. 284–300.
- KIM, J.Y., KANG, H., SANDERS, G.L., LEE, S-Y.T. 2008: Differential effects of IT investments: Complementarity and effect of GDP level, In: *International Journal of Information Management*, roč. 28, s. 508–516.
- MIYAGAWA, T., YUKITO, I., NOBUYUKI, H. 2004: The IT revolution and productivity growth in Japan. In: *J. Japanese Int. Economies*, roč. 18, s. 362–389.
- R CORE TEAM 2014: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- WELFENS, P.J.J., PERRET, J.K. 2014: Information & communication technology and true real GDP: economic analysis and findings for selected countries. In: *Int Econ Econ Policy*, roč. 11, s.5-27

Adresa autora:

Štefan Kováč, Ing.
Ekonomická fakulta, TUKE
Nemcovej 32, Košice
stefan.kovac@tuke.sk

Hodnocení úrovně kompetencí získaných studiem

Evaluation of the level of competence gained by studying

Jana Kubanová, Michal Kuban

Abstract: The article is devoted to the problems of competence of the university graduates. Competency is viewed here from three perspectives. The first aspect is the evaluation, how well are university graduates prepared for a profession. The second aspect is the evaluation of their own competencies, the third aspect are the requirements for the competence of graduate by employers.

Abstrakt: Článek je věnován problematice kompetencí absolventů vysoké školy. Na kompetence je zde pohlíženo ze tří hledisek. Prvním hlediskem je hodnocení, jak dobře absolventy připravila vysoké škola na výkon povolání, druhým hlediskem je hodnocení vlastních kompetencí, třetím hlediskem jsou požadavky zaměstnavatelů na kompetence absolventa.

Key words: competence, capabilities, skills, exercise of a profession preparation

Klíčové slová: kompetence, schopnosti, dovednosti, příprava na výkon povolání

JEL classification: I2

1. Úvod

Často diskutovaným problémem je cíl vzdělávacího procesu a s tím související otázka, jak dobře či méně dobře připravují české školy své absolventy. Cíl vzdělávacího procesu je někdy zúžen pouze na osvojení poznatků a dovedností. Souvisejícím požadavkem je i vytváření způsobilostí potřebných pro život nebo výkon povolání. Kromě osvojení poznatků musíme uvažovat i postoje a hodnotové orientace, které jsou předpokladem výkonnosti absolventa ve vymezené činnosti. Běžně užívaným pojmem je v této souvislosti kompetence, čímž je míněna schopnost nabyté znalosti a vědomosti používat, tedy se jedná o dovednosti, které jsou rozhodující pro profesní či společenské uplatnění absolventa. Cíl vzdělávacího procesu pak může být vymezen jako vybavení studentů souborem klíčových kompetencí na úrovni, která je pro ně dosažitelná, a připravit je tak na uplatnění ve společnosti.

Kompetence absolventa jsou běžně součástí informací o studijním programu, resp. oboru. V tomto článku je podán pohled opačný, to je pohled ze strany absolventa vysoké školy na to, jak byl po studiu vybaven kompetencemi, jak byl připraven na výkon povolání z hlediska znalostí, dovedností, schopností. Rovněž absolventi odhadovali svoji úroveň znalostí, dovedností a schopností. S nároky na úroveň absolventů úzce souvisí i otázka požadavků současných zaměstnavatelů.

2. Vlastní šetření

Absolventi Univerzity Pardubice hodnotili sadu 26 položek, týkajících se studia na vysoké škole. První otázka zněla, jak byli studenti na vysoké škole připraveni na výkon povolání. Každá z 26 položek byla ohodnocena na stupnici 1 – 10, přičemž hodnocení 1 znamená velmi nízká úroveň, hodnocení 10 velmi vysoká. Druhá otázka se týkala odhadu vlastní úrovně kompetencí. Třetí otázka pak mapovala názory absolventů na požadavky zaměstnavatelů.

3. Výsledky

V tabulce 1 jsou zpracovány výsledky hodnocení úrovně kompetencí získaných studiem na vysoké škole. Nejvíce studenti oceňují získanou dovednost pracovat s informacemi, dále

schopnost vzdělávat se a organizovat své učení a získání odborných teoretických a metodologických znalostí. Na druhé straně poukazují na nedostatečnou úroveň získaných jazykových dovedností, nízkou schopnost pracovat v mezinárodním prostředí a nízkou schopnost vyhledávat nové příležitosti. Nízká úroveň získaných kompetencí je zmiňována rovněž u organizačních schopností, právní způsobilosti a matematických dovedností. To i přesto, že matematika je povinným předmětem na řadě vysokých škol.

Tab. 5: Hodnocení úrovně získané studiem na VŠ

Úroveň získaná studiem na VŠ	body
Dovednost pracovat s informacemi	6,60
Schopnost vzdělávat se a organizovat své učení	6,43
Odborné teoretické a metodologické znalosti	6,41
Dovednost prezentace	6,25
Dovednost písemného projevu	6,10
Všeobecné znalosti a rozhled	6,01
.....
.....
Matematické dovednosti	5,11
Právní způsobilost	4,97
Organizace a řízení, dovednost vést kolektiv	4,82
Schopnost pracovat v interkulturním / mezinárodním prostředí	4,75
Podnikavost, mít „čich“ pro nové příležitosti	4,41
Jazykové dovednosti v cizím jazyce	4,08

Tabulka č. 2 ukazuje, jak studenti sami odhadují své kompetence. Vysoko hodnotí svoji schopnost nést odpovědnost, dovednost pracovat s informacemi, své dovednosti v mateřském jazyce a svoji schopnost samostatného rozhodování. Nízkou úroveň kompetencí sami absolventi spatřují ve schopnosti vyhledávat nové příležitosti, v jazykových a matematických dovednostech, ve své právní způsobilosti a ve schopnosti pracovat v interkulturním prostředí.

Tab. 2: Hodnocení vlastní úrovně kompetencí

Odhad vlastní úrovně kompetencí	body
Schopnost nést odpovědnost	6,81
Dovednost pracovat s informacemi	6,77
Jazykové dovednosti v mateřském jazyce	6,70
Dovednost samostatně se rozhodovat	6,65
Schopnost vzdělávat se a organizovat své učení	6,64
Dovednost identifikovat a řešit problémy	6,63
.....
.....
Organizace a řízení, dovednost vést kolektiv	5,96
Znalost podmínek pro využití odborných metod a teorií v praxi	5,77
Schopnost pracovat v interkulturním / mezinárodním prostředí	5,59
Právní způsobilost	5,37
Matematické dovednosti	5,33
Jazykové dovednosti v cizím jazyce	5,31
Podnikavost, mít „čich“ pro nové příležitosti	5,25

Požadavky zaměstnavatelů na kompetence absolventa jsou shrnuty v tabulce 3. Ukazuje se, že zaměstnavatelé kladou největší důraz na schopnost nést odpovědnost, na dovednost komunikovat s lidmi, na schopnost zvládat zátěžové situace, dovednost identifikovat a řešit problémy, samostatně se rozhodovat, přizpůsobit se změněným podmínkám, na schopnost týmové práce a dovednost tvořivého myšlení.

Mezi kompetence, které jsou vyžadovány nejméně, patří matematické dovednosti, jazykové dovednosti, schopnost vyhledávat příležitosti, schopnost pracovat v mezinárodním prostředí, právní způsobilost.

Tab. 3: Hodnocení úrovně kompetencí vyžadované v zaměstnání

úroveň kompetencí vyžadovaná v současném zaměstnání	body
Schopnost nést odpovědnost	4,68
Dovednost komunikovat s lidmi, vyjednávat	4,61
Schopnost zvládat zátěžové situace a překážky	4,60
Dovednost identifikovat a řešit problémy	4,56
Dovednost samostatně se rozhodovat	4,52
Schopnost přizpůsobit se změněným okolnostem, podmínkám	4,49
.....
.....
Organizace a řízení, dovednost vést kolektiv	3,78
Právní způsobilost	3,55
Schopnost pracovat v interkulturním / mezinárodním prostředí	3,49
Podnikavost, mít „čich“ pro nové příležitosti	3,36
Jazykové dovednosti v cizím jazyce	3,27
Matematické dovednosti	2,90

V závěru jsme hodnotili, jaký je vzájemný vztah mezi mírou kompetencí získanou při studiu na vysoké škole, vlastním hodnocením kompetencí a požadavky zaměstnavatelů. Tyto vztahy jsou vyjádřeny pomocí tří korelačních koeficientů, Spearmanova (tab.4), korelace gamma (tab. 5) a Kendallova korelačního koeficientu tau. Všechny vypočítané hodnoty korelačních koeficientů jsou statisticky významné (tj. $p \neq 0$) na hladině významnosti 0,05.

Hodnota korelačního koeficientu ukazující na vztah mezi mírou kompetencí získanou při studiu na vysoké škole, a požadavky zaměstnavatelů je však velmi nízká, ukazuje na málo těsný vztah těchto veličin.

Tab. 4: Spearmanův korelační koeficient

	studium	vlastní hodn.	zaměstnavatel
studium	1,00	0,52	0,40
vlastní hodn.	0,52	1,00	0,78
zaměstnavatel	0,40	0,78	1,00

Tab. 5: Korelace gamma

	studium	vlastní hodn.	zaměstnavatel
studium	1,00	0,35	0,28
vlastní hodn.	0,35	1,00	0,59
zaměstnavatel	0,28	0,59	1,00

Tab. 6: Kendallův korelační koeficient tau

	studium	vlastní hodn.	zaměstnavatel
studium	1,00	0,35	0,28
vlastní hodn.	0,35	1,00	0,59
zaměstnavatel	0,28	0,59	1,00

4. Závěr

Při podrobné analýze hodnocení položek, odpovídajících na otázku, jak byli studenti na vysoké škole připraveni na výkon povolání a položek týkajících se požadavků zaměstnavatelů na úroveň kompetencí dochází k řadě nesrovnalostí. Zatímco absolventi nejvýše hodnotí přínos svého vysokoškolského vzdělání v položkách dovednost pracovat s informacemi, schopnost vzdělávat se a organizovat své učení, odborné teoretické a metodologické znalosti, dovednost prezentace a dovednost písemného projevu, z pohledu potřeb zaměstnavatelů jsou tyto položky až na jednu ve druhé polovině požadavků na úroveň kompetencí.

Současní zaměstnavatelé vyžadují především schopnost nést odpovědnost, dovednost komunikovat s lidmi, schopnost zvládat zátěžové situace, dovednost identifikovat a řešit problémy, samostatně se rozhodovat, přizpůsobit se změněným podmínkám, schopnost týmové práce a dovednost tvořivého myšlení. Tyto položky se nevyskytují v první třetině nejvýše hodnocených kompetencí získaných studiem. Zaměstnavatelé upřednostňují spíše osobnostní charakteristiky, nežli kvalitní vědomosti a znalosti.

5. Literatura

REFLEX 2010. Zaměstnatelnost a uplatnění absolventů vysokých škol na pracovním trhu. [online]. Praha: SVP Univerzita Karlova, 2010.

www.strediskovzdelavacipolitiky.info/default.asp?page=svp&KID=85.

KOUCKÝ, J. - RYŠKA, R. 2013. Projekt REFLEX 2013: 4. monitorovací zpráva. Zaměstnatelnost a uplatnění absolventů vysokých škol na pracovním trhu

a hodnocení získaného vysokoškolského vzdělání. Praha: SVP Univerzita Karlova, 2013.

<http://www.strediskovzdelavacipolitiky.info/download/REFLEX%202013%20Monitorovaci%20zprava%2004.%20SVP%20PedF%20UK%202013-08-23.pdf>

KOUCKÝ, J. - RYŠKA, R. 2013. Projekt REFLEX 2013: 5. monitorovací zpráva. Zaměstnatelnost a uplatnění absolventů vysokých škol na pracovním trhu

a hodnocení získaného vysokoškolského vzdělání. Praha: SVP Univerzita Karlova, 2013.

<http://www.strediskovzdelavacipolitiky.info/download/REFLEX%202013%20Monitorovaci%20zprava%2005.%20SVP%20PedF%20UK%202013-11-30.pdf>

Poznámka:

Článek vznikl za podpory grantu SGSFES_2014003: Vědecko - výzkumné aktivity v Systémovém inženýrství a informatice

Adresy autorov:

Jana Kubanová, doc. PaedDr. CSc.
Univerzita Pardubice
Studentská 95, 53210 Pardubice
jana.kubanova@upce.cz

Michal Kuban, Ing.
Univerzita Pardubice
Studentská 95, 53210 Pardubice
michal.kuban@csobpoj.cz

Segmentace států EU podle sektorových indikátorů důvěry

Segmentation of EU according sectorial confidence indicators

Nikolay Kulbakov

Abstract: This paper is a continuation of the series of articles about the segmentation of EU28 countries by cluster analysis and economic indices. The paper contains the output from the cluster analysis based on the time series 2004-2014 years of macroeconomic confidence indices for the EU28. In research used construction, industrial, trade, retail and consumer confidence indicators.

Abstrakt: Článek je pokračováním seriálu článků o segmentaci států EU pomocí shlukové analýzy a ekonomických indikátorů. Příspěvek obsahuje komentář k výstupu shlukové analýzy provedené na základě časové řady 2004-2014 makroekonomických indikátorů důvěry pro země EU28. Použité indikátory důvěry ve stavebnictví, v průmyslu, v obchodě, ve službách a důvěry spotřebitelů.

Key words: cluster analysis, EU, macroeconomic segmentation, leading indicators, confidence indicators, economic expectations, MATLAB.

JEL classification: E01, C38

1. Úvod

Tento článek je pokračováním seriálu článků o segmentaci států EU pomocí shlukové analýzy a makroekonomických indikátorů, viz také Kulbakov (2013) a Kulbakov (2014). Primárním cílem tohoto výzkumu je, za využití metod shlukové analýzy popsaných v Rezanková, H., & Snásel, V. (2009), Löster, T. (2011) a dalších, sestavit představu o rozdělení států EU do dvou skupin, a to „optimistů“ a „pesimistů“, pomocí sektorových indikátorů důvěry. Sekundárním cílem výzkumu je vývoj řady veřejně přístupných nástrojů v prostředí MATLAB, které budou pomáhat autorovi a všem zájemcům o podobnou problematiku provádět snadno a rychle shlukovou analýzu územních celků dle ekonomických indikátorů.

2. Použitá data

Pro účely analýzy byla použita data z databáze Eurostat pro 28 zemí Evropské Unie. Indexy obsahují data za státy EU28 a roky 2004-2013 a prvních 10 měsíců roku 2014. Data vstupující do analýzy jsou normalizovaná dle charakteru dat.

Aby bylo možno sledovat celkovou hospodářskou činnost, existuje širší indikátor ekonomického sentimentu (ESI /BS-ESI-I/ Economic sentiment indicator) a ten se vypočítává jako kompozitní indikátor, který obsahuje pět sektorových indikátorů důvěry s různými váhami:

1. indikátor důvěry v průmyslu /BS-ICI-BAL/ Industrial confidence indicator
2. indikátor důvěry ve službách /BS-SCI-BAL/ Services Confidence Indicator
3. indikátor důvěry spotřebitelů /BS-CSMCI-BAL/ Consumer confidence indicator
4. indikátor důvěry ve stavebnictví /BS-CCI-BAL/ Construction confidence indicator
5. indikátor důvěry v obchodě /BS-RCI-BAL/ Retail confidence indicator

Právě tyto indikátory byly dále použity ve výzkumech. Indikátory důvěry představují aritmetické průměry sald odpovědí na vybrané otázky úzce související s referenční

proměnnou, kterou sledují. Průzkumy jsou definovány v rámci Společného harmonizovaného programu EU pro podnikatelské a spotřebitelské průzkumy. Data jsou zpracována podle statistické klasifikace ekonomických činností v Evropském společenství (NACE Rev 2). (ČSÚ 2013) Jsou to indikátory zdraví evropské ekonomiky, které poskytují ucelený obraz hospodářské situace v regionu jako celku. Pomáhají zemím EU plánovat budoucnost a působí jako signál pro problémy v ekonomice. Dílčí indikátory dobře odráží vývoj jednotlivých sektorů. (European Commission 2006).

3. Shluková analýza

Podrobný popis algoritmu a odkaz na použitý programový kód MATLAB je v Kulbakov (2013). Výzkum je prováděn pomocí *hierarchické* shlukové analýzy.

Tab. 6: Shlukové rozhodnutí pro EU a roky 2004-2014 (+1 „optimisté“ a -1 „pesimisté“)

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Austria	1	-1	-1	1	1	1	-1	1	1	1	-1
Belgium	1	-1	-1	1	1	1	-1	1	1	1	-1
Bulgaria	1	1	-1	1	1	1	-1	1	1	1	-1
Cyprus	1	1	-1	1	1	1	-1	-1	-1	-1	-1
Czech Republic	1	1	1	1	1	1	-1	1	1	1	1
Germany	-1	-1	-1	1	-1	1	-1	1	1	1	-1
Denmark							-1	1	1	1	-1
Estonia	1	1	1	1	-1	-1	-1	1	1	1	-1
Greece	1	-1	-1	1	1	1	-1	-1	-1	1	-1
Spain	1	-1	-1	1	-1	1	-1	-1	1	1	-1
Finland	1	1	1	1	1	1	-1	1	1	1	-1
France	1	-1	-1	1	1	1	-1	1	1	1	-1
Croatia					1	1	-1	1	1	1	-1
Hungary	1	-1	-1	-1	-1	-1	-1	1	1	1	-1
Ireland											
Italy	1	-1	-1	1	-1	1	-1	1	1	1	-1
Lithuania	1	1	1	1	-1	-1	-1	1	1	1	-1
Luxembourg											
Latvia	1	1	1	1	-1	-1	-1	1	1	1	-1
Malta					1	1	-1	1	1	1	-1
Netherlands	1	-1	1	1	1	1	-1	1	1	1	-1
Poland	-1	-1	-1	1	1	1	-1	1	1	1	-1
Portugal	-1	-1	-1	-1	-1	1	-1	-1	-1	1	-1
Romania	1	1	-1	1	1	1	-1	1	1	1	-1
Sweden	1	1	1	1	1	1	1	1	1	1	1
Slovenia	1	1	1	1	1	1	-1	1	1	1	-1
Slovakia	1	1	1	1	1	1	-1	1	1	1	-1
UK	1	-1	-1	1	-1	-1	-1	1	1	1	1

4. Výsledky

Ve výsledcích analýzy autor dospěl k umístění 26 zemí EU do 2 shluků: 1 – „optimisté“ a -1 – „pesimisté“. Hodnota 1 reprezentuje shluk, do kterého byly umístěny optimističtější státy z hlediska souhrnu normalizovaných indexů. Shluk -1 obsahuje pesimističtější státy z hlediska sledovaných indikátorů pro příslušný rok. Distribuce do shluků se mění rok od roku a státy se pohybují mezi nimi. Výsledné roztřídění států se nachází v tabulce 1., kde vedle názvů států je vždy časová řada ukazující na příslušný shluk v letech 2004-2014.

Luxemburg a Irsko bohužel vypadli z analýzy, protože v databázi EU chyběly některé údaje. Pokud země pro příslušný rok nemá k dispozici údaje pro jeden nebo více indikátorů, pak není možné zařadit ji do žádného shluku.

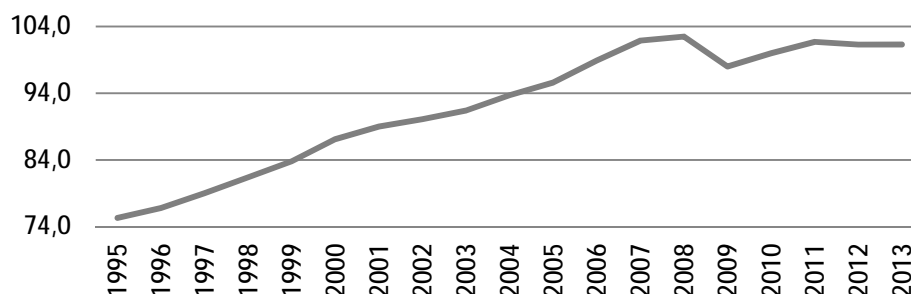
Zařazení zemí do shluků není v různých letech stálé, jenom Švédsko je vždy ve

skupině optimistů, protože vykazuje nejvyšší hodnoty indexů za sledované období.

Například v pokrizovém roce 2010 zastupuje „optimisty“ jenom Švédsko, a v roce 2013 naopak „pesimisty“ zastupuje jenom Kypr.

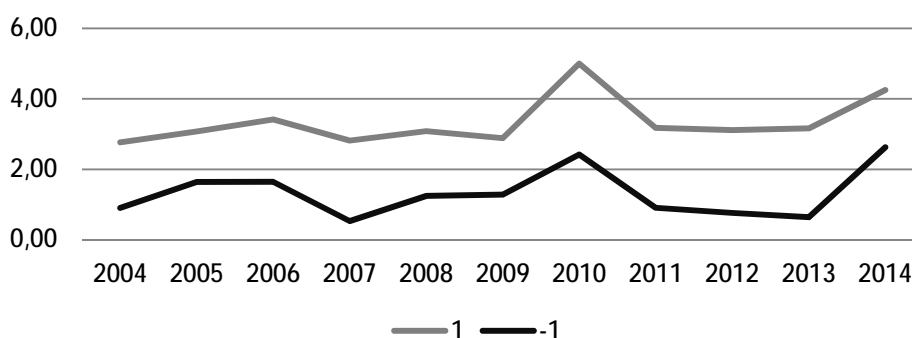
Nyní se podíváme na odvozené indikátory pro dva výše popsané shluky. Ale předtím se podíváme na vývoj nominálního HDP pro EU28, protože hodnoty indexů důvěry jsou propojeny s vývojem ekonomické situace.

Obr. 1. Nominální HDP EU v letech 1995-2013



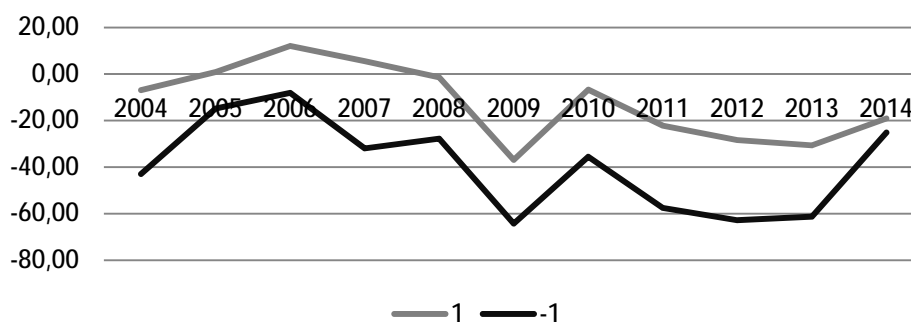
Ekonomika EU nominálně rostla od roku 1995 do 2008, ale v roce 2007 tempo růstu začalo zpomalovat, a v roce 2009 se nominální HDP EU propadlo o -4,4%. Potom v roce 2010 se HDP vrátilo téměř na hodnotu roku 2007 a ekonomika se od té doby nachází ve stagnaci. Indikátory důvěry pro shluky a rating většinou popisují vývoj přírůstu HDP.

Obr. 2: žebříček shluků dle sumy normalizovaných průměrů rozhodovacích hodnot pro země EU rozdělených do dvou shluků



Hodnoty pro shluky +1 a -1 souhrnu pro všech pět indexů rostou do roku 2006, potom vytuší krizi a začínají klesat až do roku 2010, kdy prudce narůstají vlivem optimismu způsobenému pokrizovým oživením, a nadále ukazatele opakují stagnaci a ožívují se jenom v roce 2014.

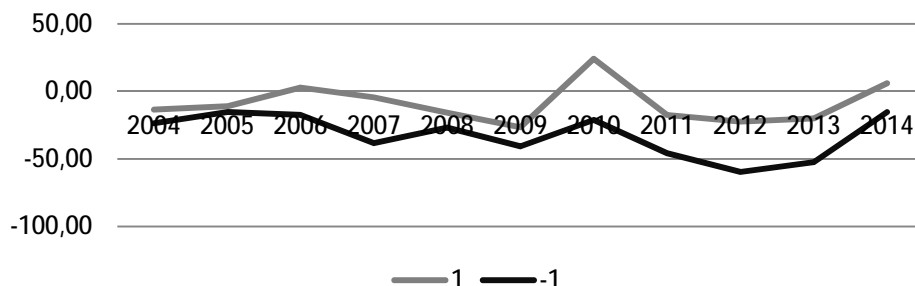
Obr. 3. Indikátor důvěry ve stavebnictví pro dva shluky zemí EU



Indikátor důvěry ve stavebnictví je průměr dvou ukazatelů: hodnocení celkové poptávky a očekávaný vývoj zaměstnanosti. Ve sledovaném období se pesimistický shluk nachází skoro stále v pesimismu, hodnoty průměru jsou většinou pod nulou. Optimistický shluk během krize

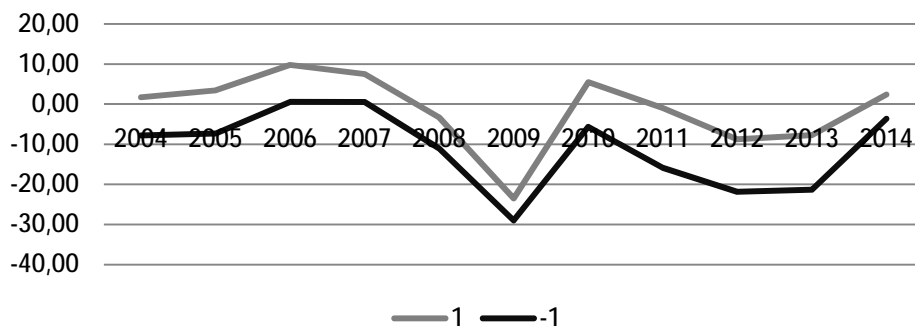
2009 a po roce 2011 také vykazuje pesimistickou náladu. Předpokládám, že také vlivem protikrizových fiskálních balíčků a zvýšením investování států do staveb vykazoval sektor stavebnictví v roce 2010 oživení.

Obr. 4. Indikátor důvěry spotřebitelů pro dva shluky zemí EU



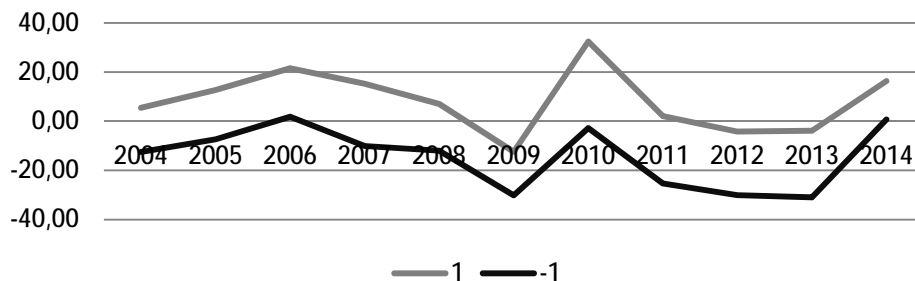
Indikátor důvěry spotřebitelů je složen ze čtyř ukazatelů: očekávaná finanční situace spotřebitele, očekávaná celková ekonomická situace, očekávaná celková nezaměstnanost (s opačným znaménkem) a očekávané úspory spotřebitele v příštích 12 měsících. Ve sledovaném období je nálada spotřebitelů převážně pesimistická pro oba shluky.

Obr. 5. Indikátor důvěry v průmyslu pro dva shluky zemí EU

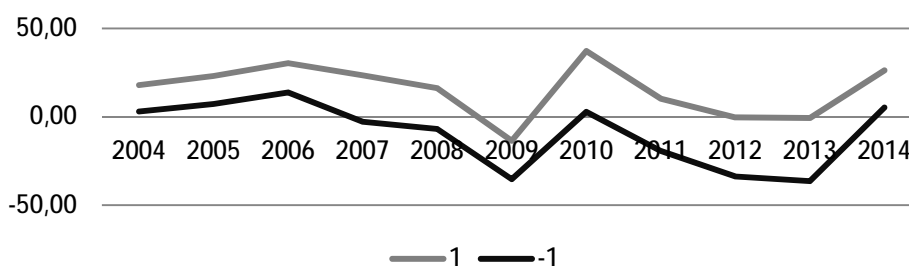


Indikátor důvěry v průmyslu je průměr sald tří ukazatelů: hodnocení celkové poptávky, zásoby hotových výrobků s opačným znaménkem a očekávaný vývoj výrobní činnosti. Na obrázku č.5. je vidět jak průměry pro oba shluky klesají od roku 2006 a od roku 2010, což popisuje celkovou náladu výrobců v recesi. Nedostatečná poptávka po krizovém roce 2009 nevrátila ukazatel na předkrizovou úroveň. V roce 2014 pozorujeme oživení.

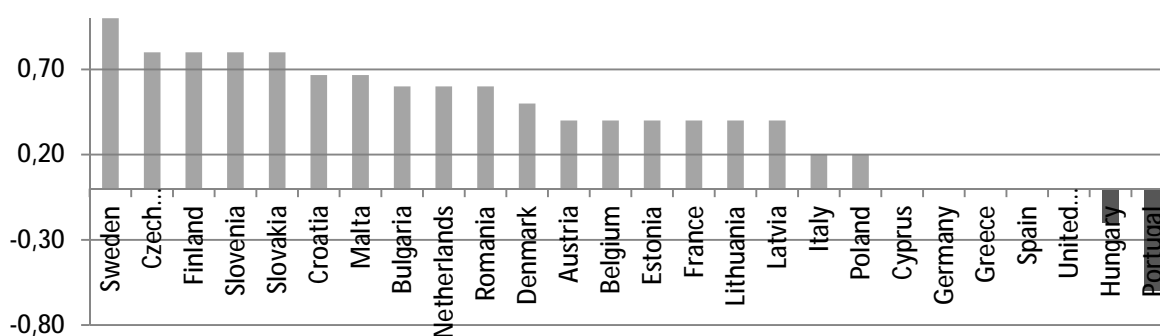
Obr. 6. Indikátor důvěry v obchodě pro dva shluky zemí EU



Indikátor důvěry v obchodě je průměr tří ukazatelů: hodnocení ekonomické situace, hodnocení současného stavu zásob s opačným znaménkem a očekávaný vývoj ekonomické situace. Graf č.6 představuje cyklus souběžného vývoje průměrů indikátorů pro oba shluky se špičkami v roce 2006, kdy ekonomika EU vykazovala expanzi a v roce 2010, kdy HDP začal znovu růst po velkém propadu.

Obr. 7. Indikátor důvěry ve službách pro dva shluky zemí EU

Indikátor důvěry ve vybraných odvětvích služeb je průměr tří ukazatelů: hodnocení ekonomické situace, hodnocení poptávky a očekávaná poptávka. V sektoru služeb začíná hodnota ukazatele klesat v obou shlucích od roku 2006 a klesá do krizového roku 2009. Je vidět prudký nárůst představující maximum za 10 let v roce 2010, kdy HDP začal vykazovat kladný růst. Dále, se stagnací produktu, začíná ukazatel klesat a ožívuje se jen v roce 2014.

Obr. 8. Rating optimismu zemí EU za posledních 10 let (2004-2013)

Pokud vypočítáme průměr pro hodnoty shluků z tabulky č.1., pak můžeme seřadit země EU podle toho, jak často se nacházely ve shluku optimistů nebo pesimistů. Neoptimističtější zemí z hlediska hodnot vykazujících v sektorových indikátorech důvěry je Švédsko, nejpesimističtější je Portugalsko. Česká republika a Slovensko v porovnání s ostatními zeměmi jsou spíše optimističtější.

5. Závěr

Indikátory důvěry zemí EU velmi korelovaly mezi sebou a jsou také závislé na předchozím vývoji HDP, i když se považují za předstihové indikátory. Toto snižuje kvality a přínosy shlukové analýzy podobného druhu indikátorů. Ekonomika EU je značně propojená a otevřená, a proto vývoj indikátorů důvěry je pro různé země ve značné míře synchronní a napodobuje „světový“ nebo „evropský“ hospodářský cyklus. Když pokles v průměru vykazují pesimistické země, tak ho vykazují i optimistické, rozdíl je jenom ve velikosti hodnot, ale marginální vývoj pro shluky zemí je velice podobný.

Acknowledgment

This article was created with the help of the Internal Grant Agency of University of Economics in Prague No. 6/2013 under the title „Evaluation of results of cluster analysis in Economic problems.”

Literatura

- Rezankova, H., & Loster, T. (2013). Shlukova analyza domacnosti charakterizovanych kategoriálními ukazateli. *E+M. Ekonomie a Management*, 16(3), 139-147. ISSN: 1212-3609
- Baker, F., & Hubert, L. (1975). Measuring Power of Hierarchical Cluster-Analysis. *Journal of the American Statistical Association*, 70(349), 31–38. doi:10.2307/2285371
- Berkhin, P. (2006). A Survey of Clustering Data Mining Techniques. In J. Kogan, C. Nicholas, & M. Teboulle (Eds.), *Grouping Multidimensional Data* (pp. 25–71). Springer Berlin Heidelberg. Retrieved from
- Cattinelli, I., Valentini, G., Paulesu, E., & Borghese, N. A. (2013). A Novel Approach to the Problem of Non-uniqueness of the Solution in Hierarchical Clustering. *Ieee Transactions on Neural Networks and Learning Systems*, 24(7), 1166–1173. doi:10.1109/TNNLS.2013.2247058
- Cernakova, V., & Hudec, O. (2012). Quality of Life: Typology of European Cities Based on Cluster Analysis. *E & M Ekonomie a Management*, 15(4), 34–48.
- Eurostat. (2014). *Eurostat research database*. Retrieved from http://epp.eurostat.ec.europa.eu/NavTree_prod/everybody/BulkDownloadListing
- Faber, V. (1994). Clustering and the continuous k-means algorithm. *Los Alamos Science*, 22, 138–144.
- Löster, T. (2011). *Hodnocení výsledků metod shlukové analýzy*. FIS VŠE. Retrieved from <https://isis.vse.cz/auth/lide/clovek.pl?id=8340;zalozka=7;studium=92375>
- Rezanková, H., & Snásel, V. (2009). *Shluková analýza dat*. Praha: Professional Publishing.
- S&P. (2013, May). *Sovereigns Rating*. Retrieved from <http://www.standardandpoors.com/ratings/sovereigns/ratings-list/en/us?sectorName=null&subSectorCode=39>
- Kulbakov, N.(MSED 2013) *Segmentation of EU/EA Countries via Cluster Analysis of Macroeconomics Indicators*. Retrieved from <http://msed.vse.cz/files/2013/196-Kulbakov-Nikolay-paper.pdf> or <http://www.ilovecz.ru/research/matlab01.zip>
- Kulbakov, N.(VS 2013) *Segmentation of EU27 into four groups and their dynamics*. Retrieved from <http://www.ilovecz.ru/research/matlab02.zip>
- Kulbakov, N. (MSED 2014) *SEGMENTATION OF THE EU USING LEADING INDICATORS*. Retrieved from <http://www.ilovecz.ru/research/matlab03.zip>
- Kulbakov, N. (VS 2014) *Source for MATLAB cluster analysis, input and output data*. Retrieved from <http://www.ilovecz.ru/research/matlab04.zip>
- ČSÚ. (2013a). *Indikátor ekonomického sentimentu*. Retrieved from <http://apl.czso.cz/pll/eutab/html.h?ptabkod=teibs010>
- ČSÚ. (2013b). *Metodika - Konjunkturální průzkumy*. Retrieved from http://www.czso.cz/csu/redakce.nsf/i/konjunkturalni_pruzkum
- European Commission. (2006). *The joint harmonised EU programme of business and consumer surveys*. Luxembourg: Office for Official Publications of the European Communities.

Adresa autora:

Nikolay Kulbakov
University of Economics, Prague
W. Churchill Sq. 4, 130 67 Prague 3
Czech Republic
kulbakov@gmail.com

Príjmová nerovnosť na Slovensku, 2004 – 2010

Income inequality in Slovakia, 2004-2010

Viera Labudová

Abstract: This article aims to evaluate the contribution of different characteristics of the population to income inequality in Slovakia. The analysis uses data EU SILC to study the components of income inequality and its change during the period 2004 - 2010. The Theil Index is chosen for this analysis as it is decomposable into between-group and within-group components.

Abstrakt: Cieľom tohto článku je zhodnotiť vplyv rôznych charakteristík obyvateľstva na veľkosť príjmovej nerovnosti Slovensku. Na základe údajov zo zisťovania EU SILC sú analyzované zložky príjmovej nerovnosti a ich zmeny v období rokov 2004 až 2010. V analýze je použitý Theilov index, ktorý je možné jednoducho rozložiť na medziskupinovú a vnútroskupinovú zložku.

Key words: Inequality, Theil's T index, EU SILC, decomposition

Kľúčové slová: nerovnosť, Theilov T index, EU SILC, rozklad

JEL classification: I32, I33, D63

1. Úvod

Nerovnosť rozdelenia príjmov sa veľmi často využíva ako indikátor ekonomickej vyspelosti krajiny, funkčnosti sociálnej politiky štátu a správnosti opatrení aplikovaných v sociálnej oblasti. Kvantifikácia nerovnosti príjmov býva často súčasťou širšej analýzy týkajúcej sa chudoby a blahobytu. Na meranie príjmovej nerovnosti sme v príspevku použili Theilov T index, ktorý patrí do skupiny mier generalizovanej entropie. Výber tejto miery bol motivovaný našou snahou identifikovať tie sociálnoekonomické charakteristiky, ktoré majú najväčší vplyv na nerovnosť v rozdelení príjmov. Rozklad Theilovho indexu na tzv. medziskupinovú a vnútroskupinovú zložku nám umožnil naplniť sledovaný cieľ.

2. Zdroje údajov, príprava údajov

Použili sme súbory individuálnych údajov, ktoré pochádzajú z výberového štatistického zisťovania EU SILC 2005 (ŠÚ SR: UDB verzia 12.07.06) a zisťovania EU SILC 2011 (ŠÚ SR: UDB verzia 22/11/2012).

Výberové zisťovanie EU SILC 2005 bolo realizované na vzorke 6016 domácností a zisťovanie EU SILC 2011 na vzorke 5 801 domácností. Premenné sú v 4 rôznych súboroch: Register domácností (D_súbor), Register osôb (R_súbor), Údaje za domácnosť (H_súbor), Osobné údaje (P_súbor). Register H obsahuje údaje o tých domácnostiach z registra D, v ktorých opytovanie prebehlo úspešne. Register osôb (R_súbor) obsahuje záznam za každú osobu, ktorá v čase zisťovania žila v domácnosti akceptovanej do databázy alebo bola dočasne neprítomná. Súbor osobných údajov (P_súbor) obsahuje záznam za každú osobu vo veku 16 rokov a viac, ktorá odpovedala na osobný dotazník. Na spojenie registra D a registra H boli použité premenné DB030 a HB030 (*identifikácia domácnosti*), na spojenie registrov P a R sme použili premenné PB030 a RB030 (*osobná identifikácia*), prepojenie medzi registrami domácností a registrami osôb bola použitá premenná RX030 (PX030) – *identifikácia domácnosti*. Pred spojením registrov bolo potrebné uskutočniť selekciu záznamov. V súbore D sme ponechali iba tie domácnosti, v ktorých bolo opytovanie za danú domácnosť akceptované do databázy (DB135=1), čo znamená, že domácnosť bola

kontaktovaná (DB120=11) a zrealizovalo sa opytovanie v súbore, ktorý obsahuje dáta za domácnosť (DB130=11). Z databázy UDB verzia 12.07.06 boli z registra H vylúčené záznamy o domácnostiach, v ktorých premenná HY020 (*celkový disponibilný príjem domácnosti*) nadobudla záporné hodnoty (0,15 % domácností).

V súbore R sme ponechali iba tie záznamy, ktoré sa vzťahovali k osobám vo veku 16 rokov a starším (RX010 – *vek na konci príjmového referenčného obdobia*).

3. Metodika

V príspevku sme sa zamerali na posúdenie vplyvu vybraných sociálnoekonomických charakteristík domácnosti na nerovnosť v rozdelení ekvivalentného disponibilného príjmu. Využili sme pritom vlastnosť rozložiteľnosti pre miery zo skupiny generalizovanej entropie. Vlastnosť rozložiteľnosti umožňuje vypočítať veľkosť miery nerovnosti pre celú populáciu I_{TOTAL} na základe jej hodnôt pre podmnožiny, na ktoré je možné celú populáciu rozložiť bezo zvyšku $I_{TOTAL} = I_{BETWEEN} + I_{WITHIN}$, kde I_{TOTAL} je hodnota miery na celej populácii, $I_{BETWEEN}$ je medziskupinová hodnota miery a I_{WITHIN} je vnútroskupinová hodnota miery.

V práci sme použili mieru Theilovo T, pre ktorú možno medziskupinovú a vnútroskupinovú zložku vyjadriť (Bourguignon 1979, Cowell 1980) takto:

$$\text{Theil } T = \text{Theil } T_w + \text{Theil } T_b \quad (1)$$

$$\text{Theil } T = \sum_{j=1}^k \left(\frac{y_j}{y} \right) T_j + \sum_{j=1}^k \left(\frac{y_j}{y} \right) \log \frac{y_j/y}{n_j/n} \quad (2)$$

kde j označuje populačnú skupinu, n_j je absolútna početnosť j -tej skupiny (podskupiny), y_j predstavuje celkový príjem j -tej skupiny, T_j je hodnota Theilovho T v j -tej skupine.

Prvý člen výrazu vyjadruje nerovnosť vnútri populačných skupín, druhý člen vyjadruje nerovnosť medzi populačnými skupinami.

Podrobnosti o mierach nerovnosti a ich vlastnostiach možno nájsť v (Labudová, 2013).

4. Vytvorenie populačných podskupín

Na vytvorenie populačných podskupín v spojených registroch P a R boli použité tieto premenné: **PB150**: *pohlavie*, (1 – muž, 2 – žena), **PE040**: *najvyššia dosiahnutá úroveň podľa ISCED*, (0 – predprimárne vzdelanie, 1 – primárne vzdelanie, 2 – nižšie sekundárne vzdelanie, 3 – (vyššie) sekundárne vzdelanie, 4 – post-sekundárne vzdelanie, nie terciálne vzdelanie, 5 – prvý stupeň terciálneho vzdelania (nevedie priamo k získaniu vedeckej kvalifikácie), 6 – druhý stupeň terciálneho vzdelania (vedie priamo k získaniu vedeckej kvalifikácie)), **RB210**: *status základnej ekonomickej aktivity* (1 – pracujúci, 2 – nezamestnaný, 3 – dôchodca, 4 – iná neaktívna osoba), **HT** (EU SILC 2011 – register R), **HX020** (EU SILC 2005 – register H): *typ domácnosti* (5 – jednočlenná domácnosť, 6 – domácnosť 2 dospelých bez závislých detí – obaja vo veku pod 65 rokov, 7 – domácnosť 2 dospelých bez závislých detí – aspoň jeden dospelý vo veku 65 rokov a viac, 8 – ostatné domácnosti bez závislých detí, 9 – domácnosť s 1 rodičom a s 1 alebo viac závislými deťmi, 10 – domácnosť 2 dospelých s 1 závislým dieťaťom, 11 – domácnosť 2 dospelých s 2 závislými deťmi, 12 – domácnosť 2 dospelých s 3 alebo viac závislými deťmi, 13 – ostatné domácnosti so závislými deťmi, 16 – iné domácnosti (nie je možné určiť typ).

V populačných skupinách, ktoré boli vytvorené na základe obmien (kategórií) týchto premenných, sme merali veľkosť nerovnosti v rozdelení ekvivalentného disponibilného

príjmu (EDP). Hodnoty ekvivalentného disponibilného príjmu boli pre rok 2004 vypočítané pomocou hodnôt premenných HY020 – *celkový disponibilný príjem domácnosti* a HX050 – *ekvivalentná veľkosť domácnosti*, nachádzajúcich sa v registri H, a následne priradené jednotlivým členom domácnosti. Hodnoty tejto premennej pre rok 2010 boli uvedené v súbore R ako premenná EQ_INC20.

Pre každú populačnú podskupinu bola vypočítaná hodnota Theilovho T a pomocou rozkladu Theilovho indexu, meraného na celej populácii osôb vo veku 16 rokov a viac sme urobili závery o vplyve danej premennej na veľkosť príjmovej nerovnosti. Na vyvodenie záverov týkajúcich sa základného súboru osôb 16 a viacročných sme použili premennú **PB040** (*osobné prierezové váhy všetkých členov domácnosti vo veku 16 rokov a viac*).

5. Výsledky analýzy

Theilov T index ekvivalentného disponibilného príjmu (EDP) osôb 16 a viacročných na Slovensku mal v roku 2004 hodnotu Theil T = 0,12558, v roku 2010 bola jeho hodnota Theil T = 0,12682. Rozkladom indexu podľa kategórií *statusu základnej ekonomickej aktivity* sa ukázalo, že najvyššia príjmová nerovnosť v rozdelení EDP bola v oboch rokoch v skupine nezamestnaných, najnižšia príjmová nerovnosť bola v roku 2005 v populácii dôchodcov a v roku 2010 v populácii pracujúcich (Tab.1).

Tabuľka 1: Rozklad Theilovho indexu EDP podľa premennej statusu základnej ekonomickej aktivity

Názov	Poznámka	Hodnoty	Theilovo T	
			2004	2010
Status základnej ekonomickej aktivity	pracujúci	1	0,126034768	0,100619
	nezamestnaný	2	0,132183541	0,185615
	dôchodca	3	0,084638275	0,139733
	iná neaktívna osoba	4	0,130732329	0,115167
Rozklad indexu				
Názov zložky	2004		2010	
	Hodnoty zložiek	Percentuálne vyjadrenie	Hodnoty zložiek	Percentuálne vyjadrenie
Medziskupinová zložka T_B	0,008504	6,77 %	0,011029	8,70 %
Vnútro skupinová zložka T_W	0,117071	93,23 %	0,115791	91,30 %

Ďalším rozkladom Theilovho indexu sme chceli zistiť, aký vplyv má *typ domácnosti*, z ktorej osoba pochádza na celkovú príjmovú nerovnosť. Najnižšia príjmová nerovnosť v oboch rokoch pozorovania bola zaznamenaná v skupine bezdetných osôb žijúcich vo dvojiciach, pričom aspoň jeden z dvojice je mladší ako 65 rokov. Theilov index v tejto skupine mal v roku 2004 hodnotu o 25 % nižšiu, v roku 2010 o takmer 70 % nižšiu ako bola jeho hodnota zameraná na celej populácii. Najvyššie hodnoty nerovnosti v rozdelení EDP boli v skupine jednočlenných rodín samostatne žijúcich rodičov, ktorí sa starajú aspoň o jedno závislé dieťa (Tab.2). Použitím premennej *najvyššia dosiahnutá úroveň podľa ISCED* sme vytvorili šesť populačných podskupín (kategória 0 – predprimárne vzdelanie sa v roku 2004 nevyskytovala, v roku 2010 ju obsahoval jeden záznam, pre tento prípad bola zlúčená s kategóriou 1 – primárne vzdelanie). Pre obidve porovnávané obdobia bola najvyššia príjmová nerovnosť v skupine osôb s ukončeným druhým stupňom terciárneho vzdelávania.

Najnižšiu hodnotu mal Theilov index v roku 2004 v skupine osôb s primárnym vzdelaním v roku 2010 v skupine s post-sekundárnym vzdelaním.

Veľmi malé rozdiely sú medzi príjmovou nerovnosťou v skupine žien a v skupine mužov. Prejavilo sa to aj v hodnote medziskupinovej zložky, ktorá v roku 2004 predstavovala 0,32 % a v roku 2010 len 0,19 % príjmovej nerovnosti meranej pomocou Theilovho indexu.

Tab. 2: Rozklad Theilovho indexu EDP podľa premennej typ domácnosti

Názov	Poznámka	Hodnoty	Theilovo T	
			2004	2010
Typ domácnosti	Jednočlenná domácnosť	5	0,114902	0,29819
	Domácnosť 2 dospelých bez závislých detí – obaja vo veku pod 65 rokov	6	0,113774	0,107134
	Domácnosť 2 dospelých bez závislých detí – aspoň jeden dospelý vo veku 65 rokov a viac	7	0,094261	0,038346
	Ostatné domácnosti bez závislých detí	8	0,100033	0,097286
	Domácnosť s 1 rodičom a s 1 alebo viac závislými deťmi	9	0,319214	0,117299
	Domácnosť 2 dospelých s 1 závislým dieťaťom	10	0,13295	0,115136
	Domácnosť 2 dospelých s 2 závislými deťmi	11	0,102731	0,100357
	Domácnosť 2 dospelých s 3 alebo viac závislými deťmi	12	0,11499	0,14462
	Ostatné domácnosti so závislými deťmi	13	0,122061	0,101149
	Iné domácnosti (nie je možné určiť typ)	16	0,060974	-
	Rozklad indexu			
Názov zložky	2004		2010	
	Hodnoty zložiek	Percentuálne vyjadrenie	Hodnoty zložiek	Percentuálne vyjadrenie
Medziskupinová zložka T_B	0,010669	8,50 %	0,011026	8,69 %
Vnútroskupinová zložka T_W	0,114906	91,50 %	0,115795	91,31 %

Najväčší vplyv na veľkosť príjmovej nerovnosti v rozdelení ekvivalentného disponibilného príjmu mal v roku 2004 *typ domácnosti*, v roku 2010 to bola *základná ekonomická aktivita jednotlivca*. Najviac zosilnel vplyv premennej *status základnej ekonomickej aktivity*, medziskupinová zložka Theilovho T vzrástla o 1,93 p.b. Zaznamenali sme aj zväčšenie rozdielov v úrovni príjmovej nerovnosti medzi populačnými skupinami, na ktoré bola celá

populácia osôb rozdelená podľa kategórií premennej *typ domácnosti*. Najmenší vplyv na príjmovú nerovnosť malo v oboch rokoch pohlavie (Tab.3).

Tabuľka 3: Vývoj hodnôt zložiek Theilovho T EDP medzi rokmi 2004 a 2010

Premenná	Theilov T index			
	zložky	2004	2010	Zmena
najvyššia dosiahnutá úroveň podľa ISCED	T _B	7,95 %	8,02 %	0,07 p.b.
	T _W	92,05 %	91,98 %	-0,07 p.b.
pohlavie	T _B	0,32 %	0,19 %	- 0,13 p.b.
	T _W	99,68 %	99,81 %	0,13 p.b.
status základnej ekonomickej aktivity	T _B	6,77 %	8,70 %	1,93 p.b.
	T _W	93,23 %	91,30 %	-1,93 p.b.
typ domácnosti	T _B	8,50 %	8,69 %	0,19 p.b.
	T _W	91,50 %	91,31 %	-0,19 p.b.
Celá populácia	T	0,12558	0,12682	0,99 %

6. Záver

V príspevku sme sa venovali analýze nerovnosti v rozdelení ekvivalentného disponibilného príjmu osôb vo veku 16 rokov a starších a zmenám v rozdelení, ku ktorým došlo medzi rokmi 2004 a 2010. Na meranie veľkosti nerovnosti bol v práci použitý Theilov T index, ktorý má vlastnosť rozložiteľnosti na vnútrokupinovú a medziskupinovú zložku. Táto vlastnosť Theilovho indexu umožnila vypočítať veľkosť jeho zložiek na populačných podmnožinách a na základe týchto výsledkov určiť tie sledované charakteristiky domácnosti, ktoré majú najväčší vplyv na celkovú príjmovú nerovnosť.

Literatúra

- BOURGUIGNON, F. 1979. Decomposable Income Inequality Measures. In: *Econometrica*. roč. 47, č. 4, s. 901-920.
- COWELL, F. A. 1980. On the Structure of Additive Inequality Measures. In: *Review of Economic Studies*. roč. 47, č. 3. s. 521-531.
- LABUDOVÁ, V. 2013. *Meranie príjmovej nerovnosti*. Bratislava: EKONÓM.
- SIPKOVÁ, Ľ. 2004. Prehľad teoretických východísk merania príjmovej nerovnosti. In: *Slovenská štatistika a demografia*. roč. 14, č. 3, s. 36-49.
- ŽELINSKÝ, T. 2010. Nerovnosť rozdeľovania príjmov v krajoch Slovenskej republiky. In: *Slovenská štatistika a demografia*. roč. 20, č. 1, s. 49-60.

Príspevok bol vytvorený s podporou vedeckovýskumného projektu VEGA 1/0761/12 Alternatívne prístupy k meraniu sociálno - ekonomického rozvoja (v kontexte Stratégie 2020 a ponaučení z globálnej finančnej krízy) a projektu VEGA 1/0127/11 Priestorová distribúcia chudoby v Európskej únii.

Adresa autora:

Viera Labudová, doc. RNDr. PhD.
Fakulta hospodárskej informatiky,
Ekonomická univerzita v Bratislave
Dolnozemska cesta 1, 852 35 Bratislava
viera.labudova@eubal.sk

Úspěšnost studentů v bakalářském stupni studia The success rate of students in the bachelor study

Bohdan Linda, Michal Kuban

Abstract: Graduates of almost all types of secondary schools study at universities nowadays. The paper analyzes the learning outcome of students of bachelor degree from the perspective of a type of graduated secondary school.

Abstrakt: V současné době na vysokých školách studují absolventi prakticky všech typů středních škol. Příspěvek se zabývá analýzou studijních výsledků studentů bakalářského stupně studia z pohledu absolvovaného typu střední školy.

Key words: bachelor degree, study results, the type of graduated secondary school

Klíčové slová: bakalářské studium, studijní výsledky, typ absolvované střední školy

JEL classification: I21, I29

1. Úvod

V polistopadovém období prošel a stále prochází reformou celý školský systém. V důsledku bezbřehé liberalizace všech stupňů a typů studia započaté v devadesátých letech minulého století se rapidně zhoršila úroveň a kvalita absolventů. V poslední době se konečně začínají ozývat i rozumné hlasy, volající po nápravě. Na rozdíl od období před rokem 1989, kdy na vysokých školách studovalo přibližně 15% populačního ročníku, dnes toto číslo převyšuje 81%. V té době většinu vysokoškoláků tvořili absolventi středních škol gymnaziálního typu, malou část tvořili absolventi středních odborných škol. Absolventi učňovských oborů se dostali na vysokou školu jen výjimečně po doplnění maturity. Dnes na většině vysokých škol je tento poměr absolventů gymnázií a odborných středních škol obrácen a nemalou část dokonce tvoří absolventi učňovských škol. Příspěvek se zabývá rozbořením studijních výsledků studentů bakalářského stupně studia z pohledu absolvovaného typu střední školy

2. Vymezení zkoumaného statistického souboru

Statistické šetření probíhalo na souboru tvořeném studenty, kteří nastupovali do bakalářského studia na Fakultu ekonomicko-správní UPa v roce 2010. K dispozici jsme měli, a tedy i zpracovány byly jejich studijní výsledky počínaje zimním semestrem školního roku 2010/201 a konče zimním semestrem školního roku 2012/2013. Studenti byli rozděleni do tří skupin

- absolventi gymnázií
- absolventi středních odborných škol
- absolventi středních odborných učilišť

Tyto skupiny byly dále děleny na studenty, přijaté v prvním řádném kole přijímacího řízení a studenty, přijaté v dodatečném přijímacím řízení. V prvním kole byli studenti přijímáni na základě výsledků SCIO - testů¹, ve druhém kole na základě studijních výsledků na střední škole. Oba tyto způsoby přijímání studentů ke studiu nemají relevantní vztah ke studijním výsledkům.

¹ O vztahu výsledků SCIO – testů (tzv. předpokladů ke studiu) a studijních výsledků viz článek Kubanová, J.-Linda, B.: Learning Potential Tests and Study Success. In: *Proceedings of 9th International Conference Efficiency and Responsibility in Education*. Praha 2012, s. 279 - 286.

Statistický soubor tvořilo 920 studentů, z nichž bylo:

246 absolventů gymnázií přijatých na základě výsledků SCIO testů

32 absolventů gymnázií přijatých na základě studijních výsledků

455 absolventů středních odborných škol přijatých na základě výsledků SCIO testů

107 absolventů středních odborných škol přijatých na základě studijních výsledků

61 absolventů odborných učilišť přijatých na základě výsledků SCIO testů

19 absolventů odborných učilišť přijatých na základě studijních výsledků

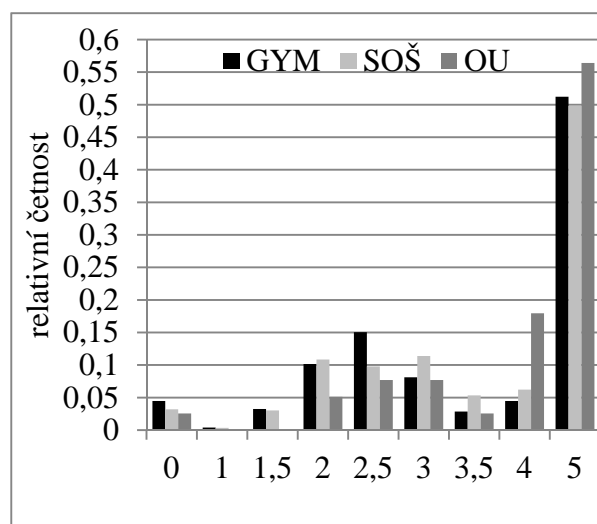
Studenti, přijatí na základě studijních výsledků ze středoškolského studia absolvovali dodatečné přijímací řízení.

3. Výsledky

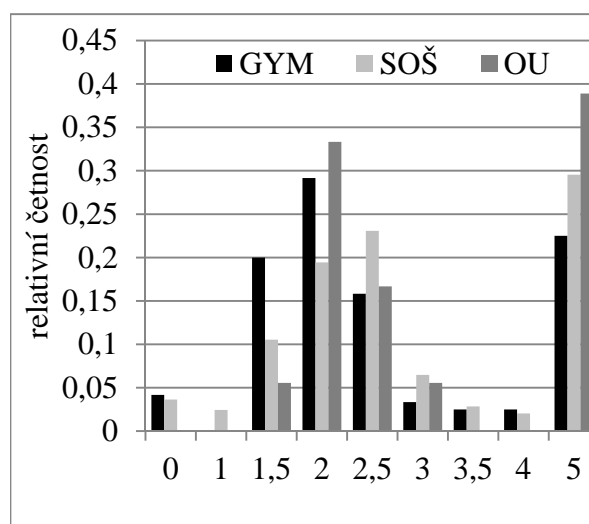
Výsledky jsou prezentovány formou histogramů relativních četností, srovnávajících studijní průměry studentů v daném semestru. V jednom obrázku jsou vždy zobrazeny histogramy studijních výsledků studentů všech tří skupin. Studijní průměry jsou zařazeny do kategorií 0; 1; 1,5; 2; 2,5; 3; 3,5; 4; 5. Kategorie 0 představuje studenty, kteří v daném semestru nevykonali žádnou zkoušku, ale z nějakých důvodů buď mohli pok

+račovat ve studiu, anebo v daném semestru ukončili studium. Vzhledem k tehdy používanému systému vedení studijní agendy se studenti, kteří v daném roce ukončili studium (bez ohledu na semestr) objeví až v zimním semestru příštího roku, kde jsou zařazeni v kategorii 5. To je zaznamenáno v histogramech zimních semestrů. Ostatní kategorie představují studijní průměry, zaokrouhlené na dané číslo (např. kategorie 2,5 představuje průměry 2,26 – 2,75). Z úsporných důvodů jsou v příspěvku uvedeny pouze histogramy zimních semestrů, ve kterých vystupují i studenti, kteří po ukončení předchozího školního roku zanechali studia (tj. zimní semestr 2. ročníku a zimní semestr 3. ročníku).

Obrázky 1 – 2 představují histogramy relativních četností studijních průměrů studentů jednotlivých skupin, přijatých do studia v řádném přijímacím řízení.



Obr. 12: Studijní průměry 2. roč. ZS
Řádné přijímací řízení

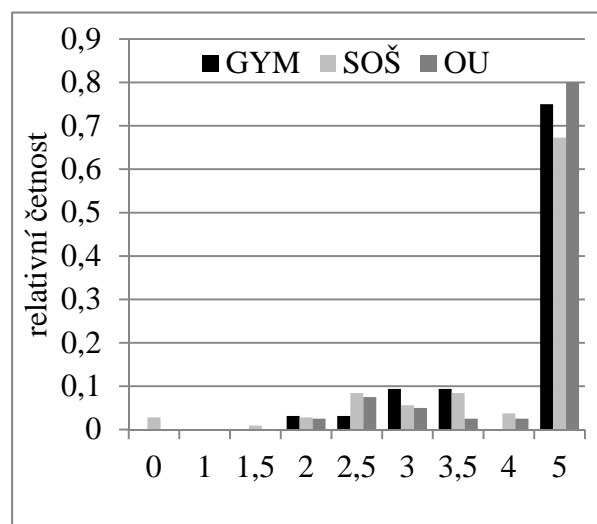


Obr. 2: Studijní průměry 3. roč. ZS
Řádné přijímací řízení

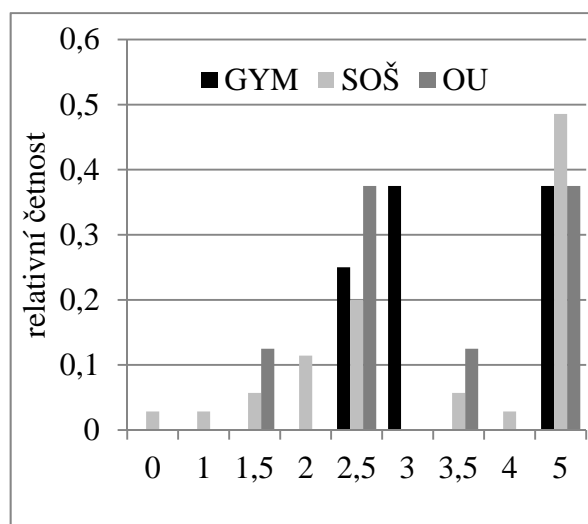
Uvedené grafy podávají zajímavé, bohužel ne potěšující informace. Obrázek 1. nám říká, že v průběhu 1. ročníku zanechá studium přibližně polovina studentů. Konkrétně u absolventů gymnaziálního studia je to 51,2%, u absolventů středních odborných škol 49,8% a u

absolventů odborných učilišť 56,4%. Zajímavé na těchto výsledcích je, že podíl studentů, kteří byli nuceni zanechat studium po prvním roku studia je ve všech skupinách přibližně stejný. Dalo by se totiž očekávat, že tento podíl bude nejmenší u absolventů gymnázií a největší u absolventů odborných učilišť. Důvodů může být několik, jako například, že se podstatně snížila připravenost gymnazistů na vysokoškolské studium. Absolventům všech typů škol dělají problémy předměty teoretického základu, především matematika a statistika. Tento jev je však všeobecným problémem a proto jsou snahy opětovně zavést matematiku jako povinný maturitní předmět. Další důvod může být ten, že tito studenti nemají předpoklady zvládnout vysokoškolské studium. Jistou část mezi těmito studenty představují „studenti“, kteří nepřišli na vysokou školu studovat, ale získat pouze statut studenta a s ním všechny výhody pro podnikání, které provozují jako svoji hlavní činnost. Obrázek druhý nás informuje, že ještě v průběhu druhého ročníku zanechá studia poměrně velký počet studentů, avšak podstatně menší, než po prvním roku studia. Zde se již projevují větší rozdíly mezi studenty, kteří absolvovali různé typy středních škol. Z absolventů gymnázií zanechává po druhém ročníku studium 22,5% studentů, u absolventů středních odborných škol je to 29,6% a u absolventů odborných učilišť 38,9%. Tyto rozdíly lze částečně vysvětlit tím, že do konce druhého ročníku studenti musí úspěšně absolvovat zkoušky z matematiky a zde se může projevit i skutečnost, že absolventi gymnázií v průběhu středoškolského studia mohli získat lepší matematické dovednosti. Co se týče studijních průměrů, v prvních třech semestrech není podstatný rozdíl mezi absolventy gymnázií a středních odborných škol. Horších studijních výsledků však dosahují absolventi učilišť. Významnější rozdíly se začínají projevovat od letního semestru druhého ročníku. V těchto semestrech absolventi gymnázií dosahují zřetelně lepších výsledků (viz. např. obr. 2).

Obrázky 3. a 4. představují stejné histogramy jako obrázky 1. a 2. s tím rozdílem, že se týkají studentů přijatých v dodatečném přijímacím řízení.



Obr. 3: Studijní průměry 2.roč. ZS
Dodatečné přijímací řízení



Obr.3: Studijní průměry 3.roč. ZS
Dodatečné přijímací řízení

Z obrázků 3. a 4. je patrné, že ze studentů, přijatých do studia v dodatečném přijímacím řízení většina zanechala studia v průběhu prvního ročníku (75% absolventů gymnázií, 67,3% absolventů středních odborných škol a 80% absolventů odborných učilišť). Po druhém roce studia tyto podíly jsou 37,5% u gymnazistů, 48,5% u absolventů středních odborných škol a 37,5% u absolventů učilišť. Taktéž studijní průměry u těchto studentů jsou podstatně horší, než u studentů přijatých na základě řádného přijímacího řízení.

Vzhledem k tomu, že statistický soubor tvořilo 920 studentů, lze závěry pokládat za relevantní. Ostatně podobné výsledky získali autoři i ze statistického souboru tvořeného studenty, kteří nastoupili ke studiu v roce 2009.

4. Závěr

Z uvedené analýzy plyne, že bude potřebné provést reformu celého školského systému. Na vysokou školu se hlásí studenti, které střední škola nepřipravila dostatečně na vysokoškolské studium. To se projevuje ve špatném studijním průměru, který u většiny studentů je v rozmezí 2 - 3 a ve vysokém počtu studentů, zanechávajících studia pře jeho dokončením. Především u studentů přijatých v dodatečném přijímacím řízení jsou tyto počty alarmující. Ze 178 takto přijatých studentů se do posledního semestru studia dostalo pouze 51 studentů. Bohužel systém financování vysokých škol v České republice nutí fakulty přijímat i takovéto studenty.

Literatura

KUBANOVÁ, J. 2008. Statistické metody pro ekonomickou a technickou praxi. Bratislava: Statis.

KUBANOVÁ, J. - LINDA, B. 2012. Learning Potential Tests and Study Success. In: *Proceedings of 9th International Conference Efficiency and Responsibility in Education*. Praha, s. 279 - 286.

Potvrzení

Příspěvek byl vypracován v rámci projektu Studentské grantové soutěže Univerzity Pardubice SGSFES_2014003

Adresy autorů:

Bohdan Linda, doc. RNDr. CsC.
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 95, 532 10 Pardubice
bohdan.linda@upce.cz

Michal Kuban
Fakulta ekonomicko-správní
Univerzita Pardubice
Studentská 95, 532 10 Pardubice
bohdan.linda@upce.cz

Modelovanie pozitívnych vzťahov na báze stochastického agentovo orientovaného modelu

Positive relationships modelling based on stochastic agent-based model

Radka Lipovská, Peter Tóth

Abstract: Stochastic Actor Oriented Model (SAOM) allows for modelling dynamics of social networks. This study presents basic functioning of SAOM on an example of positive friendship ties between peers in school classroom. Aim of the study is to identify significant effects in dynamics of these ties. We consider modeling of relationship dynamics as important, because of its potential and eventual importance also in economics.

Abstrakt: Stochastický agentovo orientovaný model (Stochastic Actor Oriented Model - SAOM) umožňuje modelovať dynamiku sociálnych sietí. Táto práca popisuje základné fungovanie SAOM na príklade pozitívnych priateľských vzťahov medzi žiakmi v školských triedach. Cieľom práce je identifikovať významné efekty, ktoré pôsobia na dynamiku týchto vzťahov. Modelovanie dynamiky vzťahov považujeme za dôležité, pretože má potenciál a môže mať veľký význam aj v ekonomike.

Key words: social networks, network modelling, friendship dynamics, school classroom

Kľúčové slová: sociálne siete, modelovanie siete, dynamika priateľských väzieb, školská trieda

JEL classification: C01, C12, C13

1. Úvod

Medziľudské vzťahy nikdy neboli nemenné, neustále sa vyvíjajú. Aj medzi ekonomickými subjektmi dochádza k rôznym interakciám, preto má význam modelovať dynamiku vzťahov. V tejto práci modelujeme dynamiku priateľských vzťahov pomocou stochastického agentovo orientovaného modelu – SAOM (Snijders et al., 2010). Naším cieľom je tento model priblížiť a predviesť jeho fungovanie na príklade priateľských vzťahov v školských triedach. Hoci dynamika priateľských vzťahov je skúmaná často (Snijders et al., 2010; Knecht, 2008; Espelage et al., 2007), v našich podmienkach nepatrí medzi často využívané metódy. V ekonomike dochádza k rôznym interakciám medzi ekonomickými subjektmi, preto môže táto metóda nájsť široké uplatnenie na ich modelovanie.

SAOM pracuje s panelovými dátami – pozorovania jednej siete v určitých časových intervaloch. Prostredníctvom nich skúma v sociálnej sieti štrukturálne efekty. Umožňuje hodnotiť vplyv endogénnych a exogénnych efektov, ktoré sú modelované v priebehu času. Endogénne efekty vypovedajú o preferencii štrukturálnych efektov (napr. či bude vo vzťahoch preferovaná reciprocita) a exogénne efekty berú do úvahy napr. to, ako súvisia charakteristiky aktérov s dynamikou siete (Ripley et al., 2014). Dôležité je, že jednotlivé pozorovania sú považované za diskrétno pozorovania procesu, ktorý prebieha kontinuálne. V tomto procese aktéri robia zmeny, ktoré nie sú pozorované medzi jednotlivými meraniami a zároveň tým vytvárajú premenlivé prostredie jeden pre druhého. Všetky zmeny, ktoré sa v sieti udejú sú rozložené na veľmi malé *mini kroky*. V jednom mini kroku aktér buď vytvorí alebo ukončí práve jeden vzťah (alebo stav zachová a nič nezmení). Evolúcia siete je teda modelovaná ako rozhodnutia (voľby) aktérov vo vzťahoch, ktoré vysielajú ku druhým. Každý jeden ďalší mini krok zmení stav siete. Práve tieto zmeny sú simulované, nie je možné ich pozorovať. Zmenu, ktorú je možné pozorovať, resp. merať je stav siete, ktorý získame ďalšou vlnou meraní (Ripley et al., 2014). Pre tieto modely platia určité pravidlá, ktoré sú podrobne

opísané v Snijders et al. (2010). Podrobný návod celej metódy poskytuje manuál, ktorý je priebežne aktualizovaný (Ripley et al., 2014)

2. Vzťahy v triede a sebahodnotenie

V školskej triede vytvárajú deti skupinu najprv na základe pôsobenia vonkajších faktorov (Vágnerová, 2005). Postupom času sa interakcie menia a vzniká štruktúra neformálnych vzťahov. Vývoj neformálnych vzťahov podlieha rôznym vplyvom a je možné odhaliť rôzne tendencie, na základe ktorých sa vzťahy vyvíjajú. Pomocou SAOM bolo zistené, že priateľské vzťahy v triede sú opätované a platí pre nich lokálna hierarchická štruktúra. Okrem toho v tejto vekovej kategórii platí, že deti si vyberajú najlepších kamarátov na základe podobnosti v rode (napr. Snijders et al., 2010; Knecht, 2007, Ojanen, 2013). Naším zámerom je zistiť, či aj naše dáta poskytujú dôkaz o pôsobení niektorých štruktúrnych efektov (sú podrobnejšie opísané v 4. kapitole) na vývoj priateľských vzťahov. Taktiež zisťujeme, či rod bude zohrávať významnú úlohu v dynamike vzťahov. Poslednú premennú, ktorú sme do modelu pridali, je sebahodnotenie žiakov.

3. Model

Dynamiku sociálnych sietí v triedach modelujeme pomocou programu na štatistickú analýzu sieťových dát SIENA (Simulation Investigation for Empirical Network Analysis), ktorý je v prostredí R implementovaný v balíku RSiena. Cieľovú funkciu vyjadruje rovnica (1), kde $f_i^{net}(x)$ predstavuje cieľovú funkciu, β_k^{net} sú parametre a $s_{ik}^{net}(x)$ sú efekty zahrnuté do cieľovej funkcie.

$$f_i^{net}(x) = \sum_k \beta_k^{net} s_{ik}^{net}(x) \quad (1)$$

V našom modeli odhadujeme parametre pre 5 štruktúrnych efektov. Prvým použitým efektom je *reciprocity effect*, ktorý je definovaný ako počet recipročných vzťahov a matematicky je zapísaný pomocou vzťahu (2), kde $x_{ij} = 1$ ak existuje vzťah medzi i-tou a j-tou osobou a $x_{ij} = 0$, ak takýto vzťah neexistuje. Druhým je *transitive triplets effect*, ktorý je definovaný počtom trojíc. Efekt vypovedá o tendencii ku tranzitivite, alebo zhlukovaniu. Inými slovami, priatelia našich priateľov sa stanú našimi priateľmi. V prípade orientovaných sietí je možné tento efekt zapísať pomocou vzťahu (3). Tretím efektom je *number of three-cycles*, ktorý je považovaný za zovšeobecnenú verziu reciprocity a vyjadruje ho rovnica (4). Obidva efekty reprezentujú uzavreté štruktúry, avšak *transitive triplets effect* poukazuje na hierarchické usporiadanie a *three-cycles* na anti-hierarchické. Ďalším použitým efektom je *transitive ties effect*, ktorý vyjadruje počet osôb, ku ktorým sa i-ta osoba môže dostať priamo alebo nepriamo. Podobne ako *transitive triplets effect*, aj tento efekt meria tranzitivitu. Matematicky je zapísaný pomocou vzťahu (5). Do modelu sme ďalej zahrnuli *in-degree related popularity effect*. Tento efekt vyjadruje popularitu osôb, s ktorými má i-ta osoba vzťah. Popularita je vyjadrená ako počet osôb, ktoré majú vzťah s osobami, s ktorými má vzťah i-ta osoba. Matematicky je vyjadrený rovnicou (6).

$$s_{i2}^{net}(x) = \sum_j x_{ij} x_{ji} \quad (2)$$

$$s_{i3}^{net}(x) = \sum_{j,h} x_{ij} x_{ih} x_{jh} \quad (3)$$

$$s_{i6}^{net}(x) = \sum_{j,h} x_{ij} x_{jh} x_{hi} \quad (4)$$

$$s_{i9}^{net}(x) = \sum_j x_{ij} \max_h (x_{ih} x_{hj}) \quad (5)$$

$$s_{i18}^{net}(x) = \sum_j x_{ij} x_{+j} = \sum_j x_{ij} \sum_h x_{hj} \quad (6)$$

V modeli ďalej odhadujeme parametre pre 3 efekty súvisiace s nezávislými premennými. Prvým efektom je *covariate-related popularity*, nazývaný taktiež *ego effect*. Predstavuje sumu hodnôt nezávislej premennej všetkých osôb, s ktorými má i-ta osoba vzťah. Meria to, či aktéri

s vyššou hodnotou danej nezávislej premennej, majú tendenciu nominovať ostatných aktérov siete častejšie. Matematicky je tento efekt vyjadrený pomocou rovnice (7). Ďalej sme použili efekt *covariate-related activity*, alebo ináč *alter effect*. Je definovaný ako počet vzťahov i -tej osoby vážený jej hodnotou danej nezávislej premennej. Zisťuje, či aktéri s vyššou hodnotou budú nominovaní ostatnými aktérmi siete častejšie. Tento efekt vyjadruje rovnica (8). Posledným efektom je *covariate-related similarity*, alebo jednoducho *similarity effect*. Vyjadruje do akej miery si agenti vyberajú partnerov s podobnými vlastnosťami. Efekt je definovaný ako suma centrovaných podobností sim_{ij}^v medzi i -tym agentom a iným j -tymi agentmi, s ktorými má vzťah. Počíta sa použitím vzťahov (9), kde Δ predstavuje rozsah pozorovaných hodnôt nezávislej premennej.

$$s_{i49}^{net}(x) = \sum_j x_{ij} v_j \quad (7)$$

$$s_{i51}^{net}(x) = v_i x_{i+} \quad (8)$$

$$s_{i52}^{net}(x) = \sum_j x_{ij} (sim_{ij}^v - \widehat{sim}^v), \quad sim_{ij}^v = \frac{\Delta - |v_i - v_j|}{\Delta}, \quad \Delta = \max_{ij} |v_i - v_j| \quad (9)$$

Okrem uvedených efektov je v modeli automaticky testovaný efekt *out-degree effect*, ktorý je vyjadrený ako počet vzťahov i -teho agenta a počíta sa pomocou rovnice (10). Vyjadruje tendenciu mať vzťahy (Snijders et al., 2010).

$$s_{i1}^{net}(x) = x_{i+} = \sum_j x_{ij} \quad (10)$$

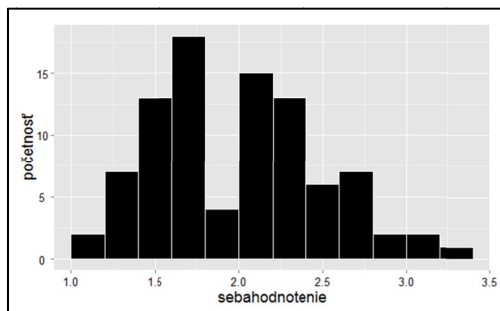
Model je potom možné vyjadriť rovnicou (11), kde efekty sú počítané pomocou vzťahov (2), (3), (4), (5), (6), (7), (8), (9) a (10).

$$f_i^{net}(x) = \beta_1^{net} s_{i1}^{net} + \beta_2^{net} s_{i2}^{net}(x) + \beta_3^{net} s_{i3}^{net}(x) + \beta_6^{net} s_{i6}^{net}(x) + \beta_9^{net} s_{i9}^{net}(x) + \beta_{18}^{net} s_{i18}^{net}(x) + \beta_{49}^{net} s_{i49}^{net}(x) + \beta_{51}^{net} s_{i51}^{net}(x) + \beta_{52}^{net} s_{i52}^{net}(x) \quad (11)$$

4. Dáta

Pre odhad modelu boli použité dáta z piatich tried žiakov piateho ročníka. Vzorku tvorilo celkovo 103 žiakov, z čoho je 47 chlapcov a 56 dievčat. Pozorovania pozostávali z 3 vln: t_1 = október piateho ročníka, t_2 = apríl piateho ročníka, a t_3 = október šiesteho ročníka. Dáta boli zbierané v triede počas školskej hodiny.

Siete pozitívnych vzťahov. Údaje o pozitívnych väzbách boli získané pomocou grafického sociometrického dotazníka. Úlohou detí bolo zapísať do dotazníka mená všetkých svojich spolužiakov podľa toho, ako veľmi sa kamarátia, resp. ako veľmi sú si blízki. Najlepší kamaráti boli najbližšie. Z týchto údajov sme následne vytvorili maticu susedností nasledovným spôsobom: 1 = najlepší kamarát, 0 = nie je prítomný kamarátsky vzťah.



Obr. 13: Histogram sebahodnotenia žiakov

Sebahodnotenie. Súčasťou dát je aj sebahodnotenie žiakov, ktoré sme získali pomocou Rosenbergovej škály sebahodnotenia. Žiaci jednotlivým výrokom priradzovali váhu od 1 po 4 (1 najvyššie/pozitívne sebahodnotenie, 4 najnižšie/negatívne sebahodnotenie). Následne sa vypočítala priemerná hodnota pre každého žiaka. Stredná hodnota sebahodnotenia je 1,991. Výskyt jednotlivých hodnôt sebahodnotenia je zobrazený na obrázku Obr.1.

V tabuľke Tab. 1 sú uvedené základné štatistické ukazovatele použitých konštantných nezávislých premenných. Z tabuľky vidíme, že v troch triedach je prevaha dievčat, v jednej triede je počet chlapcov a dievčat vyrovnaný a v jednej triede prevládajú chlapci. Hodnoty, ktoré poukazujú na najnižšie zistené priemerné sebahodnotenie boli zistené u žiakov v triede A (1=vysoké sebahodnotenie; 4=nízke sebahodnotenie). Súčasne sú v tejto triede najväčšie rozdiely v sebahodnotení medzi žiakmi. Hodnoty, ktoré vypovedajú o najvyššom priemernom sebahodnotení boli namerané u žiakov v triede D. Najmenší rozptyl, tzn. žiaci sú si najviac podobný bol v triede B. Zo šikmosti vidíme, že v triedach A, C a D sa väčšina hodnôt nachádza naľavo od priemeru a v triedach B a E má väčšina žiakov nižšie sebahodnotenie ako je stredná hodnota. Vo všetkých prípadoch je sebahodnotenie špicatejšie ako normálne rozdelenie.

Tab. 7: Deskriptívna štatistika konštantných nezávislých premenných

Trieda	Pohlavie		Sebahodnotenie			
	Chlapci	Dievčatá	Stredná hodnota	Rozptyl	Šikmost'	Špicatosť
A	7	7	2,148	0,396	0,511	1,996
B	6	10	2,132	0,167	-0,046	2,307
C	9	19	1,880	0,184	0,643	2,931
D	11	18	1,876	0,237	0,392	2,638
E	14	2	2,120	0,313	-0,212	2,033

V tabuľke Tab. 2 je uvedená hustota skúmaných sietí v troch pozorovaných obdobiach T1, T2 a T3. Hustota je vyjadrená ako podiel existujúcich vzťahov k všetkým možným vzťahom. V triede D existovalo najviac vzťahov v období T2, v ostatných prípadoch to bolo v období T1. Reciprocita uvedená v tabuľke Tab. 2 meria podiel vzájomne opätovaných vzťahov v triedach a dosahuje vysoké hodnoty od 76 % až do 96 %.

Tab. 2: Deskriptívna štatistika skúmaných sietí

Trieda	Hustota			Reciprocita		
	T1	T2	T3	T1	T2	T3
A	0,206	0,199	0,197	0,846	0,763	0,873
B	0,278	0,241	0,253	0,812	0,962	0,770
C	0,142	0,123	0,141	0,906	0,887	0,880
D	0,129	0,152	0,133	0,872	0,893	0,888
E	0,383	0,272	0,220	0,812	0,955	0,871

5. Výsledky

Predpokladáme, že parametre pre jednotlivé triedy sú rovnaké, preto sme matice susedností v jednotlivých triedach spojili do jednej matice susedností. Využili sme štrukturálne nuly, ktoré zabezpečia, že žiaci z rôznych tried medzi sebou nemôžu vytvárať vzťahy. Efekty spojené s konštantnými premennými odhadujeme pre 2 konštantné premenné: pohlavie a sebahodnotenie. Aby sme spresnili odhad parametrov, odhad sme zrealizovali dvakrát, tzn. že parametre odhadnuté prvýkrát sme použili ako počiatočné hodnoty pre druhý odhad.

Pre skúmanie sociálnych sietí v priebehu času je potrebná aj určitá stabilita vo vzťahoch. O nej vypovedá Jaccardov index. Jeho hodnoty majú byť väčšie než 0.3 (Ripley et al., 2014). V našom prípade dosahovali 0,459 (medzi 1. a 2. vlnou) a 0,340 (medzi 2. a 3. vlnou).

V tabuľke Tab.3 sú uvedené odhadnuté parametre pre jednotlivé efekty. Miera rýchlosti poukazuje na príležitosti ku zmene. Výsledok 6,2482 poukazuje na to, že odhadovaný počet príležitostí urobiť zmenu pre každého aktéra medzi 1. a 2. vlnou merania je 6,24. Medzi 2. a 3. vlnou je to 5,29. Pri zvyšných efektoch si ďalej všimáme mieru konvergencie, ktorá nemá presahovať hodnotu 0.1 v absolútnej hodnote. Toto kritérium spĺňajú všetky efekty.

T-test uvedený v tabuľke Tab.3, nazývaný aj Wald-type test, testuje nulovú hypotézu o tom, že daný parameter je rovný nule. Testovacia charakteristika je definovaná ako podiel odhadnutého parametra a smerodajnej odchýlky parametra. V prípade ak je testovacia charakteristika t-testu v absolútnej hodnote väčšia ako 2, na hladine významnosti 0,05 nulovú hypotézu zamietneme, čo znamená, že daný parameter sa nerovná nule a teda je štatisticky významný (Ripley et al.,2014). Všetky parametre okrem 8., 10., 11., a 12. sú významné.

Tab. 3: Odhadnuté parametre, štandardné odchýlky, miera konvergencie a t-test

P. č.	Efekt	Odhadnuté parametre	Štandardná odchýlka	Miera konvergencie t-ratio	t-test
0.1	Miera rýchlosti - perióda 1	6,2482	0,6443		
0.2	Miera rýchlosti – perióda 2	5,2932	0,5864		
1.	Outdegree (density)	-2,0466	0,2504	0,0262	-8,1746
2.	Reciprocity	1,9471	0,2679	0,0317	7,2687
3.	Transitive triplets	0,4513	0,0566	0,0388	7,9707
4.	3-cycles	-0,6002	0,1324	0,0111	-4,5323
5.	Transitive ties	1,0817	0,3925	-0,0245	2,7558
6.	Indegree - popularity	-0,1807	0,0436	-0,0106	-4,1442
7.	Gender alter	0,4058	0,1385	0,0488	2,9290
8.	Gender ego	-0,0416	0,1337	0,0216	-0,3112
9.	Gender similarity	0,6108	0,1309	-0,0172	4,6655
10.	Rosen alter	-0,1604	0,1273	0,0161	-1,2604
11.	Rosen ego	-0,0430	0,1079	-0,0575	-0,3984
12.	Rosen similarity	0,1403	0,3255	0,0281	0,4312

Na základe výsledkov sme potvrdili, že priateľské vzťahy sa nevytvárajú náhodne (negatívny *outdegree e.*). Ďalej, že tieto vzťahy sú pravdepodobnejšie, keď sú vzájomne opätované (pozitívny *reciprocity e.*). Pozitívny *transitive triplets* a *transitive ties* efekt, poukazuje na prítomnosť hierarchickej štruktúry vo vzťahoch. Podľa toho vieme, že žiaci si budú s vyššou pravdepodobnosťou vyberať priateľov svojich priateľov za priateľov. Negatívna hodnota odhadovaného parametra *3-cycles* taktiež vypovedá o hierarchickej štruktúre. Znamená to, že niektorí žiaci získali viac priateľských väzieb, než iní v rámci tej istej triády. Vytváranie vzťahov nebolo štatisticky významne ovplyvnené sebahodnotením žiakov. Čo sa týka rodu, pozitívny *gender similarity* efekt potvrdil, že chlapci si s vyššou pravdepodobnosťou vyberú za najlepších kamarátov chlapcov a dievčatá si vyberú dievčatá. Nakoniec pozitívny *gender alter* efekt indikuje, že chlapci sú častejšie vyberaní druhými žiakmi za kamarátov než dievčatá.

Odhadnutú funkciu vyjadruje rovnica (12), kde index G predstavuje efekty pre pohlavie a index R pre sebahodnotenie.

$$\begin{aligned}
f_i^{net}(x) = & -2,0466s_{i1}^{net}(x) + 1,9471s_{i2}^{net}(x) + 0,4513s_{i3}^{net}(x) - 0,6002s_{i6}^{net}(x) + \\
& + 1,0807s_{i9}^{net}(x) - 0,1807s_{i18}^{net}(x) + 0,4058s_{i49G}^{net}(x) - 0,0416s_{i51G}^{net}(x) + \\
& + 0,6108s_{i52G}^{net}(x) - 0,1604s_{i49R}^{net}(x) - 0,0430s_{i51R}^{net}(x) + 0,1403s_{i52R}^{net}(x)
\end{aligned} \quad (12)$$

Na základe výsledkov t-testu uvedených v tabuľke Tab.3 sme odstránili efekty, ktoré majú hodnotu v absolútnej hodnote nižšiu ako 2 a zopakovali sme odhad. Rovnica (13) obsahuje nové odhadnuté parametre, pre ktoré platí, že majú t-ratio v absolútnej hodnote menšie ako 0,1 a hodnotu t-testu v absolútnej hodnote vyššiu ako 2, tzn. že sú štatisticky významné.

$$\begin{aligned}
f_i^{net}(x) = & -2,0511s_{i1}^{net}(x) + 1,9148s_{i2}^{net}(x) + 0,4417s_{i3}^{net}(x) - 0,5947s_{i6}^{net}(x) + \\
& + 1,1302s_{i9}^{net}(x) - 0,1805s_{i18}^{net}(x) + 0,3337s_{i49G}^{net}(x) + 0,6117s_{i52G}^{net}(x)
\end{aligned} \quad (13)$$

6. Záver

V príspevku sme popísali základné fungovanie stochastického agentovo orientovaného modelu na príklade evolúcie sociálnej siete pozitívnych vzťahov. Pomocou modelovania sme ukázali, ktoré efekty sú významné v sieti priateľských vzťahov. V súčasnosti sa výskum v oblasti analýzy sociálnych sietí zameriava najmä na ko-evolúciu vzťahov v sieti a rôzneho správania (Veenstra et al., 2013). Ide o náročnejšie postupy, však môžu nájsť svoje využitie v rôznych vedných disciplínach.

Literatúra

- ESPELAGE, D.L. – GREEN, H.D.Jr. – WASSERMAN, S. 2007. Statistical Analysis of Friendship Patterns and Bullying Behaviors among Youth. In: *New Dir Child Adolesc Dev*, roč. 118, s. 61-75.
- KNECHT, A. 2008. *Friendship Selection and Friends' Influence*. Dynamics of Networks and Actor Attributes in Early Adolescence. Doctoral dissertation, University of Utrecht.
- OJANEN, T. – SIJTSEMA, J.J. – RAMBARAN, A.J. 2013. Social Goals and Adolescent Friendships: Social Selection, Deselection, and Influence. In: *Journal of Research on Adolescence*, roč. 23, č. 3, s. 550-562.
- RIPLEY, R.M. – SNIJDERS, TOM A.B. – BODA, Z. – VÖRÖS, A. – PRECIADO, P. 2014. *Manual for RSiena version 4.0*. Oxford: University of Oxford, Department of Statistics, Nuffield College. http://www.stats.ox.ac.uk/~snijders/siena/RSiena_Manual.pdf
- SNIJDERS, TOM A.B. - VAN DE BUNT, G.G. – STEGLICH, CH.E.G. 2010. Introduction to Stochastic Actor-Based Models for Network Dynamics. In: *Social Networks*, roč. 32, s.44 – 60.
- STADTFELD, CH. 2012. *Events in Social Networks. A Stochastic Actor-Oriented Framework for Dynamic Even Processes in Social Network*. Karlsruhe, KIT Scientific Publishing.
- VÁGNEROVÁ, M. 2005. *Vývojová Psychologie*. Praha, Portál.
- VEENSTRA, R. - DIJKSTRA, J.K. - STEGLICH, CH. - VAN ZALK, M.H.W. 2013. Network–Behavior Dynamics. In: *Journal of Research on Adolescence*, roč.23, č.3, s. 399-412.

Adresa autora (-ov):

Radka Lipovská, Mgr.
Spoločenskovedný ústav
Slovenská akadémia vied v Košiciach
Karpatská 5, 040 13 Košice
lipovska@saske.sk

Peter Tóth, Ing.
Ekonomická fakulta
Technická univerzita v Košiciach
Němcovej 32, 040 01 Košice
peter.toth@tuke.sk

Odhad správného počtu shluků pro překrývající se data

Estimating the Correct Number of Clusters for Overlapping Data

Elena Makhalova, Kornélia Cséfalvaiová, Iva Pecáková

Abstract: Clustering techniques are mostly unsupervised methods that can be used to organize data into groups based on similarities or dissimilarities among the individual data items. Many clustering algorithms have been proposed, namely, hierarchical clustering, fuzzy clustering, k-means, k-modes, k-prototypes and others. Regardless of the algorithm we choose, there are two important problems, which are still deciding by scientists. The first is to choose the measure of distance and the second one is to estimate the correct number of clusters. In this paper, we would like to introduce the results of our study about estimating the correct number of clusters. The algorithm, we used in our research is a fuzzy clustering algorithm with Euclidean distance. We used generated dataset with different number of clusters.

Abstrakt: Většina shlukovacích metod patří mezi metody «učení bez učitelů» a slouží k uspořádání dat do skupin na základě jejich podobnosti nebo odlišnosti. Existuje mnoho shlukovacích metod, například: hierarchické shlukování, fuzzy shlukování, metoda k-průměru, k-medoidů, k-modů a další. Bez ohledu na zvolenou metodu shlukování existují dvě podstatné otázky, jimiž se stále vědci zabývají. První otázkou je zvolit míru vzdálenosti a druhou je odhadnout správný (vhodný) počet shluků. V tomto článku bychom rádi představili výsledky našeho výzkumu o koeficientech, hodnotících správný počet shluků. Shlukovací metoda, která byla použita v tomto výzkumu, je fuzzy shlukování s Euklidovskou mírou vzdáleností. Pro zkoumání koeficientů byly použity generované datové soubory s různým počtem shluků.

Key words: fuzzy clustering, coefficients, correct number of clusters

Klíčová slova: fuzzy shlukování, koeficienty, vhodný počet shluků

JEL classification: C38

1. Introduction

Clustering has become a widely accepted synonym of a broad array of activities of exploratory data analysis and model development in science, engineering, life sciences, business and economics, defense, and biological and medical disciplines (Valente de Oliveira; J., 2007). Clustering techniques can be used to organize data (numerical or categorical or a mixture of both) into groups based on similarities among the individual data items. In other words, clustering techniques is a tool for discovering previously hidden structure in a data set. There are many clustering algorithms, which can be based on hard or on fuzzy clustering. We choose clustering algorithm with regard to what type of data we have. For instance, most real data, by nature, is overlapping. What does it mean? Using cluster analysis we cannot define the data point with one hundred percent probability for certain cluster. Mostly it happens that the object has the properties as one cluster, but even the properties others clusters, so you could include this object in more than one cluster. And fuzzy clustering allows to do this. Fuzzy clustering algorithm is used for overlapping data. Regardless of what clustering methods we used, hard or fuzzy clustering, we always need to determine the correct number of clusters. There are many different coefficients for estimating the correct number of clusters. Each of these coefficients has its strengths and weaknesses. In our paper we deal with several coefficients for estimating the correct number of clusters (for fuzzy clustering techniques).

Those coefficients are: silhouette coefficient, Dunn coefficient and standardized Dunn coefficient. Before we go to the coefficients, let's look at the basic notions of fuzzy clustering.

2. Fuzzy clustering

Fuzzy clustering allows the data points to belong to several clusters simultaneously, with different membership degrees. Substantially, fuzzy clustering is more natural than hard clustering. Fuzzy clustering is based on the concept of the fuzzy sets, which was introduced by L. Zadeh in 1965. The idea of fuzzy clustering is following: each data point belongs to every cluster with some degree that is specified by a membership grade. It is the main advantage of fuzzy clustering, it allows gradual memberships of data points to clusters measured as degrees in $\langle 0,1 \rangle$. This gives the flexibility to express that data points can belong to more than one cluster, for example:

$$U = \begin{bmatrix} u_{ij} & u_{ij} & u_{ij} & u_{ij} & u_{ij} \\ 0,7 & 0,1 & 0,05 & 0,05 & 0,1 \\ 0,2 & 0,2 & 0,2 & 0,3 & 0,1 \\ 0,3 & 0,13 & 0,05 & 0,02 & 0,5 \\ 0,5 & 0,3 & 0,03 & 0,07 & 0,1 \end{bmatrix}$$

The matrix U whose factors are the ones taken from the membership functions presents the data for fuzzy clustering. In our case the data point can belong to more than one cluster with degree of membership between data points and the centers of clusters.

Membership degrees can also express how ambiguously or definitely a data point should belong to a cluster. So, this partitioning is carried out through an iterative optimization of the objective function J_m , with the update of membership u_{ij} and the cluster centers c_j . This function is shown below:

$$J_m = \sum_{i=1}^N \sum_{j=1}^c u_{ij}^m \|x_i - c_j\|^2, 1 \leq m \leq \infty. \quad (1)$$

Larger membership values indicate higher confidence in the assignment of the pattern to the cluster. Fuzzy partitioning is carried out through an iterative optimization of the objective function (1), with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad (2)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}. \quad (3)$$

This iteration will stop when $\max_{ij} \{ |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \} < \epsilon$, where ϵ is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum of J_m .

3. Coefficients for determining the correct number of clusters

Coefficients for determining the correct number of clusters (cluster validity indexes) can help us to evaluate how good clustering results are. There are many cluster validity indexes that have been proposed in the literature for evaluating fuzzy and other clustering techniques.

In this part of the paper, we will research some of these indexes on the generated data sets, their pros and cons.

The basic validity index associated with the fuzzy clustering is Dunn's coefficient defined as

$$PC = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^c u_{ij}^2. \quad (4)$$

Here $u_{ij} \in \langle 0,1 \rangle$ is a membership degree of an object i to a cluster j . This index assumes only the compactness measurement for each cluster and for the data structure. It is obviously by a lack connection with the geometric structure of data. According to the theory of fuzzy sets, the sum of memberships of every object to all clusters is 1. Consequently, with the increasing number of clusters, the single degrees of cluster membership have a decreasing value. Squaring the membership's degrees, we obtain even smaller values. So, with the increasing number of clusters the value of this coefficient is decreasing.

The next validity index was proposed by Dave as a modification of the previous one:

$$PC_{mod}(k) = 1 - \frac{k}{k-1} (1 - PC(k)) \quad (5)$$

This index can take values from $\langle 0,1 \rangle$, where k is the appropriate number of clusters. Obviously, with increasing number of clusters PC_{mod} is on the decrease. This cluster number k is defined by solving of

$$\max_{2 \leq k \leq n} PC_{mod}(k). \quad (6)$$

When the variability in clusters is small, this normalized Dunn's coefficient (PC_{mod}) usually determined the number of clusters correctly (Řezanková H., Húsek D., 2012). When the cluster variability is greater, the normalized Dunn's coefficient usually achieved its highest value for the highest possible number of clusters. (Řezanková H., Húsek D., 2012).

The total average silhouette coefficient SC is the most complicated validity index for fuzzy clustering mentioned here. This coefficient can determine the compactness and separation degree for the whole data structure, not only for each cluster. The silhouette coefficient for each point determines how that point is similar to points in its own cluster compared to points in other clusters; it ranges from -1 to $+1$. The SC coefficient combines ideas of both *cohesion*, which is the sum of the weight of all links within a cluster, and *separation*, which is the sum of the weights between points in the cluster and points outside the cluster, but for individual points, as well as clusters.

For an individual point i , the silhouette coefficient is based on two measures: on the average distance between i and every points in its cluster A_{iC_i} and on the minimal average distance between i and points in other clusters (Rezaee, B., (2010)). The total average silhouette coefficient is than

$$SC = \frac{1}{n} \sum_{i=1}^n \frac{\min(A_{ij}, j \in C_{-i}) - A_{iC_i}}{\max(\min(A_{ij}, j \in C_{-i}), A_{iC_i})}. \quad (7)$$

Here C_{-i} denotes cluster labels which do not include the case i as a member, while C_i is the cluster label which includes the case i . If $\max(\min(A_{ij}, j \in C_{-i}), A_{iC_i})$ equals 0, the silhouette coefficient of case i is not used in the average operations. The average SC over all data of a cluster is a measure of how tightly grouped all the data in the cluster are. Thus the average SC over all data of the entire dataset is a measure of how appropriately the data has been clustered. If there are too many or too few clusters, as may occur when a poor choice of k is used in the k-means algorithm for example, some of the clusters will typically display much narrower silhouettes than the rest. The silhouette coefficient is typically between 0 and 1. When the silhouette coefficient is closer to 1, it means the best clustering result. When the

variability in clusters is small, the *SC* usually helps to determine the number of clusters correctly (Řezanková H., Húsek D., 2012).

As we can see, all indexes have drawbacks with evaluation of clustering results in a large number of clusters and with increasing variability of data. They do not solve the problem of identifying of the correct number of clusters.

These drawbacks are evident also in results of the analysis of the generated datasets. (For example, both of the indexes (7) and (4) – *SC* and Dunn's coefficient – for the Iris dataset show that the best number of clusters is two (instead of the correct number of three). The silhouette coefficient evaluates one-element clusters by zero. That is why the silhouette coefficient achieves its highest value for two clusters in analyzed data (Zalik, K.R., 2011).

This is due to the fact that the clusters are overlapping and Dunn's coefficient is not able to recognize the correct structure of the clusters. A similar situation is observed in the evaluation of data clustering Vowel and Glass, where the coefficients do not determine the number of clusters correctly. And like we can observe PC_{mod} not always correctly identifies the right numbers of clusters. (Řezanková, Húsek, 2012)

Our task is to propose an alternative coefficient that will work better with the increasing variability of data and with the different number of clusters.

We introduce a next modified approach: we combine two components into one index; in doing so, we use the harmonic mean. Why we use the harmonic mean will be explained below. One of these components is based on fuzzy clustering theory and the other one is based on hard clustering theory. The theory of fuzzy clustering is based on the assumption that each object belongs to each cluster with a membership degree u_{ij} . The hard clustering theory is based on the assumption that each object belongs to one cluster, the average distance from the cluster centre and points of this cluster should be minimal.

Joining two elements based on different approaches into one index helps to reduce disadvantages of both. Let him be the first element Dunn's coefficient (4).

We can distinguish two extreme situations:

- 1) completely fuzzy clustering, where all $u_{ij} = 1/c$ and then $PC = 1/c$;
- 2) hard clustering, where one $u_{ij} = 1$, all others $u_{ij} = 0$ and then $PC = 1$.

The second element is based on the hard clustering theory: we calculate the ratio of the distance minimum in case of c clusters to the distance minimum in case of a single cluster,

$$C = \frac{\sum d_{c,\min}}{d_{1,\min}}. \quad (8)$$

Here $\sum d_{c,\min}$ is the minimal sum on the Euclidean distances between points in the case of c clusters; and $d_{1,\min}$ is the minimal sum on the Euclidean distances between points in the case of a single cluster (when the dataset is one cluster, it means before clustering). The sum of $d_{c,\min}$ should be minimal for the best clustering, it means the minimal value of C should achieve the minimum for the best clustering.

We want to combine two parts in the aggregate function: one of them is Dunn's coefficient (the maximum value for the best clustering), and the second one should also strive to maximum; that's why we use $1 - C = N$ (which achieves the maximum value for the best clustering).

And now the optimization problem is to solve. It can be represented in the following way:

$$E = 1 - \frac{2}{\frac{1}{PC} - \frac{1}{N}} \rightarrow \min \quad (9)$$

This function tends to its minimum for the best clustering, because the inverse values of the indexes of PC and N receive its maximum for the best clustering.

An optimization problem consists of minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function. The allowed set we can describe as:

$$E = \left\{ \begin{array}{l} \frac{1}{PC} + \frac{1}{N} \neq 0 \\ PC \neq 0 \\ N \neq 0 \\ N \neq 1 \\ PC \in \langle 0,1 \rangle, N \in (0,1) \end{array} \right. \quad (10)$$

In the case that $N = 0$ we have a dataset as a single cluster.

4. Simulations of the overlapping data

Now, we present how the coefficients work in six generated datasets with different number of clusters. Every cluster contains 70 data points (thus, the total number of data points depends on the number of clusters). The number of attributes is five for every dataset (numeric, predictive attributes). The clusters are overlapping. The dataset with 6 and 8 clusters have greater degree of overlap than others in our research.

The task of our experiments is to estimate the correct number of clusters in the actual datasets with the help of mentioned coefficients. We applied the k-means fuzzy clustering algorithm with Euclidean distance on these six different data sets. And we applied fuzzy k-means clustering to calculate the degree of belonging to clusters for every case. The reason to choose these approaches is because they are simple and well known by the scientific community.

For finding out the correct number of clusters, we calculated the PC (4) and PC_{mod} (5) coefficients, which is grounded on the membership degrees, the SC coefficient (7) based on the hard clustering technique and E coefficient (9), which includes hard clustering (distances between points) and fuzzy clustering (memberships degrees) basis.

Now the results of the modified approach were compared with SC , PC_{mod} and PC coefficient (tables 1 – 6).

As we can see with increasing number of clusters coefficients SC i PC no longer indicate the correct number of clusters. Coefficient SC indicates that the right number of clusters is 3 instead the 4 is the right (table 2), in table 4 – SC indicates 5 clusters, instead the right number is 6, in table 6 indicates 7 instead of 8 is correct. PC coefficient does not work correctly on the whole: incorrectly identified clusters for the next cases: table 2 (instead 4, indicated 2), table 3 (instead 5, indicated 2), table 4 (instead 6, indicated 2) and table 6 (instead 8, indicated 2). Coefficients PC_{mod} i E determine the correct number for all presented cases.

5. Conclusions

On the basis of exploring cases with the different number of clusters we can provide the following conclusions:

1. For the cases with small number of clusters all the coefficients can usually define the right number of clusters correctly. With increasing number of clusters the estimation accuracy decreases and mostly SC and PC coefficients define the number of clusters incorrectly.
2. As can be seen from the experiments SC and PC coefficients work correctly only with a small number of clusters. PC_{mod} and E coefficients work correctly with different number of clusters.
3. It follows from the surveys, with simultaneously increasing degree of overlap and increasing number of clusters SC and PC coefficients define the number of clusters

incorrectly and PC_{mod} and E coefficients still define the number of clusters correctly. Further surveys of those coefficients will continue in future articles.

Tab. 8: Dataset with 3 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,5468	0,7878	0,5756	0,5141
3	0,6953	0,8469	0,7704	0,2588
4	0,5173	0,6839	0,5786	0,3227
5	0,518	0,5967	0,4958	0,3623
6	0,5073	0,4485	0,3382	0,4565
7	0,3385	0,3607	0,2541	0,5217
8	0,1723	0,3756	0,2864	0,5055
9	0,1622	0,3379	0,2551	0,5384
10	0,1891	0,3206	0,2451	0,5531

Tab. 2: Dataset with 4 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,2630	0,7182	0,4364	0,7188
3	0,4697	0,6514	0,4772	0,4708
4	0,3637	0,6888	0,5850	0,4041
5	0,4902	0,5672	0,4590	0,4075
6	0,4824	0,4872	0,3846	0,4513
7	0,3511	0,4062	0,3072	0,5044
8	0,3385	0,3679	0,2776	0,5326
9	0,2825	0,3306	0,2469	0,5586
10	0,2495	0,2956	0,2173	0,5908

Tab. 3: Dataset with 5 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,3910	0,6872	0,3745	0,6785
3	0,4660	0,6582	0,4873	0,4804
4	0,5703	0,6712	0,5616	0,4009
5	0,5767	0,6619	0,5773	0,3990
6	0,5030	0,5702	0,4843	0,4350
7	0,4915	0,0381	0,1223	0,9286
8	0,4017	0,4391	0,3589	0,5006
9	0,3840	0,3837	0,3067	0,5380
10	0,2399	0,3436	0,2707	0,5628

Tab. 4: Dataset with 6 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,3456	0,7261	0,4523	0,6659
3	0,3799	0,6022	0,4033	0,5289
4	0,3611	0,5994	0,4659	0,6376
5	0,4605	0,5163	0,3954	0,4596
6	0,3867	0,4976	0,4971	0,4466
7	0,3496	0,4289	0,3338	0,5099
8	0,3321	0,3794	0,2907	0,5276
9	0,3234	0,3419	0,2596	0,5562
10	0,3234	0,3027	0,2253	0,5891

Tab. 5: Dataset with 7 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,2541	0,5328	0,0656	0,7554
3	0,3616	0,4742	0,2113	0,6350
4	0,4638	0,5012	0,3349	0,5421
5	0,4088	0,5517	0,4397	0,4883
6	0,5504	0,5648	0,4778	0,4148
7	0,5814	0,5691	0,4973	0,3859
8	0,4308	0,5019	0,4307	0,4801
9	0,4480	0,4477	0,3786	0,4599
10	0,4490	0,4105	0,3450	0,4873

Tab. 6: Dataset with 8 clusters

The name of the coefficient	SC	PC	PC _{mod}	E
The number of clusters				
2	0,4146	0,6749	0,3498	0,6369
3	0,3952	0,6056	0,4084	0,5510
4	0,4495	0,5571	0,4095	0,4875
5	0,2928	0,5776	0,4721	0,5068
6	0,4381	0,6142	0,5271	0,4002
7	0,5761	0,5972	0,5300	0,3527
8	0,5103	0,5411	0,5355	0,3444
9	0,4272	0,4891	0,4252	0,4310
10	0,4506	0,4378	0,3754	0,4546

References

- GOJUN, G.– CHAOGUN, M., JIANHONG, W. (2007): *Data Clustering theory, algorithms, and applications*. : Siam, Society for Industrial and Applied Mathematics
- HÖPPNER, F., KLWONN, F., RUNKLER, T. (1999): *Fuzzy Cluster Analysis*. London: Wiley
- KRUSE, R., DÖRING, C., LESOT, M., J. (2007): *Fundamentals of fuzzy clustering, advances in fuzzy clustering and its applications*. London: Wiley.
- VALENTIE DE OLIVIERA, J. (2007): *Advances in Fuzzy clustering and its applications*. London: Wiley
- BRODOWSKI, S. (2011): A Validity Criterion for Fuzzy Clustering. *Computational collective intelligence: technologies and applications*. Berlin: Springer-Verlag Berlin
- LÖSTER, T., PAVELKA, T. (2013): Evaluating of the Results of Clustering in Practical Economic Tasks. *International Days of Statistics and Economics*. Slaný: Melandrium, 804–818 p.
- LÖSTER, T. (2012): Kritéria pro hodnocení výsledků shlukování se známým zařazením do skupin založená na konfuzní matici. *Forum Statisticum Slovacum*, 8/7, 85–89 p.
- REZAEI, B., (2010): A cluster validity index for fuzzy clustering. *Fuzzy sets and systems*. Amsterdam: Elsevier Science BV
- ŘEZANKOVÁ, H. – HÚSEK, D. (2012): Fuzzy Clustering: determining the number of clusters. *Computational Aspects of Social Networks (CASoN)*, 277–28p.
- XIE, NN. (2011): A Classification of Cluster Validity Indexes Based on Membership Degree and Applications. *Web information systems and mining*. Berlin: Springer-Verlag Berlin
- ZADEH, L.A. (1965). Fuzzy Sets. *Information and control* 8, 338-353p.
- ZALIK, K.R. (2011). Validity index for clusters of different sizes and densities. *Computer science, artificial intelligence*. Amsterdam: ELSEVIER SCIENCE BV

Acknowledgments

This paper was created with the help of the Internal Grant Agency of University of Economics in Prague No. 6/2013 under the title „Evaluation of results of cluster analysis in Economic problems.”

Adresa autorov:

Elena Makhalova, Mgr.
Vysoká škola ekonomická
nám. W. Churchilla 4, 130 67 Praha
elena.makhalova@vse.cz

Kornélia Cséfalvaiová, Ing.
Vysoká škola ekonomická
nám. W. Churchilla 4, 130 67 Praha
xcsek00@vse.cz

Iva Pecáková, doc. Ing. CSc.
Vysoká škola ekonomická
nám. W. Churchilla 4, 130 67 Praha
pecakova@vse.cz

**Dôsledky porušenia predpokladu nezávislosti
v prípade jednovýberového Studentovho t-testu**
**Consequences of Independence Assumption Violation
Regarding One-Sample Student's t-test**

Tomáš Marcinko

Abstract: Nearly all classical statistical hypothesis tests are derived under a few fundamental assumptions, which may or may not be met in real world applications, and the classical Student's t-test is no exception. The main aim of this article is to study consequences of the assumption violation concerning independence of the observations and mainly its effect on type I and type II errors when dealing with one-sample t-test. Based on a simulation study the consequences of positive and negative autocorrelation were examined for various sample sizes. From the results of the conducted simulation study it can be inferred that the Student's t-test tends to be substantially liberal in case of positive autocorrelation and substantially conservative in case of negative autocorrelation. Therefore, in case of autocorrelated observations alternative statistical techniques should always be preferred.

Abstrakt: Takmer všetky klasické testy štatistických hypotéz sú odvodené na základe malého počtu základných predpokladov, ktoré môžu, ale nemusia byť splnené v reálnych aplikáciách, a klasický Studentov t-test o strednej hodnote nie je žiadnou výnimkou. Hlavným cieľom tohto článku je preskúmanie dôsledkov porušenia predpokladu nezávislosti pozorovaní, a to predovšetkým ich vplyv na pravdepodobnosti chýb prvého a druhého druhu jednovýberového t-testu. Na základe simulačnej štúdie bol priamo kvantifikovaný vplyv kladnej a zápornej autokorelácie pre rôzne rozsahy výberu, pričom z výsledkov štúdie môžeme dedukovať, že Studentov t-test má tendenciu byť výrazne liberálny v prípade kladnej autokorelácie a výrazne konzervatívny v prípade zápornej autokorelácie. Z tohto dôvodu by v prípade akejkolvek sériovej závislosti medzi pozorovaniami mali byť preferované alternatívne štatistické metódy.

Key words: one-sample t-test, independence assumption violation, generalized least squares, HAC estimators.

Kľúčové slová: jednovýberový t-test, porušenie predpokladu nezávislosti, zovšeobecnená metóda najmenších štvorcov, HAC odhady.

JEL classification: C12, C15

1. Úvod

Testy hypotéz o parametre polohy patria bezpochyby medzi najdôležitejšie štatistické testy, ktoré sa využívajú za účelom rozhodnutia, či je parameter polohy rovný danej konštante. Parametrom polohy nášho záujmu býva najčastejšie stredná hodnota, aj keď v niektorých prípadoch môže byť preferovaný medián alebo iná miera polohy.

V tomto článku sa zameriame na základný štatistický problém, ktorým je jednovýberový test o strednej hodnote, a to v situácii, kedy je populačný rozptyl neznámy. Najvyužívanejším štatistickým testom, ktorý je určený pre testovanie hypotézy o strednej hodnote, je určite jednovýberový Studentov t-test. Jedná sa však o parametrický test, ktorý je odvodený na základe niekoľkých zjednodušujúcich predpokladov, ktoré nemusia byť (a často nie sú) splnené v reálnych aplikáciách. Konkrétne, tento test predpokladá, že máme k dispozícii náhodný výber z populácie, ktorá sa riadi normálnym rozdelením, tzn. že jednotlivé pozorovania sú nezávislé a rovnako rozdelené náhodné veličiny majúce normálne rozdelenie so zhodnou strednou hodnotou a zhodným rozptylom.

Cieľom tohto článku je spracovanie simulačnej štúdie v programovom prostredí **R**, ktorá preskúma dôsledky porušenia predpokladu nezávislosti pozorovaní, a to hlavne ich negatívny vplyv na pravdepodobnosti chýb prvého a druhého druhu Studentovho t-testu. Na základe výsledkov simulačnej štúdie určíme, aký vplyv má kladná a záporná autokorelovanosť pozorovaní pre rôzne rozsahy výberu, pričom výsledky získané klasickým t-testom budú pre názornosť porovnané s niektorými metódami, ktoré boli navrhnuté pre lineárne modely s autokorelovanou náhodnou zložkou.

2. Parametrický t-test o strednej hodnote

Pri štatistickom testovaní hypotézy o strednej hodnote náhodnej veličiny sú bezpochyby najčastejšie využívané klasické parametrické metódy, pričom jeden z najdôležitejších testov v jednorozmerných úlohách je tzv. Studentov t-test. Tento test je, rovnako ako každý parametrický test, založený na niekoľkých teoretických predpokladoch, ktoré sú kľúčové pre platnosť pravdepodobnostných úsudkov vyplývajúcich z tohto testu.

Predpokladajme, že máme k dispozícii jednorozmerný náhodný výber o rozsahu n (tzn. postupnosť nezávislých a rovnako rozdelených náhodných veličín Y_1, Y_2, \dots, Y_n) z normálneho rozdelenia so strednou hodnotou m a rozptylom s^2 , pričom potrebujeme rozhodnúť, či stredná hodnota m je rovná danej hodnote m_0 . Pre test hypotézy

$$H_0 : m = m_0 \quad (1)$$

môžeme využiť štatistiku

$$t = \frac{\bar{y} - m_0}{s} \sqrt{n}, \quad (2)$$

kde \bar{y} je výberový priemer a s je výberová smerodajná odchýlka, ktorá má za platnosti nulovej hypotézy Studentovo t rozdelenie s $n - 1$ stupňami voľnosti. Pri splnení všetkých predpokladov Studentovho t-testu sa dá ukázať, že sa jedná o rovnomerne najsilnejší nestranný test hypotézy (1) proti obojstrannej alternatíve (Lehmann a Romano, 2005)

$$H_0 : m \neq m_0. \quad (3)$$

V ďalšej časti s výhodou využijeme lineárny model v tvare

$$y_i = m + e_i, \quad i = 1, 2, \dots, n, \quad (4)$$

kde e_1, e_2, \dots, e_n vyjadrujú náhodnú zložku tvorenú nezávislými a normálne rozdelenými náhodnými veličinami s nulovou strednou hodnotou a zhodným rozptylom s^2 .

Použitím klasickej metódy najmenších štvorcov na model (4) bude odhadom parametra m opäť výberový priemer \bar{y} a napríklad na základe Gaussovej-Markovovej vety sa dá jednoducho overiť, že výberový priemer \bar{y} je v takomto prípade najlepším nestranným lineárnym odhadom parametra m . Navyše, po nahradení pozorovaní y_i hodnotami $y_i - m_0$ v lineárnom modeli (4) je klasický Studentov t-test nulovej hypotézy (1) ekvivalentný s t-testom významnosti regresného parametra m v tomto lineárnom modeli, ktorý je súčasťou výstupu základnej funkcie `summary.lm` programového prostredia **R**.

3. Možnosti riešenia autokorelovanosti náhodnej zložky

V praxi sa však pomerne často stáva, že náhodné veličiny Y_1, Y_2, \dots, Y_n nie sú navzájom nezávislé. Pre názornosť budeme teda predpokladať, že platí lineárny model v tvare

$$y_i = m + e_i, \quad i = 1, 2, \dots, n, \quad (5)$$

pričom pre náhodnú zložku platí autoregresný vzťah prvého radu v tvare

$$e_i = r e_{i-1} + e_i. \quad (6)$$

Ak pre parameter r platí $r > 0$, potom sú vzhľadom k platnosti modelu (5) náhodné veličiny Y_1, Y_2, \dots, Y_n kladne autokorelované, naopak v prípade platnosti $r < 0$ hovoríme o zápornej autokorelácii.

Z teórie vieme (Cipra, 2008, ods. 4.3), že autokorelovanosť náhodnej zložky má negatívny vplyv na klasické odhady modelu (5) získané metódou najmenších štvorcov. Konkrétne, napriek tomu, že odhad získaný metódou najmenších štvorcov zostáva neustranným a konzistentným odhadom parametra m , nejedná sa už o najlepší neustranný lineárny odhad. Navyše, odhad štandardnej chyby tohto odhadu je pri výskyte autokorelácie vychýlený a z tohto dôvodu sa štatistika (2) nebude za platnosti nulovej hypotézy (1) riadiť Studentovým t rozdelením.

Existuje však pomerne veľké množstvo alternatívnych metód, ktoré sa pokúšajú negatívny vplyv autokorelovanosti obmedziť alebo eliminovať.

Optimálnym riešením pri znalosti parametra r , ktoré vedie k najlepšiemu neustrannému lineárnemu odhadu regresných parametrov, je využitie Aitkenovho odhadu (zovšeobecnená metóda najmenších štvorcov) vo všeobecnom tvare

$$(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}^{-1}\mathbf{y}, \quad (7)$$

kde $\mathbf{y} = (y_1, \dots, y_n)'$ je vektor pozorovaní, \mathbf{X} je matica modelu a $\text{var}\mathbf{e} = \mathbf{s}^2\mathbf{W}$. Pre kovariančnú maticu odhadu (7) navyše platí

$$\mathbf{s}^2(\mathbf{X}'\mathbf{W}^{-1}\mathbf{X})^{-1}. \quad (8)$$

V nami uvažovanom jednorozmernom prípade platí vzhľadom k platnosti modelu (5), že $\mathbf{X} = (1, \dots, 1)'$, a zároveň vzhľadom k predpokladu (6) možno maticu \mathbf{W} vyjadriť v tvare

$$\mathbf{W} = \begin{bmatrix} 1 & r & \mathbf{L} & r^{n-1} \\ r & 1 & \mathbf{L} & r^{n-2} \\ \mathbf{M} & \mathbf{M} & \mathbf{O} & \mathbf{M} \\ r^{n-1} & r^{n-2} & \mathbf{L} & 1 \end{bmatrix}. \quad (9)$$

Zjavným nedostatkom tohto postupu je, že v praxi často nepoznáme hodnotu parametra r . Jedným z priamočiarych riešení je nahradenie parametra r v matici \mathbf{W} konzistentným odhadom autokorelačného koeficientu prvého radu.

Inú možnosť riešenia poskytuje *Cochranova-Orcuttova metóda*, ktorá vychádza z tzv. Koyckovej transformácie $y_i^* = y_i - r y_{i-1}$, ktorá vedie na transformovaný model

$$y_i^* = y_i - r y_{i-1} = m(1-r) + (e_i - r e_{i-1}) = m(1-r) + e_i, \quad i = 2, 3, \dots, n \quad (10)$$

ktorý možno odhadnúť klasickou metódou najmenších štvorcov. Z dôvodu neznalosti parametra r sa však využíva nasledovný iteratívny algoritmus:

- 1) výpočet reziduí \hat{e}_i v modeli (5) metódou najmenších štvorcov;
- 2) výpočet konzistentného odhadu \hat{r} s využitím reziduí \hat{e}_i (napríklad ako odhad regresného koeficientu v modeli (6));
- 3) výpočet odhadu strednej hodnoty \hat{m} v modeli (10) s využitím odhadu \hat{r} ;
- 4) výpočet nových reziduí $\hat{e}_i = y_i - \hat{m}$;
- 5) možná iterácia krokov 2) – 4) až do stabilizácie odhadov \hat{m} a \hat{r} .

Vyššie uvedený algoritmus obsahuje **R**-ková procedúra `cochrane.orcutt` z balíku `orcutt`.

Ďalšou možnosťou, ako odhadnúť lineárny model (5) za predpokladu autoregresného vzťahu (6), je simultánny odhad regresného parametra m a autokorelačného koeficientu r , ktorý možno získať napríklad metódou maximálnej vierohodnosti. K tomuto účelu je možné využiť **R**-kovú funkciu `gls` z balíku `nlme`, ktorá okrem maximalizácie logaritmickéj vierohodnosti, skrátene ML (maximum likelihood), umožňuje i tzv. metódu REML (restricted maximum likelihood), ktorá potláča efekt rušivých parametrov, čím je možné získať menej vychýlené odhady kovariančnej matice (8). Detaily k tejto metóde uvádzajú napríklad Pinheiro a Bates (2000).

Ako poslednú možnosť, ktorá sa pokúša riešiť problém s autokoreláciou v lineárnom modeli, si uvedieme využitie tzv. HAC (heteroskedasticity and autocorrelation consistent) odhadu kovariančnej matice odhadu regresných parametrov, ktorý navrhli Newey a West (1987), prípadne Andrews (1999). Testy významnosti regresných parametrov s využitím HAC odhadu možno získať na základe **R**-kovej funkcie `coefest` z balíku `lmtest`, pričom odhad kovariančnej matice získame funkciou `NeweyWest`, resp. `vcovHAC` z balíku `sandwich`.

Na záver poznamenajme, že všetky uvedené metódy umožňujú po nahradení pozorovaní y_i hodnotami $y_i - m_0$ testovať nulovú hypotézu (1), a to formou testu významnosti príslušného regresného parametra v odhadnutom lineárnom modeli (5).

4. Simulačná štúdia

Cieľom uvedenej simulačnej štúdie je Monte Carlo odhad pravdepodobností chýb prvého a druhého druhu v prípade testu hypotézy o strednej hodnote (1) pri výbere z normálneho rozdelenia v situácii, kedy jednotlivé pozorovania môžu byť vzájomne sériovo závislé. Budeme teda predpokladať platnosť lineárneho modelu (5) za predpokladu, že pre náhodné zložky platí autoregresný vzťah prvého radu (6).

Za účelom získania Monte Carlo odhadu silofunkcie Studentovho t-testu v prípade porušenia predpokladu nezávislosti pozorovaní, ako aj analogických odhadov silofunkcií alternatívnych testových procedúr, bol nasimulovaný dostatočne veľký počet dátových súborov ($N = 10\,000$) za platnosti lineárneho modelu (5), a to pre rôzne hodnoty autokorelačného koeficientu r a rôzne rozsahy výberu n (bez ujmy na interpretáciu výsledkov boli uvažované parametre $m = 100$ a $s^2 = 100$).

Základné výsledky simulačnej štúdie sú uvedené v tabuľke 1, ktorá obsahuje odhady pravdepodobností chýb prvého druhu Studentovho t-testu, ako aj alternatívnych testových procedúr založených na Aitkenovom odhade, Cochranovej-Orcuttovej iteratívnej metóde, simultánnom odhade získanom pomocou metód maximálnej vierohodnosti ML a REML a na testoch využívajúcich HAC odhady rozptylu odhadu strednej hodnoty m .

Z uvedených výsledkov je zrejmé, že v prípade nulovej autokorelácie (tzn. v prípade vzájomnej nezávislosti pozorovaní) je očakávané optimálne použitie Studentovho t-testu, pričom všetky alternatívne procedúry, ktoré predpokladajú možný výskyt autokorelovaných pozorovaní, majú miernu tendenciu byť liberálne (tzn. majú pravdepodobnosť chyby prvého druhu väčšiu ako je zvolená hladina významnosti $\alpha = 0,05$). V prípade výskytu čo len malej autokorelácie sa však pravdepodobnosť chyby prvého druhu Studentovho t-testu výrazne líši od zvolenej hladiny významnosti – v prípade kladnej autokorelácie pozorovaní je Studentov t-test značne liberálny, a naopak v prípade zápornej autokorelácie je výrazne konzervatívny (tzn. má pravdepodobnosť chyby prvého druhu menšiu ako je zvolená hladina významnosti).

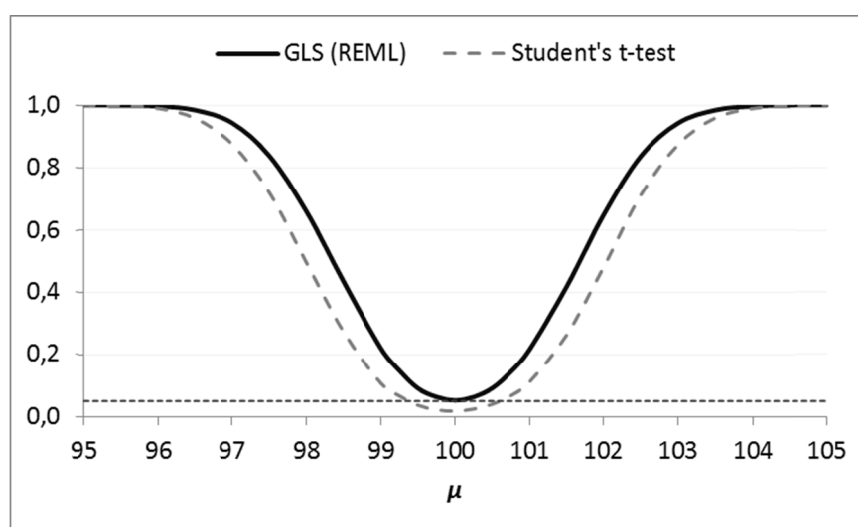
Naopak z alternatívnych testových procedúr (v prípade autokorelácie dávajú všetky uvažované procedúry lepšie výsledky ako Studentov t-test) najlepší dopadol test založený na simultánnom REML odhade modelu (5) pomocou **R**-kovej funkcie `gls`, aj keď v prípade značnej kladnej autokorelácie má aj táto metóda tendenciu byť mierne liberálna.

Tab. 9: Odhad pravdepodobností chýb prvého druhu pre rôzne veľkosti autokorelácie

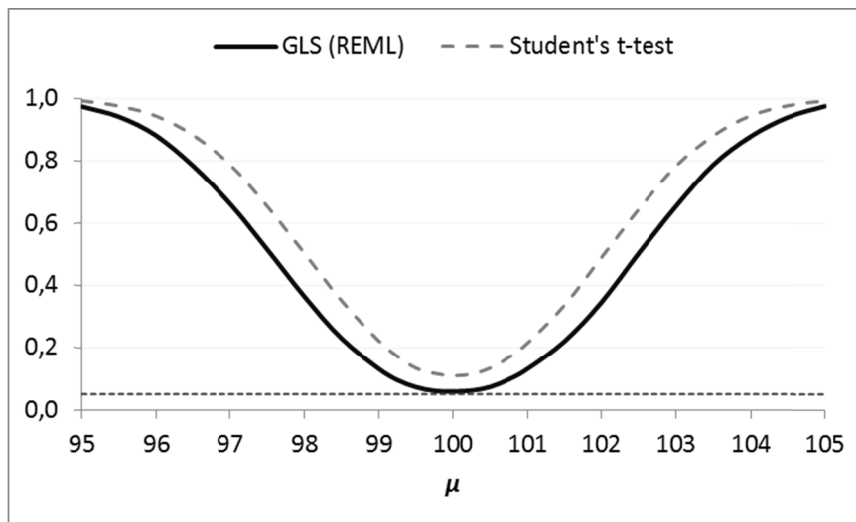
n	ρ	Student	Aitken	CO	GLS _{ML}	GLS _{REML}	HAC _{NW}	HAC _A
$n = 30$	-0,8	0,0002*	0,0595*	0,0483	0,0486	0,0475	0,1168*	0,0376*
	-0,5	0,0021*	0,0545*	0,0517	0,0520	0,0489	0,1187*	0,0395*
	-0,2	0,0174*	0,0600*	0,0599*	0,0603*	0,0555*	0,1153*	0,0451*
	0,0	0,0487	0,0653*	0,0670*	0,0653*	0,0590*	0,1109*	0,0587*
	0,2	0,1021*	0,0739*	0,0738*	0,0736*	0,0632*	0,1078*	0,0826*
	0,5	0,2487*	0,0998*	0,1031*	0,0999*	0,0819*	0,1167*	0,1378*
	0,8	0,5278*	0,1818*	0,2001*	0,1851*	0,1322*	0,1917*	0,2997*
$n = 50$	-0,8	0,0000*	0,0560*	0,0510	0,0512	0,0510	0,0883*	0,0435*
	-0,5	0,0017*	0,0560*	0,0548*	0,0542	0,0530	0,0874*	0,0435*
	-0,2	0,0184*	0,0607*	0,0593*	0,0605*	0,0569*	0,0848*	0,0474
	0,0	0,0505	0,0650*	0,0632*	0,0647*	0,0600*	0,0844*	0,0592*
	0,2	0,1093*	0,0705*	0,0691*	0,0707*	0,0642*	0,0849*	0,0806*
	0,5	0,2576*	0,0877*	0,0883*	0,0888*	0,0764*	0,0920*	0,1210*
	0,8	0,5284*	0,1440*	0,1549*	0,1479*	0,1132*	0,1495*	0,2293*
$n = 100$	-0,8	0,0000*	0,0532	0,0501	0,0505	0,0502	0,0753*	0,0436*
	-0,5	0,0018*	0,0515	0,0514	0,0510	0,0505	0,0728*	0,0438*
	-0,2	0,0175*	0,0549*	0,0552*	0,0548*	0,0525	0,0725*	0,0455*
	0,0	0,0510	0,0574*	0,0582*	0,0574*	0,0551*	0,0726*	0,0543
	0,2	0,1095*	0,0616*	0,0619*	0,0615*	0,0587*	0,0733*	0,0729*
	0,5	0,2539*	0,0700*	0,0699*	0,0701*	0,0646*	0,0748*	0,0924*
	0,8	0,5179*	0,1002*	0,1054*	0,1025*	0,0859*	0,0994*	0,1534*

* pravdepodobnosť chyby prvého druhu sa významne líši od hladiny významnosti $\alpha = 0,05$ na základe exaktného testu o parametre alternatívneho rozdelenia (Clopper a Pearson, 1934)

Na záver si uvedieme dva grafy, ktoré porovnávajú vplyv kladnej a zápornej autokorelácie na silofunkcie Studentovho t-testu a testu s využitím REML odhadu modelu (5).



Obr. 1: Odhad silofunkcií Studentovho t-testu a testu na základe REML ($\rho = -0,2$; $n = 100$)



Obr. 2: Odhad silofunkcií Studentovho t-testu a testu na základe REML ($\rho = 0,2$; $n = 100$)

5. Záver

Klasický jednovýberový Studentov t-test je odvodený za predpokladu, že máme k dispozícii náhodný výber, ktorý tvoria nezávislé a normálne rozdelené náhodné veličiny so zhodnou strednou hodnotou a zhodným rozptylom. Často sa však stáva, špeciálne v prípade časovo usporiadaných dát, že jednotlivé pozorovania nie sú sériovo nezávislé.

Cieľom tohto článku bolo preskúmanie vplyvu zápornej a kladnej autokorelácie na pravdepodobnosti chýb prvého a druhého druhu Studentovho t-testu. Na základe výsledkov simulačnej štúdie sa zistilo, že Studentov t-test je značne liberálny v prípade kladnej autokorelácie pozorovaní a značne konzervatívny v prípade zápornej autokorelácie. Z tohto dôvodu sa pri výskyte čo len malej sériovej závislosti medzi pozorovaniami odporúča využiť alternatívne testové procedúry, napríklad test založený na REML odhade lineárneho modelu.

Pod'akovanie: Tento príspevok vznikol za podpory projektu IGA VŠE 128/2014.

Literatúra

- ANDREWS, D. W. K. 1991. Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. In: *Econometrica*, roč. 59, s. 817 – 858.
- CIPRA, T. 2008. *Finanční ekonometrie*. Praha: Ekopress.
- CLOPPER, C. J. – PEARSON, E. S. 1934. The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial. In: *Biometrika*, roč. 26, č. 4, s. 404 – 413.
- LEHMANN, E. L. – ROMANO, J. P. 2005. *Testing Statistical Hypotheses*. London: Springer.
- NEWAY, W. K. – WEST, K. D. 1987. A Simple, Positive-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. In: *Econometrica*, roč. 55, s. 703 – 708.
- PINHEIRO, J. C. – BATES, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer.

Adresa autora:

Tomáš Marcinko, Mgr. Ing.
 Vysoká škola ekonomická v Praze
 Fakulta informatiky a statistiky
 nám. W. Churchilla 4, 130 67 Praha
xmart14@vse.cz

Vplyv vybraných faktorov na zisk životnej poisťovne pri doživotných dôchodkoch vyplácaných z úspor v II. dôchodkovom pilieri na Slovensku

The influence of selected factors on the profit of life insurance company in the case of life annuities paid from the second pillar in Slovakia

Igor Melicherčík, Gábor Szűcs

Abstract: In this article we examine selected risk factors that may affect future profit of life insurance companies when paying annuities from the second pillar. The investment strategy of insurer is constructed by applying the asset-liability matching method. The aim of this article is to analyze the longevity risk in the case of different demographic assumptions and to examine the impact of changing market yields on the present value of future profits of the insurance company.

Abstrakt: V rámci tohto článku sa skúmajú vybrané rizikové faktory, ktorým môžu čeliť životné poisťovne pri vyplácaní doživotných dôchodkov z II. dôchodkového piliera. Pri zostrojení investičnej stratégie sa používa tzv. metóda spárovania aktív a pasív životnej poisťovne. Cieľom tohto článku je vykonať analýzu rizika dlhovekosti budúcich dôchodcov pri rôznych demografických predpokladoch a analyzovať vplyv premenlivých trhových výnosov na súčasnú hodnotu zisku poisťovne v prípade vyššie uvedeného doživotného dôchodku.

Key words: asset-liability matching method, second pillar, longevity risk, yield curves.

Kľúčové slová: metóda spárovania aktív a pasív, II. dôchodkový pilier, riziko dlhovekosti, výnosové krivky.

JEL classification: G17, G22

1. Úvod

Od 1. januára 2015 sa začne vyplácanie prvých starobných dôchodkov zo starobného dôchodkového sporenia (z II. dôchodkového piliera). Je známe, že väčšina sporiteľov, ktorá dosiahla vek 62 rokov a pôjde do dôchodku po spomínanom dátume, si kúpi za svoje nasporené peniaze doživotný dôchodok v životnej poisťovni. Z hľadiska životných poisťovní to znamená uzatvorenie nových kontraktov na doživotnú dôchodkovú ochranu klienta s pevným jednorazovo plateným poistným. Ročnú výšku dôchodkovej dávky poisťovňa počíta pri pevnej technickej úrokovej miere (podľa Zákona č. 8/2008 Z. z. o poisťovníctve¹), ktorá môže mať maximálnu výšku $i = 1,9\%$ p. a. (vid' Opatrenie č. 3/2013 Národnej banky Slovenska²). Na jednej strane teda poisťovne kalkulujú výšku dávky na základe danej úrokovej miery, na druhej strane však investujú inkasované poistné **pri premenlivých trhových výnosoch**. Podľa Zákona o poisťovníctve musia dodržiavať zásadu diverzifikácie a určité limity pri umiestnení prostriedkov technických rezerv. V praxi životné poisťovne najväčšiu časť prostriedkov technických rezerv investujú do bezpečných dlhopisov, hypotekárnych záložných listov a bankových vkladov a len niekoľko percent alokujú do rizikovejších aktív. Práve preto sme stanovili v rámci našich modelových výpočtov umelý predpoklad, že poisťovne investujú výlučne len do dlhopisov pri daných (jednotných) úrovniach výnosov. Pri modelovaní výnosov dlhopisov sme používali tri modely výnosových

¹ Úplné znenie Zákona č. 8/2008 Z. z. o poisťovníctve je dostupné na adrese:

<http://www.nbs.sk/_img/Documents/Legislativa/UplneZneniaZakonov/Z0082008.PDF>.

² Opatrenie Národnej banky Slovenska z 25. júna 2013 o maximálnej výške technickej úrokovej miery je dostupné na adrese: <http://www.nbs.sk/_img/Documents/Legislativa/Vestnik/OPAT3-2013.pdf>.

kriviek: výnosovú krivku odvodenú od Coxovho-Ingersollovho-Rossovho modelu pre krátkodobé úrokové miery (Cox et al., 1985), Nelsonov-Siegelov model (Nelson, Siegel, 1987) a Svenssonov model (Svensson, 1994).

Ďalším rizikovým faktorom, ktorému životné poisťovne budú čeliť pri vyplácaní dôchodkov, je **dlhovkosť dôchodcov**. V klasických poisťno-matematických výpočtoch sa obvykle používajú úmrtnostné tabuľky opisujúce situáciu k určitému časovému bodu (kalendárnemu roku). Táto skutočnosť môže viesť k nepresnostiam pri takých dlhodobých poisťných produktoch, akými sú napríklad doživotné dôchodky. Poisťovne totiž „naplánujú“ vyplácanie dôchodkov pri aktuálnych podmienkach, no postupom času sa tieto podmienky môžu zmeniť, napr. môže prísť k zníženiu mier úmrtnosti a skupina poistencov (dôchodcov) sa môže dožiť dlhšieho veku ako sa predpokladalo. Jedným z našich cieľov bolo poukázať na efekty rizika dlhovkosti dôchodcov a v rámci modelových výpočtov ilustrovať zmeny v ziskovosti poisťovní pri zmenených predpokladoch o úmrtnosti.

Ako sme už spomínali, životné poisťovne sú z určitého aspektu súborom aktív a pasív. V tomto kontexte aktíva poisťovne sú napr. dlhopisy, akcie, termínované vklady a pod., teda prostriedky, do ktorých poisťovňa investuje inkasované poistné v mene poistených. Na druhej strane pasíva poisťovne predstavujú jej záväzky voči poistencom vyplývajúce z poisťných kontraktov (nazývame ich aj technické rezervy). Princíp spárovania aktív a pasív životnej poisťovne spočíva v tom, že poisťovňa ku každému záväzku (budúcemu poistnému plneniu) pripojí vhodné aktívum s rovnakou (budúcou) hodnotou a totožným termínom splatnosti. Pri našich analýzach a modelových výpočtoch sme využívali práve tento princíp na definovanie investičnej stratégie životných poisťovní pri spomínaných doživotných dôchodkoch.

V rámci tohto článku ponúkame statickú analýzu ziskovosti životných poisťovní pri rôznych predpokladoch o trhových výnosoch a demografickom vývoji v prípade anuit vyplácaných z úspor v II. dôchodkovom pilieri. Pod pojmom „statická analýza“ rozumieme jednak to, že poisťovňa hneď po inkasovaní poistného ho rozloží, a potom už nespraví ďalšie investičné rozhodnutia a aj to, že predpokladáme len jednorazový príchod poistencov (namiesto pravidelného prúdenia sporiteľov odchádzajúcich do dôchodku). V druhej kapitole príspevku sú uvedené tri modely výnosov dlhopisov, postupy odhadov ich parametrov a špecifikácia historických dátových súborov úrokových mier a výnosov dlhopisov. Tretia časť obsahuje označenia a vzťahy z oblasti aktuárskej demografie a tri modely úmrtnosti, resp. prežívania poistených osôb. Štvrtá kapitola článku sa zaoberá metódou spárovania aktív a pasív a definovaním pojmov ako čistá akumulovaná hodnota poistenia alebo diskontovaná hodnota budúceho zisku. V poslednej, piatej kapitole sú uvedené výsledky a závery výpočtov ohľadom odhadu zisku poisťovne pri vyplácaní doživotných dôchodkov.

2. Modelovanie trhových výnosov

Ako sme už spomínali v úvodnej časti, pri modelovaní vyplácania dôchodkových dávok a odhadu zisku poisťovne kľúčovú úlohu zohráva adekvátny odhad trhových výnosov, pri ktorých poisťovne investujú inkasované poistné. Pri kalibrácii modelov sme používali okamžité úrokové miery (*short rate*) na Slovensku podľa webovej stránky Organizácie pre hospodársku spoluprácu a rozvoj (*Organisation for Economic Co-operation and Development*, OECD), krátkodobé úrokové miery (trojmesačný, šesťmesačný a dvanásťmesačný EURIBOR) z Európskeho inštitútu peňažného trhu (*European Money Market Institute*, EMMI) a ročné výnosy slovenských vládnych dlhopisov (VSVD) s dvoj-, päť-, resp. desaťročnou splatnosťou podľa záznamov Národnej banky Slovenska (NBS). Všetky historické dáta boli vyjadrené pri ročnej časovej jednotke, t. j. *per annum* (p. a.). Keďže všetky tri modely výnosových kriviek uvedené v úvodnej časti sú definované pre prípad spojitého úročenia, predpokladali sme, že aj historické úrokové miery a výnosy dlhopisov sú vlastne intenzita úrokovania, t. j. ich podkladové aktíva (dlhopisy) prinášajú

spojité príjmy. V Tab. 1 ponúkame prehľad o historických úrokových mierach, resp. výnosoch dlhopisov.

Tab. 1: Prehľad historických dát používaných pri kalibrácii modelov výnosových kriviek

	označenie úrokových mier, resp. výnosov	obdobie (od - do)	počet dátových bodov	frekvencia dát	vyjadrenie	zdroj
okamžitá úroková miera	short rate	jan. 2011 – sept. 2014	45	mesačná	p. a.	OECD ¹
úrokové miery	EURIBOR 3M	jan. 2011 – aug. 2014	44	mesačná	p. a.	EMMI ²
	EURIBOR 6M		44			
	EURIBOR 12M		44			
výnosy slovenských vládnych dlhopisov	VSVD 2Y	jan. 2011 – aug. 2014	44	mesačná	p. a.	NBS ³
	VSVD 5Y		44			
	VSVD 10Y		44			

Naším cieľom pri modelovaní vyplácania dôchodkov z úspor v II. pilieri a odhade zisku poisťovne bolo to, aby sme vytvorili čo najrealistickejší model pre podmienky na Slovensku. Práve preto sme zvolili okamžité úrokové miery pre Slovenskú republiku, krátkodobé úrokové miery EURIBOR viazané na euro a výnosy slovenských dlhopisov.

Pri modelovaní trhových výnosov bezkupónových dlhopisov s nominálnou hodnotou jedna peňažná jednotka (tiež ich nazývame *štandardné dlhopisy*) používame:

- Coxov-Ingersollov-Rossov model kalibrovaný na základe dát z obdobia od januára 2011 do augusta 2014 (viď Tab. 1), označíme ako *CIR2011 model*,
- Coxov-Ingersollov-Rossov model kalibrovaný na základe dát z obdobia od januára 2014 do augusta 2014 (viď Tab. 1), označíme ako *CIR2014 model*,
- Nelsonov-Siegelov model kalibrovaný len na základe dát z augusta 2014 (EURIBOR 3M, EURIBOR 6M, EURIBOR 12M, VSVD 2Y, VSVD 5Y, VSVD 10Y), označíme ako *NS2014 model*,
- Svenssonov model kalibrovaný tiež len na základe posledných dát z augusta 2014 (viď bod c.), označíme ako *SV2014 model*.

V nasledujúcej časti tejto kapitoly sú zavedené označenia, ktoré sú dôležité pri modelovaní výnosov dlhopisov a odhadu budúceho zisku životnej poisťovne. Uvažujme časové body $t \in [0; \infty)$; $T \in [0; \infty)$; $t \leq T$ zadané pri danom časovom kroku. Označme symbolom r_t okamžitú úrokovú mieru v čase t vyjadrenú p. a. Ďalej označme zápisom $R(t, T)$ ročný výnos dlhopisu v čase t , pričom splatnosť dlhopisu (maturita) je v čase T . Platí vzťah: $r_t = R(t, t)$ pre všetky $t \in [0; \infty)$. Symbolom $P(t, T)$ označujeme hodnotu v čase t štandardného dlhopisu s maturitou T , pričom pri spojitom modeli úročenia platí (viď napríklad v knihe Melicherčík et al., 2005):

$$P(t, T) = \exp(-R(t, T) \times (T - t)) \Rightarrow R(t, T) = -\frac{\ln P(t, T)}{T - t} \quad (1)$$

Uvažujme teraz klasický trojparametrový Coxov-Ingersollov-Rossov (CIR) model pre okamžitú úrokovú mieru podľa článku Cox, Ingersoll, Ross (1985) v tvare stochastickej diferenciálnej rovnice:

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t,$$

¹ Historické dáta o okamžitej úrokovej miere (short rate) sú dostupné na adrese <http://stats.oecd.org/index.aspx?querytype=view&queryname=86>.

² Historické dáta o úrokových mierach EURIBOR (3-mesačné, 6-mesačné, 12-mesačné) sú dostupné na adrese: <http://www.emmi-benchmarks.eu/euribor-org/euribor-rates.html>.

³ Historické dáta o ročných výnosoch slovenských vládnych dlhopisov sú dostupné na adrese: http://www.nbs.sk/img/Documents/STATIST/US/VDSVD_CR.xls.

kde $\theta > 0, \kappa > 0, \sigma > 0$ sú reálne parametre CIR modelu a W_t je štandardný Brownov pohyb. Od vyššie uvedeného podkladového short rate CIR procesu sa dajú odvodiť hodnoty základných dlhopisov a platí:

$$P(r_0, t, T) \equiv P(t, T) = A(T) \times \exp(-B(T)r_0(t)), \quad (2)$$

kde funkcie $A(T) \equiv A(T, \theta, \kappa, \sigma, \lambda)$ a $B(T) \equiv B(T, \kappa, \sigma, \lambda)$ sú definované napríklad v knihe Kwok (1998), parameter λ vyjadruje tzv. trhovú cenu rizika (*market price of risk*) a $r_0(t)$ je začiatočná okamžitá úroková miera v čase t .

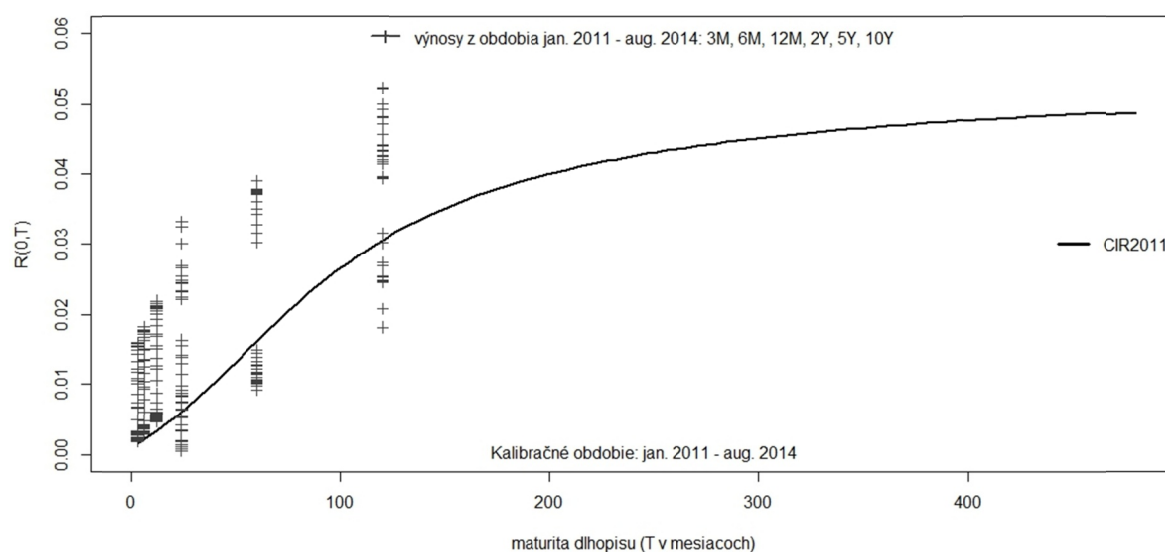
Kľúčovou úlohou bola správne odhadnúť neznáme parametre $\theta, \kappa, \sigma, \lambda$ z príslušných dátových súborov pri modeloch *CIR2011*, resp. *CIR2014*. Pri kalibrácii parametrov podkladových short rate procesov sme postupovali podľa článku Ševčovič, Urbánová Csajková (2005). Zvolili sme mesačný časový krok a dvojfázovú optimalizáciu. Pre *CIR2011 model* sme používali dáta z obdobia od januára 2011 do augusta 2014 a dostali sme nasledovné odhadnuté parametre:

$$\hat{\theta} = 0,00321477; \hat{\kappa} = 0,10401951; \hat{\sigma} = 0,01943799; \hat{\lambda} = -0,12910498.$$

Pomocou vyššie uvedených parametrov a rovnice (2) sme vypočítali hodnoty štandardných dlhopisov v čase 0 (teda v našom prípade v čase podpisu poisťnej zmluvy), pričom pre začiatočnú hodnotu (short rate procesu) sme zvolili okamžitú úrokovú mieru zo septembra 2014 (podľa OECD, vid' Tab. 1): $r_0(0) = 0,001$ p. a. Následne pomocou vzťahu (1) sme vyrátali ročné výnosy dlhopisov s maturitou T (*spot rate* $R(0, T)$) pre $T \in \{3, 4, 5, \dots, 480\}$ uvedený v mesiacoch¹.

Po kalibrácii parametrov a odhade ročných výnosov nás zaujímala presnosť *CIR2011 modelu* voči originálnym historickým výnosom. Pri posúdení kvality fitu sme postupovali podľa článku Ševčovič, Urbánová Csajková (2005). Pre ukazovateľ *maximum likelihood ratio* (MLR) sme dostali hodnotu $MLR(CIR2011) = 0,94463$. Vidíme, že hodnota ukazovateľa $MLR(CIR2011)$ je blízko k jednej. To znamená, že viazaná maximálna hodnota funkcie vierohodnosti je blízko ku globálnej maximálnej hodnote funkcie vierohodnosti. Inými slovami, podkladový *CIR2011 model* nakalibrovaný dvojfázovou optimalizačnou metódou by mal byť použiteľný pre odhad celej časovej štruktúry výnosov bezkupónových dlhopisov. V prípade ďalšieho indikátora kvality modelu, nelineárneho koeficientu determinácie (*nonlinear R² ratio*), sme dostali: $R^2(CIR2011) = 0,69233$. Môžeme skonštantovať, že hodnota $R^2(CIR2011)$ je bližšie k hodnote 1 ako k hodnote 0. To znamená, že nakalibrovaná *CIR2011* výnosová krivka dostatočne dobre fituje reálne výnosy, teda opäť by sme ju mohli použiť na modelovanie skutočnej výnosovej krivky. Kvalitu fitu nakalibrovaného *CIR2011 modelu* môžeme vizuálne posúdiť aj pomocou Obr. 1.

¹ Pri výpočte ročných výnosov dlhopisov sme zvolili časovú hranicu 480 mesiacov, teda 40 rokov. Presne takéto obdobie sme používali aj pri plánovanom vyplácaní dôchodkov z II. piliera, vid' kapitoly 4 a 5.



Obr. 14: Grafické posúdenie kvality fitu nakalibrovaného CIR2011 modelu a znázornenie pôvodných úrokových mier (3M, 6M, 12M) a výnosov dlhopisov (2Y, 5Y, 10Y) z obdobia od januára 2011 do augusta 2014

Pri kalibrácii CIR2014 modelu sme postupovali analogicky ako pri odhade parametrov CIR2011 modelu, len pri tejto druhej kalibrácii sme používali dáta z iného obdobia: od januára 2014 do augusta 2014. Dostali sme nižšie uvedené odhady parametrov podkladového CIR2014 modelu:

$$\tilde{\theta} = 0,00071059; \tilde{\kappa} = 0,08483863; \tilde{\sigma} = 0,01003501; \tilde{\lambda} = -0,11825065.$$

V ďalšom kroku sme opäť vypočítali odhady hodnôt štandardných dlhopisov $P(0, T)$, resp. ročné výnosy $R(0, T)$; $T \in \{3, 4, 5, \dots, 480\}$ pri začiatočnej hodnote okamžitej úrokovej miery $r_0(0) = 0,001$ p. a. Pri posudzovaní presnosti odhadnutých výnosov v porovnaní s historickými výnosmi sme dostali hodnoty $MLR(CIR2014) = 0,97401$, resp. $R^2(CIR2014) = 0,67785$, čo znamená, že aj CIR2014 model by mal byť použiteľný pre odhad celej časovej štruktúry výnosov bezkupónových dlhopisov. Odhadnuté výnosy a kvalitu fitu pri CIR2014 modeli sme znázornili na Obr. 2.

Uvažujme teraz Nelsonov-Siegelov model (Nelson, Siegel, 1987) pre výnosovú krivku v čase $t = 0$ definovanú predpisom

$$R(t = 0, T) = \beta_0 + \beta_1 \frac{1 - \exp(-\xi T)}{\xi T} + \beta_2 \left(\frac{1 - \exp(-\xi T)}{\xi T} - \exp(-\xi T) \right),$$

kde $\beta_0, \beta_1, \beta_2, \xi$ sú reálne parametre modelu a premenná T je maturita dlhopisu. Pri odhade parametrov sme využívali poslednú šesticu dát z nášho dátového súboru (t. j. dáta z augusta 2014 pre EURIBOR 3M, EURIBOR 6M, EURIBOR 12M, VSVD 2Y, VSVD 5Y, VSVD 10Y). Samotnú kalibráciu sme spravili v prostredí štatistického softvéru R (R Core Team, 2013) pomocou balíku YieldCurve (Guirrieri, 2010; Consiglio, Guirrieri, 2011) a dostali sme odhady

$$\hat{\beta}_0 = 0,0403697; \hat{\beta}_1 = -0,03712724; \hat{\beta}_2 = -0,04663509; \hat{\xi} = 0,0284603.$$

Vizuálny fit nakalibrovaného NS2014 modelu môžeme posúdiť pomocou Obr. 2.

Poznámka. Nelsonova-Siegelova výnosová krivka je určená štyrmi parametrami. Vyššie uvedené NS2014 model sme kalibrovali pomocou šiestich reálnych výnosov (použili sme dáta

z augusta 2014), takže sme očakávali dobrú kvalitu fitu, čo sa aj potvrdilo: pre *NS2014 model* sme dostali koeficient determinácie $R^2 = 0,95588$.

Naším posledným modelom, ktorý sme používali na modelovanie výnosov, bol Svenssonov model (Svensson, 1994) definovaný vzťahom

$$R(t=0, T) = \alpha_0 + \alpha_1 \frac{1 - \exp\left(-\frac{T}{\lambda_1}\right)}{\frac{T}{\lambda_1}} + \alpha_2 \left(\frac{1 - \exp\left(-\frac{T}{\lambda_1}\right)}{\frac{T}{\lambda_1}} - \exp\left(-\frac{T}{\lambda_1}\right) \right) + \alpha_3 \left(\frac{1 - \exp(-T/\lambda_2)}{T/\lambda_2} - \exp(-T/\lambda_2) \right),$$

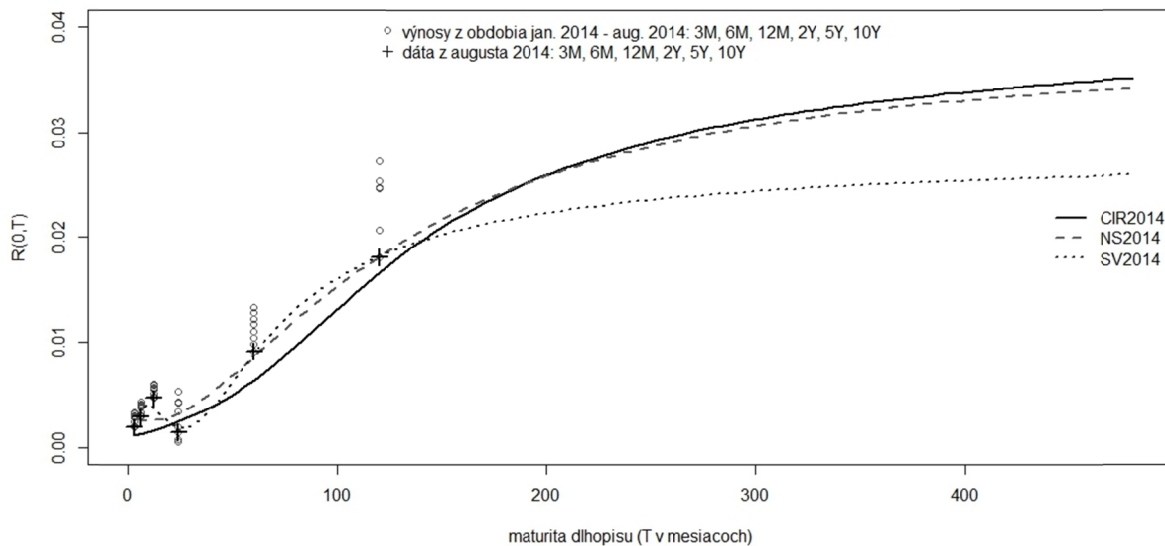
kde $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \lambda_1, \lambda_2$ sú reálne parametre modelu a premenná T je maturita dlhopisu. Pri odhade parametrov sme postupovali analogicky ako v prípade *NS2014 modelu*: používali sme dáta z augusta 2014 a aplikovali sme balík štatistického softvéru R s názvom *YieldCurve*. Po vykonaní postupu sme dostali nasledovných šesť parametrov *SV2014 modelu*:

$$\hat{\alpha}_0 = 0,02853997; \hat{\alpha}_1 = -0,03230188; \hat{\alpha}_2 = -27344,19;$$

$$\hat{\alpha}_3 = 27344,15; \hat{\lambda}_1 = 10,03748; \hat{\lambda}_2 = 10,03746.$$

Kvalitu odhadnutého *SV2014 modelu* môžeme vizuálne zhodnotiť pomocou Obr. 2.

Poznámka. Svenssonova výnosová krivka pre spot rate je určená šiestimi parametrami. Naš *SV2014 model* sme kalibrovali pomocou šiestich reálnych výnosov, takže sme očakávali skoro dokonalú kvalitu fitu, čo sa aj potvrdilo: pre *SV2014 model* sme dostali koeficient determinácie $R^2 = 0,99152$.



Obr. 2: Grafické posúdenie kvality fitu nakalibrovaného CIR2014 modelu, NS2014 modelu a SV2014 modelu a znázornenie pôvodných úrokových mier (3M, 6M, 12M) a výnosov dlhopisov (2Y, 5Y, 10Y) z obdobia od januára 2014 do augusta 2014

3. Demografické modely vývoja úmrtnosti a prežívania

V rámci našich výpočtov sme predpokladali, že výška ročnej dôchodkovej dávky sa počíta pri spoločných úmrtnostných tabuľkách pre mužov a ženy Štatistického úradu Slovenskej republiky za rok 2011 (*ÚT ŠÚSR 2011 spolu*) (viď ŠÚSR, online). Označme skratkou **DEM1** základný demografický model založený na *ÚT ŠÚSR 2011 spolu* (podrobnejšie v 5. kapitole).

Poisťovňa pri výpočte svojho budúceho zisku v súvislosti s doživotným dôchodkom by mala brať do úvahy aj efekt dlhovekosti budúcich dôchodcov. Práve preto sme definovali aj ďalšie modely, označené skratkami **DEM2**, resp. **DEM3**, ktoré sú z pohľadu životnej poisťovne pesimistickejšie, pretože berú do úvahy možnosť zníženia úmrtnostných mier, a tak aj možnú dlhovekosť poistených osôb.

Nech q_x značí pravdepodobnosť, že osoba vo veku x rokov zomrie do jedného roka. Symbolom $p_x = 1 - q_x$ označujeme pravdepodobnosť, že x -ročná osoba prežije najbližší rok. Faktory prežitia k rokov označíme ${}_k p_x$ pre $k = 1, 2, \dots$ a definujeme ich vzťahom

$${}_k p_x = \prod_{j=0}^{k-1} p_{x+j}. \quad (3)$$

Nech T_x , resp. K_x označuje budúcu dĺžku života, resp. skrátenú budúcu dĺžku života x -ročnej osoby. Symbolmi $E[T_x]$, resp. $E[K_x]$ značíme očakávanú budúcu dĺžku života, resp. skrátenú očakávanú budúcu dĺžku života osoby vo veku x rokov. Za platnosti určitých predpokladov približne platí: $E[T_x] \approx E[K_x] + \frac{1}{2}$. Skrátenú očakávanú budúcu dĺžku života môžeme počítať podľa formuly

$$E[K_x] = \sum_{k=1}^{100-x} {}_k p_x.$$

Očakávaná dĺžka života novonarodenej osoby v *ÚT ŠÚSR 2011 spolu* je $E[T_0] = 75,85$ rokov. Podľa prognózy Inštitútu informatiky a štatistiky (INFOSTAT, online) očakávaná budúca dĺžka novonarodenej ženy v roku 2050 bude 84,01 rokov (podľa stredného variantu), kým očakávaná budúca dĺžka novonarodeného muža v roku 2050 bude 77,13 rokov (tiež podľa stredného variantu). Ako odhad očakávanej budúcej dĺžky novonarodenej osoby (bez špecifikácie pohlavia) v roku 2050 sme zobrali aritmetický priemer vyššie uvedených hodnôt: $E[T'_0] = 80,57$ rokov a skrátenú očakávanú budúcu dĺžku novonarodenej osoby sme odhadli (podľa vyššie uvedeného približného vzťahu) hodnotou $E[K'_0] = 80,07$ rokov. Pravdepodobnosti úmrtia, resp. pravdepodobnosti prežitia sme upravili podľa postupu uvedeného v článku Pastor, Raučinová (1991): predpokladali sme, že pravdepodobnosti úmrtia sa znížia v každom veku rovnakým koeficientom $u \in (0; 1)$. Nové pravdepodobnosti úmrtia, resp. pravdepodobnosti prežitia sme počítali nasledovne:

$$q'_x = (1 - u)q_x, \quad p'_x = 1 - q'_x = 1 - (1 - u)q_x \quad \text{pre } x = 0, 1, \dots, 100,$$

kde za q_x sme zobrali hodnoty z *ÚT ŠÚSR 2011 spolu* (t. j. zo základného modelu **DEM1**). Nové faktory prežitia pre novonarodenú osobu ${}_k p'_0$ sme počítali dosadením pravdepodobností prežitia p'_x do vzťahu (3). Neznámy parameter u sme nastavili tak, aby platilo

$$E[K'_0] = \sum_{k=1}^{100} {}_k p'_0 = \sum_{k=1}^{100} \prod_{j=0}^{k-1} p'_j = \sum_{k=1}^{100} \prod_{j=0}^{k-1} (1 - (1 - u)q_j) = 80,07.$$

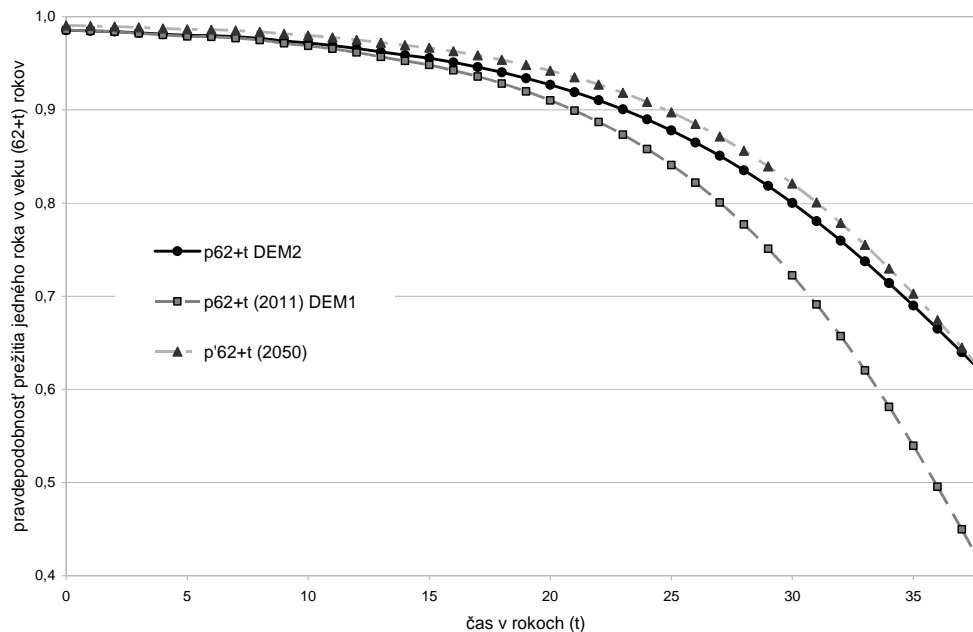
Postupným skúšaním rôznych hodnôt parametra u sme dostali vhodnú hodnotu $u = 35,47\%$. Získali sme tak sady pravdepodobností $p'_0, p'_1, \dots, p'_{100}$, resp. ${}_0 p'_0, {}_2 p'_0, {}_3 p'_0, \dots, {}_{100} p'_0$, ktoré sú z hľadiska poisťovne pesimistickejšie, pretože sú väčšie ako pôvodné pravdepodobnosti zo základného modelu **DEM1**.

V súlade s cieľmi uvedenými v úvodnej časti článku sme v ďalšom postupe uvažovali poistenú osobu, budúceho dôchodcu, vo veku $x = 62$ rokov. V našom modeli **DEM2** sme skonštruovali postupný prechod od pôvodných pravdepodobností $p_{62}, p_{63}, \dots, p_{100}$ k novým pravdepodobnostiam $p'_{62}, p'_{63}, \dots, p'_{100}$ a získali sme tak finálne pravdepodobnosti $\tilde{p}_{62}, \tilde{p}_{63}, \dots, \tilde{p}_{100}$ (viď Obr. 3):

$$\tilde{p}_{62+t} = p_{62+t} + \frac{t}{38} (p'_{62+t} - p_{62+t})$$

$$\tilde{p}_{62+t} = p_{62+t} + \frac{t}{38} (1 - (1 - u)(1 - p_{62+t}) - p_{62+t})$$

$$\tilde{p}_{62+t} = p_{62+t} + \frac{t}{38} u(1 - p_{62+t}) \quad \text{pre } t = 0, 1, \dots, 38.$$



Obr. 3: Znáznornenie pravdepodobností prežitia \tilde{p}_{62+t} v modeli DEM2 (stredná krivka): postupný prechod od pravdepodobností prežitia p_{62+t} k hodnotám p'_{62+t}

V treťom demografickom modeli, ktorý sme označili skratkou **DEM3**, sme používali odporúčanie Európskeho orgánu pre poisťovníctvo a dôchodkové poistenie zamestnancov (*European Insurance and Occupational Pensions Authority, EIOPA*). V odporúčaných stresových testoch organizácie EIOPA pre dlhovekosť poistených osôb sa obvykle používa koeficient zníženia úmrtnosti 10%, resp. 18% s okamžitou platnosťou, takzvaný jednorazový šok poklesu pravdepodobností úmrtia (viď EIOPA, online). V predpokladoch **DEM3** sme zvolili vyššiu zo spomínaných hodnôt, t. j. predpokladali sme, že pravdepodobnosť úmrtia sa zníži v každom veku o 18%. Modifikované pravdepodobnosti úmrtia, resp. pravdepodobnosti prežitia sme počítali nasledovne:

$$\bar{q}_x = (1 - 0,18)q_x, \quad \bar{p}_x = 1 - \bar{q}_x = 1 - (1 - 0,18)q_x \quad \text{pre } x = 0, 1, \dots, 100,$$

kde za q_x sme zobrali hodnoty zo základného demografického modelu **DEM1**.

4. Metóda spárovania aktív a pasív životnej poisťovne

Ako sme už uviedli v úvodnej časti, pri uzatvorení zmluvy o vyplácaní doživotných predlehotných dôchodkov z úspor v II. pilieri poisťovňa inkasuje jednorazové poistné. V našich modelových výpočtoch sme predpokladali, že poisťovňa v čase 0 investuje do τ -ročných dlhopisov presne toľko peňazí, koľko bude potrebovať na vyplatenie očakávanej dôchodkovej dávky¹ v čase τ , pre $\tau \in \{1, 2, 3, \dots, n\}$, kde n je maximálna doba vyplácania dôchodku. To znamená, že pri tejto zjednodušenej (statickej) investičnej stratégii poisťovňa

¹ Pre jednoduchosť predpokladáme, že ak poisťovňa nemá k dispozícii peniaze na vyplatenie dôchodku, tak si dokáže požičať na trhu pri rovnakej úrokovej miere, pri ktorej by sama investovala.

počas trvania poistnej doby už nespraví žiadne investičné rozhodnutia (jedine v čase podpisu poistnej zmluvy poisťovňa rozloží celé poistné do dlhopisov s rôznymi maturitami a predpokladá, že jej príjmy z investícií presne pokryjú očakávané dôchodkové dávky), preto nazývame našu analýzu statickou. K tomu je potrebné dodať aj ďalšie modelové predpoklady:

- 1) na trhu existuje plná škála dlhopisov z hľadiska doby splatnosti, t. j. dlhopisy s 1-ročnou, 2-ročnou, 3-ročnou, ..., n -ročnou splatnosťou,¹
- 2) v poisťovni uzatvorí zmluvu o doživotnom dôchodkovom poistení veľký počet poistencov, a tak poisťovňa môže používať princíp veľkého (fiktívneho) súboru a pracovať na báze očakávaných dávok,
- 3) poisťovňa vypláca predlehotný dôchodok raz ročne, v rovnakej výške.²

Z vyššie uvedenej jednoduchej investičnej schémy vyplýva, že v tomto prípade sa dá používať zjednodušená (statická) verzia metódy spárovania aktív a pasív. Poisťovňa k záväzku splatnému v čase τ (očakávanému poistnému plneniu) pripojí aktívum (štandardný dlhopis) s rovnakou budúcou hodnotou a dobou splatnosti. Označme symbolom c_τ ; $\tau \in \{1, 2, 3, \dots, n\}$ relatívnu časť jednorazového poistného P , ktoré sa investuje do τ -ročných štandardných dlhopisov. Nech S označuje ročnú výšku dôchodkovej dávky. Potom platí:

$$c_\tau \times (P - S) = S \times {}_\tau p_x \times \exp(-R(0, \tau) \times \tau),$$

kde ${}_t p_x$ je faktor prežitia τ rokov (vo výpočtoch získaný z modelov **DEM1**, **DEM2**, resp. **DEM3**) a $R(0, \tau)$ je výnos τ -ročného štandardného dlhopisu (odhadovaný modelmi *CIR2011*, *CIR2014*, *NS2014*, resp. *SV2014*). Poznamenáme, že poisťovňa v čase 0 síce inkasuje od poistenca čiastku P , no hneď nato musí vyplatiť poistencovi prvú predlehotnú dôchodkovú dávku S , teda ostane jej suma $P - S$, ktorú následne rozloží do investícií. Po vyjadrení hľadaných investičných percent sme dostali

$$c_\tau = \frac{S \times {}_\tau p_x \times \exp(-R(0, \tau) \times \tau)}{P - S} \quad \text{pre } \tau = 1, 2, \dots, \tau_0,$$

ak platí hraničná podmienka $\sum_{h=1}^{\tau_0} c_h \leq 1$. Ak $\tau_0 = n - 1$, tak to znamená, že pri daných demografických predpokladoch a danom vývoji investičných výnosov inkasované poistné postačí na pokrytie očakávaných dôchodkových dávok a v hraničnom čase n poisťovňa nebude vykazovať stratu. V takomto prípade nadbytok jednorazového poistného investuje poisťovňa v čase 0 do n -ročných dlhopisov, pretože tie majú najvyšší ročný výnos (aspoň podľa našich štyroch modelov, viď v 2. kapitole). Platí teda: $c_n = 1 - \sum_{\tau=1}^{n-1} c_\tau$.

V opačnom prípade, teda ak $\tau_0 < n - 1$, jednorazové poistné P by nestačilo na vyrovnanie záväzkov a v posledných rokoch poistenia by si poisťovňa musela požičať. Predpokladali sme, že v takomto prípade si poisťovňa dokáže požičať na trhu pri rovnakej úrokovej miere, pri ktorej by ona sama investovala. Označme premennou $\tau_1 \in \mathbb{N}$, $\tau_1 < n - 1$ časový bod, pre ktorý platí $\sum_{\tau=1}^{\tau_1} c_\tau < 1$ a zároveň $\sum_{\tau=1}^{\tau_1+1} c_\tau \geq 1$. Následne položíme $c_{\tau_0+1} = 1 - \sum_{\tau=1}^{\tau_0} c_\tau$ a $c_{\tau_0+2} = c_{\tau_0+3} = \dots = c_{n-1} = 0$. Takouto úpravou sme vlastne zabezpečili, aby aj v tomto druhom prípade platilo: $\sum_{\tau=1}^{\tau_0+1} c_\tau = \sum_{\tau=1}^n c_\tau = 1$.

Čistá akumulovaná hodnota poistenia (*net accumulated value*, *NAV*) v čase $t \in \{1, 2, 3, \dots, n\}$ v prípade predlehotného dôchodku s jednorazovo plateným poistným sa definuje vzťahom

$$NAV_t \equiv NAV_t(c_t) = -S \times {}_t p_x + (P - S) \times c_t \times \exp(R(0, t) \times t).$$

¹ V realite takúto konštrukciu je možné dosiahnuť vhodnou kombináciou kupónových dlhopisov.

² V praxi starobné dôchodky sa vyplácajú mesačnou frekvenciou, v našich modelových výpočtoch sme však predpokladali ročný interval vyplácania dôchodkových dávok.

Hodnota NAV_t neobsahuje všetky peňažné toky súvisiace s predlehotným dôchodkom, len tie, ktoré majú „splatnosť“ práve v čase t . Všimnime si, že ak $\tau_0 = n - 1$, tak $NAV_1 = NAV_2 = \dots = NAV_{n-1} = 0$, pretože v každom celočíselnom časovom bode (vo výročiach uzatvorenia poistnej zmluvy) investičné príjmy poisťovne presne vyrovnajú očakávané výdavky (vyplývajú z investičnej stratégie skonštruovanej pomocou spárovania aktív a pasív) a $NAV_n \geq 0$, t. j. zisk sa koncentruje do maximálneho času ukončenia poistného kontraktu. V druhom prípade, keď $\tau_0 < n - 1$, platí $NAV_1 = NAV_2 = \dots = NAV_{\tau_1} = 0$, a zároveň $NAV_{\tau_1+1} \leq 0$, $NAV_{\tau_1+2} < 0$, ..., $NAV_{n-1} < 0$, teda poisťovňa si bude musieť požičiavať na vyplatenie dôchodkových dávok a utrpí stratu.

Nech $DVFP$ označuje diskontovanú hodnotu budúceho zisku poisťovne (*discounted value of future profits*) a nech platí

$$DVFP = \sum_{t=1}^n NAV_t \times \exp(-R(0, t) \times t), \quad (4)$$

kde $R(0, t)$ znamená aj v tomto prípade (odhadovaný) ročný výnos t -ročného štandardného dlhopisu. V rámci našich štúdií sme práve pomocou veličiny $DVFP$ odhadovali súčasnú hodnotu zisku/straty poisťovne (viď v 5. kapitole).

5. Výsledky výpočtov

Pri našich výpočtoch sme aplikovali vyššie definované modely, predpoklady a označenia. Uvažovali sme teda poistenca (dôchodcu) vo veku $x = 62$ rokov s úsporami v II. pilieri vo výške $P = 15000$ eur. Ako sme to už uviedli vo 4. kapitole, v našich výpočtoch sme predpokladali ročné intervaly vyplácania dôchodkových dávok (namiesto mesačnej frekvencie vyplácania aplikovanej v praxi). Pomocou tohto predpokladu sme zjednotili frekvencie platenia dávok a maturovania dlhopisov.¹ Maximálnu dobu vyplácania dôchodku sme obmedzili na $n = 40$ rokov. Ročnú výšku dôchodkovej dávky sme počítali pomocou klasického netto-princípu životného poistenia (viď napríklad v knihe Potocký, 2012):

$$P = S \times \ddot{a}_{62} = S \times \ddot{a}_{62:\overline{40}|i} = S \times \sum_{t=0}^{39} {}_t p_x \times (1 + i)^{-t}, \quad (5)$$

kde \ddot{a}_{62} označuje súčasnú hodnotu predlehotného dôchodku, $i = 0,019$ p. a. je (maximálna) technická úroková miera a ${}_t p_x$ pre $t = 0, 1, 2, \dots, 39$ sú faktory prežitia počítané podľa *ÚT ŠÚSR 2011 spolu* (t. j. na základe modelu **DEM1**). Po vykonaní výpočtov sme dostali ročnú dávku vo výške $S = 954,22$ eur. V ďalších krokoch sme postupne vypočítali očakávané dávky na začiatku jednotlivých rokov poistenia, aplikovali metódu spárovania aktív a pasív, kalkulovali investičné percentá c_t a počítali čistú akumulovanú hodnotu poistenia. V Tab. 2 sme uviedli výsledky štúdií: odhadnutú diskontovanú hodnotu budúceho zisku poisťovne ($DVFP$) pri rôznych demografických nastaveniach modelu a rôznych predpokladoch ohľadom výnosov dlhopisov.

Tab. 2: Hrubý odhad diskontovanej hodnoty budúceho zisku poisťovne (uvedená v eurách) v prípade doživotného dôchodku pri rôznych modeloch výnosových kriviek, rôznych demografických predpokladoch a jednorazovom poistnom vo výške 15000 eur

	<i>DVFP (CIR2011)</i>	<i>DVFP (CIR2014)</i>	<i>DVFP (NS2014)</i>	<i>DVFP (SV2014)</i>
DEM1	1911,30	360,76	477,48	119,41
DEM2	1523,75	-132,59	-41,93	-416,25
DEM3	1257,96	-402,64	-303,28	-645,14

¹ V prípade sofistikovaného dynamického modelu s pravidelným prúdením sporiteľov by sme mohli aplikovať aj mesačnú frekvenciu vyplácania dôchodkových dávok. V zásade by však aj tento model dával podobné výsledky ako náš jednoduchý statický model.

Z Tab. 2 vidíme, že pri základnom demografickom modeli **DEM1** by poisťovňa skončila so ziskom pri všetkých štyroch modeloch výnosov dlhopisov. Je to hlavne kvôli tomu, že všetky naše modely výnosov dlhopisov predpovedajú vyššie výnosy slovenských dlhopisov ako je technická úroková miera $i = 0,019$ p. a., pri ktorej sa kalkulovala ročná výška dôchodkovej dávky. Najoptimistickejší odhad pre výnosy dlhopisov a veličinu *DVFP* nám dal *CIR2011 model*, pretože kalibračný súbor tohto modelu obsahoval aj výnosy z optimistického obdobia 2011-2012. Pri pesimistickejších demografických predpokladoch z **DEM2** jedine v prípade *CIR2011 modelu* by poisťovňa dosiahla zisk v maximálnom čase ukončenia poistného kontraktu, pri ostatných troch modeloch by skončila v strate. Na základe výsledkov uvedených v Tab. 2. môžeme skonštatovať, že najhorším scenárom z hľadiska poisťovne je **DEM3**, pričom globálne najhoršou kombináciou je **DEM2** + *SV2014 model*, pri ktorej by poisťovňa utrpela stratu vo výške 645,14 peňažných jednotiek. Z toho vyplýva, že poisťovní by nemuselo stačiť jednorazové poistné inkasované z II. dôchodkového piliera na vyplácanie doživotných dôchodkových dávok počítaných pri základných demografických predpokladoch a na základe klasického netto-princípu.

6. Záver

Vyplácanie doživotných dôchodkov z úspor v II. dôchodkovom pilieri môže prinášať pre životné poisťovne viaceré riziká. Okrem demografických, trhových a úrokových rizík, ktoré sme skúmali v rámci tohto príspevku, to môžu byť aj rôzne politické, menové alebo sociálne riziká. V každom prípade, životné poisťovne pri uzatvorení zmlúv o doživotných anuitách by mali brať do úvahy, že sa jedná o dlhodobý poistný kontrakt ovplyvnený viacerými rizikovými faktormi a zodpovedne stanoviť výšku dohodnutých dôchodkových dávok. V našich štúdiách sme analyzovali vplyv premenlivých trhových výnosov a rôznych demografických predpokladov na zisk poisťovne. V úvodnej časti sme ponúkli motiváciu na výskum problematiky a stanovili sme základné predpoklady. V druhej, resp. tretej kapitole sme uviedli kalibráciu štyroch modelov pre výnosy dlhopisov, resp. troch demografických modelov. V predposlednej kapitole sme definovali metódu spárovania aktív a pasív. V záverečnej piatej kapitole sme zhrnuli výsledky našich analýz a výpočtov.

PodĎakovanie

Príprava článku bola podporovaná grantom APVV-0465-12.

Literatúra

- CONSIGLIO, A. - GUIRRERI, S. S. 2011. Simulating the Term Structure of Interest Rates with arbitrary marginals. In: *International Journal of Risk Assessment and Management*, roč. 15, č. 4, September 2011.
- COX, J. - INGERSOLL, J. - ROSS, S. 1985. A theory of the term structure of interest rates. In: *Econometrica*, roč. 53, č. 2, s. 385-407.
- EIOPA [online]. EIOPA Stress Test 2014. [cit. 11.11.2014]. Dostupné na adrese: <https://eiopa.europa.eu/fileadmin/tx_dam/files/activities/financial_stability/insurance_stress_test_2014/eiopa-14-215_stress_test_2014_specifications.pdf>
- GUIRRERI, S. S. 2010. Simulating the Term Structure of Interest Rates with arbitrary marginals. University of Palermo - Department of Statistics and Mathematics "S. Vianelli".
- INFOSTAT [online]. Prognóza vývoja obyvateľstva SR do roku 2050. In: *Edícia: Akty*, Bratislava, november 2002. [cit. 21.10.2014]. Dostupné na adrese: <<http://www.infostat.sk/vdc/pdf/prognóza2050vdc2.pdf>>.
- KWOK, Y. K. 1998. *Mathematical Models of Financial Derivatives*. New York, Heidelberg, Berlin: Springer-Verlag.

- MELICHERČÍK, I. - OLŠÁROVÁ, L. - ÚRADNÍČEK, V. 2005. *Kapitoly z finančnej matematiky*. Epos, Bratislava. ISBN 80-8057-651-3.
- NELSON, C. R. – SIEGEL, A. F. 1987. Parsimonious Modeling of Yield Curve. In: *Journal of Business* 1987, roč. 60, č. 4, s. 473-489.
- PASTOR, K. - RAUČINOVÁ, E. 1991. Parametrizácia úmrtnostných tabuliek a jej použitie v regionálnom plánovaní. In: *Acta Facultatis Rerum Naturalium Universitatis Comenianae*, s.119-124.
- POTOCKÝ, R. 2012. *Modely v životnom a neživotnom poistení*. Vydavateľstvo STATIS Bratislava. ISBN 978-80-85659-71-9.
- R CORE TEAM. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <<http://www.R-project.org/>>.
- SVENSSON, L. E. O. 1994 [online]. Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. In: *National Bureau of Economic Research Working Paper*, Cambridge, Massachusetts, september 1994, č. 4871. [cit. 12.11.2014]. Dostupné na adrese: <http://www.nber.org/papers/w4871.pdf?new_window=1>
- ŠEVČOVIČ, D. - URBÁNOVÁ CSAJKOVÁ, A. 2005. On a two-phase minmax method for parameter estimation of the Cox, Ingersoll, and Ross interest rate model. In: *Central European J. of Operation Research*, roč. 13, s. 169-188.
- ŠÚSR [online]. Tabuľky úmrtnosti (roky 1996 - 2013): Slovenská republika. In: *Obyvateľstvo a migrácia: Tabuľky života*. Štatistický úrad Slovenskej republiky. [cit. 21.10.2014]. Dostupné na adrese: <http://slovak.statistics.sk/wps/wcm/connect/221c1ae4-1e73-491a-ba30-744fba17f037/Umrtnostne_tabulky_SR_1996_2013.zip?MOD=AJPERES>.

Adresa autorov:

Igor Melicherčík, doc. Mgr. PhD.
Fakulta matematiky, fyziky a informatiky,
Univerzita Komenského v Bratislave
Mlynská dolina, 842 48 Bratislava
igor.melichercik@fmph.uniba.sk

Gábor Szűcs, Mgr.
Fakulta matematiky, fyziky a informatiky,
Univerzita Komenského v Bratislave
Mlynská dolina, 842 48 Bratislava
gabor.szucs@fmph.uniba.sk

International Financial Integration and Volatility of Financial Flows

Medzinárodná finančná integrácia a volatilita finančných tokov

Rajmund Mirdala, Aneta Svrčeková

Abstract: Macroeconomic stability represents a strong fundamental pillar for maintenance of a long-term economic growth. Current empirical literature highlights a necessity to observe the implications of increasing international financial integration on the macroeconomic stability of the countries. While economic effects of international financial integration obviously differs according to the overall performance of the country, implications of excessive financial flows volatility still remain disputable. In the paper we examine the relationship between international financial integration and volatility of financial flows. Examination of the international financial integration and its effects on macroeconomic volatility or stability is particularly important due to existence of generally expected positive relationship between macroeconomic volatility and economic growth, common trends of decreased macroeconomic instability worldwide and occurrence of negative sides of financial integration - financial crises. Following our results we suggest that relationship between financial integration and volatility of financial flows is positive, however not significant. Moreover the relationship is stronger in case of developing countries.

Abstrakt: Makroekonomická stabilita predstavuje výrazný fundamentálny pilier udržania dlhodobého ekonomického rastu. Súčasná empirická literatúra zdôrazňuje potrebu skúmania vplyvu narastajúcej medzinárodnej finančnej integrácie na makroekonomickú stabilitu krajín. Zatiaľ čo sa efekty medzinárodnej finančnej integrácie líšia v závislosti od celkovej výkonnosti krajiny, dopady nadmernej volatility medzinárodných finančných pohybov zostávajú diskutabilné. V príspevku budeme analyzovať vzťah medzinárodnej finančnej integrácie a volatilitou finančných tokov. Skúmanie medzinárodnej finančnej integrácie a jej pôsobenia na makroekonomickú volatilitu či stabilitu je prínosné aj v kontexte očakávania pozitívnej previazanosti medzi makroekonomickou volatilitou a ekonomickým rastom, dlhodobých trendov klesajúcej makroekonomickej nestability v celosvetovom meradle, avšak na strane druhej výskytu sprievodných negatívnych fenoménov finančnej integrácie - finančnými krízami. Na základe našich výsledkov možno konštatovať, že vzťah medzi finančnou integráciou a volatilitou finančných tokov je pozitívny, avšak nie výrazný. Identifikovaný vzťah je pritom výraznejší v rozvíjajúcich sa ekonomikách.

Key words: financial integration, volatility of financial flows, macroeconomic volatility

Kľúčové slová: finančná integrácia, volatilita finančných tokov, makroekonomická volatilita

JEL classification: C32, E31, F41

1. Introduction

Empirical literature provides rich evidence about the effects of international financial integration on economic growth. Moreover, many authors examined the causal relationship between international financial integration and economic growth. Most of the empirical studies suggest that, on average, more financially integrated economies performs better than less financially open economies, in terms of improvements in per capita output as a measure of standards of living. However, many of studies also come to the conclusion that relationship between financial integration and economic growth is not always strong or robust. There has also been a rigorous debate about the relationship between macroeconomic volatility and financial integration. However, empirical evidence on the effects of international financial integration on volatility is far more limited. While the key features together with the main and

side implications of international financial integration represents the crucial topic of economic research for decades, the results of many empirical studies provides different or even biased conclusions. The main objective of the paper is to investigate effects of international financial integration on macroeconomic volatility that enables us to highlight some key stylized facts about their mutual relationship.

2. Overview of the Literature

Investigation of international financial integration, its crucial patterns, trends and effects on macroeconomic volatility is still quite challenging area of economic research. Following the results of many research studies, examination of mutual relationship between international financial integration and macroeconomic volatility, leads to different empirical conclusions. This fact is not surprising, given that economic theory does not provide comprehensive and clear view of how the increasing international financial openness and international financial integration should affect the volatility of macroeconomic outcomes and others macroeconomic variables. The ambiguity of results proposed by large number of empirical studies is caused by a relative variety in methodology and econometric techniques and variability of samples of countries and time horizons included in the most of regression models. Moreover, different threshold effects that determine the growth effects of international financial integration also contributed to the relative diversity of empirical results. Kose et al. (2006) highlights the composition of capital flows, domestic financial development, institutional development and discipline, macroeconomic discipline and trade openness as the main determinant affecting the impact of international financial integration on macroeconomic volatility. Composition of international capital flows is probably the most crucial determinant affecting the macroeconomic volatility. Taylor and Sarno (1999) revealed that FDI flows are more stable and persistent than other groups of international capital flows. These results are confirmed by Hausmann and Fernandez-Arias (2000), who confirmed that although the volatility of FDI flows followed increasing trend in last few decades, it still remains much lower than volatility of other types of capital flows.

Fischer and Reisen (1992), Bekaert, Harvey and Lundblad (2006), IMF (2007) and Herrera and Vincent (2008) revealed significantly negative relationship between financial openness and macroeconomic volatility. Bekaert et al. (2004) analyzed effects of stock market liberalization and capital account openness on the volatility of real consumption growth rate over the 20 years period. Their results show a strong negative correlation between international financial liberalization and consumption volatility. Bekaert et al. (2002) show, that overall capital account openness has weaker influence on output and consumption volatility. Therefore, the authors suggest that it is the integration of stock markets worth reducing the output volatility. Herrera and Vincent (2008) show similar results. According to research by the authors mentioned above, greater integration of financial markets to international capital markets is associated with lower macroeconomic volatility.

On the other hand, Kose et al. (2003) provides a comprehensive analysis of changes in macroeconomic volatility in the large group of industrial and developing countries over 50 years (76 countries, period 1960-1999). Authors divided developing countries into two groups (more (MFI) and less (LFI) financially integrated countries). Their results show a positive though insignificant impact of international financial integration on the volatility of domestic output and domestic consumption. In case of the relative consumption volatility, expressed as the ratio of the consumption and output volatility, they confirmed that international financial integration has a positive and significant impact. Mendoza (1992) suggested that the volatility of output increases with increasing degree of international financial integration when countries experience large and long-term shocks. However the relationship is insignificant

too. Baxter and Crucini (1995) partially confirm the Mendoza's findings about the output volatility though they rejected results about consumption volatility. Volatility of real and relative consumption is decreasing by growing international financial integration. Study of Gavin and Hausmann (1996) shows that capital account openness is a significant potential channel of macroeconomic volatility causing domestic output fluctuations in Latin America. Study is focused on the period 1970-1992. Islam and Stiglitz (2000) confirmed that financial openness significantly contributes to the volatility of GDP growth per capita in developed and developing OECD countries. Evans and Hnatovska (2006) suggest that there exist a nonlinear relationship between macroeconomic volatility and international financial integration. International financial integration causes an increase in the output and consumption volatility initially, but this relationship is gradually replaced by a much larger decrease in the volatility of macroeconomic variables. The final effect of changes in the macroeconomic volatility is positive for the country. On the other hand, the lack of empirical evidence about the existence of relationship between international financial integration and macroeconomic volatility is presented by Razin and Ros (1994), Butch, Döpke and Pierdzioch (2002) and others.

3. Data and Methodology

In our study we employ data from the database of Lane and Milesi-Ferretti (2007) which consist of comprehensive data on foreign financial assets and liabilities for a large sample of countries for the period 1970-2011. For measure of financial integration (openness) we use modification of conventional trade openness indicator and calculate financial integration as sum of capital inflows and outflows divided by GDP. We also calculate international financial integration using financial flows subcategories - foreign direct investments, portfolio investments and debt investments. We use median of standard deviation to measure cross-sectional volatility of variables applied in this paper Kose et al. (2003). We calculate standard deviation for whole period observed, and then we compute volatility for 10 years periods, namely 1970-1979, 1980-1989, 1990-1999, and 2000-2009 to review changes in the volatility through the time. We define a measure of overall and decomposed macroeconomic volatility in few ways. First, we use volatility of real GDP per capita growth for evaluating the overall macroeconomic volatility. Then we calculate decomposed volatility as real private consumption per capita growth. Because the cyclical components of government consumption may affect households' consumption, we also use the constant final consumption expenditure per capita growth to measure consumption volatility. This could be particularly important for less developed economies as well as more open economies that tend to have higher ratios of government consumption to output. Then we calculate ratio of final consumption volatility to output volatility to evaluate efficacy of consumption smoothing. This ratio should be significantly lower in industrial countries in comparison with developing countries. Finally we calculate volatility of real total investment growth per-capita. Suggested decomposition enables us to reveal different effects of international financial integration on households and business sector. Total investment growth is measured as gross fixed capital formation.

For evaluating the extent of variations in effects of international financial integration it is useful to begin with rough classification of the developing countries into two groups based on ranking according to the average measures of the financial openness over the last four decades. Our sample of countries is divided in two groups (developed and developing countries) following the classification provided by International Monetary Fund. In order to examine the effects of international financial integration we follow Kose et al. (2003) and subdivide developing economies into two groups: more financially integrated (MFI) and less financially integrated (LFI). The criterion used to organize countries into these two groups is cross-sectional median of financial openness representing the value 1.02831 for the whole

sample of developing countries over the period 1970-2011. Following this procedure we have obtained 23 developed countries, 32 MFI countries and 48 LFI countries.

4. Trends in Financial and Macroeconomic Volatility

4.1 Volatility of International Financial Flows

Financial openness increased in the most countries all over the world during the period of the last four decades. Intensified international financial integration followed by higher financial openness caused an increase in the volatility of international financial flows. Due to existing differences in the dynamics of international financial integration and associated volatility of international financial flows, many authors focused on examination of the relationship among financial integration, volatility of financial flows and macroeconomic volatility. In Tab. 1 we summarize the evolution of international financial flows volatility over time.

Our results confirm that the volatility of international financial flows is higher in developed countries in comparison with developing countries. In addition, the volatility of international financial flows rises with higher international financial openness. Financial volatility of LFI economies is clearly smaller than of those from the group of MFI economies.

Volatility of financial flows (median)					
Standard deviation ¹	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	1.501	0.092	0.201	0.407	0.646
MFI	0.607	0.196	0.396	0.248	0.311
LFI	0.320	0.077	0.137	0.158	0.164

**Tab. 1 Cross-sectional Volatility of Financial Flows
Measured by Median of Standard Deviation**

Source: Authors calculations.

Tab. 2 summarizes estimated results for volatility of financial flows measured by cross-sectional mean of standard deviation. While the volatility of LFI economies remained at the same levels despite the changed measurement, we observed a substantial increase in the financial volatility in the group of developed and MFI economies. Our results also indicate that the volatility of financial flows in MFI economies is even higher than the volatility in developed countries revealing higher diversity of calculated results in the countries from this group. Moreover, coefficients of volatility of financial flows significantly drop to from 2.136 to 0.875 (for the whole period) if the outliers are excluded from the analysis. After this adjustment, the results are more consistent with previous table.

Tables 1 and 2 also summarize estimated volatilities of international financial flows decomposed into individual decades (columns 2-5) within the whole period to examine the evolution of main trends. It seems that median volatility rises over the observed period in all three groups of countries. As we expected, the most dynamic changes in terms of a relative increase in the median of standard deviation experienced developed economies. We suggest that this trend was caused by the rapid growth of the financial openness in developed economies over the examined period. Quite similar results are reported by the Tab. 2. In general, increasing financial openness is obviously accompanied by the excessive growth of volatility of financial flows. This relation is observable mainly in developed countries and

¹ Median of standard deviation for each group of countries

MFI economies. The overall dynamics in volatility of financial flows is thus positively affected by financial openness.

Volatility of financial flows (mean)					
Standard deviation ¹	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	2.091	0.125	0.246	0.591	1.192
MFI	2.136	0.387	0.944	0.745	1.162
LFI	0.331	0.089	0.174	0.164	0.191
Developed economies ²	1.743	0.120	0.250	0.481	1.005
MFI ³	0.875	0.409	0.640	0.363	0.377

**Tab. 2 Cross-sectional Volatility of Financial Flows
Measured by Mean of Standard Deviation**

Source: Authors calculations.

Examination of volatility of international financial flows revealed interesting implications of international financial integration according to the relative differences in the overall performance of the countries. Subsequent decomposition of international capital flows into key components provides additional information about the sources of volatility of international financial flows according to the typology of capital movements (foreign direct investment (FDI) and debt investments (portfolio debt investment included)⁴ are concerned). In Tab. 3 we investigate cross-sectional differences in the volatility of particular financial flows. This approach will be helpful in the next section of the paper to examine the relative importance of individual financial flows in determining the overall financial flows volatility.

Following our results we suggest that the contribution of debt investment flows is clearly more important in determining overall financial flows volatility. Tab. 3 shows the volatility of FDI and debt investment as ratio to GDP. It seems that FDI flows represents less volatile component of private financial flows given their long-term character and relatively stable nature. This is consistent with the paper Taylor and Sarno (1999). Authors conclude that FDI flows are more stable and persistent than other groups of financial flows. These results are also confirmed by Hausmann and Fernandez-Arias (2000), who shows that even volatility of FDI flows has been continuously growing over last few decades, it still remains generally much lower than the volatility of other types of international financial flows. Authors highlight that the most volatile flows of FDI can be found in MFI economies. However, the difference from developed countries is relatively small. Debt financial flows tend to be far more volatile and sensitive to sudden reversals than FDI. This conclusion also results from our analysis. In addition, according to Tab. 3 we suggest that FDI and debt flows volatility remains relatively stable in each group of the countries. However, an increments between individual periods has risen slightly in most cases.

¹ Mean of standard deviation for each country group

² Ireland excluded as outlier

³ Bahrain, Liberia, Mauritius and Singapore excluded as outliers (the group of four the most financially opened and volatile countries)

⁴ Portfolio equity investment are excluded from analysis due to lack of data available

Volatility of capital flows					
FDI/GDP					
Median of Standard deviation	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	0.148	0.008	0.014	0.035	0.066
MFI	0.186	0.040	0.029	0.054	0.095
LFI	0.086	0.011	0.009	0.027	0.052
Mean of Standard deviation	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	0.192	0.015	0.023	0.052	0.108
MFI	0.464	0.072	0.056	0.177	0.333
LFI	0.089	0.022	0.014	0.035	0.055
Debt investment/GDP					
Median of Standard deviation	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	0.451	0.052	0.093	0.099	0.242
MFI	0.470	0.136	0.205	0.188	0.180
LFI	0.223	0.062	0.103	0.106	0.163
Mean of Standard deviation	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	0.681	0.074	0.102	0.153	0.460
MFI	0.851	0.238	0.475	0.449	0.357
LFI	0.227	0.068	0.129	0.116	0.169

**Tab. 3 Cross-sectional Volatility of Different Capital Flows
Measured by Standard Deviation¹**

Source: Authors calculations.

As we already proposed, the volatility of financial flows is affected by increasing financial openness. We provide some stylized facts concerning financial flows volatility. Fig. 1 outlines the volatility of financial flows according to the measure of international financial integration for the full sample of countries as well as individual subsamples of countries. It is clear that countries with higher financial openness obviously experience higher financial volatility. Fig. 1 confirms the results from the previous sections. Positive relationship is presented in all countries group. However, the contribution of international financial integration in determining financial volatility differs for each group of countries.

Countries group	Correlation coefficient
All countries	0.851
Developed economies	0.930
MFI	0.816
LFI	0.503

Tab. 4 Correlation between Financial Integration and Volatility of Financial Flows

Source: Authors calculations.

In **Chyba! Nenašiel sa žiaden zdroj odkazov.** we provide results of a simple correlation analysis between international financial integration and volatility of financial flows. It seems that financial integration has the highest influence on the volatility of financial flows in the group of developed countries (the coefficient of correlation is 0.930). Slightly reduced effect

¹ FDI and debt investments measured as total FDI inflows and total debt inflows

of international financial integration we observed in the group of MFI countries (the coefficient of correlation is 0.816). LFI group of countries experienced lower interconnection between both variables (the coefficient of correlation is 0.503). We suggest that the relative importance of financial openness in determining financial volatility in LFI economies is clearly marginal and doesn't affect the financial volatility as the primary factor.

We conclude that increasing financial volatility induced by rising financial openness is followed by higher macroeconomic volatility. If this presumption is correct then macroeconomic volatility in developed countries should be the highest from all groups of countries.

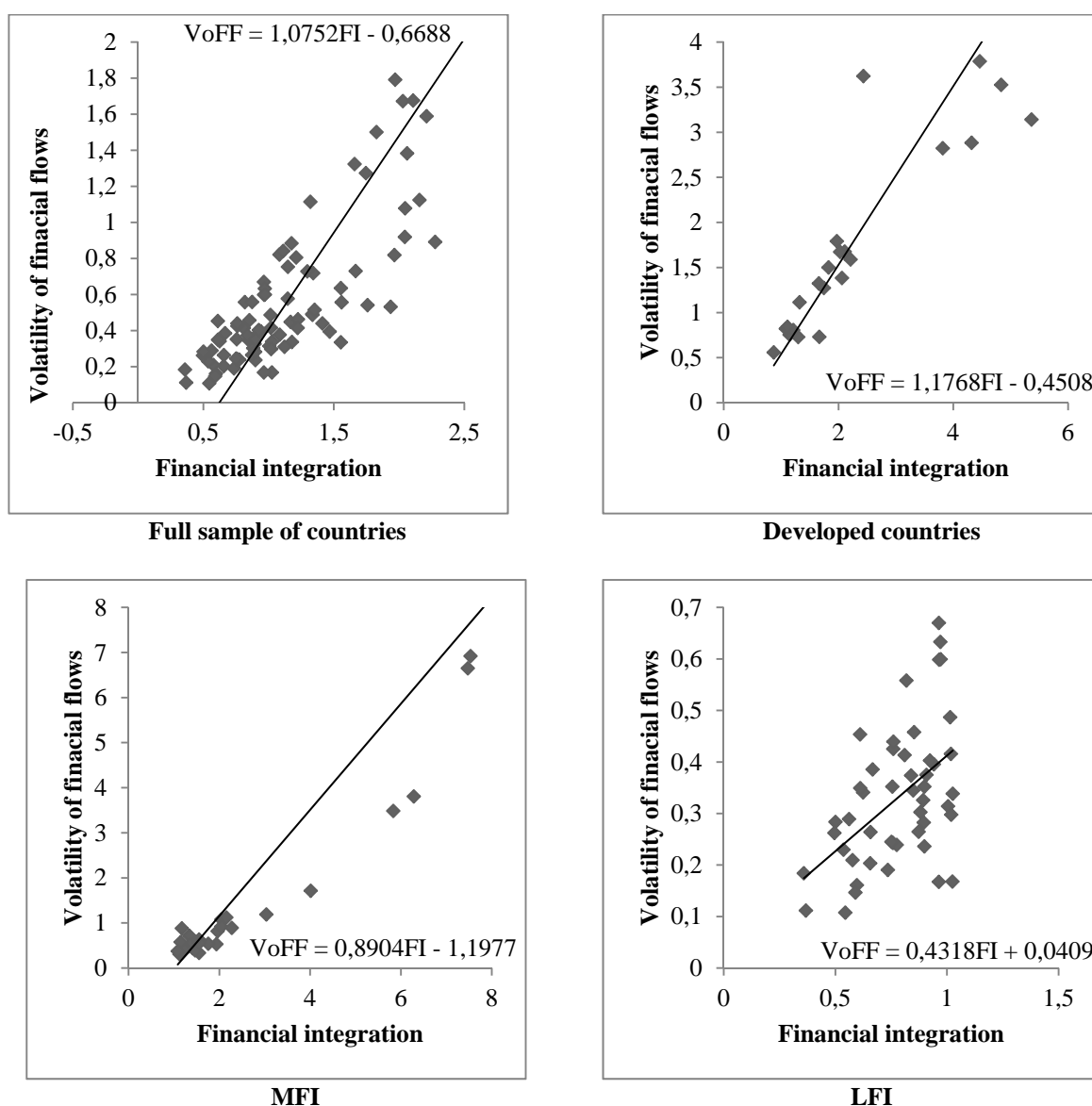


Fig. 1 Financial Integration and Volatility of Financial Flows

Note: VoFF - volatility of financial flows, FI - financial integration.

Source: Authors calculations.

4.2 Macroeconomic Volatility

In this part of the paper we investigate the macroeconomic volatility in all three groups of countries. Tab. 5 (column 1) shows the cross-sectional medians of the volatility of output,

consumption and investment growth over the whole period of four decades. Developed countries maintained substantially lower macroeconomic volatility than other two groups of countries. Despite higher volatility of international financial flows and higher financial openness it seems that developed countries were able to maintain generally low levels of the overall macroeconomic volatility. It seems that international financial integration and rising volatility of international financial flows was not associated with corresponding increase in the macroeconomic volatility in developed countries.

On the other hand, our results revealed different results for the group of developing countries (both MIF and LFI). In general, developing countries tend to experience more intensive fluctuations in macroeconomic variables in comparison with developed countries. Following the results from Tab. 5 (column 1) we suggest that MFI economies experienced substantially higher rates of volatility in macroeconomic variables than LFI countries. Moreover, in case of investment volatility, the gap is even much higher. As a result, increasing financial openness seems to be associated with undesirable effects on the macroeconomic stability in developing economies. Therefore, it is important to investigate more details about trends in macroeconomic volatility during the subsequent periods.

Macroeconomic volatility of growth rates of selected variables¹					
Output	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	2.200	2.486	1.771	3.360	2.167
MFI	5.017	4.669	4.416	4.062	2.354
LFI	4.173	3.499	4.109	4.081	2.437
Private consumption	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	2.085	2.444	1.978	1.730	1.499
MFI	8.542	8.517	7.920	5.972	4.463
LFI	5.867	4.966	5.059	4.377	3.607
Final consumption	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	1.725	1.948	1.368	1.279	1.010
MFI	6.778	6.735	6.483	4.836	4.127
LFI	5.182	4.558	4.549	4.050	3.142
Ratio of total consumption to output	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	0.167	0.235	0.742	0.879	0.783
MFI	0.468	0.593	0.859	0.867	0.856
LFI	0.385	0.545	0.915	0.925	0.883
Investment	1970-2011	1970-1979	1980-1989	1990-1999	2000-2009
Developed economies	6.326	6.158	6.149	5.403	5.986
MFI	20.996	16.577	16.219	16.541	11.116
LFI	14.670	12.560	13.733	12.968	8.825

Tab. 5 Macroeconomic Volatility

Source: Author's calculations.

Tab. 5 (columns 2-5) highlights changes in macroeconomic volatility of growth rates of selected variables during all four decades. In general, developing countries experienced decreasing trend in macroeconomic volatility over time. Moreover, this trend was obvious in both MFI and LFI economies. However, the decrease is more dynamic in the group of MFI

¹ Median of standard deviation for each group of countries

countries. We suggest that increasing financial openness associated with higher volatility of international financial flows was associated with higher macroeconomic stability in developing countries, especially MFI countries. This conclusion is obvious even though the overall macroeconomic volatility in MFI countries exceeded macroeconomic volatility in LFI countries in each individual decade.

Intensive international financial integration initially causes macroeconomic instability in developing countries. However, increasing openness to global financial markets causes significant economic improvement over time due to associated positive implications. We suggest that increasing financial openness affects the macroeconomic volatility in the positive way. In other words the relationship between examined variables is negative.

5. Conclusion

In the paper we analyzed main trends in international financial integration, volatility of financial flows and macroeconomic volatility in the group of developed and developing countries to examine mutual relationship between all indicators in countries with fundamentally different economic background. We provided some stylized facts and empirical support according to the main trends in all countries during last four decades. Following our results it seems that rising financial openness causes increase in the volatility of financial flows especially in developed and more financially integrated economies. We also provided supportive evidence that volatility of financial flows is increasing over the observed time. Finally, increasing financial openness associated with higher volatility of financial flows was accompanied by excessive macroeconomic volatility though empirical support for this relationship doesn't seem to be strong.

References

- AGHION, P., BANERJEE, A., PIKETTY, T. (1999) Dualism and Macroeconomic Volatility. *The Quarterly Journal of Economics*, 114(4): 1359-1397
- BECK, T., KUNT, A., LEVINE, R. (2001) Legal Theories of Financial Development, *Oxford Review of Economic Policy*, 17(4): 483-501
- BEKAERT, G., HARVEY, C., LUNDBLAD, CH. (2006) Growth Volatility and Financial Liberalization, *Journal of International Money and Finance*, 25(3): 379-406
- BEKAERT, G., HARVEY, C., LUNDBLAD, CH. (2002) Growth Volatility and Equity Market Liberalization. working Paper
- BUCH, C., DÖPKE, J., PIERDZIOCH, CH. (2002) Financial Openness and Business Cycle Volatility, [Kiel Working Paper, no. 1121], Kiel, Kiel Institute for the World Economy, 37 p.
- EOZENOU, P. (2008) Financial Integration and Macroeconomic Volatility: Does Financial Development Matter. MPRA Working Paper No. 12738, 35 p.
- ERA, D.N., NARAPONG, S. (2012) Revisiting the Link between Finance and Macroeconomic Volatility, [IMF Working Paper, No. 29/2013], Washington D.C., International Monetary Fund, 36 p.
- EVANS, M., HNATKOVSKA, V. (2006) Financial Integration, Macroeconomic Volatility and Welfare, *Journal of the European Economic Association*, 5(2-3): 500-508
- FISCHER, B., REISEN, H. (1992) Towards Capital Account Convertibility. [OECD Development Centre Policy Briefs, No. 4/1992], Paris, Organization for Economic Cooperation and Development, 28 p.
- FRIEDRICH, CH., SCHNABEL, I., ZETTELMEYER, J. (2010) Financial Integration and Growth: Is Emerging Europe Different, [EBRD Working Paper, No. 123/2010], London, European Bank for Reconstruction and Development, 42 p.
- GLICK, R., GUO, X., HUTCHISON, M. (2004) Currency Crises, Capital Account Liberalization, and Selection Bias, *The Review of Economics and Statistics*, 88(4): 698-714

- HAUSMANN, R., GAVIN, M. (1996) Securing Stability and Growth in a Shock Prone Region: The Policy Challenge for Latin America, [IADB Research Department Publication, No. 4020], Washington, Inter-American Development Bank, 42 p.
- HERRERA, S., VINCENT, B. (2008) Public Expenditure and Consumption Volatility, [WB Policy Research Working Paper, No. 4633], Washington, World Bank, 25 p.
- IMF Research Department (2007) Reaping the Benefits of Financial Globalization, [IMF Occasional Paper, No. 264/2007], Washington D.C., International Monetary Fund, 52 p.
- KAMINSKY, G., REINHART, C. (1999) The Twin Crises, *American Economic Review*, 89(3): 473-500
- KOSE, A., PRASAD, E., ROGOFF, K., WEI, S. (2010) Postscript to Financial Globalization and Economic Policies, book chapter in RODRICK, D., ROSENZWEIG, M.R., (eds.) 2010, Handbook of Development Economics, Book Vol. 5 (No. 6), Elsevier Science Pub., New York, 1072 p.
- KOSE, A., PRASAD, E., TERRONES, M. (2003) Financial Integration and Macroeconomic Volatility, [IMF Working Paper, No. 03/50], Washington D.C., International Monetary Fund, 27 p.
- KOSE, A., PRASAD, E., TERRONES, M. (2006) How do Trade and Financial Integration Affect the Relationship between Growth and Volatility, *Journal of International Economics*, 69(1): 176-202.
- LANE, P.R., MILESI-FERRETTI, G.M. (2007) The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970–2004 , *Journal of International Economics*, 73(2): 223-250
- MIRDALA, R. et al. (2011) Economic Aspects of Capital Flows Liberalization in the European Transition Economies, University Press, Technical University of Košice, First Edition, 230 p.
- MENDOZA, E. (1992) Robustness of Macroeconomic Indicators of Capital Mobility, [IMF Working Paper, No. 92/111], Washington D.C., International Monetary Fund, 40 p.
- RAZIN, A., ROSE, A. (1994) Business Cycle Volatility and Openness: An Exploratory Cross-Section Analysis, [NBER Working Paper, no. 4208], Cambridge, National Bureau of Economic Research, 31 p.
- RUMLER, F., SCHARLER, J. (2009) Labor Market Institutions and Macroeconomic Volatility in a Panel of OECD Countries, [ECB Working Paper, no. 1005], Frankfurt am Main, European Central Bank, 27 p.
- SEDIK, T., SUN, T. (2012) Effects of Capital Flow Liberalization: What is the Evidence from Recent Experiences of Emerging Market Economies?, [IMF Working Paper, No. 12/275], Washington D.C., International Monetary Fund, 26 p.
- SOUZA, L.V. (2004) Financial Liberalization and Business Cycles: The Experience of Countries in the Baltics and Central Eastern Europe, [DB Discussion Paper, No. 23/2004], Frankfurt am Main, Deutsche Bundesbank, 44 p.
- SUTHERLAND, A. (1996) Financial Market Integration and Macroeconomic Volatility, *Scandinavian Journal of Economics*, 98(4): 521-539

Acknowledgement

This paper was written in connection with scientific project VEGA no. 1/0892/13. Financial support from this Ministry of Education's scheme is also gratefully acknowledged.

Authors:

Rajmund Mirdala, doc. Ing. PhD.
Faculty of Economics
Technical University of Košice
Němcovej 32
04001 Košice
Slovak republic
rajmund.mirdala@tuke.sk

Aneta Svrčeková, Ing.
Faculty of Economics
Technical University of Košice
Němcovej 32
04001 Košice
Slovak republic
aneta.avrcekova@tuke.sk

Využití shlukové analýzy ke srovnání dopadů celosvětové finanční krize na trh práce zemí Evropské unie

The use of cluster analysis for a comparison of the global financial crisis effects on the labour market of the European Union

Tomáš Pavelka, Tomáš Löster

Abstract: The overall economic situation of EU member states is strongly connected with their labour market. The aim of the article is using a cluster analysis to find groups of countries with similar characteristics. The paper use indicators of the labour market, the overall unemployment rate and the incidence of long-term unemployment. The article compare the situation before the start of the economic recession with the situation after the start of the economic recession. It turns out that the economic recession had a significant impact on the unemployment rate and the incidence of long-term unemployment. The same development in all countries, however, cannot be always found.

Abstrakt: Celková ekonomická situace jednotlivých členských zemí Evropské unie je silně propojena s jejich trhem práce. Cílem článku je pomocí shlukové analýzy nalézt skupiny zemí, které mají podobné charakteristiky. V článku jsou využity ukazatele trhu práce, a to celková míra nezaměstnanosti a míra výskytu dlouhodobé nezaměstnanosti. V článku je srovnávána situace před vypuknutím ekonomické recese se situací po vypuknutí ekonomické recese. Ukazuje se, že ekonomická recese měla značné dopady na míru nezaměstnanosti a míru výskytu dlouhodobé nezaměstnanosti. Nelze však u všech zemí nalézt vždy stejný vývoj.

Key words: unemployment, long-term unemployment, labour market, economic recession

Klíčové slová: nezaměstnanost, dlouhodobá nezaměstnanost, trh práce, ekonomická recese

JEL classification: J01, J08, J64

1. Úvod

Ekonomiky členských států Evropské unie byly v roce 2009 zasaženy ekonomickou recesí, jejíž kořeny lze nalézt ve Spojených státech. Jedinou zemí, která se vyhnula samotné recesi, bylo Polsko. Následné problémy s veřejnými financemi se u některých zemích projeví v dalším zhoršení ekonomické situace. V některých zemích, např. v České republice, které sice neměly tak závažné problémy s veřejným dluhem, však k prodloužení recese přispěla nevhodná hospodářská politika.

Ekonomická recese se projevila samozřejmě i na trhu práce. Cílem tohoto článku je zhodnotit dopad recese na míru nezaměstnanosti a míru výskytu dlouhodobé nezaměstnanosti. Pomocí shlukové analýzy budou porovnávány shluky členských zemí Evropské unie. Sledována jsou dvě období, předkrizové období 2004 – 2008 a pokrizové období 2009 – 2013.

Data pro analýzu (např. míra nezaměstnanosti i míra výskytu dlouhodobé nezaměstnanosti) pocházejí z databáze Eurostatu, a to vždy pro věkovou skupinu 15 – 74 let. Dlouhodobá nezaměstnanost je definována jako nezaměstnanost, která trvá déle než jeden rok. V tomto článku bude využívána mj. míra výskytu dlouhodobé nezaměstnanosti, která vyjadřuje procento dlouhodobě nezaměstnaných na celkovém počtu nezaměstnaných. Jde tedy o jiný

ukazatel než je míra dlouhodobé nezaměstnanosti, který udává procento dlouhodobě nezaměstnaných na pracovní síle.¹

Evropská unie se vyznačuje v průměru dlouhodobě vyšší mírou nezaměstnanosti a i vyšší dlouhodobou nezaměstnaností. Jednou z všeobecně přijímaných příčin je tzv. hystereze na trhu práce (viz. např. BLANCHARD, 2006 a BLANCHARD a SUMMERS, 1986, PAVELKA, 2012). Také v České republice představuje dlouhodobá nezaměstnanost značný sociální i národohospodářský problém (viz. PAVELKA, LÖSTER, MAKOVSKÝ, LANGHAMROVÁ, 2011).

2. Shluková analýza

Pro vytvoření shluků objektů byla použita hierarchická shluková analýza. Kritérium pro spojování shluků vychází z myšlenky, aby v každém kroku shlukování došlo k minimálnímu přírůstku Wardova kritéria, jehož výpočet je uveden v (GAN, MA, WU, 2007). Používá se pouze ve spojení se čtvercem Euklidovy vzdálenosti, viz (GAN, MA, WU, 2007).

Pro stanovení optimální počtu shluků bylo využito několik postupů. Jednak na základě znalostí ekonomické teorie, jednak na základě vhodných kritérií a jednak na základě dendrogramu. Na základě dendrogramu a na základě hodnotících koeficientů a na základě praktických zkušeností byly jako optimální stanoveny 4 shluky. Jejich počet byl stanoven na základě CHF koeficientu, RMSSTD koeficientu a Davies-Bouldinova koeficientu. K výpočtu koeficientů byl využit systém SYSTAT. Podrobný výpočet těchto hodnotících koeficientů je uveden např. v (GAN, MA, WU, 2007). Ke shlukování a analýze zemí byl využit systém IBM SPSS, verze 20.

Shluková analýza může využívat řadu ukazatelů. V tomto článku budou následně hodnoceny dva vybrané ukazatele, o kterých již bylo zmíněno výše. Prvním je celková míra nezaměstnanosti a druhá je míra výskytu dlouhodobé nezaměstnanosti. Výsledné shluky zemí jsou vytvořeny na základě různých ukazatelů, a to v období před vypuknutím recese, tedy v období 2004 – 2008, a v období po vypuknutí recese, tedy období 2009 – 2013. Rozdělení zemí do jednotlivých shluků zachycuje tabulka č. 1. Informace o shlucích lze nalézt jednak v obrázcích č. 1 a 2 a jednak v tabulkách č. 2 a 3.

Tab. 1: Rozdělení zemí do jednotlivých shluků

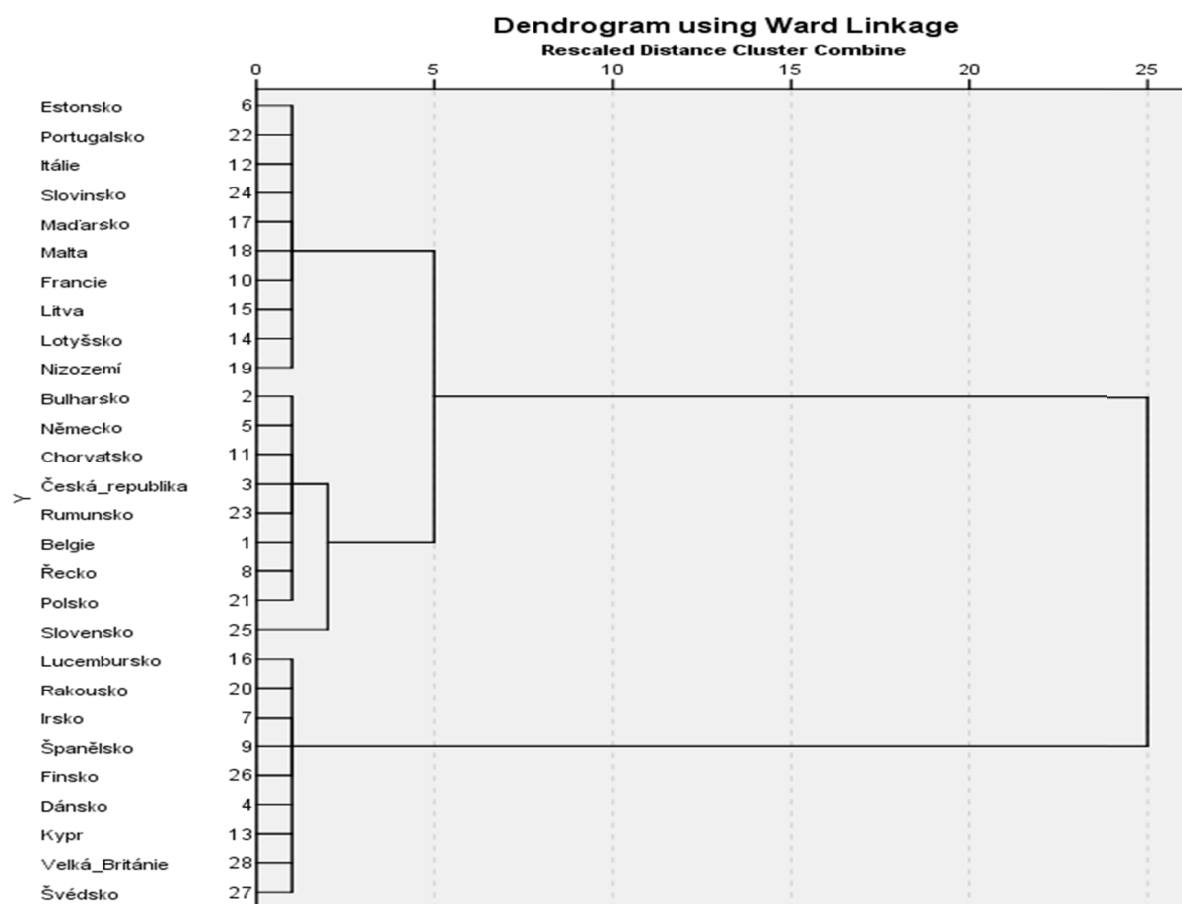
1		2		3		4	
2004-08	2009-13	2004-08	2009-13	2004-08	2009-13	2004-08	2009-13
DK	DK	EE	BE	BE	BG	SK	HR
IE	CY	FR	CZ	BG	EE		SK
ES	LU	IT	DE	CZ	IE		
CY	NL	LV	FR	DE	GR		
LU	AT	LT	MT	HR	ES		
AT	PL	HU	RO	PL	IT		
FI	FI	MT	SI	RO	LV		
SE	SE	NL		GR	LT		
GB	GB	PT			HU		
		SI			PT		

Země Evropské unie byly nakonec rozděleny do čtyř shluků:

Shluk č. 1

¹ Podrobnější vysvětlení lze nalézt např. na stránkách Českého statistického úřadu (ČSÚ, 2014).

Tento shluk zahrnuje země, které se vyznačují nízkou mírou nezaměstnanosti a také nízkou mírou výskytu dlouhodobé nezaměstnanosti. Z tabulky č. 1 je zřejmé, že ekonomická recese tuto skupinu částečně změnila. Do tohoto shluku patřily před recesí i po ní všechny tři severské země, Velká Británie, Lucembursko, Kypr a Rakousko. V předkrizovém období sice do tohoto shluku patřily i Španělsko a Irsko, které však byly výrazně zasaženy ekonomickou (či dluhovou) krizí. Ekonomika Kypru sice byla také postižena výrazně dluhovou krizí a míra nezaměstnanosti vzrostla, míra výskytu dlouhodobé nezaměstnanosti se však podařilo udržet na nízké úrovni. V pokrizovém období se do prvního shluku zařadilo Nizozemí a Polsko. Polská ekonomika byla jedinou ekonomikou EU, ve které nešlo k propadu do recese. Její míra nezaměstnanosti v pokrizovém období vykazuje průměrně nižší hodnotu než v předkrizovém období. V případě Nizozemí, míra nezaměstnanosti vzrostla v pokrizovém období o jeden procentní bod (na 4,9 % - což je velmi nízká míra nezaměstnanosti), ale zároveň se mu podařilo snížit míru výskytu dlouhodobé nezaměstnanosti o téměř 7 procentních bodů na 31,2 %.



Obr. 15: Dendrogram v situaci před recesí

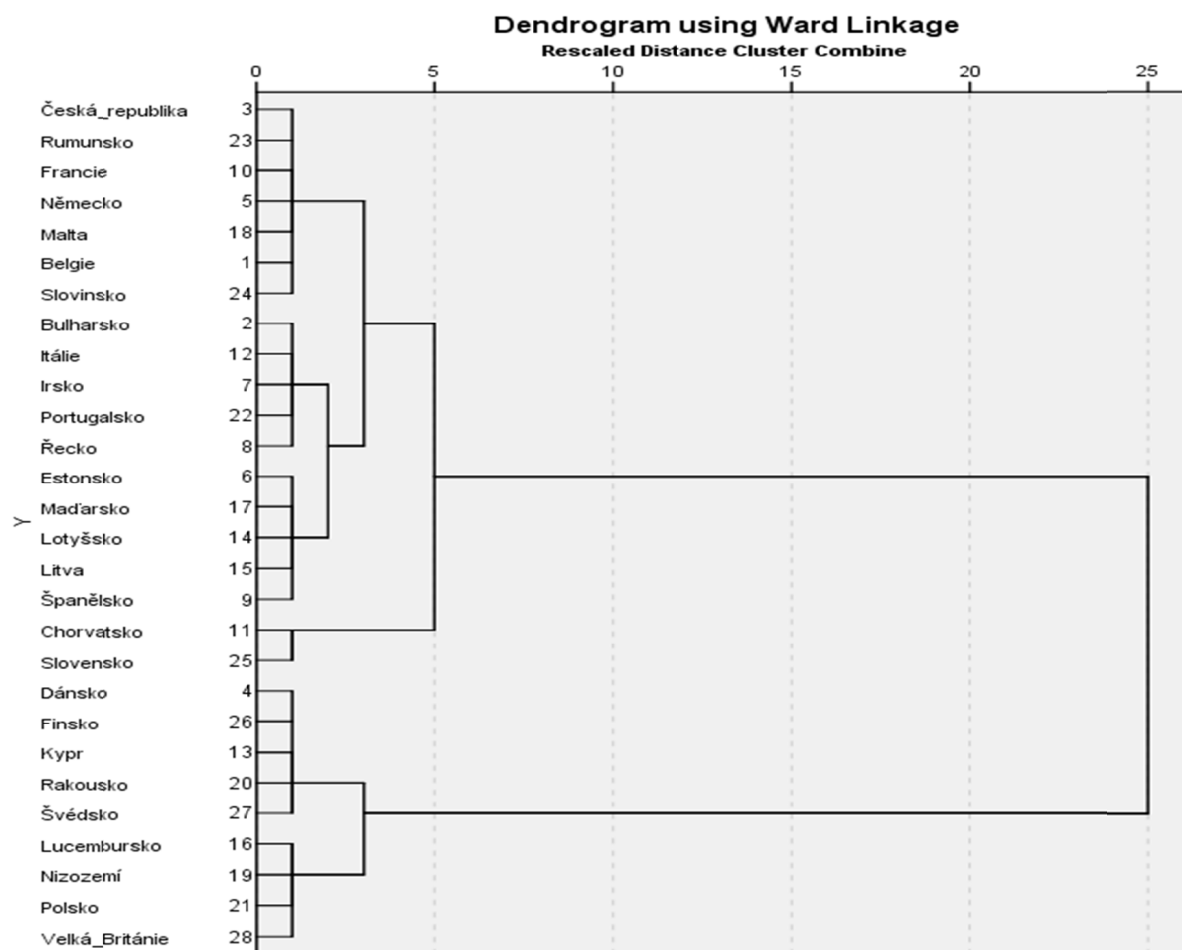
Shluk č. 2

Tento shluk zahrnuje země, které se vyznačují nižší až střední nezaměstnaností a spíše střední mírou výskytu dlouhodobé nezaměstnanosti. V předkrizovém období byl tento shluk tvořen 10 zeměmi, kdežto v pokrizovém období pouze 7 zeměmi. V předkrizovém období do tohoto shluku patřily pobaltské státy, které jsou velmi otevřenými ekonomikami, a dopad hospodářské recese na ně byl značný, když se průměrná míra nezaměstnanosti v pokrizovém období ve srovnání se situací před krizí v podstatě zdvojnásobila. Proto jsou pobaltské země

v pokrizovém období zahrnuty až do následujícího shluku. Situace na trhu práce se výrazně zhoršila také v Itálii, Portugalsku a Maďarsku, a to nejen v podobě růstu celkové míry nezaměstnanosti, ale i růstu míry výskytu dlouhodobé nezaměstnanosti. Naopak v Belgii, České republice se míra nezaměstnanosti sice v pokrizovém období mírně zvýšila, ale míra výskytu dlouhodobé nezaměstnanosti klesla. Německu klesla v pokrizovém období průměrná míra nezaměstnanosti i míra výskytu dlouhodobé nezaměstnanosti. V obou obdobích patřily do tohoto shluku pouze tři země: Francie, Malta a Slovinsko.

Tab. 2: Charakteristika shluků v situaci před recesí

Ward Method	N	Minimum	Maximum	Průměr
1 nezaměstnanost	9	4,20	9,70	5,8222
míra_dlouhodobé_nezam	9	14,70	31,20	23,1444
2 nezaměstnanost	10	3,90	8,50	6,8100
míra_dlouhodobé_nezam	10	35,60	48,50	43,5200
3 nezaměstnanost	8	6,60	13,50	9,1625
míra_dlouhodobé_nezam	8	50,10	59,30	53,3750
4 nezaměstnanost	1	13,80	13,80	13,8000
míra_dlouhodobé_nezam	1	71,20	71,20	71,2000



Obr. 2: Dendrogram v situaci po recesi

Shluk č. 3

Tento shluk zahrnuje země, které se vyznačují vyšší až vysokou mírou nezaměstnanosti a vyšší mírou výskytu dlouhodobé nezaměstnanosti. I v tomto shluku způsobila hospodářská recese výrazné změny ve složení. V předkrizovém období byl tento shluk tvořen 9 zeměmi a v pokrizovém období 10 zeměmi. O některých změnách již bylo napsáno výše. V Polsku došlo k výraznému zlepšení situace na trhu práce, a proto se z tohoto shluku v předkrizovém období přesunulo do shluku s nízkou mírou nezaměstnanosti a nízkou mírou výskytu dlouhodobé nezaměstnanosti, tedy do prvního shluku. Zlepšení ve výskytu dlouhodobé míry nezaměstnanosti znamenal přesun Belgie, České republiky, Rumunska a Německa z tohoto shluku v předkrizovém období do shluku 2 v pokrizovém období. Výrazné zhoršení míry nezaměstnanosti znamenalo přesun pobaltských zemí a Španělska do tohoto shluku v pokrizovém období. Výrazné zhoršení míry výskytu dlouhodobé nezaměstnanosti znamenalo přesun Irska do toho shluku v pokrizovém období. Zhoršení (i když ne výrazné) u obou ukazatelů způsobil přesun Itálie, Portugalska a Maďarska do tohoto shluku v pokrizovém období.

Tab. 3: Charakteristika shluků v situaci po recesi

Ward Method	N	Minimum	Maximum	Průměr
1 nezaměstnanost	9	4,50	9,60	7,2000
míra_dlouhodobé_nezam	9	17,80	36,30	26,3222
2 nezaměstnanost	7	6,30	9,50	7,5000
míra_dlouhodobé_nezam	7	39,50	46,40	42,9714
3 nezaměstnanost	10	9,50	22,00	14,0600
míra_dlouhodobé_nezam	10	39,20	52,10	47,5200
4 nezaměstnanost	2	13,50	13,60	13,5500
míra_dlouhodobé_nezam	2	61,00	64,70	62,8500

Shluk č. 4

Tento shluk je tvořen zeměmi, které se vyznačují vysokou mírou nezaměstnanosti a také vysokou mírou výskytu dlouhodobé nezaměstnanosti. Z tabulky č. 1 je zřejmé, že Slovensko patřilo do tohoto shluku jak v předkrizovém tak pokrizovém období. Průměrná míra nezaměstnanosti na Slovensku se vlivem krize sice nepatrně snížila (o 0,2 procentního bodu na 13,6 %), a také míra výskytu dlouhodobé nezaměstnanosti se snížila (o 6,5 procentního bodu na 64,7 %), přesto však zůstala míra výskytu dlouhodobě nezaměstnanosti tak vysoká, že Slovensko zůstalo ve shluku č. 4. V pokrizovém období přibýlo do tohoto shluku také Chorvatsko, které si pohoršilo u obou ukazatelů.

3. Závěr

Ekonomická recese vedla k značným změnám na trzích práce členských zemích Evropské unie. Předkrizové období bylo charakterizováno relativně vysokým tempem růstu reálného produktu (zejména u některých zemích střední a východní Evropy), které bylo spojeno s poklesem míry nezaměstnanosti často na velmi nízké hodnoty, které byly pod úrovní přirozené míry nezaměstnanosti. Ekonomická recese z roku 2009, která byla v některých členských zemích Evropské unie protažena dluhovou krizí či hospodářskou politikou, vedla ve většině zemí ke zhoršení situace na trhu práce.

Shluková analýza byla zaměřena na rozdělení členských zemí Evropské unie na jednotlivé skupiny podle průměrné míry nezaměstnanosti a průměrné míry výskytu dlouhodobé nezaměstnanosti v předkrizovém a pokrizovém období. I když lze potvrdit obecný fakt, že ekonomická recese výrazně zhoršila situaci na trzích práce, mezi členskými zeměmi však

existují rozdíly. Je patrné, že země, které byly postiženy hospodářskou recesí výrazně, vykazaly následně vyšší míry nezaměstnanosti či vyšší míry výskytu dlouhodobé nezaměstnanosti. To se týká např. pobaltských států, Španělska, Portugalska, Itálie a Irska. U některých států je sice patrný nárůst průměrné míry nezaměstnanosti v pokrizovém období, ale míra výskytu dlouhodobé nezaměstnanosti se snížila¹. To platí např. u Polska, České republiky, Německa a Rumunska. Nejvyšší míru nezaměstnanosti a i nejvyšší míru výskytu dlouhodobé nezaměstnanosti vykazovala předkrizovým i pokrizovým období Slovensko.

Z provedené analýzy je zřejmé, že samotná ekonomický recese není jediným faktorem, který ovlivňuje, v jaké míře se nezaměstnaní přesouvají do skupiny dlouhodobě nezaměstnaných.

Literatúra

BLACHRD, O. – SUMMERS, L. 1986. *Hysteresis and Unemployment Problem*. Cambridge: NBER Working Paper, č. 1950.

BLANCHARD, O. 2006. European unemployment: the evolution of facts and ideas. In: *Economic Policy*, č. 21. str. 5 – 59.

ČSÚ. 2014. Online. Zaměstnanost a nezaměstnanost podle výsledků VŠPS – Metodika. Dostupné na: http://www.czso.cz/csu/redakce.nsf/i/zam_vsps

GAN, G., MA CH., WU J. 2007. *Data Clustering Theory, Algorithms, and Applications*, ASA, Philadelphia.

PAVELKA, T. – LÖSTER, T. – MAKOVSKÝ, P. – LANGHAMROVÁ, J. 2011. *Dlouhodobá nezaměstnanost v České republice*. Slaný: Melandrium.

PAVELKA, T. 2012. Long-term unemployment in the European Union during last five turbulent years. In: *Intellectual Economics*, roč. 6, č. 3, s. 293-305.

Adresa autorů:

Tomáš Pavelka, doc. Ing., Ph.D.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67 Praha 3
pavelkat@vse.cz

Tomáš Löster, Ing., Ph.D.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67 Praha 3
losterto@vse.cz

¹ Je třeba upozornit, že se jedná o průměrné hodnoty za 5 let, tudíž je eliminován fakt, že vysoký počet nově nezaměstnaných za jinak stejných podmínek snižuje míru výskytu dlouhodobé nezaměstnanosti.

Paradox imputace Imputation Paradox

David Pejčoch

Abstract: The goal of this article is to describe a situation when imputation of missing values could potentially lead to better characteristics of model built based on imputed data than in situation when there are no missing values from the beginning. This situation is usually caused by the fact that imputation modified original distribution of imputed variables. I called this situation as an Imputation Paradox. This paradox has been explored during experimental evaluation of imputation methods using different levels of missing data occurrence. Described methodology of this benchmark could be considered also as an additional value of this article.

Abstrakt: Cílem tohoto článku je popsat situaci, kdy imputace chybějících pozorování může potenciálně vést k lepším charakteristikám modelu vytvořeného na základě imputovaných dat, než za situace, kdy byl model vytvořen s využitím úplných pozorování. Tento jev je zpravidla způsoben modifikací původního rozdělení proměnných vlivem imputace. Tuto situaci jsem nazval Paradoxem imputace. Paradox byl pozorován v rámci realizovaného experimentu, majícího za cíl porovnat výsledky jednotlivých metod při různé míře výskytu chybějících pozorování. Popsaná metodika experimentu může být chápána jako další z přínosů tohoto článku.

Key words: Missing values, Imputation, Data preparation.

Klíčové slová: Chybějící pozorování, imputace, příprava dat.

JEL classification: C18, C80, C53

1. Introduction

The problems of missing values are one of the most important areas within Data Quality Management. We can hardly imagine data preparation for some analysis without including the step dealing with incomplete data. With increasing trend of collecting all possible data sources and creation of so called Big Data environment we can hardly consider that some subject would simply accept presence of missing values without any attempts to replace them by real observations.

(Dasu & Johnson, 2003) mentioned two different kinds of missing data: (1) missing records with the meaning of not filled values of selected observations and (2) the topic of not completed records with the reference to so called censored data. This article is focused on first of them. In available literature we can find a lot of different methods and their classifications how to deal with this situation. Classification published in (Pejčoch, 2014) represents compilation of these approaches. As a simplification of provided classification we can consider four basic approaches: (1) to keep status quo, (2) look for these values in some kind of reference source, (3) impute missing values without using any model (e.g. using some characteristics of level, randomly generated numbers, etc.) and (4) impute these values using explicit or implicit model. Paradox described in this work could be result of the last mentioned approach.

Application of strategies mentioned above usually depends on so called mechanisms of missingness described first in (Rubin, 1976). This article is focused on situation when MAR (Missing at Random) mechanism is expected and missing values could be potentially imputed using predictive model built on complete records of remaining variables. The main goal of this article is to describe the situation when imputation will deform original distribution of imputed variable. It provides critical view on blind using of predictive methods for imputation

even in situation when there is only weak dependency between imputed variable and used predictors.

2. Description of Experiment Methodology

Original purpose of described experiment was to evaluate different imputation methods using increasing amount of simulated missing values. This experiment compared 32 different methods. Complete list of these methods and detailed description of results is available online at: <http://www.dataquality.cz/vyzkum/Imputace/ImputaceExperimentPrehled.xlsx>.

Experiment contained from below mentioned sequence of steps:

1. Loading data to experimental framework and data preparation (encoding non-ASCII letters, transformation of categories to dummy variables, discretization of continuous variables, standardization and cleansing);
2. Creation of reference data model on complete data using SAS GENMOD procedure;
3. Simulation of missing values across attributes used as a part of reference model (from 5 % to 50 % by 5 % as a step);
4. Application of selected methods for imputation of missing values using additional attributes from original data set;
5. Integration of imputed data sets;
6. Rounding continuous values of imputed binary variables to {0, 1};
7. Creation of models on imputed data sets using the same variables as in reference model;
8. Integration of selected characteristics of model quality into single data set, their evaluation and visualization.

The reason for using SAS GENMOD procedure to create reference model was its relative flexibility to cover different kinds of so called General Linear Model theoretically described e.g. in (Nelder & Wedderburn, 1972). Using simple change in its parameters it is possible to use the same procedure for linear regression, logistic regression, etc.

As evaluating criteria of model quality I used standard characteristics provided on the output of this procedure:

- § Likelihood L of model (preferred model maximizes L);
- § Akaike Information Criterion (AIC) derived from Likelihood using the formula $AIC = 2k - 2\ln(L)$, where k represents the number of parameters in model. For practical usage it seems to be more suitable to transform this formula to the form of relative Likelihood introduced e.g. in (Burnham & Anderson, 2002) and defined as $\exp((AIC_{min} - AIC_j)/2)$. Preferred model minimizes AIC.
- § Bayesian Information Criterion (BIC), which is for enough number of observation possible to approximate using the formula $-2\ln(L) + k \cdot \ln(n)$, where k represents the number of parameters in model and n represents the number of observations. Preferred model minimizes BIC.

3. Description of Data

As an input I used twelve different data sets from different domains. Their list is published in Table 1 below. Detailed description of data is available in (Pejcoch, 2014). In phase of data preparation I made several transformations described below:

- § From ds6, ds7, ds8 and ds9 data sets I randomly selected 10 % of original observations using simple random sampling. The reason was enormous runtime of some methods when applied on these data (several hours for single variant of data set);

- § Continuous variables have been discretized to intervals of similar length. For derivation of length so called Sturges rule has been used;
- § Names of categories have been encoded to numerical values because of requirements of method used for simulation of missing values. Apart from this transformation these variables remained to be considered as nominal in metadata. For predictive methods which work only with numerical variables they have been later transformed to dummy {0,1} variables;
- § Based on initial exploratory analysis, selected continuous variables have been transformed using logarithm function. The purpose was to come as much closer to ideal normal distribution of these variables as it is possible. In case of zero values small constant 0.000001 was added before the transformation. This kind of modification has been used in case of data sets ds1 (attributes a2, a3, a7, a10, a13, a14), ds2 (attributes a2, a5, a13), ds3 (attribute a10), ds4 (attributes a6, a15, a16), ds6 (attribute a8), ds7 (attributes a1, a3, a11), ds8 (attributes a1, a12, a15), ds10 (attributes a5, a6, a11), ds11 (attributes a4, a9) a ds12 (attributes a5, a6, a7, a8, a9, a10, a11, a12, a13, a14, a15, a16, a19).

After data preparation phase data sets contained from original categories and continuous variables and also from derived discretized and binarized variables.

Tab. 10: Data Sets Used in Experiment

ID	Description	Number of Obs.	Number of Var.	Number of Bins***	Type of Target Variable**
ds1	Australian Credit Approval (STATLOG)	690	15	10	C
ds2	German Credit (STATLOG)	1000	20	11	N
ds3	Heart Disease (STATLOG)	270	13	9	C
ds4	Vehicle Silhouettes (STATLOG)	846	18	11	C
ds5	Landsat Satellite (STATLOG)	4435	36	13	C
ds6	Shuttle (STATLOG)	4350*	9	13	C
ds7	Adult	3250*	15	13	N
ds8	Bank Marketing	4521*	17	13	C
ds9	Nursery	1296*	9	11	C
ds10	Wine Quality	4898	12	13	N
ds11	Auto MPG Data Set	398	9	10	N
ds12	Parkinsons Telemonitoring Data Set	5875	22	10	N

* Original number of observations has been reduced based on simple random sampling. ** N = numerical continuous variable, C = category. *** Number of bins derived according to Sturges rule.

4. Process of Experiment

During experiment 32 different methods or their combinations have been evaluated. Most of these methods have been theoretically described e.g. in (Pejcoch, 2014). For most of them the Rapid Miner tool (version 5.3.013) has been used as an application environment. Especially because this tool enables running automatically generated data flows based on XML (eXtensible Markup Language). The list of all method is:

- § M1: Simple mean
- § M2: Conditioned mean (Buck method)
- § M3: Midrange
- § M4: Combination of simple mean and mode
- § M5: Combination of median and mode
- § M6: Hot-deck imputation
- § M7: Nearest neighbour
- § M10: Linear regression
- § M11: Naive Bayes classifier
- § M12: Bayesian Network
- § M13: Neural Network (MLP)
- § M14: Neural Network (RBF)
- § M15: Multinomial Logistic Regression
- § M16: Support Vector Machine
- § M17: Support Vector Machine with Particle Swarm Optimization
- § M18: Linear Discriminant Analysis
- § M19: Quadratic Discriminant Analysis
- § M20: Expectation Maximization with MCMC
- § M21: Parametric Regression
- § M22: Method of Propensity Score
- § M25: ID3
- § M26: CHAID
- § M27: Random Forrest
- § M29: Decision Stump
- § M30: Conjunctive Rules
- § M31: M5 Rules
- § M32: Jrip
- § M33: Listwise / Pairwise
- § M34: Decision Tree (universal procedure implemented in Rapid Miner)
- § M36: Median
- § M37: Mode

If we multiplied number of methods by number of all used data sets and modifications corresponding to different levels of missing values we would get 3,720 as a total number of imputed data sets. This kind of experiment wouldn't be possible to realize without some kind of automation.

Time of imputation of single data set differed from 1 s to 12,291 s (approx. 3.5 hours). For some method it wasn't even possible to finish imputation in reasonable time. The worst

problems from time perspective have been observed for M14: RBF, M15: Multinomial Logistic Regression and both Support Vector Machines (M16, M17). Total time spent on successful processing was 40 hours of CPU time. However, this information is based on logs. Real time of processing was approximately double of reported time.

5. Results of Experiment

It is possible to observe general trend of decreasing time of imputation with increasing number of imputed values when using methods based on a model. This phenomenon could be explained by decreasing number of observations which are used as a source for model. Detailed view on all times of processing related to different methods can be found on <http://www.dataquality.cz/vyzkum/Imputace/ImputaceExperimentPrehled.xlsx>.

From the perspective of selected characteristics of model quality it would be useful to look for methods whose implementation led to only small increase of AIC, BIC and slight decrease in L with increasing number of imputed values. At the same time it would be reasonable to filter methods to those with approximately the same level of selected criteria on data sets with imputed 5 % of missing values and complete data.

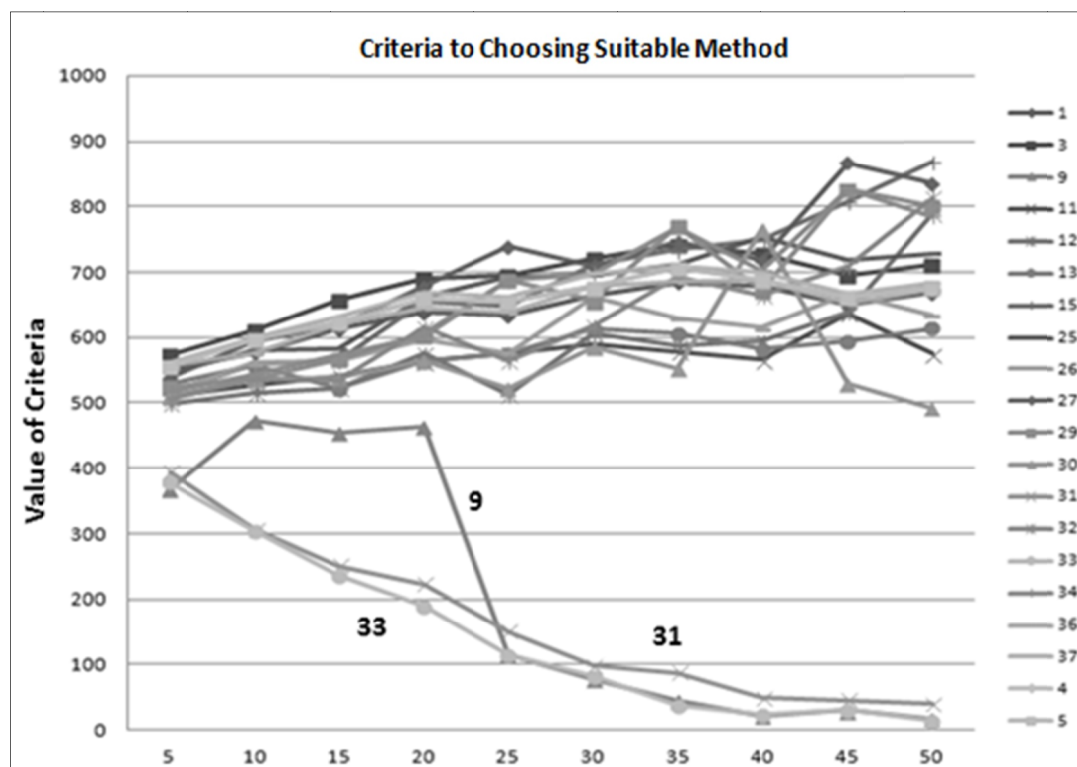


Fig. 16: Decreasing trend of AIC for models built on data with increasing rate of imputed missing values

It is necessary to be careful in case of methods which lead to radical decrease of AIC or BIC. This decrease could be caused by three reasons: (1) reduction of models parameters as a result of remaining missing values, (2) weak dependency of imputed variable and used predictors which deforms original distribution and sometimes make it even more suitable for the model and (3) another kind of model weakness usually caused by too many rows with missing values. First reason was typical for application of M33: Listwise method (exclusion of records with missing values). Second reason occurred usually in implementation of M31: M5Rules.

Fig. 1 shows examples of AIC measurements related to ds1. Results of M9: Generalized Nearest Neighbour (at the end excluded from original list of methods) show how AIC looks

like for methods for which only few of imputed data sets were available (only from 5 % to 20 % of originally missing values): first there is an increase in AIC (as expected) and then we can observe decrease to the level of M33.

6. Conclusion

Based on results of described experiment it seems to be critical to always evaluate real dependency between variables when imputing missing values using model and decide whether to use or not to use a model. This opens question how much can analyst trust to algorithms implemented in some tools for Knowledge Discovery in Databases which provides automated imputation based on model as a black box.

When choosing the best method for imputation it is important to consider also time factor of implementation. Data preparation for purpose of some methods and later implementation of some methods itself significantly extends time spent on analysis and could have critical impact to timeliness of delivering results. From time perspective the worst problems were caused by implementation of RBF, Multinomial Logistic Regression and Support Vector Machine methods. Described experiment on purpose didn't consider different "wild" combinations of methods such as described e.g. in (Marwala, 2009) which could lead even to more time spent on modelling with just small increase of precision.

Generally, it could be subject of discussion where is the threshold of missing values percentage above which it doesn't make any sense to try to impute them. From results of experiment described above it seems to be somewhere between 20 – 25 % of missing values. From this threshold most of methods started to behave abnormally.

References

- BURNHAM, K. P., ANDERSON, D. R. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.). Springer-Verlag: 2002. ISBN 0-387-95364-7.
- DASU, T., JOHNSON, T. *Exploratory Data Mining and Data Cleansing*. New Jersey: Wiley & sons, 2003.
- MARWALA, Tshilidzi. *Computational Intelligence for Missing Data Imputation, Estimation, and Management: Knowledge Optimization Techniques*. Information Science Reference, 2009. ISBN: 1605663360.
- NELDER, J.A., WEDDERBURN, R.W.M., Generalized Linear Models. In: Journal of the Royal Statistical Society. Series A (General). Vol. 135, No. 3 (1972), pp. 370-384, Wiley.
- PEJČOCH, D. Benchmark metod pro doplňování chybějících pozorování. *dataqualitycz* [online]. 2014-10-18. [cit. 2014-11-15]. Dostupné pod: <http://www.dataquality.cz/index.php?ID=4&SUBID=2>
- RUBIN, D.B (1976). Inference and Missing Data (with Discussion). In: *Biometrika* 63, pp.581-592.

Adresa autora:

David Pejčoch, Ing.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 67 Praha 3
david@pejcoch.com

Shluky nefinančních podniků evropských zemí dle metodiky národních účtů

Clusters of the non-financial organizations of European union countries by the national account methodology

Tomáš Pivoňka, Tomáš Löster

Abstract: This paper apply the method of cluster analysis on data from national accounts of particular countries in European Union. Balance items together with other items competing the picture of situation on the market of non-financial organizations are used as a input of the analysis. Four clusters were used. The first cluster contain the biggest number of countries especially in 2012, when the cluster contain 17 countries. This cluster was the most affected by current economic crisis, when the values of indicators falled in 2012 substantially. On the other hand, there is cluster with much better values of indicators, which contain Baltic countries.

Abstrakt: Článek využívá metody shlukové analýzy, která je aplikovaná na data z finančních účtů jednotlivých zemí Evropské unie. Jako vstupní data jsou použity bilanční položky několika účtů spolu dalšími položkami dokreslujícími situaci nefinančních podniků v zemích Evropské Unie. Byly vytvořeny shluky, kdy první shluk obsahuje nejvíce zemí a najdeme zde zástupce ze západní Evropy. Tento shluk je nejvíce postižen ekonomickou krizí propuknuvší v roce 2008, kdy střední hodnoty ukazatelů zaznamenaly výrazný pokles. Do tohoto shluku se také v roce 2012 přesunuly další státy. V roce 2012 čítal první shluk 17 států. Na druhou stranu velice dobře dopadly pobaltské země, které se nacházely ve shluku s výrazně lepšími průměrnými hodnotami.

Key words: National accounts, balance items, cluster analysis, economic crisis, non-financial organizations.

Klíčová slova: Národní účty, bilanční položky, shluková analýza, ekonomická krize, sektor nefinančních podniků

JEL classification: C38, E22

1. Úvod

V tomto článku je použita metoda shlukové analýzy, která je aplikovaná na data z národních účtů nefinančních podniků jednotlivých zemí Evropské unie. Jako vstupní data analýzy posloužily spolu s čistou přidanou hodnotou bilanční položky účtů tvorby důchodů, druhotného rozdělení důchodu a dále pak některé další položky dokreslující situaci na trhu nefinančních podniků v jednotlivých zemích. Jelikož se jedná o data, která vycházejí z celkového výstupu sektoru, má celková ekonomická situace země významný vliv na hodnoty těchto ukazatelů. Shlukování bylo prováděno v letech 2004, 2008 a 2012. Cílem článku je potom zhodnotit složení jednotlivých shluků a také popsat vývoj vstupních dat mezi těmito třemi obdobími. Rok 2004 reprezentuje období před ekonomickou krizí, které je spojeno s velkým rozšířením Evropské Unie. V roce 2008 se začala projevovat ekonomická krize, tudíž tento rok reprezentuje období těsně před jejím propuknutím. V roce 2012 pak můžeme hodnotit dopad krize na sektor nefinančních podniků, což je dalším cílem této práce.

Problematikou nefinančních podniků z pohledu národního účetnictví se zabýval například V. Pošta a M. Nečadová, kde autoři zkoumali ziskovost podniků v agregátním měřítku v České republice.

Tento článek je jedním ze série příspěvků, které se snaží popsat důsledky celosvětové finanční krize na země Evropské unie pomocí metody shlukové analýzy s využitím různých vstupních dat. V první části se krátce věnujeme samotné shlukové analýze. V následující části pak obsahuje stručný popis vstupních proměnných analýzy, jejíž výsledky jsou zařazeny za touto částí. Poslední kapitolou článku je závěr obsahující shrnutí výsledků.

2. Popis použité metody

Pro vytvoření skupin zemí EU byla použita oblíbená vícerozměrná statistická metoda - shluková analýza. Tato metoda je používána ke tvorbě skupin objektů v mnoha disciplínách a je využívána v mnoha odborných statích, viz například [LÖSTER], [ŘEZANKOVÁ, LÖSTER], [MEGYESIOVÁ], atd. Použili jsme hierarchické shlukování pomocí Wardovy metody, která řeší proces tvorby shluků, na rozdíl od jiných metod shlukování, pomocí minimalizace heterogenity shluků. Kritérium pro spojování shluků vychází z myšlenky, aby v každém kroku shlukování došlo k minimálnímu přírůstku tzv. Wardova kritéria, jehož výpočet je uveden například v [LÖSTER]. Wardova metoda má tendenci odstraňovat malé shluky a tvořit shluky přibližně stejné velikosti. Používá se pouze ve spojení se čtvercem Euklidovy vzdálenosti, viz [GAN].

Pro stanovení optimální počtu shluků jsme využili jednak dendrogram a jednak koeficienty pro stanovení optimálního počtu shluků (CHF, RMSSTD, DB, atd.), jejichž podrobný výpočet je uveden například v [LÖSTER]. Tyto koeficienty jsou aplikovány v systému SYSTAT, verze 12, který byl využit k jejich stanovení. Na základě dendrogramu a uvedených koeficientů jsme vybrali 4 shluky. K analýze shluků byl využit systém IBM SPSS, verze 20.

3. Popis proměnných

Proměnné pocházejí z národních účtů jednotlivých zemí, kdy u prvních čtyř se jedná o bilanční položky odpovídajících účtů, další ukazatele dokreslují situaci na trzích nefinančních podniků. Národní účty jsou rozděleny na dvě strany, zdroje a užití. Strana zdrojů obsahuje (jak je z názvu patrné) prostředky, ze kterých se čerpá. Na straně užití potom najdeme způsob, jak a na co jsou finanční prostředky vynaloženy. Z konstrukce jednotlivých ukazatelů je patrné, že jsou silně ovlivněny celkovým produktem ekonomiky. Veškeré ukazatele jsou vztaženy k HDP, aby je bylo možno porovnávat mezi jednotlivými zeměmi. Vzhledem k tomu, že je tu jasná vazba na celkovou produkci, bude možné shluky v jednotlivých letech porovnat v závislosti na projevu ekonomické krize, která propukla v roce 2008, viz např. [SPĚVÁČEK].

- Čistá přidaná hodnota (Net value added): vznik odečtením meziprodukce od produkce při zohlednění spotřeby fixního kapitálu
- Čistý provozní přebytek (Net entrepreneur income): nachází se na účtu tvorby důchodů, kde se vytvořený důchod rozděluje mezi výrobní faktory a vládní instituce (daně z výroby a dovozu a dotace). Účet tvorby důchodů zaznamenává, jakým způsobem je podnik schopen pokrýt náhrady zaměstnanců a daně z přidané hodnoty.
- Čistý disponibilní důchod (Net disposable income) : je vyrovnávací položkou účtu druhotného rozdělení důchodu. Z národního důchodu, tedy důchodu všech rezidentů bez ohledu na zemi, kde k důchodu přišel, dochází k přerozdělení mezi sektory prostřednictvím běžných daní, sociálních dávek a příspěvků. To co podniku zbude po tomto přerozdělení je čistý disponibilní důchod. V sektoru nefinančních podniků můžeme tuto položku ztotožnit s čistými úsporami nacházejícími se na účtu užití disponibilního důchodu. Čisté úspory jsou rozdílem disponibilního důchodu a výdajů

na konečnou spotřebu. A jelikož podniky nemají výdaje na konečnou spotřebu, tak se čisté úspory a disponibilní důchod rovnají. Tato položka může být považována za proxy čistých zisků podniků.

- Čisté půjčky/ výpůjčky (Net lending/borrowing): je součástí účtů akumulace, které navazují na předchozí tzv. běžné účty. Čisté půjčky/výpůjčky jsou vyrovnávací položkou účtu pořízení nefinančních aktiv a vychází z čistých úspor (resp. z disponibilního důchodu) podniků. Ukazuje, zda subjekt financuje nákup těchto aktiv z vlastních zdrojů (čisté úspory) nebo si na financování nákupů půjčuje. Jinými slovy je zde prezentována informace, zda je subjekt na trhu věřitelem či dlužníkem.
- Tvorba hrubého fixního kapitálu: jedná se v podstatě o investice firem do nefinančních aktiv a je jednou z položek na straně užití na účtu akumulace zmíněným výše.
- Domácí poptávka: je celkové množství peněz vynaložených na nákup statků a služeb v ekonomice. Jedná se o velikost kupní síly subjektů na domácím trhu a pro sektor nefinančních podniků je tato položka velmi důležitá. Opět má silnou vazbu na stav celkové ekonomiky země.
- Kompenzace zaměstnancům: doprovodný údaj zohledňující nejen mzdu, ale také veškeré odvody na zaměstnance.
- Dotace podnikům: jakým způsobem stát zasahuje do fungování trhu formou peněžních prostředků pro podniky
- Důchody z vlastnictví: zde se jedná o čisté důchody, tedy po zohlednění jak strany aktiv, tak pasiv.
- Běžné daně: daně z příjmu, majetku

4. Aplikace metody shlukové analýzy

Pomocí metody shlukové analýzy jsme vytvořili 4 různé shluky evropských zemí ve třech letech, a sice v roce 2004, 2008 a 2012. V následující části jsou prezentovány dvě tabulky, kdy první tabulka obsahuje složení shluků a druhá potom nese informace o charakteristice jednotlivých shluků v jednotlivých letech.

Z první tabulky je patrné, že shluk číslo jedna je nejpočetnější. V roce 2004 a 2008 obsahuje 11 zemí a v roce 2012 pak 17 zemí. V roce 2012 tedy došlo k poměrně velkému přesunu zemí do prvního shluku. Z dalších shluků zde můžeme zmínit, že pobaltské státy se drží víceméně při sobě a tvoří téměř samostatný shluk číslo tři. K tomu, abychom mohli posoudit shluky z hlediska situace v sektoru nefinančních podniků, přejdeme k tabulce číslo dvě, která obsahuje střední hodnoty jednotlivých ukazatelů v rámci shluků. Komentář k tabulce se nachází pod tabulkou.

Tab. 1 Příslušnost zemí k jednotlivým shlukům.

1			2			3			4		
2004	2008	2012	2004	2008	2012	2004	2008	2012	2004	2008	2012
BEL	BEL	BEL	BUL	BUL	BUL	EST	CZE	EST	GRE	IRL	IRL
CZE	DEN	CZE	SPA	SPA	DEN	IRL	EST	LAT	IT	GRE	LIT
DEN	GER	GER	CRO	CRO	GRE	LAT	LAT	HOL	CYP	IT	
GER	FRA	SPA	HUN	POR	POL	LIT	LIT		POL	CYP	
FRA	HUN	FRA	POR	SLO	SK		ROM			POL	
MAL	MAL	CRO	ROM							SK	
HOL	HOL	IT	SLO								
AUS	AUS	CYP	SK								
FIN	FIN	HUN									
SWE	SWE	MAL									
UK	UK	AUS									
		POR									
		ROM									
		SLO									
		FIN									
		SWE									
		UK									

Tab. 2 Charakteristiky shluků

	2004				2008				2012			
	1	2	3	4	1	2	3	4	1	2	3	4
Čistá přidaná hodnota	0,45	0,38	0,50	0,35	0,46	0,39	0,50	0,37	0,42	0,35	0,52	0,54
Čistý provozní přebytek	0,16	0,10	0,22	0,14	0,17	0,07	0,18	0,15	0,13	0,14	0,20	0,31
Čistý disponibilní důchod	0,04	0,03	0,05	0,02	0,04	0,00	0,07	0,01	0,01	0,06	0,11	0,08
Čisté půjčky/výpůjčky	0,02	-0,04	-0,05	0,01	0,00	-0,11	-0,04	-0,03	0,00	0,05	0,06	0,05
Tvorba hrubého fixního kapitálu	0,11	0,15	0,16	0,09	0,12	0,18	0,19	0,10	0,11	0,11	0,13	0,08
Domácí poptávka	0,12	0,17	0,18	0,09	0,13	0,20	0,20	0,11	0,11	0,12	0,15	0,09
Kompenzace zaměstnancům	0,32	0,28	0,27	0,21	0,32	0,30	0,32	0,22	0,31	0,23	0,32	0,25
Dotace	0,01	0,01	0,00	0,00	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00
Důchody z vlastnictví	-0,06	-0,05	-0,15	-0,07	-0,06	-0,06	-0,09	-0,09	-0,06	-0,03	-0,07	-0,17
Běžné daně	0,02	0,02	0,02	0,03	0,03	0,02	0,03	0,03	0,02	0,02	0,01	0,02

Shluk číslo 1:

Jedná se o nejpočetnější shluk zemí EU, který v roce 2012 obsahuje většinu zemí. Jedná se o shluk, ve kterém je poměrně vysoká přidaná hodnota, Během třech sledovaných let došlo k poklesu hodnoty čistých úspor, především v roce 2012. Vzhledem k tomu že průměr hodnoty přidaných hodnot je více méně stabilní, podnikům tak zbude méně ve formě čistých úspor. Jsou zde zastoupeny hlavně vyspělejší země, kde je v průměru poměrně stabilní míra investic k HDP. Domácí poptávka je v porovnání s ostatními shluky.

Shluk číslo 2:

V tomto shluku se nachází státy, ve kterých je poměrně nízká přidaná hodnota. Podívali se na čisté úspory firem, pak tyto hodnoty vychází podobně jako v případě shluku 1 v roce 2004. V dalším období pak čisté úspory klesly, aby v roce 2012 poměrně výrazně narostly. Opět zde můžeme shrnout, že dopadem je různé složení shluků, kdy v roce 2012 došlo přesunu zemí do 3 shluku. Investice jsou v porovnání s ostatními shluky v jednotlivých letech na vyšší úrovni.

Shluk číslo 3:

Ve třetím shluku najdeme zejména pobaltské státy. Shluk bychom mohli charakterizovat poměrně vysokou přidanou s vyššími čistými úsporami. Větší nárůst čistých investic je patrný v roce 2012. To je přesný opak toho, co se stalo ve shluku číslo jedna. Investice ve formě hrubého fixního kapitálu jsou v tomto shluku na vyšší úrovni, ačkoliv vlivem krize došlo k poklesu poměru investic k HDP. Tyto země mají také vyšší úroveň domácí poptávky. Lze tedy shrnout, že se jedná o shluk složený ze zemí, kde se nefinanční podniky vyvíjí lépe než v ostatních shlucích.

Shluk číslo 4:

Poslední shluk obsahuje země, kde se sektoru nefinančních podniků nevedlo příliš dobře. Je charakterizován nízkou přidanou hodnotou a nízkými čistými úsporami. Stejně tak je nízká tvorba hrubého kapitálu a domácí poptávky. Zde je možné nalézt země z jihu Evropy a některé další méně vyvinuté země z východní Evropy.

Pohyby zemí napříč shluky

Shluk 1 obsahuje vyspělejší země, kde se ekonomická krize negativně podepsala na vývoji sektoru nefinančních podniků. Pozice tohoto shluku se v porovnání s ostatními zhoršila v roce 2012. Můžeme tedy popsat pohyb zemí z a do shluku 1 v roce 2012 a postihnout tak dopad krize na jednotlivé země.

- Země stále v prvním shluku: BEL, GER, FRA, HUN, MAL, RAK, FIN, SWE, UK
 - Jedná se o vyspělé západoevropské země, výjimkou Maďarska a Malty. Je ovšem nutné poznamenat, že v případě Malty se setkáváme s problémem nedostatku některých dat.
- Přesun do prvního shluku: CZ, SPA, CRO, IT, CYP, POR, RUM, SLO
 - Vlivem negativního ekonomického vývoje po roce 2008 se tyto země dostaly do prvního shluku. Najdeme zde státy z jihu Evropy ale také Českou republiku. V těchto zemích se tedy negativně projevila krize na vývoj sektoru nefinančních podniků
- Přesun z prvního shluku: DEN, HOL
 - Zde naopak došlo v porovnání s ostatními zeměmi k lepšímu vývoji v sektoru nefinančních podniků. Vývojem se tedy významně odlišovaly od zbytku zemí v prvním shluku.
- Země stále mimo první shluk: BUL, EST, IRL, GRE, LAT, LIT, POL, SK
 - Najdeme zde již zmíněné pobaltské země, a také Polsko a státy z jihu Evropy. Zatímco pobaltské státy jsou charakterizované spíše pozitivním vývojem, další

státy na tom jsou po celou dobu hůře. Opět zde však narážíme na problém s dostupností dat tentokrát u Řecka.

5. Závěr

Metoda shlukové analýzy je zde využita k popisu situace v sektoru nefinančních podniků Evropský zemí. Jako vstupní data byly využity položky z národního účetnictví. Shluková analýza rozdělila země do čtyř shluků v letech 2004, 2008 a 2012. Tyto roky reprezentují situace při rozšíření Evropské unie v roce 2004, dále pak situaci těsně před vypuknutím ekonomické krize v roce 2008 a situaci po krizi v roce 2012.

Nejpočetnějším shlukem se stal shluk 1, kde najdeme zástupce ze západní Evropy. Tento shluk se v roce 2012 rozšířil na 17 členů. Vzhledem k tomu, že došlo ke zhoršení středních hodnot sledovaných ukazatelů, může tvrdit, že se ekonomická krize významně projevila na situaci v zemích shluku. Analýza dále odhalila, že situace v pobaltských státech je výrazně lepší než ve zbytku Evropy. Tyto státy tvoří téměř samostatný shluk ve všech sledovaných letech.

Poděkování

Tento článek byl vytvořen s pomocí IGS Vysoké školy ekonomické v Praze č. 6/2013 pod názvem „Hodnocení výsledků shlukové analýzy v ekonomických problémech“

Literatura

GAN, G., MA CH., WU J. 2007. Data Clustering Theory, Algorithms, and Applications, ASA, Philadelphia.

LÖSTER, T. 2014. Metody shlukové analýzy a jejich hodnocení. 1. vyd. Slaný: Melandrium.

MEGYESIOVÁ, S. 1999. Nezamestnanosť na Slovensku a v okolitých krajinách. In Acta oeconomica Cassoviensia No 3. Podnikovohospodárksa fakulta EU so sídlom v Košiciach.

POŠTA, V., NEČADOVÁ, M. 2010. Analysis of Non-Financial Corporations Based on National Accounts, Ekonomika a Management.

ŘEZANKOVÁ, H., HÚSEK, D., a SNÁŠEL, V. 2009. Shluková analýza dat, Prague: Professional Publishing.

ŘEZANKOVÁ, H., LOSTER, T. 2013. Shlukova analyza domacnosti charakterizovanych kategoriálními ukazateli. E+M. Ekonomie a Management.

SPĚVÁČEK, V. VINTROVÁ R., ZAMRAZILOVÁ E., ŽDÁREK V., ROJÍČEK. M. 2010. Makroekonomická analýza-Magisterský kurz. Praha, VŠEM.

STANKOVIČOVÁ, I., VOJTKOVÁ, M. 2007: Viacrozmerné štatistické metódy s aplikáciami,Ekonomía, Bratislava.

Adresy autorů:

Tomáš Pivoňka, Ing.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 97 Praha 3
Pivonkat@gmail.com

Tomáš Löster , Ing, PhD.
Vysoká škola ekonomická v Praze
nám. W. Churchilla 4, 130 97 Praha 3
tomas.loster@vse.cz

Prieskum závislosti medzi finančnými možnosťami študentov generácie Y a nakupovaním na zľavových portáloch

Survey on the Dependence between the Financial Options of Generation Y Students and Shopping at Discount Portals

Ivana Polakevičová, Edita Szabová, Tatiana Kamenská

Abstract: In this contribution, we focus on the phenomenon of university students shopping at discount portals. 343 students from universities in Nitra participated in the survey. We examined whether the shopping at discount portals differs considering how many Euros per month is available to the student. We found out that most students actively use discount portals, irrespective of what financial means they may have at their disposal monthly.

Abstrakt: V príspevku sa zameriavame na fenomén nakupovania na zľavových portáloch u študentov univerzít. Prieskumu sa zúčastnilo 343 študentov nitrianskych univerzít. Zisťovali sme, či sa využívanie nákupov na zľavových portáloch líši vzhľadom na to, koľko eur mesačne má študent k dispozícii. Zistili sme, že väčšina študentov aktívne využíva zľavové portály, a to nezávisle na tom, akým finančným obnosom mesačne môže disponovať.

Key words: discount portals, student finances, independence testing.

Kľúčové slová: zľavové portály, študentské financie, testovanie nezávislosti.

JEL classification: M31

1. Úvod

Súčasná digitálna éra, vyznačujúca sa vývojom a napredovaním nových technológií, prináša množstvo inovácií založených na globálnej platforme internetu. Ten popri základnej komunikačnej funkcii spĺňa i funkciu obchodnú. Obchodovanie na internete tak otvorilo nové podnikateľské možnosti, ktoré so sebou priniesli i nové formy nakupovania. Jednou z nich sú zľavové portály poskytujúce vybrané tovary a služby vo vysokej zľave. On-line nakupovanie prostredníctvom zľavových portálov sa postupne rozšírilo do celého sveta, pričom sa v jednotlivých krajinách prispôbilo domácemu trhu a vytvorilo vlastné modifikácie, aby čo najlepšie uspokojilo potreby zákazníkov, predajcov a v neposlednom rade aj poskytovateľov zliav. Tento obchodný model neobišiel ani Slovensko a v roku 2010 sa dostal do povedomia širšej verejnosti. Vznikli desiatky zľavových portálov, z ktorých sú najúspešnejšie Zlavadna.sk a Zlavomat.sk. Vzhľadom na technologický pokrok a neustále stúpajúci počet internetových užívateľov je možné v posledných rokoch pozorovať zmenu v spotrebiteľskom správaní a spôsobe nakupovania. Pre súčasnú mladú generáciu, ktorú z marketingového hľadiska reprezentuje zákaznícky segment generácie Y, predstavujú zľavové portály pomerne obľúbenú formu nákupu. Na základe výsledkov celosvetového prieskumu realizovaného spoločnosťou Insight Express v roku 2012 by sme formu on-line nakupovanie súčasnej mladej generácie mohli nazvať globálnym trendom. Výsledky prieskumu dokazujú, že on-line nakupovanie je medzi generáciou Y veľmi obľúbené (Cisco, 2012). Reflektovali skutočnosť, že deväti z desiatich opýtaných generácie Y tvrdia, že využívajú on-line nakupovanie. Takmer traja z piatich (58%) sa pri on-line nákupoch pravidelne spoliehajú na zákaznícke hodnotenia. Ďalších 28% si pozerá on-line hodnotenia príležitostne. Takmer traja z piatich opýtaných (57%) sú ochotní poskytnúť svoju e-mailovú adresu obchodom a on-line stránkam, aby dostávali upozornenia na zľavy a výpredaje. Ale pri zdieľaní ďalších informácií sú už opatrnejší – len málokto z nich sú ochotní poskytnúť telefónne čísla, adresy domov či iné osobné údaje. Nakoľko generácia Y tvorí takmer štvrtinu svetovej populácie a jej veľkosť je

výrazná, viacero odborníkov (Lincényi, et al., 2011; Světlík, 2012; Džupina, Fandelová, et al. 2013; Szabová, Szabo, 2013) upozorňuje, že sú jednou z najdôležitejších kohort dnešného trhu. Keďže podľa J. Krotza (In: Mendoza, 2002) sa bude ich kúpna sila zvyšovať prostredníctvom jednotlivých štádií života, marketingoví pracovníci nemôžu túto skutočnosť ignorovať, ale práve naopak, mali by tomuto segmentu adekvátne porozumieť a charakterizovať ho aj napriek tomu, že výskum v tejto oblasti sa vzťahuje na horizont 20 rokov. Podľa Sedlákovej et al. (2007) predstavujú mladí ľudia na Slovensku novú generáciu slovenských spotrebiteľov po politických zmenách v 90. rokoch minulého storočia. Označuje ju za potenciálnych multiplikátorov, ktorí budú v budúcnosti výrazne ovplyvňovať spotrebiteľské správanie a rozhodovanie pri nákupe tovarov a služieb. Zákaznícky segment generácie Y sa teda v priebehu nasledujúcich rokov stane atraktívnou cieľovou skupinou pre všetkých obchodníkov a marketérov. Vzhľadom na vyššie uvedené skutočnosti, prezentujeme parciálne výsledky prieskumu, ktorý sa zaoberal fenoménom zľavovej politiky v kybernetickom priestore vo vzťahu k nákupnému správaniu generácie Y.

2. Materiál a metódy

Prieskumu sa zúčastnilo 343 študentov nitrianskych univerzít, z toho žien bolo 288 a mužov 55. Údaje boli zbierané formou elektronického dotazníka. Dotazník bol zameraný na fenomén zľavovej politiky v kybernetickom priestore. V dotazníku študenti uvádzali, aký finančný obnos majú k dispozícii počas jedného mesiaca, ako dlho a ako často používajú internet, koľko času denne trávia na internete, či majú nejaké skúsenosti s online nakupovaním prostredníctvom zľavových portálov, v prípade, že áno, ako často využívajú služby zľavových portálov, odkiaľ sa o nich dozvedeli, čo je pre nich rozhodujúce pri výbere zľavového portálu, pre koho nakupujú a čo a aká je ich hlavná motivácia nakupovania na zľavových portáloch.

V príspevku sme sa zamerali na to, koľko eur mesačne majú študenti nitrianskych univerzít k dispozícii a či nakupujú na zľavových portáloch. Vzhľadom na hlavný cieľ prieskumu, ktorým bolo zmapovať súčasný stav on-line nakupovania na zľavových portáloch u vysokoškolských študentov v Nitre, sme formulovali nasledujúci predpoklad: Predpokladáme, že on-line nakupovanie prostredníctvom zľavových portálov sa bude u respondentov generácie Y líšiť vzhľadom na diferenciáciu mesačného finančného obnosu.

Pri analýze sme použili metódy popisnej a inferenčnej štatistiky (χ^2 - test nezávislosti). Výsledky sme spracovali v programe SPSS.

3. Analýza závislosti medzi mesačným finančným obnosom študentov a nakupovaním na zľavových portáloch

V položke dotazníka „Aký finančný obnos máte mesačne k dispozícii?“ sa každý zo študentov zaradil do jednej z nasledujúcich kategórií: menej ako 50 eur, 50 – 100 eur, 100 – 200 eur, 200 – 300 eur, 300 – 400 eur, viac ako 400 eur. 14% študentov uviedlo, že ich mesačný finančný obnos je menší ako 50 eur, od 50 do 100 eur má k dispozícii 34,4% študentov, od 100 do 200 eur má k dispozícii 28,6% študentov, od 200 do 300 eur má k dispozícii 10,5% študentov, od 300 do 400 eur mesačne má 3,2% študentov a viac ako 400 eurami mesačne môže disponovať 9,3% študentov. Najpočetnejšiu skupinu študentov tvoria tí, ktorí môžu minúť mesačne od 50 do 100 eur. Celkovo na zľavových portáloch nakupuje 63% študentov, 20,7% študentov uviedlo, že na zľavovom portáli ešte nenakupovali, ale o takomto nákupe uvažujú, 16,3% študentov na zľavovom portáli nenakupujú a ani v najbližšej dobe neplánujú.

Kontingenčná tabuľka financií študentov a ich nákupného správania je zobrazená v tab. 1.

Tab. 1: Kontingenčná tabuľka

			Nakupujete na zľavových portáloch?			Spolu
			Áno	Nie, ale uvažujem o takomto nákupe.	Nie, ani v najbližšej dobe neplánujem takýto nákup	
Mesačný obnos	Menej ako 50 EUR	počet	26	9	13	48
		riadkové %	54,2%	18,8%	27,1%	100,0%
		% celkovo	7,6%	2,6%	3,8%	14,0%
	50 – 100 EUR	počet	74	28	16	118
		riadkové %	62,7%	23,7%	13,6%	100,0%
		% celkovo	21,6%	8,2%	4,7%	34,4%
	100 – 200 EUR	počet	60	22	16	98
		riadkové %	61,2%	22,4%	16,3%	100,0%
		% celkovo	17,5%	6,4%	4,7%	28,6%
	200 – 300 EUR	počet	22	8	6	36
		riadkové %	61,1%	22,2%	16,7%	100,0%
		% celkovo	6,4%	2,3%	1,7%	10,5%
	300 – 400 EUR	počet	8	0	3	11
		riadkové %	72,7%	0,0%	27,3%	100,0%
		% celkovo	2,3%	0,0%	0,9%	3,2%
400 EUR a viac	počet	26	4	2	32	
	riadkové %	81,3%	12,5%	6,3%	100,0%	
	% celkovo	7,6%	1,2%	0,6%	9,3%	
Spolu		počet	216	71	56	343
		% celkovo	63,0%	20,7%	16,3%	100,0%

Pre potreby ďalšej analýzy sme zlúčili niektoré kategórie mesačných finančných prostriedkov študentov. Vytvorili sme tri kategórie – prvá kategória študentov mesačne disponuje finančnými prostriedkami do 100 eur a tvoria 48,4% počtu študentov. Druhá kategória zahrňuje študentov, ktorí majú mesačne k dispozícii od 100 do 300 eur, títo študenti tvoria 39,1% z celkového počtu respondentov. Nakoniec, študenti, ktorí majú mesačne k dispozícii viac ako 300 eur, tvoria 12,5% celkového počtu respondentov. Zo študentov, ktorí disponujú mesačne financiami do 100 eur, 60,2% využíva zľavové portály, 22,3% študentov nevyužíva nákupy na zľavových portáloch, ale uvažuje o takom nákupe, a 17,5% študentov zľavové portály nevyužíva a ani ich v blízkej dobe neplánuje využívať. Veľmi podobný stav je v skupine študentov s mesačnými finančnými prostriedkami od 100 do 300 eur, 61,2% využíva zľavové portály, 22,4% ich zatiaľ nevyužíva, ale blízkej dobe plánuje, a 16,4% zľavové portály nevyužíva. Študenti, ktorí majú mesačne k dispozícii viac ako 300 eur, zľavové portály prevažne využívajú – uviedlo tak 79,1% z nich. 9,3% študentov s obnosom viac ako 300 eur nenakupuje zatiaľ na zľavových portáloch, ale plánuje takýto nákup v blízkej dobe uskutočniť, a 11,6% študentov takýto nákup nezvažuje (tab. 2).

Tab. 2: Kontingenčná tabuľka – zlúčené kategórie, riadkové percentá

			Nakupovanie na zľavových portáloch			Spolu
			Áno	Nie, ale uvažujem o takomto nákupe	Nie, ani v najbližšej dobe neplánujem takýto nákup	
Mesačný obnos	do 100 EUR	počet	100	37	29	166
		%	60,2%	22,3%	17,5%	100,0%
	100 – 300 EUR	počet	82	30	22	134
		%	61,2%	22,4%	16,4%	100,0%
	viac ako 300 EUR	počet	34	4	5	43
		%	79,1%	9,3%	11,6%	100,0%
Spolu		počet	216	71	56	343
		%	63,0%	20,7%	16,3%	100,00%

Na základe týchto výsledkov usudzujeme, že mesačný finančný obnos, ktorý má k dispozícii študent nitrianskej vysokej školy, nemá vplyv na to, či využíva alebo nevyužíva zľavové portály, resp. či ich plánuje v dohľadnej dobe využiť. Platnosť tejto hypotézy sme overovali aj pomocou χ^2 -testu nezávislosti. Zvolili sme hladinu významnosti $\alpha = 0,05$. Výsledok testovania je uvedený v tab. 3.

Tab. 3: Výsledok χ^2 -testu nezávislosti

	Hodnota testovacej štatistiky χ^2	Počet stupňov voľnosti	p – hodnota
Pearson Chi-Square	5,834	4	0,212
Likelihood Ratio	6,461	4	0,167
N of Valid Cases	343		

Hodnota testovacej štatistiky $\chi^2 = 5,834$. Výsledok testu vyhodnotíme na základe p – hodnoty. Je to pravdepodobnosť chyby, ktorej sa dopustíme, keď zamietneme testovanú hypotézu. V tomto prípade $p = 0,212$. Keďže p – hodnota je vyššia ako zvolená hladina významnosti $\alpha = 0,05$, testovanú hypotézu nemôžeme na hladine významnosti $\alpha = 0,05$ zamietnuť. To znamená, že realizácia nákupov na zľavových portáloch nezávisí od toho, do akej kategórie mesačných finančných prostriedkov spadá študent niektorej z nitrianskych univerzít.

4. Záver

V príspevku sme zamerali na fenomén nakupovania na zľavových portáloch u študentov univerzít. Prieskumu sa zúčastnilo 343 študentov dvoch nitrianskych univerzít: Univerzity Konštantína Filozofa v Nitre a Slovenskej poľnohospodárskej univerzity, reprezentujúci zákaznícky segment generácie Y. V realizovanom prieskume sme zisťovali, či sa využívanie on-line nákupov na zľavových portáloch líši vzhľadom na rôznorodosť finančného obnosu, koľko má študent mesačne k dispozícii. Z výsledkov, ktoré sme nadobudli, vyplýva, že väčšina študentov aktívne využíva zľavové portály, a to nezávisle na tom, akým finančným obnosom mesačne môže disponovať. Uvedené skutočnosti však nemožno paušalizovať na celé zákaznícke spektrum generácie Y na Slovensku. Výsledky prieskumu sa vzťahujú len jednu kohortu tejto cieľovej skupiny. V praxi by sme nadobudnuté zistenia mohli uplatniť v oblasti marketingu a marketingovej komunikácie. Ponúkajú totiž náhľad na charakter nákupného

správania generácie Y realizovaného na internete prostredníctvom zľavových portálov, čo umožňuje marketingovým pracovníkom efektívnejšie zacieliť komunikačné stimuly, ktoré podnecujú ku kúpe produktu alebo služby. Zároveň ponúkajú možnosť, ako uplatniť nové obchodné postupy za využitia elektronických platforiem pre realizáciu podnikania v spoločnosti.

Nakoľko súčasná generácia Y vstupuje na trh práce, stáva sa súčasťou zarábajúcej časti obyvateľstva. Vzhľadom na ich masívne zrenie ako demografického segmentu je nevyhnutné zaoberať sa ich nákupným správaním, ktoré už v dnešnej dobe nastoľuje a udáva marketingové trendy.

Literatúra

DŽUPINA, M. – FANDELOVÁ, E. et al. 2013. Vnímanie sociálnej zodpovednosti a vek. In: *Analýza a výskum v marketingovej komunikácii*, roč. 1, č. 1, s. 41 – 51.

CISCO PRESS RELEASE. *Toothpaste, Toilet Paper, and Texting– Say Good Morning to Gen Y*. [online]. 2012. Dostupné na internete: <<http://newsroom.cisco.com/release/1114955>>.

LINCÉNYI, M. et al. 2011. *Marketing – Vybrané kapitoly*. Trenčín: TnUAD.

MENDOZA, J. 2002. *Marketing Insight: Talkin' about Y Generation*. [online]. North Parramatta – Australia: The Australian Leadership Foundation, 2004. Dostupné na internete: <<http://www.grafica.com/pdf/YGeneration.pdf>>.

SEDLÁKOVÁ, J. et al. 2007. Spotrebiteľské správanie sa mladých slovenských konzumentov – vplyv etnocentrizmu a pôvodu potravín. In: *Acta oeconomica et informatica*, roč. 10, č. 2, s. 45 – 48.

SVĚTLÍK, J. 2012. O podstatě reklamy. Bratislava: Eurokódex.

SZABOVÁ, E. – SZABO, P. 2013. Použitie štatistických metód v projektoch študentov marketingovej komunikácie a reklamy. In: *Forum statisticum slovacum*, roč. 9, č. 3, s. 129 – 134.

Adresy autorov:

Ivana Polakevičová, Mgr., PhD.
Katedra masmediálnej komunikácie
a reklamy
Filozofická fakulta
Univerzita Konštantína Filozofa v Nitre
Dražovská 4, 949 74 Nitra
ipolakevicova@ukf.sk

Edita Szabová, PaedDr.
Katedra matematiky
Fakulta prírodných vied
Univerzita Konštantína Filozofa v Nitre
Tr. A. Hlinku 1, 949 74 Nitra
edita.szabova@ukf.sk

Tatiana Kamenská, Mgr.
Katedra masmediálnej komunikácie
a reklamy
Filozofická fakulta
Univerzita Konštantína Filozofa v Nitre
Dražovská 4, 949 74 Nitra
kamenska.tatiana@gmail.com

Grantová podpora:

Príspevok bol podporený projektom UGA III / 10/2014 The Perception of Advertisement by Generation Y within the Interpretative Framework of Transactional Analysis.

Řešení problému polynomické kalibrace metodou maximální věrohodnosti Solving the problem of polynomial calibration by the maximum likelihood method

Petra Širůčková

Abstract: Calibration is important part of statistics frequently used in metrology. It is a set of tasks which gives relationship between the reference device and calibrated device (if some special conditions are realized). This relationship is describe by transformation function and represented by transformation curve. In this paper we focuse on polynomial calibration model, which is solved by the maximum likelihood method.

Abstrakt: Kalibrace je důležitou součástí statistiky, často využívaná v metrologii. Je to soubor úkonů, které udávají vztah mezi referenčním přístrojem a kalibrovaným přístrojem (za předpokladu, že požadované podmínky jsou splněny). Tento vztah je popsán transformační funkcí a reprezentován transformační křivkou. V tomto příspěvku se zaměříme na polynomický kalibrační model, který popíšeme pomocí metody maximální věrohodnosti.

Key words: Polynomial calibration, transformation function, maximum likelihood method.

Klíčové slová: polynomická kalibrace, transformační funkce, metoda maximální věrohodnosti.

JEL classification: C130, C150

1. Úvod

V práci se budeme zabývat řešením problému polynomické kalibrace s využitím metody maximální věrohodnosti. Předpokládejme, že měříme m různých objektů pomocí dvou různých přístrojů (přístroje A a přístroje B). Dále budeme předpokládat, že přístroj A je „méně přesný“ (kalibrovaný) a přístroj B je „přesnější“ (etalonový). Navíc budeme předpokládat, že naměřené hodnoty na obou přístrojích mají normální rozdělení a pro každý z m objektů platí, že při měření přístrojem A jsou bezchybně měřené hodnoty $\mu = (\mu_1, \dots, \mu_m)$, a při měření přístrojem B jsou bezchybně měřené hodnoty $v = (v_1, \dots, v_m)$, kde $v_i = a_k \mu_i^k + \dots + a_2 \mu_i^2 + a_1 \mu_i + a_0$, $k \in \mathbb{N}$, $a_0, a_1, \dots, a_k \in \mathbb{R}$, $i = 1, 2, \dots, m$. V příspěvku popíšeme odhad parametrů transformační polynomické funkce, nejistoty určení těchto parametrů, pás spolehlivosti kolem transformační funkce a algoritmus měření kalibrovaným přístrojem. Transformační funkce je v příspěvku polynom stupně k a vyjadřuje (teoretickou) závislost bezchybně měřených hodnot etalonu na bezchybně měřených hodnotách kalibrovaným přístrojem.

2. Odhad parametrů transformační funkce metodou maximální věrohodnosti

Mějme tedy m různých objektů, označme je z_1, z_2, \dots, z_m , na kterých provádíme měření pomocí dvou přístrojů A a B . Předpokládejme, že měření na přístrojích A , B jsou nezávislá.

Označme x_{ij} realizaci náhodné proměnné X_{ij} , která představuje j -té měření veličiny z_i přístrojem A a předpokládejme, že $X_{ij} \sim N(\mu_i, \sigma_x^2)$. Dále označme y_{ij} realizaci náhodné proměnné Y_{ij} , která představuje j -té měření veličiny z_i přístrojem B , kde $Y_{ij} \sim N(v_i, \sigma_y^2)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$.

První měření v našem modelu necht' je vektor $(X_{11}, X_{21}, \dots, X_{m1}, Y_{11}, Y_{21}, \dots, Y_{m1})$. Měření provedeme n krát. Tedy n -té měření bude mít tvar $(X_{1n}, X_{2n}, \dots, X_{mn}, Y_{1n}, Y_{2n}, \dots, Y_{mn})$. Všechna měření jsou nezávislá.

Naším cílem je najít odhad vektoru parametrů $\Theta = (\sigma_x^2, \sigma_y^2, \mu_1, \dots, \mu_m, a_0, a_1, \dots, a_k)'$. Protože známe rozdělení náhodných veličin X_{ij} , Y_{ij} pro $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$ a předpokládáme, že tyto veličiny jsou nezávislé, můžeme určit jejich hustotu a k výpočtu neznámého parametru Θ použít odhad, který získáme metodou maximální věrohodnosti.

Označíme-li $\varphi = (\varphi_1, \dots, \varphi_n)$, kde φ_i je hustota i -tého měření, pak:

$$\begin{aligned} \varphi_1(x_{11}, \dots, y_{m1}; \Theta) &= \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x_{i1}-\mu_i)^2}{2\sigma_x^2}} \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y_{i1}-a_k\mu_i^k - \dots - a_1\mu_i - a_0)^2}{2\sigma_y^2}} \\ &\quad \vdots \\ \varphi_n(x_{1n}, \dots, y_{mn}; \Theta) &= \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x_{in}-\mu_i)^2}{2\sigma_x^2}} \prod_{i=1}^m \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{(y_{in}-a_k\mu_i^k - \dots - a_1\mu_i - a_0)^2}{2\sigma_y^2}} \end{aligned} \quad (1)$$

Označme $p_i = a_k\mu_i^k + \dots + a_1\mu_i + a_0$ a derivaci $dp_i = \frac{\partial p_i}{\partial \mu_i} = k \cdot a_k\mu_i^{k-1} + \dots + 2a_2\mu_i + a_1$.

Jestli-že x_{11}, \dots, y_{mn} jsou naměřené hodnoty, získáme věrohodnostní funkci:

$$L(x_{11}, \dots, y_{mn}; \Theta) = (2\pi)^{-mn} \sigma_x^{-\frac{2-mn}{2}} \sigma_y^{-\frac{2-mn}{2}} e^{-\frac{1}{2\sigma_x^2} \sum_{i=1}^m \sum_{j=1}^n (x_{ij}-\mu_i)^2} e^{-\frac{1}{2\sigma_y^2} \sum_{i=1}^m \sum_{j=1}^n (y_{ij}-p_i)^2}. \quad (2)$$

Logaritmická věrohodnostní funkce je:

$$\begin{aligned} l(x_{11}, \dots, y_{mn}; \Theta) &= \ln L(x_{11}, \dots, y_{mn}; \Theta) = -mn \ln(2\pi) - \frac{mn}{2} \ln \sigma_x^2 - \frac{mn}{2} \ln \sigma_y^2 - \\ &\quad - \frac{1}{2\sigma_x^2} \sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \mu_i)^2 - \frac{1}{2\sigma_y^2} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - p_i)^2. \end{aligned} \quad (3)$$

Počítáme maximálně věrohodný odhad Θ^* , tedy musí platit: $L(X_{11}, \dots, Y_{mn}; \Theta) \leq L(X_{11}, \dots, Y_{mn}; \Theta^*)$ pro $\forall \Theta$, přičemž Θ^* je odhad parametru Θ získaný metodou maximální věrohodnosti.

Odhad Θ^* parametrů Θ získáme s využitím softwaru Matlab. Všechny funkce, které jsou v této práci použity, jsou k dispozici na webových stránkách <http://www.math.muni.cz/xsirucko>. Parametry transformační funkce odhadneme pomocí funkce `n_iterace.m`, vstupem této funkce je počáteční odhad transformační funkce, který získáme s využitím funkce `n_odhad.m` (na vstupu jsou matice $X \dots$ matice $n \times m$ získaná měřením pomocí kalibrovaného přístroje, $Y \dots$ matice $n \times m$ získaná měřením pomocí etalonového přístroje, $st \dots$ stupeň odhadovaného polynomu).

3. Konfidenční oblast pro vektor parametrů

Nyní odvodíme konfidenční oblast pro odhad Θ^* , zejména pro odhad parametrů a_0, a_1, \dots, a_k .

Podle [2, str. 160] je Θ^* asymptoticky nestranný odhad Θ , tedy pro dostatečně velké n :

$$\Theta^* \sim N\left(\Theta, \frac{1}{n} J(\Theta)^{-1}\right),$$

kde $J(\Theta)$ je Fisherova informační matice i -tého měření a platí: $\{J(\Theta)\}_{ij} = E_{\Theta} \left(-\frac{\partial^2 \ln \varphi_1(X_{11}, \dots, Y_{mn}; \Theta)}{\partial \theta_i \partial \theta_j} \right)$.

Nyní odvodíme tvar Fisherovy informační matice i -tého měření. Nejdříve spočítáme první derivace logaritmické věrohodnostní funkce podle parametru Θ a položíme je rovny nule:

$$\begin{aligned}
\frac{\partial l}{\partial a_0} &= \frac{1}{\sigma_y^2} \sum_{i=1}^m (y_{i,j} - p_i) = 0 \\
\frac{\partial l}{\partial a_1} &= \frac{1}{\sigma_y^2} \sum_{i=1}^m (y_{i,j} - p_i) \mu_i = 0 \\
\frac{\partial l}{\partial a_2} &= \frac{1}{\sigma_y^2} \sum_{i=1}^m (y_{i,j} - p_i) \mu_i^2 = 0 \\
&\vdots \\
\frac{\partial l}{\partial a_k} &= \frac{1}{\sigma_y^2} \sum_{i=1}^m (y_{i,j} - p_i) \mu_i^k = 0 \\
\frac{\partial l}{\partial \sigma_x^2} &= -\frac{m}{\sigma_x^2} + \frac{1}{2\sigma_y^4} \sum_{i=1}^m (x_{i,j} - \mu_i)^2 = 0 \quad (4) \\
\frac{\partial l}{\partial \sigma_y^2} &= -\frac{m}{\sigma_y^2} + \frac{1}{2\sigma_x^4} \sum_{i=1}^m (y_{i,j} - p_i)^2 = 0 \\
\frac{\partial l}{\partial \mu_1} &= \frac{x_{1,j} - \mu_1}{\sigma_x^2} + \frac{(y_{1,j} - p_1) dp_1}{\sigma_y^2} = 0 \\
&\vdots \\
\frac{\partial l}{\partial \mu_m} &= \frac{x_{m,j} - \mu_m}{\sigma_x^2} + \frac{(y_{m,j} - p_m) dp_m}{\sigma_y^2} = 0
\end{aligned}$$

Označme $d_{ij} = \frac{\partial^2 \ln \varphi_1(x_{11}, \dots, y_{mn}; \Theta)}{\partial \Theta_i \partial \Theta_j}$, kde $\Theta_0 = a_0, \dots, \Theta_{k+m+2} = \mu_m$, pak:

pro $l \in \{0, 1, \dots, k\}$:

$$\begin{aligned}
d_{l,0} &= -\frac{1}{\sigma_y^2} \sum_{i=1}^m \mu_i^l & d_{k+2,1} &= -\frac{1}{\sigma_y^4} \sum_{i=1}^m (y_{i,1} - p_i) \mu_i \\
&& &\vdots \\
d_{l,1} &= -\frac{1}{\sigma_y^2} \sum_{i=1}^m \mu_i^{l+1} & d_{k+2,k} &= -\frac{1}{\sigma_y^4} \sum_{i=1}^m (y_{i,1} - p_i) \mu_i^k \\
d_{l,2} &= -\frac{1}{\sigma_y^2} \sum_{i=1}^m \mu_i^{l+2} & d_{k+2,k+1} &= 0 \\
&\vdots & d_{k+2,k+2} &= \frac{m}{2\sigma_y^4} - \frac{1}{\sigma_y^6} \sum_{i=1}^m (y_{i,1} - p_i)^2 \\
d_{l,k} &= -\frac{1}{\sigma_y^2} \sum_{i=1}^m \mu_i^{2l} & d_{k+2,k+3} &= -\frac{(y_{1,1} - p_1) dp_1}{\sigma_y^4} \\
d_{l,k+1} &= 0 & &\vdots \\
d_{l,k+2} &= -\frac{1}{\sigma_y^4} \sum_{i=1}^m (y_{i,1} - p_i) \mu_i^l & d_{k+2,k+m+2} &= -\frac{(y_{m,1} - p_m) dp_m}{\sigma_y^4} \\
d_{l,k+3} &= \frac{l(y_{11} - p_1) \mu_1^{l-1} - dp_1 \mu_1^l}{\sigma_y^2} & & \text{pro } m+k+2 \geq i \geq k+3 \\
&\vdots & d_{i,0} &= -\frac{dp_i}{\sigma_y^2} \quad (5) \\
d_{l,k+m+2} &= \frac{l(y_{m1} - p_m) \mu_m^{l-1} - dp_m \mu_m^l}{\sigma_y^2} & d_{i,1} &= \frac{y_{i-k-2,1} - p_{i-k-2} - dp_{i-k-2} \mu_{i-k-2}}{\sigma_y^2} \\
&& &\vdots \\
d_{k+1,0} &= d_{k+1,1} = \dots = d_{k+1,k} = 0 & d_{i,k} &= \frac{k(y_{i-k-2,1} - p_{i-k-2}) \mu_{i-k-2}^{k-1} - dp_{i-k-2} \mu_{i-k-2}^k}{\sigma_y^2} \\
d_{k+1,k+1} &= \frac{m}{2\sigma_x^4} - \frac{1}{\sigma_x^6} \sum_{i=1}^m (x_{i,1} - \mu_i)^2 & d_{i,k+1} &= -\frac{x_{i-k-2,1} - \mu_{i-k-2}}{\sigma_x^4} \\
d_{k+1,k+2} &= 0 & d_{i,k+2} &= -\frac{(y_{i-k-2,1} - p_{i-k-2}) dp_{i-k-2}}{\sigma_y^4} \\
d_{k+1,k+3} &= -\frac{x_{1,1} - \mu_1}{\sigma_x^4} & d_{i,i} &= -\frac{1}{\sigma_x^2} + \frac{(y_{i-k-2,1} - p_{i-k-2}) dp_{i-k-2} - dp_{i-k-2}^2}{\sigma_y^2} \\
&\vdots & & \text{pro } m+k+2 \geq i, o \geq k+3, i \neq o \\
d_{k+1,k+m+2} &= -\frac{x_{m,1} - \mu_m}{\sigma_x^4} & d_{i,o} &= 0 \\
d_{k+2,0} &= -\frac{1}{\sigma_y^4} \sum_{i=1}^m (y_{i,1} - p_i)
\end{aligned}$$

Označme $j_{ij} = E \left(-\frac{\partial^2 \ln \varphi_1(X_{11}, \dots, Y_{m1}; \Theta)}{\partial \Theta_i \partial \Theta_j} \right)$, pak $J(\Theta) = \begin{pmatrix} j_{0,0} & \dots & j_{0,m+k+2} \\ \vdots & \ddots & \vdots \\ j_{m+k+2,0} & \dots & j_{m+k+2,m+k+2} \end{pmatrix}$

pro $l \in \{0, 1, \dots, k\}$:

$$\begin{aligned}
 j_{l,0} &= -d_{l,0}, \dots, j_{l,k+1} = -d_{l,k+1} \\
 j_{l,k+2} &= \frac{1}{\sigma_y^4} \sum_{i=1}^m \int \dots \int (y_{i,1} - p_i) \mu_i^l f(x_{11}) \dots f(y_{m1}) dx_{11} \dots dy_{m1} = 0 \\
 j_{l,k+3} &= \frac{dp_1 \mu_1^l}{\sigma_y^2} \\
 &\vdots \\
 j_{l,k+m+2} &= \frac{dp_m \mu_m^l}{\sigma_y^2} \\
 j_{k+1,0} &= j_{k+1,1} = \dots = j_{k+1,k} = 0 \\
 j_{k+1,k+1} &= \frac{m}{2\sigma_x^4} \\
 j_{k+1,k+2} &= 0 \\
 j_{k+1,k+3} &= \dots = j_{k+1,k+m+2} = 0 \\
 j_{k+2,0} &= \dots = j_{k+2,k+1} = 0 \\
 j_{k+2,k+2} &= \frac{m}{2\sigma_y^4} \\
 j_{k+2,k+3} &= \dots = j_{k+2,k+m+2} = 0
 \end{aligned} \tag{6}$$

pro $m+k+2 \geq i, 0 \geq k+3, i \neq 0, l \in \{0, 1, \dots, k\}$

$$\begin{aligned}
 j_{i,l} &= \frac{dp_{i-k-2} \mu_{i-k-2}^l}{\sigma_y^2} \\
 j_{i,k+1} &= j_{i,k+2} = 0 \\
 j_{i,i} &= \frac{1}{\sigma_x^2} + \frac{dp_{i-k-2}^2}{\sigma_y^2} \\
 j_{i,k} &= 0.
 \end{aligned}$$

Dostali jsem matici $J(\Theta)$:

$$\begin{pmatrix}
 \frac{\sum_{i=1}^m \mu_i^0}{\sigma_y^2} & \frac{\sum_{i=1}^m \mu_i^1}{\sigma_y^2} & \dots & \frac{\sum_{i=1}^m \mu_i^k}{\sigma_y^2} & 0 & 0 & \frac{dp_1}{\sigma_y^2} & \frac{dp_2}{\sigma_y^2} & \dots & \frac{dp_m}{\sigma_y^2} \\
 \frac{\sum_{i=1}^m \mu_i^1}{\sigma_y^2} & \frac{\sum_{i=1}^m \mu_i^2}{\sigma_y^2} & \dots & \vdots & 0 & 0 & \frac{dp_1 \mu_1}{\sigma_y^2} & \frac{dp_2 \mu_2}{\sigma_y^2} & \dots & \frac{dp_m \mu_m}{\sigma_y^2} \\
 \vdots & \vdots & \ddots & \vdots & 0 & 0 & \vdots & \vdots & \dots & \vdots \\
 \frac{\sum_{i=1}^m \mu_i^k}{\sigma_y^2} & \frac{\sum_{i=1}^m \mu_i^{k+1}}{\sigma_y^2} & \dots & \frac{\sum_{i=1}^m \mu_i^{2k}}{\sigma_y^2} & 0 & 0 & \frac{dp_1 \mu_1^k}{\sigma_y^2} & \frac{dp_2 \mu_2^k}{\sigma_y^2} & \dots & \frac{dp_m \mu_m^k}{\sigma_y^2} \\
 0 & 0 & 0 & 0 & \frac{m}{2\sigma_x^4} & 0 & 0 & 0 & \dots & 0 \\
 0 & 0 & 0 & 0 & 0 & \frac{m}{2\sigma_y^4} & 0 & 0 & \dots & 0 \\
 \frac{dp_1}{\sigma_y^2} & \frac{dp_1 \mu_1}{\sigma_y^2} & \dots & \frac{dp_1 \mu_1^k}{\sigma_y^2} & 0 & 0 & \frac{1}{\sigma_x^2} + \frac{dp_1^2}{\sigma_y^2} & 0 & \dots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 & \ddots & \ddots & 0 \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\
 \frac{dp_m}{\sigma_y^2} & \frac{dp_m \mu_m}{\sigma_y^2} & \dots & \frac{dp_m \mu_m^k}{\sigma_y^2} & 0 & 0 & 0 & \dots & 0 & \frac{1}{\sigma_x^2} + \frac{dp_m^2}{\sigma_y^2}
 \end{pmatrix}$$

Abychom získali Fisherovu informační matici, musíme spočítat $J(\Theta)^{-1}$. K výpočtu inverzní matice použijeme vzorec (viz [4, str. 66]):

$$\begin{pmatrix} A & B \\ B' & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}B')^{-1} & -(A - BD^{-1}B')^{-1}BD^{-1} \\ -D^{-1}B'(A - BD^{-1}B')^{-1} & D^{-1} + D^{-1}B'(A - BD^{-1}B')^{-1}BD^{-1} \end{pmatrix},$$

kde

$$A = \begin{pmatrix} j_{0,0} & \cdots & j_{0,k+2} \\ \vdots & \ddots & \vdots \\ j_{k+2,0} & \cdots & j_{k+2,k+2} \end{pmatrix}, \quad B = \begin{pmatrix} j_{0,k+3} & \cdots & j_{0,m+k+2} \\ \vdots & \ddots & \vdots \\ j_{k+2,k+3} & \cdots & j_{k+2,m+k+2} \end{pmatrix},$$

$$D = \begin{pmatrix} j_{k+3,k+3} & \cdots & j_{k+3,m+k+2} \\ \vdots & \ddots & \vdots \\ j_{m+k+2,k+3} & \cdots & j_{m+k+2,m+k+2} \end{pmatrix}.$$

Zaměřme se nyní na odhad parametrů a_0, a_1, \dots, a_k . Označme tento odhad $\widehat{\Theta}^k = \begin{pmatrix} \widehat{a}_0 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_k \end{pmatrix}$

a skutečnou hodnotu parametrů a_0, a_1, \dots, a_k označme Θ^k , pak podle [2, str. 160]

$$\sqrt{n}(\widehat{\Theta}^k - \Theta^k) \xrightarrow[n]{D} N\left(0, J(\Theta^k)^{-1}\right) \quad (7)$$

(konvergence v distribuci). Tedy pro dostatečně velké n můžeme psát

$$\widehat{\Theta}^k \approx N\left(\begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_k \end{pmatrix}, \frac{1}{n}J(\Theta^k)^{-1}\right), \quad (8)$$

kde $\{J(\Theta^k)\}_{i,j} = \{(A - BD^{-1}B')\}_{i,j}$, $i = 0, 1, 2, \dots, k$, $j = 0, 1, 2, \dots, k$.

Počítejme postupně:

$$D^{-1} = \begin{pmatrix} \frac{\sigma_x^2 \cdot \sigma_y^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\sigma_x^2 \cdot \sigma_y^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} \end{pmatrix},$$

$$D^{-1}B = \begin{pmatrix} \frac{\sigma_x^2 \cdot \sigma_y^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\sigma_x^2 \cdot \sigma_y^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} \end{pmatrix} \begin{pmatrix} \frac{dp_1}{\sigma_y^2} & \frac{dp_1 \mu_1}{\sigma_y^2} & \frac{dp_1 \mu_1^2}{\sigma_y^2} & \cdots & \frac{dp_1 \mu_1^k}{\sigma_y^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{dp_m}{\sigma_y^2} & \frac{dp_m \mu_m}{\sigma_y^2} & \frac{dp_m \mu_m^2}{\sigma_y^2} & \cdots & \frac{dp_m \mu_m^k}{\sigma_y^2} & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\sigma_x^2 dp_1}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \frac{\sigma_x^2 dp_1 \mu_1}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \frac{\sigma_x^2 dp_1 \mu_1^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \cdots & \frac{\sigma_x^2 dp_1 \mu_1^k}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\sigma_x^2 dp_m}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \frac{\sigma_x^2 dp_m \mu_m}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \frac{\sigma_x^2 dp_m \mu_m^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \cdots & \frac{\sigma_x^2 dp_m \mu_m^k}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & 0 & 0 \end{pmatrix},$$

$$BD^{-1}B' =$$

$$= \begin{pmatrix} \frac{dp_1}{\sigma_y^2} & \dots & \frac{dp_m}{\sigma_y^2} \\ \frac{dp_1\mu_1}{\sigma_y^2} & \dots & \frac{dp_m\mu_m}{\sigma_y^2} \\ \frac{dp_1\mu_1^2}{\sigma_y^2} & \dots & \frac{dp_m\mu_m^2}{\sigma_y^2} \\ \vdots & \vdots & \vdots \\ \frac{dp_1\mu_1^k}{\sigma_y^2} & \dots & \frac{dp_m\mu_m^k}{\sigma_y^2} \\ 0 & \dots & 0 \\ 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} \frac{\sigma_x^2 dp_1}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \frac{\sigma_x^2 dp_1 \mu_1}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \frac{\sigma_x^2 dp_1 \mu_1^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & \dots & \frac{\sigma_x^2 dp_1 \mu_1^k}{\sigma_y^2 + \sigma_x^2 \cdot dp_1^2} & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\sigma_x^2 dp_m}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \frac{\sigma_x^2 dp_m \mu_m}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \frac{\sigma_x^2 dp_m \mu_m^2}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & \dots & \frac{\sigma_x^2 dp_m \mu_m^k}{\sigma_y^2 + \sigma_x^2 \cdot dp_m^2} & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^2}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \dots & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^k}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & 0 & 0 \\ \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^2}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^3}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \vdots & \vdots & 0 & 0 \\ \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^2}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^3}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^4}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 & 0 \\ \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^k}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & \dots & \dots & \dots & \sum_{i=1}^m \frac{\sigma_x^2 dp_i^2 \mu_i^{2k}}{\sigma_y^2(\sigma_y^2 + \sigma_x^2 \cdot dp_i^2)} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{B}' =$

$$= \begin{pmatrix} \sum_{i=1}^m \frac{1}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \dots & \sum_{i=1}^m \frac{\mu_i^k}{\sigma_y^2 + \sigma_x^2 dp_i^2} & 0 & 0 \\ \sum_{i=1}^m \frac{\mu_i}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^3}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \vdots & \vdots & 0 & 0 \\ \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^3}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^4}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \vdots & \vdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & 0 & 0 \\ \sum_{i=1}^m \frac{\mu_i^k}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \dots & \dots & \dots & \sum_{i=1}^m \frac{\mu_i^{2k}}{\sigma_y^2 + \sigma_x^2 dp_i^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{m}{2\sigma_x^4} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{m}{2\sigma_y^4} \end{pmatrix}$$

Tedy

$$\mathbf{J}(\boldsymbol{\theta}^k) = \begin{pmatrix} \sum_{i=1}^m \frac{1}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \dots & \sum_{i=1}^m \frac{\mu_i^k}{\sigma_y^2 + \sigma_x^2 dp_i^2} \\ \sum_{i=1}^m \frac{\mu_i}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^3}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \vdots & \vdots \\ \sum_{i=1}^m \frac{\mu_i^2}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^3}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \sum_{i=1}^m \frac{\mu_i^4}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m \frac{\mu_i^k}{\sigma_y^2 + \sigma_x^2 dp_i^2} & \dots & \dots & \dots & \sum_{i=1}^m \frac{\mu_i^{2k}}{\sigma_y^2 + \sigma_x^2 dp_i^2} \end{pmatrix}$$

Označme

$$\mathbf{Var}_{\sigma_x^2, \sigma_y^2} = \frac{1}{n} \begin{pmatrix} \frac{m}{2\sigma_x^4} & 0 \\ 0 & \frac{m}{2\sigma_y^4} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2\sigma_x^4}{mn} & 0 \\ 0 & \frac{2\sigma_y^4}{mn} \end{pmatrix}$$

a kovarianční- matici

$$\Sigma^k = \frac{1}{n} \mathbf{J}(\Theta^k)^{-1}.$$

Dostáváme asymptotickou $(1 - \alpha) \cdot 100\%$ - konfidenční oblast pro vektor parametrů $(a_0, a_1, \dots, a_k)'$

$$*C_{(1-\alpha)}^1 = \left\{ \left(\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} : \begin{pmatrix} \widehat{a}_0 - a_0 \\ \widehat{a}_1 - a_1 \\ \dots \\ \widehat{a}_k - a_k \end{pmatrix}' \right) (\Sigma^k)^{-1} \begin{pmatrix} \widehat{a}_0 - a_0 \\ \widehat{a}_1 - a_1 \\ \dots \\ \widehat{a}_k - a_k \end{pmatrix} \leq \chi_{k+1}^2(1 - \alpha) \right\}, \quad (9)$$

kde $\chi_{k+1}^2(1 - \alpha)$ je $(1 - \alpha)$ -kvantil chí - kvadrát rozdělení s $k+1$ stupni volnosti. Matici $(\Sigma^k)^{-1} = n\mathbf{J}(\Theta^k)$ neznáme, ale parametry a_0, a_1, \dots, a_k nahradíme odhady $\widehat{a}_0, \widehat{a}_1, \dots, \widehat{a}_k$, tedy $(\widehat{\Sigma}^k)^{-1} = n\mathbf{J}(\widehat{\Theta}^k)$. Takto dostaneme přibližnou asymptotickou $(1 - \alpha)$ - konfidenční-oblast pro vektor parametrů $(a_0, \dots, a_n)'$

$$C_{(1-\alpha)}^1 = \left\{ \left(\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix} : \begin{pmatrix} \widehat{a}_0 - a_0 \\ \widehat{a}_1 - a_1 \\ \dots \\ \widehat{a}_k - a_k \end{pmatrix}' \right) \mathbf{J}(\widehat{\Theta}^k) \begin{pmatrix} \widehat{a}_0 - a_0 \\ \widehat{a}_1 - a_1 \\ \dots \\ \widehat{a}_k - a_k \end{pmatrix} \leq \frac{\chi_{k+1}^2(1-\alpha)}{n} \right\}. \quad (10)$$

Pomocí $C_{(1-\alpha)}^1$, můžeme použitím simulací měření zjistit, zda je náš odhad vhodný (s jakou empirickou četností pokrývá $C_{(1-\alpha)}^1$ skutečný parametr). K simulacím využijeme funkce `n_konfidence5.m` a `n_intervaly5.m` naprogramované v softwaru Matlab. Pro polynomy stupňů 2, 3 a 4 jsou výsledky simulace k dispozici na webové stránce <http://www.math.muni.cz/xsirucko>. Tato simulační analýza je předmětem jiného příspěvku.

Nyní necht' bezchybné měření přístrojem A je x , budeme se zabývat konfidenční oblastí pro odhad $a_0 + a_1x + \dots + a_kx^k$, tedy \widehat{v}_x .

Zkoumejme vlastnosti odhadu $a_0 + a_1x + \dots + a_kx^k$:

$$a_0 + a_1x + \dots + a_kx^k = (1 \quad x \quad \dots \quad x^k) \cdot \begin{pmatrix} \widehat{a}_0 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_k \end{pmatrix}. \quad (11)$$

Označme $L'_x = (1 \quad x \quad \dots \quad x^k)$, pak

$$L'_x \widehat{\Theta}^k \approx N \left((1 \quad x \quad \dots \quad x^k) \cdot \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_k \end{pmatrix}, L'_x \widehat{\Sigma}^k L_x \right),$$

$$L'_x \widehat{\Theta}^k - L'_x \Theta^k \approx N(0, L'_x \widehat{\Sigma}^k L_x),$$

$$\frac{L'_x \widehat{\Theta}^k - L'_x \Theta^k}{\sqrt{L'_x \widehat{\Sigma}^k L_x}} \approx N(0, 1),$$

$$\frac{(L'_x \widehat{\Theta}^k - L'_x \Theta^k)^2}{L'_x \widehat{\Sigma}^k L_x} \approx \chi_1^2,$$

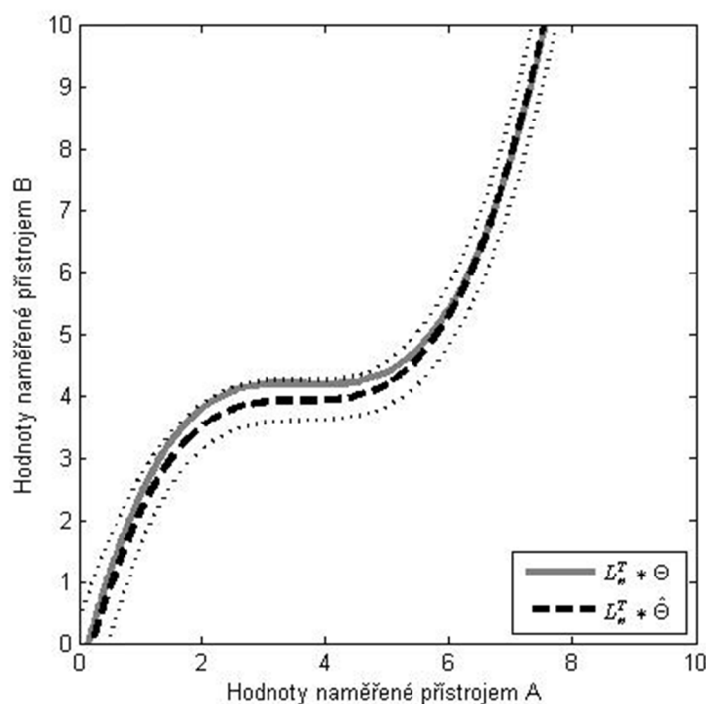
$$P \left\{ \frac{(L'_x \widehat{\Theta}^k - L'_x \Theta^k)^2}{L'_x \widehat{\Sigma}^k L_x} < \chi_1^2 (1 - \alpha) \right\} = 1 - \alpha,$$

$$P \left\{ (L'_x \widehat{\Theta}^k - L'_x \Theta^k)^2 < L'_x \widehat{\Sigma}^k L_x \cdot \chi_1^2 (1 - \alpha) \right\} = 1 - \alpha,$$

$$P \left\{ |L'_x \widehat{\Theta}^k - L'_x \Theta^k| < \sqrt{L'_x \widehat{\Sigma}^k L_x \cdot \chi_1^2 (1 - \alpha)} \right\} = 1 - \alpha,$$

$$P \left\{ L'_x \widehat{\Theta}^k - \sqrt{L'_x \widehat{\Sigma}^k L_x \cdot \chi_1^2 (1 - \alpha)} < L'_x \Theta^k < L'_x \widehat{\Theta}^k + \sqrt{L'_x \widehat{\Sigma}^k L_x \cdot \chi_1^2 (1 - \alpha)} \right\} = 1 - \alpha. \quad (12)$$

Takto jsme získali přibližný $(1 - \alpha)$ -konfidenční interval pro jednu hodnotu $v_x = a_0 + a_1 x + \dots + a_k x^k$. Pokud tuto situaci zobecníme a postupně za x dosadíme jednotlivé bezchybné hodnoty měření přístrojem A a naměřené hodnoty spojíme, dostaneme pás spolehlivosti kolem transformační křivky $a_0 + a_1 x + \dots + a_k x^k$ (Obrázek 1.1).



Obrázek 1.17 (tečkovaně...pás spolehlivosti, šedě...transformační funkce, čárkovaně ... odhad transformační funkce)

Graf je vykreslen pomocí funkce `n_graf5.m`. pro transformační funkci $0.1x^3 - 1.1x^2 + 4x - 0.6$, $\mu = (0, 2.5, 5, 7.5, 10)$, $\sigma_x = 0.5$, $\sigma_y = 0.47$, $\alpha = 0.05$.

4. Měření kalibrovaným přístrojem

V této části se zaměříme na určení konfidenčního intervalu pro v_x , za předpokladu, že přístrojem A odměříme hodnotu x .

Mějme tedy kalibrovaný přístroj A , pomocí něhož provedeme jedno měření a dostaneme hodnotu x . Tedy toto měření je $X \sim N(\mu_x, \sigma_x^2)$. Podle [2, str. 72-73] platí, že pro s^2 - nestranný odhad parametru σ_x^2 nezávislý na X , přičemž $\frac{u \cdot s^2}{\sigma_x^2} \sim \chi_u^2$ je

$$s^2 \sim \frac{\sigma_x^2}{u} \chi_u^2.$$

Rozptyl s^2 je

$$D(s^2) = \frac{\sigma_x^4}{u^2} 2u = \frac{2\sigma_x^4}{u},$$

tedy

$$u = \frac{2\sigma_x^4}{D(s^2)}.$$

Platí:

$$\frac{X - \mu_x}{\sigma_x} \sim N(0,1),$$

tedy

$$\frac{\frac{X - \mu_x}{\sigma_x}}{\sqrt{\frac{\chi_u^2}{u}}} = \frac{\frac{X - \mu_x}{\sigma_x}}{\sqrt{\frac{us^2}{\sigma_x^2 u}}} = \frac{X - \mu_x}{s} \sim t_u.$$

Nahradíme s^2 odhadem $\widehat{\sigma_x^2}$ a položíme

$$D(s^2) = D(\widehat{\sigma_x^2}) = \{\mathbf{Var}_{\sigma_x^2, \sigma_y^2}\}_{1,1}.$$

Označme

$$\hat{u} = \frac{2\widehat{\sigma_x^4}}{\{\mathbf{Var}_{\sigma_x^2, \sigma_y^2}\}_{1,1}} = m.$$

Dostaneme

$$\frac{X - \mu_x}{\widehat{\sigma_x}} \sim t_m,$$

a také $(1 - \gamma) \cdot 100\%$ - konfidenční - interval pro μ_x :

$$P\left\{X - \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right) \leq \mu_x \leq X + \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right)\right\} = 1 - \gamma,$$

kde $t_m \left(1 - \frac{\gamma}{2}\right)$ je $\left(1 - \frac{\gamma}{2}\right)$ -kvantil Studentovho rozdělení s m stupňami volnosti.

Tedy pro dané naměřené x je skutečná hodnota μ_x měření kalibrovaný přístrojem A z intervalu od $x - \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right)$ do $x + \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right)$ s předepsanou pravděpodobností $1 - \gamma$.

Abychom zjistili, jaké hodnoty by pro toto naměřené x nabýval přístroj B , musíme spočítat minimum a maximum funkce $v_y = a_k y^k + \dots + a_2 y^2 + a_1 y + a_0$, kde $y \in \left(x - \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right), x + \widehat{\sigma_x} t_m \left(1 - \frac{\gamma}{2}\right)\right) = (d_x, h_x)$.

Označme:

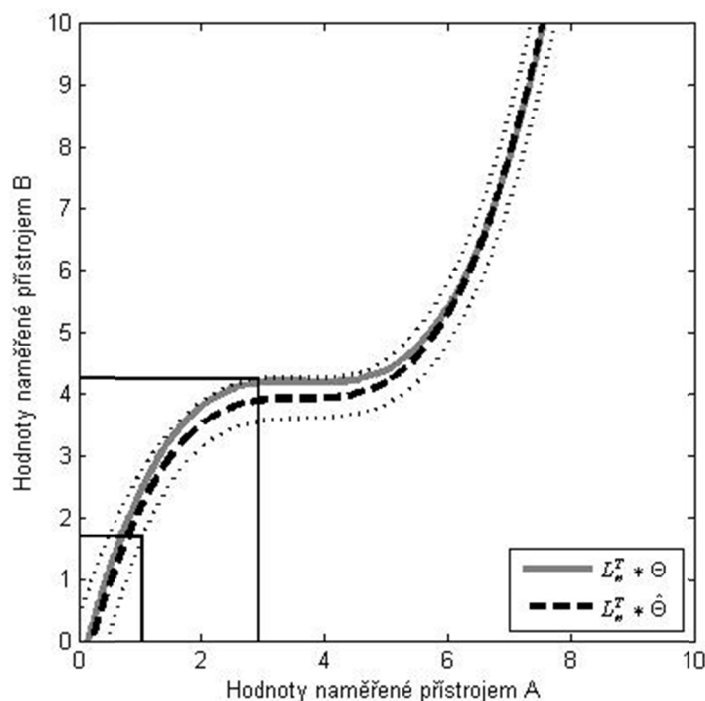
$$x_{11} = \operatorname{argmin}_{y \in (d_x, h_x)}(v_y), \quad x_{22} = \operatorname{argmax}_{y \in (d_x, h_x)}(v_y).$$

$\mathbf{x}_1 = (1, x_{11}, \dots, x_{11}^k)'$, $\mathbf{x}_2 = (1, x_{22}, \dots, x_{22}^k)'$. Pak platí (s využitím Bonferroniho nerovnosti viz [4, str. 24])

$$P\left(\mathbf{x}_1' \widehat{\Theta}^k - \sqrt{\mathbf{x}_1' \widehat{\Sigma}^k \mathbf{x}_1 \cdot \chi_1^2(1 - \alpha)} < v_x < \mathbf{x}_2' \widehat{\Theta}^k + \sqrt{\mathbf{x}_2' \widehat{\Sigma}^k \mathbf{x}_2 \cdot \chi_1^2(1 - \alpha)}\right) \geq 1 - \alpha - \gamma \quad (13).$$

Tedy na přesném přístroji bychom s pravděpodobností větší nebo rovnou $1 - \alpha - \gamma$ naměřili hodnotu v rozmezí od $x_1' \widehat{\Theta}^k - \sqrt{x_1' \widehat{\Sigma}^k x_1 \cdot \chi_1^2(1 - \alpha)}$ do $x_2' \widehat{\Theta}^k + \sqrt{x_2' \widehat{\Sigma}^k x_2 \cdot \chi_1^2(1 - \alpha)}$.

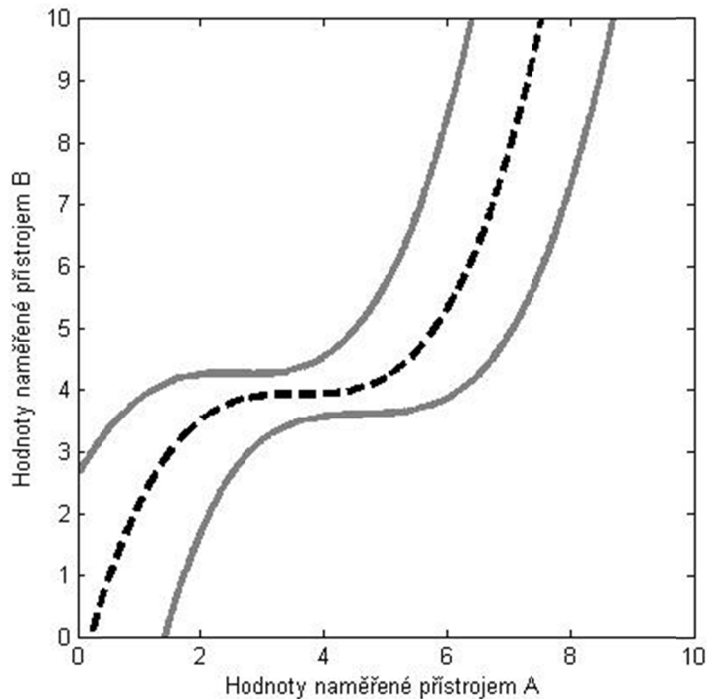
Na obrázku 1.2 je znázorněno rozmezí hodnot, kterých může nabývat přístroj B, pokud na přístroji A naměříme hodnotu x .



Obrázek 1.2 (tečkovaně...pás spolehlivosti, šedě...transformační funkce, čárkovaně ... odhad transformační funkce)

Graf na obrázku 1.2 je vykreslen pomocí funkce n_graf55.m. pro transformační funkci $0.1x^3 - 1.1x^2 + 4x - 0.6$, $\mu = (0, 2.5, 5, 7.5, 10)$, $\sigma_x = 0.5$, $\sigma_y = 0.47$, $\alpha = 0.025$, $\gamma = 0.025$, $x = 2$, $d_x = 1.05$, $h_x = 2.95$, $y \in (1.69, 4.25)$.

Pokud ke každému x z intervalu $(\min(\mu), \max(\mu))$ spočítáme hodnotu $x_1' \widehat{\Theta}^k - \sqrt{x_1' \widehat{\Sigma}^k x_1 \cdot \chi_1^2(1 - \alpha)}$, zakreslíme do grafu a získané body spojíme a zároveň ke každému x z intervalu (d_x, h_x) spočítáme hodnotu $x_2' \widehat{\Theta}^k + \sqrt{x_2' \widehat{\Sigma}^k x_2 \cdot \chi_1^2(1 - \alpha)}$, zakreslíme do grafu a získané body spojíme, získáme pás, který znázorňuje rozmezí hodnot, kterých může nabývat přístroj B, pokud na přístroji A naměříme hodnotu x z intervalu (d_x, h_x) . (Viz Obrázek 1.3.)



Obrázek 1.3 (šedě...rozptění hodnot pro dané x na přístroji B , čárkovaně ... odhad transformační funkce)

Graf je vykreslen pomocí funkce `n_pas.m` pro transformační funkci $0.1x^3 - 1.1x^2 + 4x - 0.6$, $\mu = (0, 2.5, 5, 7.5, 10)$, $\sigma_x = 0.5$, $\sigma_y = 0.47$, $\alpha = 0.025$, $\gamma = 0.025$.

Literatura

- [1] WIMMER, G. *Niektoré matematicko - štatistické metódy kalibrácie*, In ROBUST 2006, Zborník prací 14. Zimní školy, 23-27 ledna 2006, Lhota nad Rohanovem. Praha, 2006. s. 375-386, 12 s. ISBN 80-7015-073-4.
- [2] ANDĚL, J. *Základy matematické statistiky*. 1. vyd. Praha: Matfyzpress, 2005, 358 s. ISBN 8086732401.
- [3] KUBÁČEK, L., a KUBÁČKOVÁ, L. . *Statistika a metrologie*. 1. vyd. Olomouc: Univerzita Palackého, 2000, 307 s. ISBN 8024400936.
- [4] ANDĚL, J. *Matematická statistika*. 2. vyd. Praha: SNTL - nakladatelství technické literatury, Alfa, vydavatelstvo technickej a ekonomickej literatury, 1985, 346 s.

Adresa autora :

Petra Širůčková, Mgr.
 Přírodovědecká fakulta MU
 Kotlářská 2, Brno, 611 37
324037@mail.muni.cz

Analyza dát pomocou štatistického softvéru R

Analysis of data using R statistical software

Patrik Sleziaak

Abstract: The main aim of this review paper is to present some features and results that R a suitable statistical programming language for the analysis and visualization of data.

The paper intends to present an analysis of data from surveys or questionnaires and data from ggplot2 (R package). An introduction to R language is given. This is a free software programming language and environment for statistical computing and graphics. Later in this article a description of association plot is reviewed. This technique is used for analysis two variables of data from questionnaire.

Abstrakt: Hlavným cieľom práce je prezentovať výsledky, ktoré robia z R vhodný štatistický programovací jazyk pre analýzu a vizualizáciu dát.

Práca má v úmysle prezentovať analýzu dát z dotazníka a dát z ggplot2 (R package). Podali sme úvod k jazyku R. Je to voľný programovací jazyk a prostredie pre štatistické výpočty a grafiku. V ďalšej časti práce sme opísali asociačný graf. Táto technika bola použitá pre analýzu dvoch premenných v dátach z dotazníka.

Key words: R language, package ggplot2, association plot

Kľúčové slová: jazyk R, package ggplot2, asociačný graf

JEL classification: C10, F37, C3

1. Introduction

R is an open-source software for statistical computing, which means that R is freely available, so its users are free to see how it is written and improve or extend its possibilities. One of the best things about R, which is the large amount of existing add-ins (so-called packages), which are aimed at solving various tasks in different fields. R is rooted in S, a statistical computing and data visualization language, which originated at Bell Laboratories [1]. The R language provides a rich environment for working with data, especially data to be used for statistical modeling or graphics.

The R language includes [5]:

- an effective data handling and storage facility,
- a large, coherent and integrated collection of functions and tools for data analysis and for general programming tasks,
- graphic capabilities for data analysis,
- simple and effective programming language features, which include conditional statements, loops, possibilities for the definition of user-defined functions, input and output facilities, etc.

The main aim of this review paper is to present some features and results that R a suitable statistical programming language for the analysis and visualization of data. The issue focuses on the analysis and evaluation of questionnaires data and data movies from R package (ggplot2).

2. Data

The input data from a questionnaire or survey, which has two forms, printed and electronic and data movies [7] from package ggplot2 [6] were used in this study. Questionnaire consists of six groups of items – musical and movie preferences, interests, phobia types, personality and demographic items. Data was obtained with snowball sampling (students were requested to ask their families and friends to participate in survey). Samples consist of 500 people from which 301 are female and 199 male.

Data movies consist of movie information and user ratings from IMDB.com. The internet movie database is a website devoted to collecting movie data supplied by studios and fans. It claims to be the biggest movie database on the web and is run by amazon. Data set consist of 28819 rows and 24 variables. Movies were selected for inclusion if they had a known length and had been related by at least on imdb user. The data set contains the following fields: title, year, budget, length, rating, votes [7].

The data were processed in the R software language [4] using the ggplot2 package [6] for the plotting of the results.

3. Analysis of data movies from the ggplot2 (R package)

The first two figures (Figs. 1, and 2) provide information about the analysis of data movies from package ggplot 2.

From first figure (Fig. 1) we can identify that in year of 1930 was the rapid growth of films. We can see jump from 50 films per year to 500 films per year.

It can be related with:

- Creating the color film
- From silent film, sound film was created

We can notice that at the end of 1990 was very large growth of films. It can be related with the fact that video was invented and the film was more accessible for people.

The second figure (Fig. 2) gives information about quality of films depending on the year in which they were filmed. We can see that the best quality of films were at the end of 80 years.

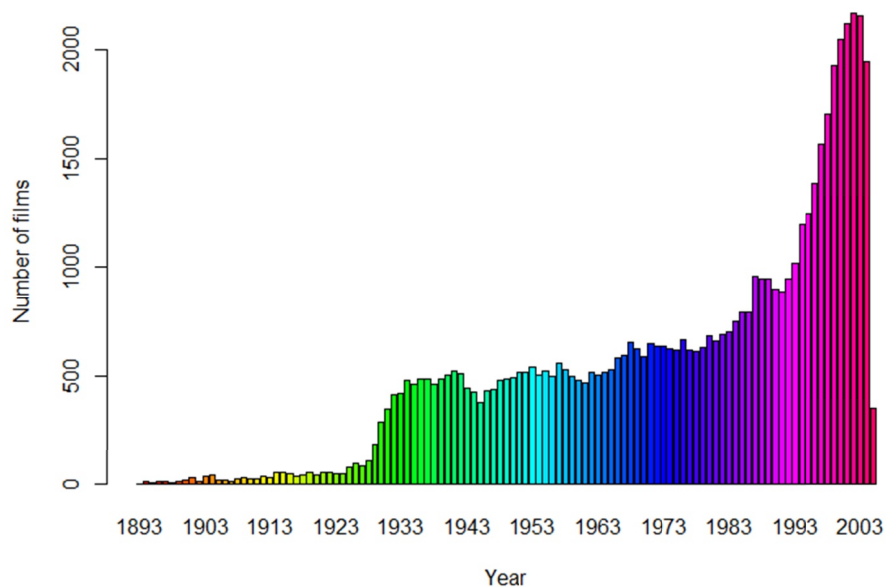


Figure 1: Plotting of number of films depending of the year in which they were filmed

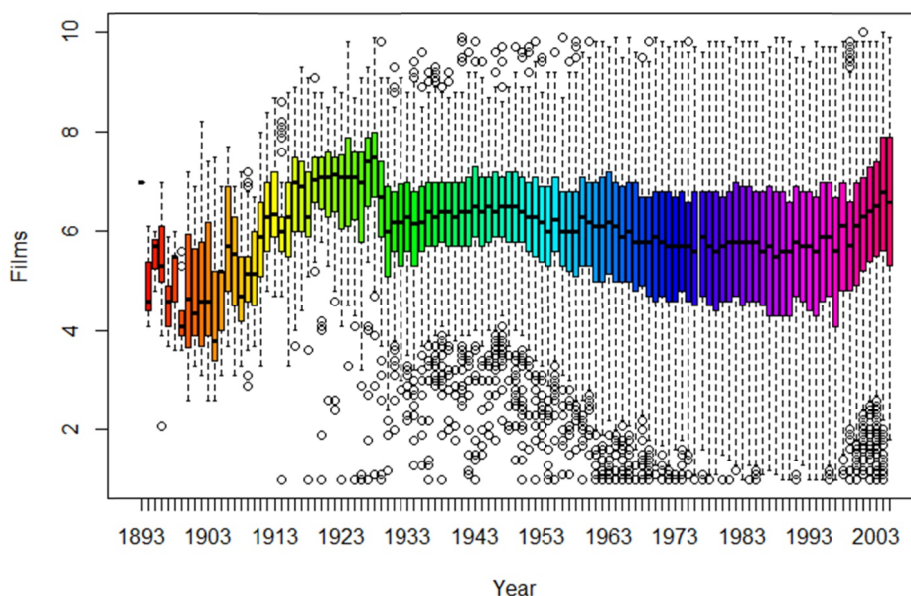


Figure 2: Plotting quality of films depending on the year in which they were filmed

4. Analysis of data from the questionnaire.

a. Descriptive statistics for two of variables

The first figure (Fig. 3) provides information about the analysis of two numerical variables (height and weight of respondents). We can identify extreme values, for example: 90 kilograms and 160 centimeters. The Pearson correlation coefficient has value 0.6119. This means that between the height and weight of respondents is strong (the number is far from 0), direct (the number is positive) linear dependence.

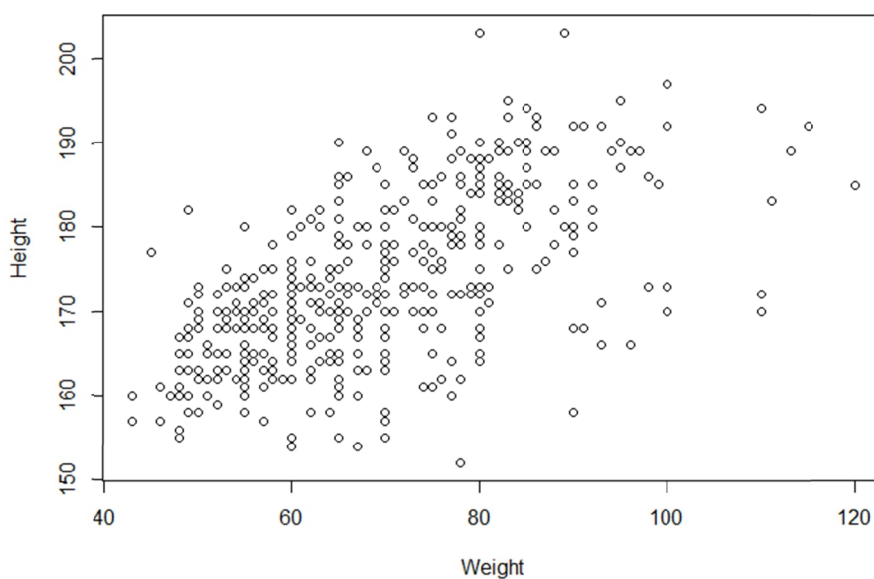


Figure 3: Plotting of two numerical variables (height and weight)

The second figure (Fig. 4) gives information about the analysis of nominal and numerical variables (gender and height of respondents). We can see that males have one extreme value (outliers) and females two extreme values. Half of the females have a height of 168 cm and half of the males have a height of 182 cm.

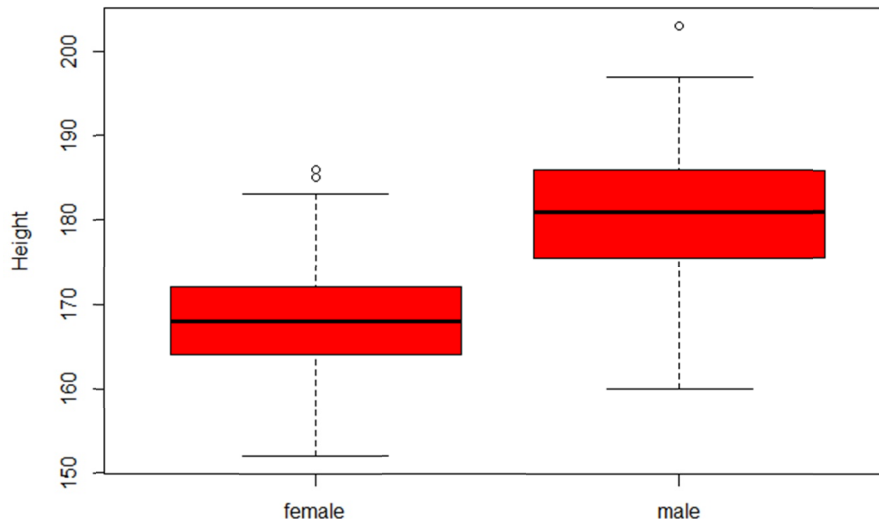


Figure 4: Plotting of nominal and numerical variables (gender and height)

The next figures (Figs. 5, 6, 7) give information about the analysis of two nominal variables (alcohol and gender, smoking and gender, alcohol and smoking).

For the analysis the variables alcohol and gender from the questionnaire were the first used (Fig. 5). We can see that females who drink alcohol is much less than male.

Association plot visualize the table of Pearson residuals: each cell is represented by a rectangle that has (signed) height proportional to the corresponding Pearson residual and width proportional to the square root of the expected counts. Thus, the area is proportional to the raw residuals. In available implementations, the sign of the residual is typically redundantly coded by the rectangle's color and its position relative to the baseline [3].

The black rectangles above the dashed lines indicate observed frequencies exceeding expected frequencies, red rectangles below the dashed lines indicate observed frequencies smaller than expected frequencies [2].

Figure 6 shows the association plot for the variables smoking and gender. We can see that females who smoke is more than males, but on the other hand females who tried smoke is much less than males.

The last figure (Fig. 7) shows the association plot for the variables alcohol and smoking. We can see strong dependence. People who smoke and simultaneous drink alcohol is more than would be expect. On the contrary, people who do not smoke and simultaneous drink alcohol is much less.

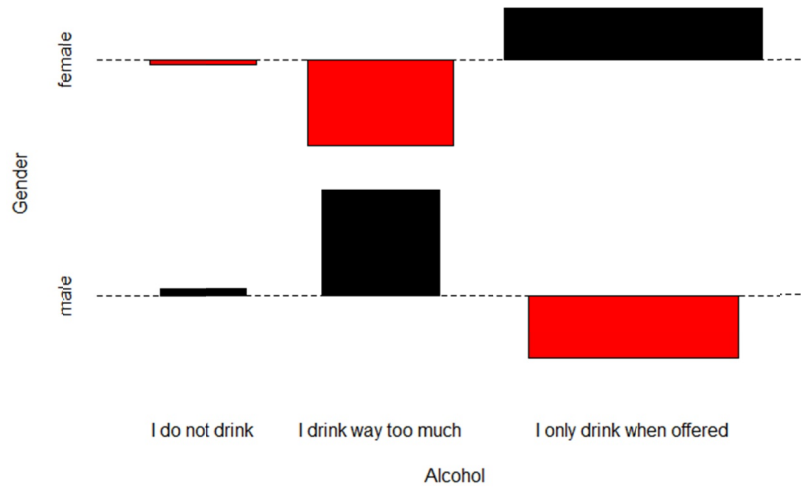


Figure 5: Association plot of two nominal variables (alcohol and gender)



Figure 6: Association plot of two nominal variables (smoking and gender)



Figure 7: Association plot of two nominal variables (alcohol and smoking)

5. Conclusion

R provides a general language for interactive computations, supported by techniques for data organization, graphics, numerical computations, model-fitting, simulation, and many other tasks.

In this paper, the application of R software to questionnaires data and data from package `ggplot2` is demonstrated. In the first part, R language is defined. In the practical part, this software is used for the analysis and visualization of movies dataset from package `ggplot2`. Later in this article a description of association plot is reviewed. This technique is used for analysis of two nominal variables from questionnaire.

The result showed that the R statistical language is appropriate for analysis of data and also useful tool for visualization of data.

References

- [1] CHAMBERS, J.M., HASTIE, T.J. (1992): Statistical Models in S. Wadsworth & Brooks/Cole, Pacific Grove, California.
- [2] FRIENDLY, M. (2000): Visualizing Categorical Data. SAS Institute, Cary, NC.
- [3] MEYER, D., ZEILEIS, A., HORNIK, K. (2003): Visualizing Independence Using Extended Association Plots. Proceedings of DSC, 9 pp.
- [4] R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.r-project.org/>
- [5] VENABLES, W.N., SMITH D.M.: An Introduction to R. [http:// www.cran.r-project.org/doc/manuals/R-intro.pdf](http://www.cran.r-project.org/doc/manuals/R-intro.pdf)
- [6] WICKHAM, H. `ggplot2` (2009): elegant graphics for data analysis. Springer New York. <http://cran.r-project.org/web/packages/ggplot2/index.html>
- [7] WICKHAM, H.: Movies dataset. <http://had.co.nz/data/movies/>

Patrik Sleziak, Ing.
Department of Land and Water Resources
Management (Slovak University of Technology)
Radlinského 11, 813 68 Bratislava
patrik.sleziak@stuba.sk

Rodové rozdiely v prevalencii špecifických fóbií

Gender differences in the prevalence of specific phobias

Patrik Sleziak, Miroslav Sabo

Abstract: This paper deals with the comparison and testing group differences in the prevalence of specific phobias. We show how social data (data from surveys or questionnaires) can be tested. As a result of testing items, significant differences were found. Later in this paper a description of a factor analysis (FA) is reviewed. The FA part of the paper mainly addresses one area: to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. Finally, some of the comparisons of the results are discussed. The results were processed in the R statistical programming language.

Abstrakt: Predložená práca sa zaoberá porovnaním a testovaním skupinových rozdielov v prevalencii špecifických fóbií. Ukážeme si ako môžu byť sociálne dáta (dáta z prieskumov a dotazníkov) testované. Ako výsledok testovania prvkov, boli zistené výrazné rozdiely. V ďalšej časti práce sme opísali faktorovú analýzu. FA je predovšetkým zameraná na jednu oblasť: opísať variabilitu medzi pozorovanými, korelovanými premennými, z hľadiska menšieho počtu nepozorovaných premenných, ktoré sa nazývajú faktory. Na záver sme jednotlivé výsledky porovnali. Výsledky boli spracované v štatistickom programovacom jazyku R.

Key words: gender differences, phobias, factor analysis

Kľúčové slová: rodové rozdiely, fóbie, faktorová analýza

JEL classification: C10, C12, C45

1. Introduction

This article was inspired by article Clustering data from surveys, where as a result of clustering items, three clusters of phobia types were found: phobias from animals (green), phobias from natural phenomena (red) and human psyche phobias (blue) (Figure 1) [2].

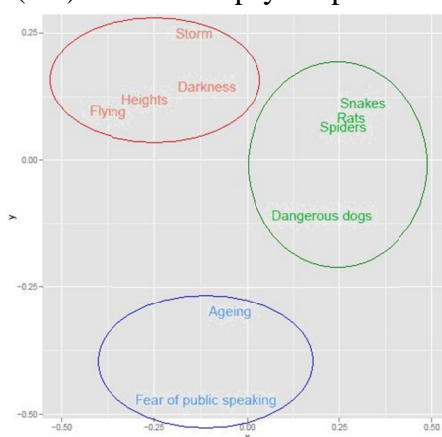


Figure 1: 2D MDS plot of 10 phobia types with cluster labels denoted by different colours

According to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV), the most phobias are classified into three categories:

1. Social phobia – fear of other people or social situations such as performance anxiety or fears embarrassment by scrutiny of others [7].

2. Specific phobias – fear of a single specific panic trigger such as spiders, snakes, dogs, heights, water, catching, flying a specific illness, etc. [8].
3. Agoraphobia – a generalized fear of leaving home or a small familiar ‘safe’ area, and of possible panic attacks that might follow [9].

In this article we show testing group differences in order to gain meaningful results from survey data. Later in this article, for describing variability among observed, correlated variables in terms of a potentially lower number of unobserved variables, the factor analysis technique is used.

2. Data description

Data used in this article comes from a questionnaire survey, which has two forms, printed and electronic. Questionnaire consists of six groups of items – musical and movie preferences, interests, phobia types, personality and demographic items. Data was obtained with snowball sampling (students were requested to ask their families and friends to participate in survey). We have worked with phobia types (snakes, dangerous dogs, spiders, rats, fear of public speaking, heights, ageing, darkness, storm, flying) and demographic items (gender, city and town or village). Samples consist of 500 people from which 301 are female and 199 male.

The data were processed in the R software language [3] using the psych package [4] for the factor analysis. The plotting of the results were processed by package ggplot2 [6].

3. Testing group differences

Gender differences in specific phobias and phobia types in city and in town or village were examined using Mann Whitney test. Tab. 1 presents comparison of phobia types in females and males (Figure 2).

It is important to note that female have larger phobia than male of all observed items. We can see that female have the largest phobia of snakes and the least of flying. On the other hand male have the largest phobia of dangerous dogs and the least of storm. According to Tab. 2, gender (female and male) have the largest phobia of dangerous dogs and the least of storm (Figure 2).

As the second example we compared phobia types in city and in town or village (Tab. 3, Figure 3). We can see that people who live in city have largest phobia of snake and the least of flying. On the contrary people who live in town or village have largest phobia of dangerous dogs and the least of flying. According to Tab. 4: Summary comparison of phobia types in city and in town or village, the largest phobia have people of dangerous dogs and the least of storm (Figure 3).

Tab. 11: Comparison of phobia types in females and males

Group: female (n=301)		Group: male (n=199)	
Phobia types	Mean	Phobia types	Mean
Snakes	3,49	Dangerous dogs	2,76
Dangerous dogs	3,41	Snakes	2,62
Spiders	3,11	Fear of public speaking	2,57
Rats	3,00	Heights	2,54
Fear of public speaking	2,85	Ageing	2,29

Heights	2,69	Spiders	2,05
Ageing	2,68	Rats	1,97
Darkness	2,50	Flying	1,88
Storm	2,30	Darkness	1,78
Flying	2,20	Storm	1,60

Tab. 2: Summary comparison of phobia types in gender (female and male)

Summary data (phobia types ~ gender)			
Phobia types	N	Mean	p-value
Dangerous dogs	500	3,15	<0,001
Snakes	500	3,14	<0,001
Fear of public speaking	500	2,73	<0,05
Spiders	500	2,68	<0,001
Heights	500	2,63	>0,05
Rats	500	2,59	<0,001
Ageing	500	2,52	<0,001
Darkness	500	2,21	<0,001
Flying	500	2,07	<0,01
Storm	500	2,02	<0,001

Tab. 3: Comparison of phobia types in city and in town or village

Group: city (n=329)		Group: town or village (n=171)	
Phobia types	Mean	Phobia types	Mean
Snakes	3,06	Dangerous dogs	3,40
Dangerous dogs	3,02	Snakes	3,32
Spiders	2,77	Heights	2,86
Fear of public speaking	2,77	Rats	2,81
Ageing	2,54	Fear of public speaking	2,67
Heights	2,51	Spiders	2,52
Rats	2,47	Ageing	2,49
Darkness	2,21	Darkness	2,22
Flying	2,01	Storm	2,22
Storm	1,91	Flying	2,20

Tab. 4: Summary comparison of phobia types in city and in town or village

Summary data (phobia types ~ city/town or village)			
Phobia types	N	Mean	p-value
Dangerous dogs	500	3,15	<0,01
Snakes	500	3,14	>0,05
Fear of public speaking	500	2,73	>0,05

Spiders	500	2,68	>0,05
Heights	500	2,63	<0,01
Rats	500	2,59	<0,05
Ageing	500	2,52	>0,05
Darkness	500	2,21	>0,05
Flying	500	2,07	>0,05
Storm	500	2,02	<0,01

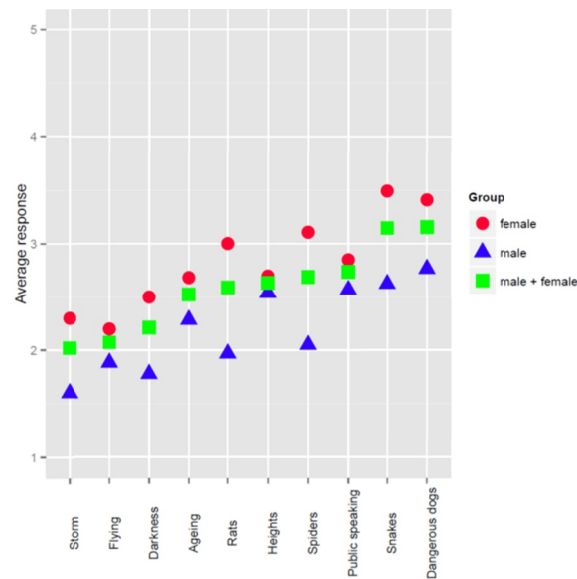


Figure 2: Phobia types in genders denoted by different colours and symbols

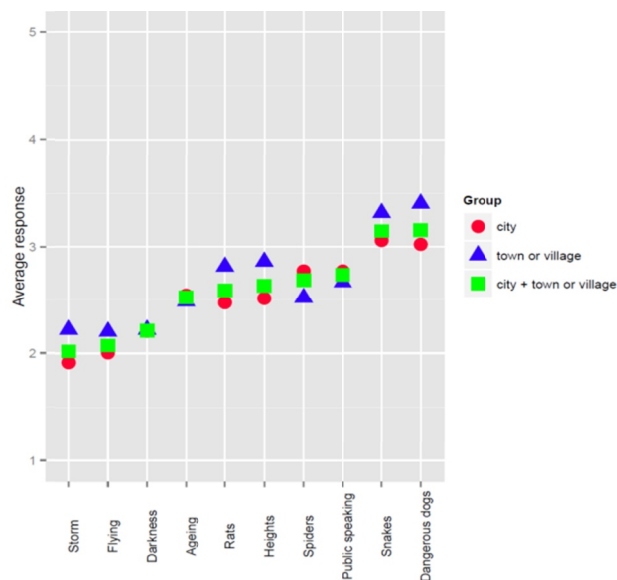


Figure 3: Phobia types in people who live in city and in town or village denoted by different colours and symbols

4. Using factor analysis

Technique of factor analysis belongs to a multivariate statistical method. This technique is used for the analysis of hidden relationships between variables and reducing the dimension variables [5].

FA is used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. This means, for example, it is possible that variations in four observed variables mainly reflect the variations in two unobserved variables. The observed are modelled as linear combinations of the potential factors. The information gained about the interdependencies between observed variables can be used later to reduce the set variables in dataset [5].

Fa is related to principal components analysis (PCA), but the two are not identical. Variable models, including FA use regression modelling techniques to test hypotheses producing error terms, while PCA is a descriptive statistical technique [1].

The fourth figure (Figure 4) provides information about the factor analysis of phobias. It is important to notice, that using this technique we reduced the variables into four factors. Purple bullet represents a positive correlation.

The first factor most positive correlate with spiders, snakes, rats. This factor could be called as phobia from animals. The second factor most positive correlate with storm, darkness. This means that this factor we can designate as phobia from natural phenomena. The third factor most positive correlate with dangerous dogs. This factor could be called as phobia from dangerous animals. The fourth factor most positive correlate with flying, heights. This means that this factor we can designate as phobia from heights.

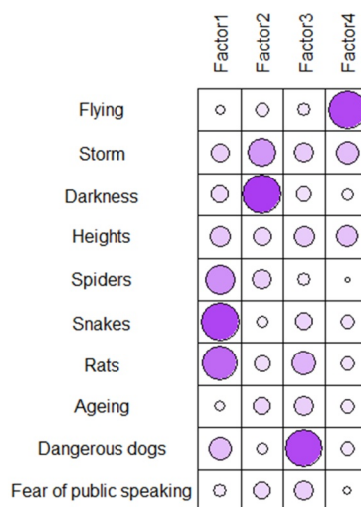


Figure 4: Factor analysis of phobias

5. Conclusion

Frederikson, Annas, Fischer and Wik in article showed that more female than male have animal and situational phobias. Phobia of animals, mainly of spiders, were higher in younger than older individuals. On the contrary phobia of lightning, heights and flying were lower in younger than older. Female reported significantly more intense phobia than male for all items.

In this paper, the testing group differences in specific phobias is demonstrated. In the first part, specific phobias are defined. In the practical part, the Mann Whitney test is used for determine gender differences in specific phobias and phobia types in city and in town or village. As a result of testing items, significant differences were found. Later in this article a

description of a factor analysis is reviewed. This technique is used for analysis of hidden relationships between variables and reducing the dimension variables. Using this technique we reduced the variables into four factors.

The result showed that the R statistical language is appropriate for analysis a questionnaire data and the factor analysis technique is also useful tool for describing variability among observed, correlated variables in terms of a potentially lower number of unobserved variables. On the other hand, all results that we obtained are well interpretable and have potential value for other researchers interested in similar surveys.

References

- [1] BROWN, J.D. (2009): Principal components analysis and exploratory factor analysis – Definitions, differences and choices. JALT Testing & Evaluation SIG Newsletter, 13 (1), p.26-30.
- [2] SABO, M. (2013): Clustering data from surveys. In Advances in architectural, civil and environmental engineering: 23rd Annual PhD Student Conference. Bratislava SR, (accepted).
- [3] R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.r-project.org/>
- [4] REVELLE, W. (2011): Psych: Procedures for psychological, psychometric, and personality research. URL <http://cran.r-project.org/web/packages/psych/index.html>
- [5] RUSCIO, J., ROCHE, B. (2012): Determining the number of factors to retain in an exploratory factor analysis using comparison data of know factorial structure. Psychological Assessment 24, p. 282-292.
- [6] WICKHAM, H. ggplot2: Elegant graphics for data analysis. Springer New York, 2009. <http://cran.r-project.org/web/packages/ggplot2/index.html>
- [7] <http://en.wikipedia.org/wiki/Phobia> [2014-10-9].
- [8] TURK, C. L. and HEIMBERG, R. G. (1998): An investigation of gender differences in social phobia: Journal of Anxiety Disorders, 12, 209-223.
- [9] FREDERIKSON, M., ANNAS, P., HÅKAN, F. and WIK, G. (1996): Gender and age differences in the prevalence of specific fears and phobias. Behav. Res. Ther., 34, 33-39.

Patrik Sleziak, Ing.
Department of Land and Water Resources
Management (Slovak University of Technology)
Radlinského 11, 813 68 Bratislava
patrik.sleziak@stuba.sk

Miroslav Sabo, Mgr.
Department of Mathematics and Constructive
Geometry (Slovak University of Technology)
Radlinského 11, 813 68 Bratislava
sabo@math.sk

Vývoj miery rizika monetárnej chudoby v regiónoch Slovenska

Trend of monetary poverty measures in Slovak regions

Iveta Stankovičová, Róbert Vlačuha

Abstract: The EU statistics on income and living conditions (EU SILC) is the reference source for comparative statistics on income distribution and social inclusion in the European Union (EU). In this paper we used Slovak EU SILC data for the years 2009 to 2013. We calculated a poverty line as 60% of the national median equivalised disposable income in Slovakia, but also in different regions of Slovakia (in 8 counties by NUTS 3). Aim of this paper is to analyse the risk of poverty calculated on the basis of national and regional poverty lines in 8 regions of Slovakia.

Abstrakt: Výberové zisťovanie o príjmoch a životných podmienkach domácností EU SILC je zdrojovou základňou pre porovnávanie distribúcie príjmov a sociálnej inklúzie na úrovni EÚ. V príspevku sme použili slovenské údaje EU SILC za roky 2009 až 2013. Vypočítali sme hranice rizika chudoby ako 60% mediánu ekvivalentného disponibilného príjmu domácností v SR, ale aj v jednotlivých regiónoch SR (v 8 krajoch podľa NUTS 3). Cieľom príspevku je analýza vývoja miery rizika chudoby v regiónoch SR vypočítanej na základe národnej a regionálnej hranice chudoby.

Key words: monetary poverty, at-risk-of-poverty threshold, at-risk-of-poverty rate, EU SILC database

Kľúčové slová: monetárna chudoba, hranica rizika chudoby, miera rizika chudoby, EU SILC databáza

JEL classification: O15, C46, I32

1. Úvod

Chudoba patrí aj v dvadsiatom prvom storočí medzi vážne problémy, ktorými je potrebné sa zaoberať. Svetová hospodárska kríza, ktorá je sprevádzaná vysokou mierou nezamestnanosti a podnikateľskej neistoty, tento jav ešte prehĺbila. Predstavitelia krajín EÚ musia čeliť rastu nezamestnanosti a súčasne aj rastu monetárnej chudoby, ktorá je alarmujúca vo vybraných vekových a sociálnych skupinách a tiež vo vybraných regiónoch.

Skúmanie chudoby a sociálneho vylúčenia je potrebné zasadiť do kontextu špecifického hospodárskeho a sociálneho vývoja na konkrétnom území. Zmenený medzinárodný a politický kontext v Európe priniesol pre Slovensko nové výzvy aj v skúmaní chudoby. Stali sme sa súčasťou spoločenstva, v ktorom sú otázky nepriaznivej životnej situácie a nerovnakých životných šancí v centre pozornosti výskumu i politického rozhodovania.

Chudoba patrí medzi mnohodimenzionálne javy. Tak ako neexistuje jediná, alebo správna definícia chudoby, neexistuje ani jediný všeobecne prijatý spôsob jej merania. Každé meranie chudoby znamená, že meriame len určitý koncept chudoby. Z údajov zisťovania EU SILC sa počíta ekvivalentný disponibilný príjem domácností a indikátory chudoby, vypočítané na základe takéhoto príjmu, odrážajú nasledovné koncepty chudoby:

- relatívny koncept – relatívna miera chudoby vychádza z hodnotenia (porovnania) príjmov k určitej strednej hodnote, v našom prípade k národnej hranici chudoby;
- nepriamy koncept – meranie chudoby sa realizuje prostredníctvom disponibilného príjmu domácností, pričom sa neberú do úvahy výdavky domácností;

- objektívny koncept – hranica chudoby sa nestanovuje subjektívne na základe skúseností a odpovedí jednotlivých členov domácností, ale sa stanovuje na základe nezávislých a objektívne vybraných kritérií.

Na základe týchto prístupov je pre medzinárodné porovnávanie chudoby stanovená vždy národná hranica chudoby (v angličtine: *poverty line* alebo *at-risk-of-poverty threshold*), ktorá predstavuje 60% mediánu ekvivalentného disponibilného príjmu domácnosti v danej krajine.

Je veľa článkov publikovaných v zborníkoch z konferencií, vo vedeckých a odborných časopisoch, ktoré sú venované analýzám mier chudoby v krajinách a regiónoch EÚ a ich vývoju. Spomeňme niektoré publikácie z posledných rokov vybraných slovenských a českých autorov: práce Bartošovej a Želinského (2013), Löstera a Langhamrovej (2012), Pacákovej, Lindu a Sipkovej (2012), Sipkovej a Sipka (2012), Stankovičovej, Vlačuhu a Ivančíkovej (2013), Tartaľovej a Želinského (2012), Želinského (2012), Želinského a Stankovičovej (2012).

2. Hranice monetárnej chudoby v rokoch 2009 až 2013

Z údajov EU SILC sme vypočítali hranice monetárnej chudoby za Slovensko v rokoch 2009 až 2013 ako 60% mediánu ekvivalentného disponibilného príjmu domácnosti. Hodnota hranice rizika chudoby je v tabuľke 1 uvedená v EUR pre ekvivalentnú domácnosť na rok a na mesiac. Pod pojmom „ekvivalentná domácnosť“ treba rozumieť 1-člennú domácnosť (t.j. domácnosť dospelého jednotlivca, ktorý má koeficient 1). Pre viacčlennú domácnosť treba túto hranicu prepočítať podľa prepočítacích koeficientov tzv. modifikovanej OECD škály. Táto škála stanovuje pre ďalších členov domácností nasledovné koeficienty:

- každému druhému a ďalšiemu dospelému členovi domácnosti a 14-ročným a starším osobám koeficient 0,5;
- každému dieťaťu mladšiemu ako 14 rokov koeficient 0,3.

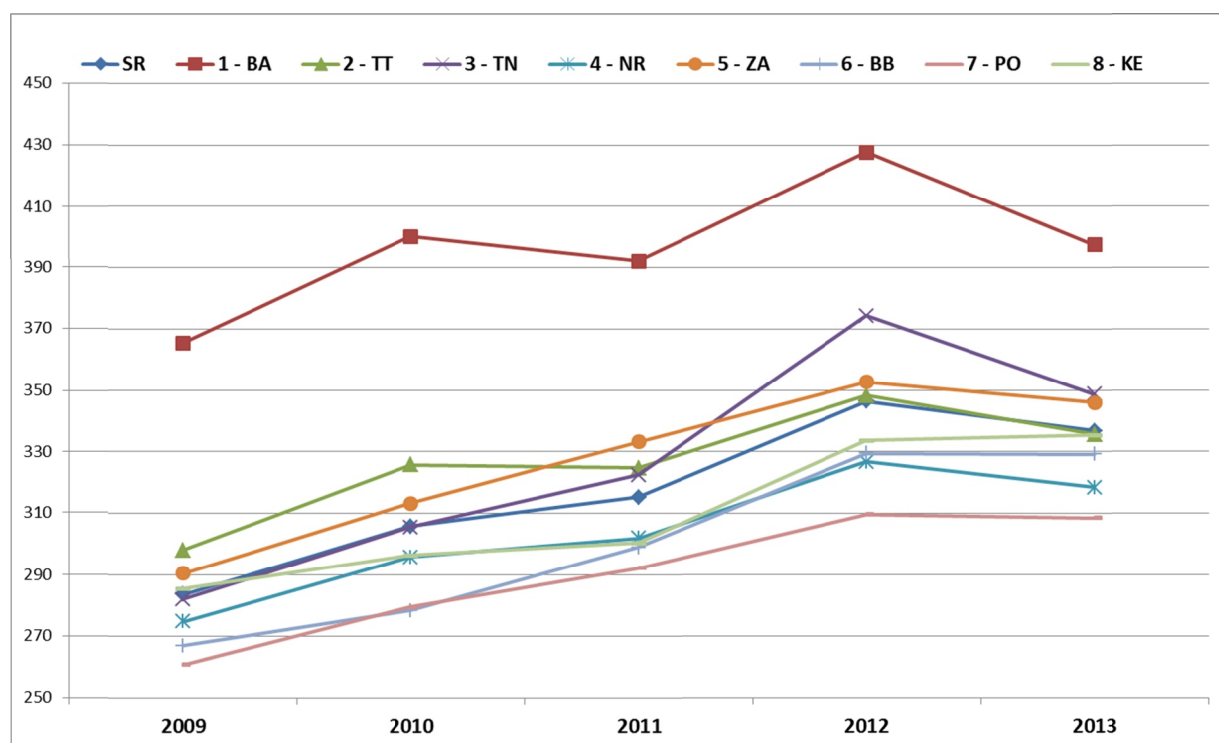
Napríklad pre 2-člennú domácnosť s dvoma dospelými členmi by bola vypočítaná mesačná hranica chudoby v SR v roku 2009 stanovená na 426 EUR ($1 \times 284 + 0,5 \times 284 = 426$).

Čísla v tabuľke 1 ukazujú vývoj národnej hranice chudoby v SR, ale aj vývoj regionálnych hraníc chudoby za 8 krajov SR podľa klasifikácie NUTS 3. Krajské hranice chudoby sú vypočítané ako 60% mediánu ekvivalentného disponibilného príjmu domácnosti v danom kraji. Porovnanie týchto hraníc poukazuje na regionálne disparity v príjmoch. Z obrázku 1 je zrejmé, že v Bratislavskom kraji sú najvyššie disponibilné príjmy, čiže aj hranice monetárnej chudoby dosahujú najvyššie hodnoty vo všetkých sledovaných rokoch 2009 až 2013. Nad národným priemerom sú aj hranice chudoby v ďalších troch krajoch: Trnavskom, Trenčianskom a Žilinskom. V štyroch krajoch sú hranice chudoby pod národným priemerom: Košickom, Banskobystrickom, Nitrianskom a Prešovskom. Celkovo najnižšie hodnoty krajských hraníc chudoby, čiže aj ekvivalentných disponibilných príjmov, boli v Prešovskom kraji.

Vývoj hraníc chudoby nie je lineárny (Obr. 1). Hlavne v roku 2013 zaznamenávame pokles hraníc rizika chudoby v porovnaní s predchádzajúcim rokom vo všetkých krajoch a aj za celé Slovensko (SR: -9 EUR/mesačne). Znamená to, že medián disponibilných príjmov sa znížil, čo predstavuje negatívny jav vo vývoji príjmov obyvateľstva SR. Výrazný nárast hranice rizika chudoby v roku 2012 nastal hlavne v Bratislavskom a Trenčianskom kraji (BA: 36 EUR/mesačne, TN: 52 EUR/mesačne), ale v oboch týchto krajoch zase v roku 2013 hranica chudoby výrazne klesla (BA: -30 EUR/mesačne, TN: -26 EUR/mesačne).

Tab. 12: Hranice rizika chudoby podľa regiónov Slovenska v rokoch 2009 – 2013 (v EUR) – ročné a mesačné

Región	SILC 2009		SILC 2010		SILC 2011		SILC 2012		SILC 2013	
	ROK	MESIAC	ROK	MESIAC	ROK	MESIAC	ROK	MESIAC	ROK	MESIAC
SR	3 403	284	3 670	306	3 784	315	4 156	346	4 042	337
Bratislavský	4 386	365	4 801	400	4 705	392	5 132	428	4 769	397
Trnavský	3 578	298	3 908	326	3 898	325	4 180	348	4 029	336
Trenčiansky	3 385	282	3 666	306	3 869	322	4 492	374	4 184	349
Nitriansky	3 296	275	3 549	296	3 623	302	3 921	327	3 821	318
Žilinský	3 487	291	3 759	313	3 999	333	4 233	353	4 152	346
Banskobystrický	3 201	267	3 340	278	3 587	299	3 954	330	3 950	329
Prešovský	3 128	261	3 354	280	3 507	292	3 717	310	3 702	309
Košický	3 426	286	3 556	296	3 604	300	4 004	334	4 025	335



Obr. 18: Vývoj mesačných krajských hraníc chudoby v porovnaní s národnou mesačnou hranicou chudoby SR (roky 2009 – 2013, v EUR)

3. Vývoj miery rizika monetárnej chudoby na Slovensku v rokoch 2009 až 2013

Chudobu ovplyvňuje množstvo faktorov. Za najvýznamnejšie faktory môžeme vo všeobecnosti považovať regionálne hľadisko, vek a typ domácnosti, pohlavie, pracovnú intenzitu domácnosti, či najčastejší status ekonomickej aktivity členov domácnosti. V tejto časti príspevku sa zameriame na analýzu vývoja miery monetárnej chudoby na Slovensku podľa jednotlivých krajov (regiónov).

Z údajov EU SILC sme pre Slovensko a jeho kraje vypočítali mieru rizika chudoby na základe jednotnej národnej hranice chudoby, ale aj na základe rozdielnych krajských hraníc chudoby.

Miera rizika chudoby po sociálnych transferoch vo všeobecnosti predstavuje podiel osôb (v percentách) v celkovej populácii, ktorých ekvivalentný disponibilný príjem sa nachádza pod hranicou rizika chudoby (t.j. 60 % mediánu národného ekvivalentného disponibilného príjmu). Vypočítané krajské hranice chudoby v tomto článku teda predstavujú podiel osôb

(v percentách) v celkovej populácii kraja (regiónu), ktorých ekvivalentný disponibilný príjem sa nachádza pod hranicou rizika chudoby kraja (t.j. 60 % mediánu krajského (regionálneho) ekvivalentného disponibilného príjmu).

Miera rizika chudoby po sociálnych transferoch na Slovensku v sledovanom období rokov 2009 až 2013 rástla. Vrchol dosiahla v roku 2012 (13,2%). V roku 2013 nepatrne klesla na 12,8%, čiže o 0,4 percentuálneho bodu.

V miere rizika chudoby v členení podľa krajov sa výrazne prejavili regionálne disparity v sledovanom období rokov 2009 až 2013 (Tab. 2, Obr. 2 a Obr. 4). Najnižšia miera rizika chudoby vypočítaná na základe národnej hranice chudoby bola zaznamenaná u obyvateľov Bratislavského kraja (5,1% v roku 2010). Pod celoslovenským priemerom sa nachádzali hlavne Trnavský, Trenčiansky a Žilinský kraj. Miera rizika chudoby v Košickom kraji v sledovanom období oscilovala okolo celoslovenského priemeru. Viac boli rizikom chudoby ohrození obyvatelia Nitrianskeho a Banskobystrického kraja a najvyššia miera rizika chudoby bola zaznamenaná v Prešovskom kraji (až 20,2% v roku 2011).

Ak vypočítame miery rizika chudoby na základe krajských hraníc, tak sa situácia zmení (Tab. 2, Obr. 3 a Obr. 5). Najnižšie miery rizika sú potom v Trnavskom alebo Trenčianskom kraji a najvyššie sú opäť v Prešovskom kraji. Bratislavský kraj je na tom zvyčajne horšie ako je celoslovenská miera, čiže dosahuje vyššiu mieru rizika chudoby. Znamená to, že v tomto kraji je vysoká miera diferenciácie príjmov, čiže sú veľké rozdiely medzi príjmami „bohatých“ a „chudobných“ rodín. Rozdelenie príjmov je výrazne zošíkmené tak, že jeho vrchol je posunutý viac k nižším príjmom a tak sa pod krajskú hranicu chudoby dostáva vyšší podiel domácností.

Naopak v Prešovskom kraji sú miery rizika chudoby vypočítané na základe krajskej hranice nižšie ako na základe národnej hranice chudoby. Znamená to, že v tomto kraji je nižšia diferenciácia príjmov, čiže nie sú až také veľké rozdiely medzi príjmami „bohatých“ a „chudobných“ rodín. Rozdelenie príjmov nie je až tak výrazne zošíkmené. Ale aj pri použití krajskej hranice chudoby pre výpočet miery rizika chudoby je na tom Prešovský kraj výrazne horšie ako ostatné kraje.

4. Záver

Predložený príspevok mal za cieľ poskytnúť prehľad o vývoji hraníc a mier rizika monetárnej chudoby v krajoch (regiónoch) Slovenska. Pri analýze sme vychádzali z dát výberového štatistického zisťovania EU SILC za roky 2009 až 2013 pre SR. Na výpočty indikátorov chudoby sme použili softvér SAS a na grafické zobrazenie výsledkov MS Excel.

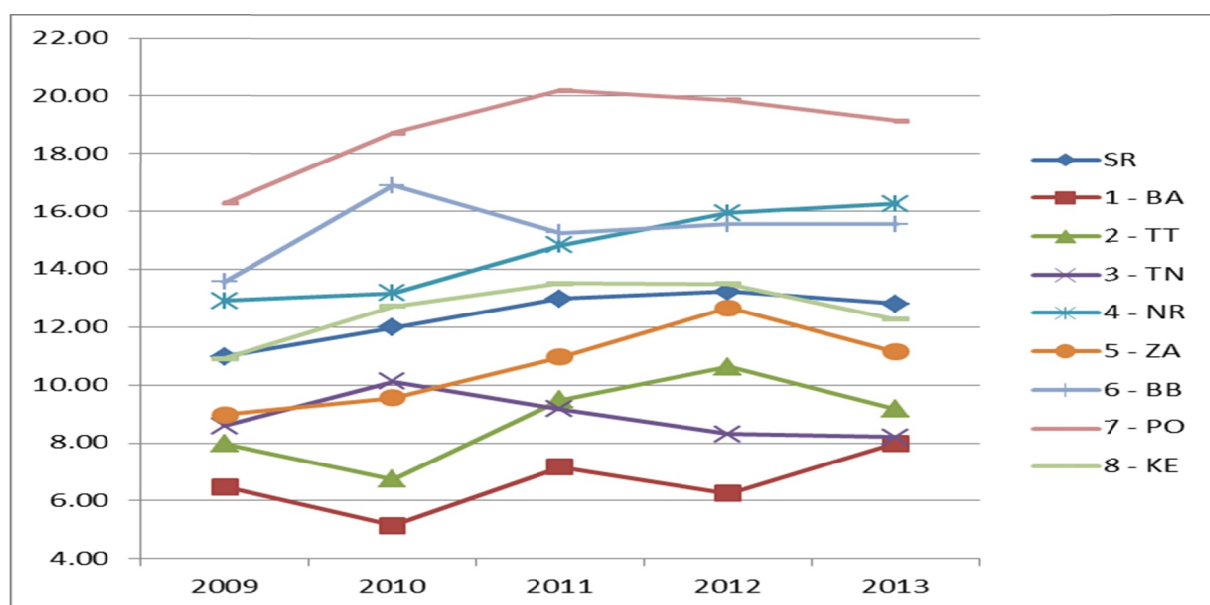
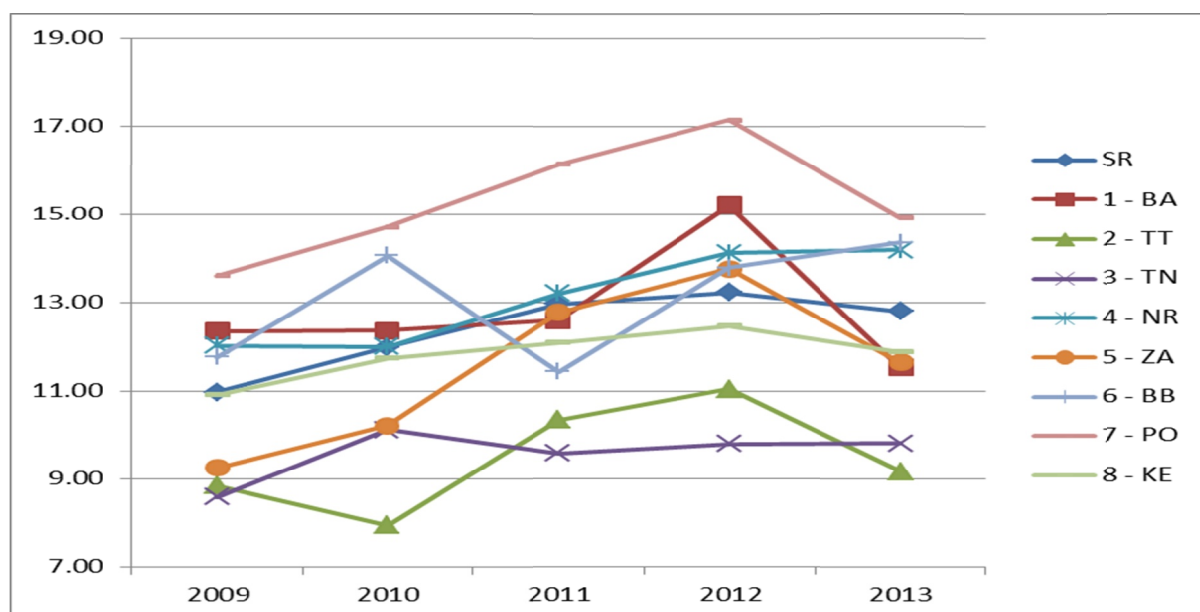
Z analýzy vyplynulo, že národné aj krajské hranice chudoby v rokoch 2009 až 2012 rástli. V roku 2013 dochádza k poklesu na národnej a aj na krajskej úrovni. Znamená to zníženie mediánu príjmov, čo môžeme vyhodnotiť ako negatívny vývoj príjmov domácností v SR.

V sledovanom období rástla aj miera rizika chudoby na Slovensku z 11,0% v roku 2009 na 13,2% v roku 2012. V roku 2013 miera rizika chudoby nepatrne poklesla (na 12,8%), čiže zaznamenali sme pokles o 0,4 percentuálneho bodu.

Analýza potvrdila veľké regionálne disparity na Slovensku v oblasti monetárnej chudoby. Najnižšie miery rizika chudoby vypočítané na základe národnej hranice za celé sledované obdobie boli v Bratislavskom kraji (5,1% – 8,0%) a najvyššie zase v Prešovskom kraji (16,3% – 20,2%). Rozdiel v hodnotách miery rizika chudoby bol v niektorých rokoch medzi Bratislavským a Prešovským krajom viac ako 3-násobný (roky 2010 a 2012). Nad celoslovenským priemerom sú aj ďalšie 2 kraje (NR a BB). Košický kraj osciluje okolo priemeru SR. Pod celoslovenským priemerom sa nachádzali aj miery rizika chudoby v Trnavskom, Trenčianskom a Žilinskom kraji.

Tab. 13: Miery rizika chudoby podľa regiónov Slovenska v rokoch 2009 – 2013 (v %)

Región		SILC 2009		SILC 2010		SILC 2011		SILC 2012		SILC 2013	
		SR	KRAJ	SR	KRAJ	SR	KRAJ	SR	KRAJ	SR	KRAJ
SR	SR	11,0	11,0	12,0	12,0	13,0	13,0	13,2	13,2	12,8	12,8
Bratislavský	1 - BA	6,5	12,4	5,1	12,4	7,2	12,6	6,3	15,2	8,0	11,5
Trnavský	2 - TT	8,0	8,9	6,7	7,9	9,5	10,3	10,6	11,0	9,2	9,2
Trenčiansky	3 - TN	8,6	8,6	10,1	10,1	9,2	9,6	8,3	9,8	8,2	9,8
Nitriansky	4 - NR	12,9	12,0	13,2	12,0	14,8	13,2	15,9	14,1	16,3	14,2
Žilinský	5 - ZA	9,0	9,3	9,6	10,2	11,0	12,8	12,7	13,8	11,1	11,7
Banskobystrický	6 - BB	13,6	11,8	16,9	14,1	15,3	11,4	15,6	13,8	15,6	14,4
Prešovský	7 - PO	16,3	13,6	18,7	14,7	20,2	16,1	19,9	17,1	19,2	14,9
Košický	8 - KE	10,9	10,9	12,7	11,7	13,5	12,1	13,5	12,5	12,3	11,9

**Obr. 19: Vývoj miery rizika chudoby podľa národných hraníc chudoby v regiónoch SR (2009 – 2013)****Obr. 20: Vývoj miery rizika chudoby podľa krajských hraníc chudoby v regiónoch SR (2009 – 2013)**

Literatúra

- BARTOŠOVÁ, J. – ŽELINSKÝ, T. 2013. The extent of poverty in the Czech and Slovak Republics 15 years after the split. In: *Post-Communist Economies*, 25(1), s. 119-131. Dostupné: <http://dx.doi.org/10.1080/14631377.2013.756704>
- BARTOŠOVÁ, J. 2013. *Finanční potenciál domácností – kvantitativní metody a analýzy*. Praha: Professional Publishing.
- LÖSTER, T. – LANGHAMROVÁ, J. 2012. Disparities between Regions of the Czech Republic for Non_business Aspects of Labour Market. In: Löster, T., Pavelka, T. (eds.): *The 6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 689–702.
- PACÁKOVÁ, V. – LINDA, B. – SIPKOVÁ, Ľ. 2012. Distribution and Factors of the Highest Wages in the Slovak Republic. In: *Ekonomický časopis*, Vol. 60, No. 9. pp. 918–934.
- SIPKOVÁ, Ľ. – SIPKO, J. 2012. Analysis of Income Inequality of Employees in the Slovak Republic. In: Löster, T., Pavelka, T. (eds.): *6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1032–1042.
- STANKOVIČOVÁ, I. – VLAČUHA, R. – IVANČÍKOVÁ, L. 2013. Trend Analysis of Monetary Poverty Measures in the Slovak and Czech Republic. In: Löster, T., Pavelka, T. (eds.): *The 7th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1334–1343.
- TARTALO VÁ, A. – ŽELINSKÝ, T. 2012. Rozdelenie príjmov na Slovensku: Analýza rozdelenia príjmov na voľbu ekvivalentnej škály. In: Pauhofová, I., Želinský, T. (eds.): *Nerovnosť a chudoba v Európskej únii a na Slovensku*. Košice: TU Košice. s. 99–106.
- ŽELINSKÝ, T. – STANKOVIČOVÁ, I. 2012. Spatial aspect of poverty in Slovakia. In: Löster, T., Pavelka, T. (eds.): *The 6th International Days of Statistics and Economics. Conference Proceedings. September 13–15, 2012. Prague, Czech Republic*. Dostupné: http://msed.vse.cz/msed_2012/en/

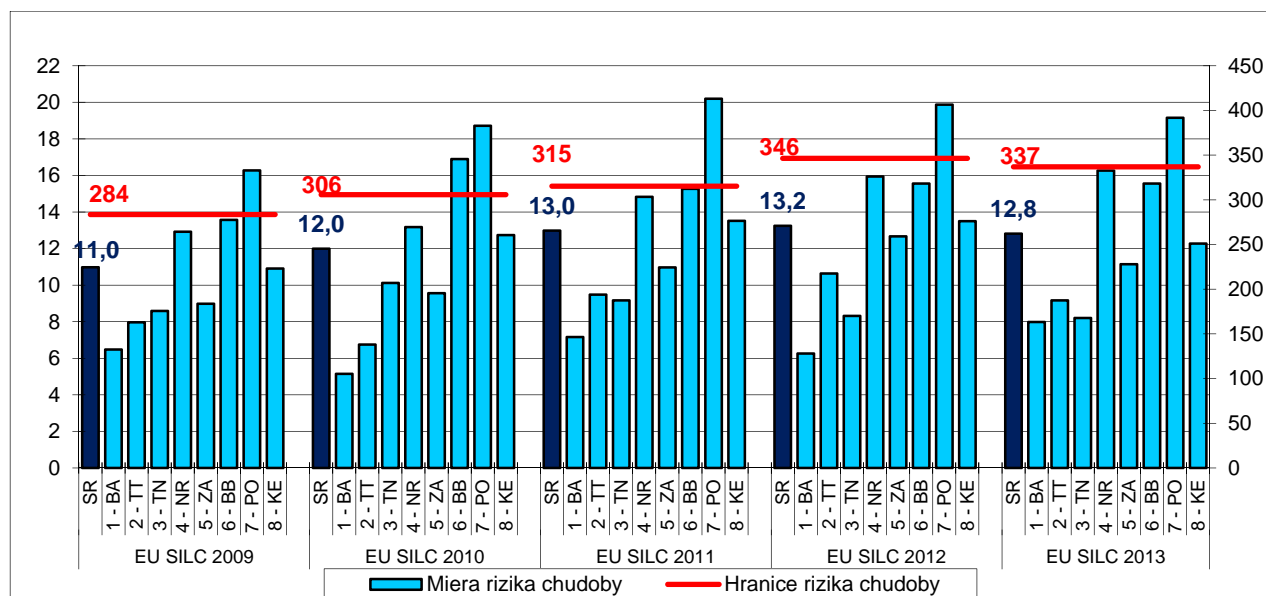
Pod'akovanie

Napísanie príspevku bolo podporené Vedeckou grantovou agentúrou MŠ SR a SAV v rámci riešenia vedecko-výskumného projektu VEGA 1/0127/11 *Priestorová distribúcia chudoby v EÚ*. V príspevku boli použité anonymizované dáta z výberového zisťovania EU SILC poskytnuté na výskumné účely na základe kontraktu č. EU-SILC/2011/33 podpísaného medzi Európskou komisiou, Eurostatom a Technickou univerzitou v Košiciach.

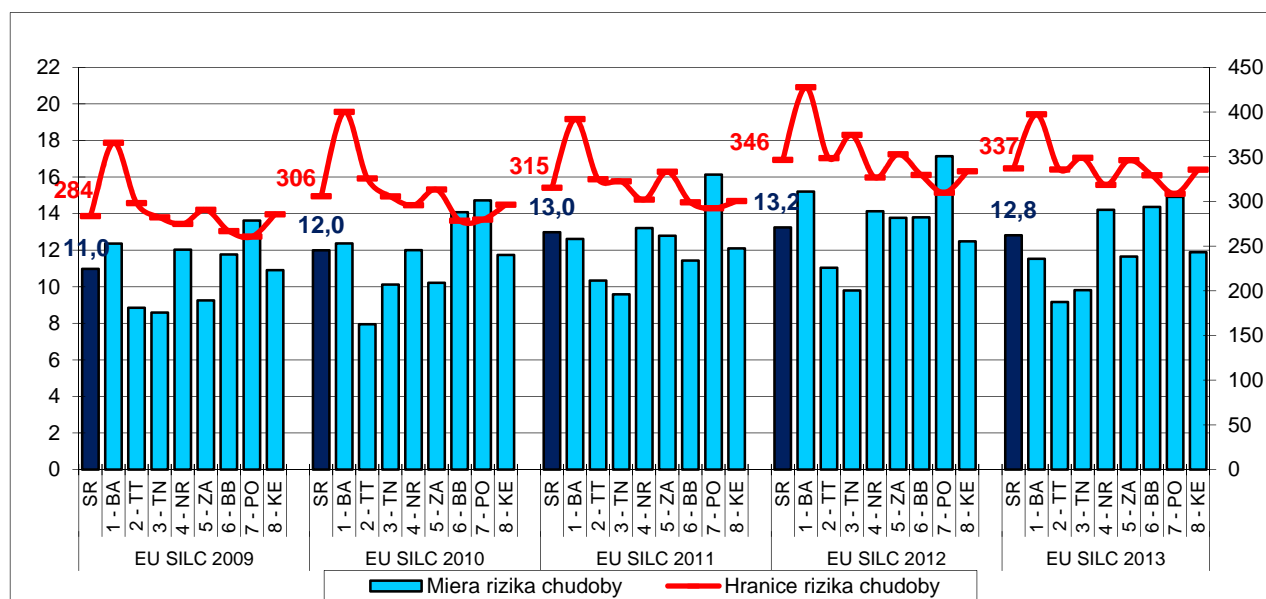
Adresy autorov:

Iveta Stankovičová
Univerzita Komenského v Bratislave
Fakulta managementu
Odbojárov 10, 820 05 Bratislava
iveta.stankovicova@fm.uniba.sk

Róbert Vlačuha
Štatistický úrad SR
Miletičova 3, 824 76 Bratislava 26
robert.vlacuha@statistics.sk



Obr. 21: Vývoj miery rizika chudoby podľa národnej hranice chudoby v SR a v regiónoch (2009 – 2013)



Obr. 22: Vývoj miery rizika chudoby podľa krajských hraníc chudoby v regiónoch SR (2009 – 2013)

McNemarov test a jeho aplikácie v ošetrovatel'stve

McNemar's Test and Its Application in Nursing

Edita Szabová, Edita Tatarová

Abstract: The paper deals with McNemar's test, which is used in research of two dichotomous dependent variables. The significance of the change after the intervention is tested. We present its application in research of efficiency of accident prevention interventions in social service facilities.

Abstrakt: Príspevok sa zaoberá McNemarovým testom, ktorý sa používa pri sledovaní dvoch závislých dichotomických premenných, pričom sa testuje významnosť zmeny po zásahu. Uvádzame jeho aplikáciu vo výskume efektívnosti protiúrazových intervencií v zariadení sociálnych služieb.

Key words: Association Table, McNemar's Test, Nursing, Intervention.

Kľúčové slová: asociačná tabuľka, McNemarov test, ošetrovatel'stvo, intervencia.

JEL classification: C14

1. Úvod

Nie vždy je cieľom štatistickej analýzy kontingenčnej tabuľky zistiť, či sú dva kvalitatívne znaky závislé. Niekedy sledujeme prítomnosť a neprítomnosť určitého javu, vykoná sa nejaká intervencia, opatrenie a pod., a je potrebné zistiť, či táto intervencia zmenila pravdepodobnosť výskytu daného javu (Anděl, 2011). Skúmame napríklad, či výučba o pozitívnom pôsobení športu na zdravie vedie k zmene postojov žiakov k športovaniu (Hendl, 2006). V pedagogickom výskume môžeme sledovať, či dve úlohy, ktoré chceme použiť v teste, majú rovnaký stupeň náročnosti (Markechová, Stehlíková, Tirpáková, 2011). V medicínskom výskume sa môže zisťovať, či majú pacienti pred podaním lieku a po podaní lieku rýchlu alebo pomalú zrážanlivosť krvi (Anděl, 2011). Vo fonetickom výskume Volín (2007) overoval významnosť rozdielov v úspešnosti rozpoznávania medzi dvomi pracovnými algoritmi. Na všetky tieto účely je možné použiť McNemarov test.

V príspevku uvedieme princíp testovania hypotéz pomocou McNemarovho testu a jeho aplikáciu vo výskume efektívnosti protiúrazových intervencií v zariadení sociálnych služieb.

2. McNemarov test

V štatistickej praxi sa často na prvkoch uvažovaného súboru sleduje dichotomická premenná dvakrát – pred pokusom (zásahom) a po ňom. Treba zistiť, či prišlo k štatisticky významnej zmene v rozdelení tejto premennej.

Pomer početností v oboch kategóriách dichotomickej premennej sa bude medzi obomi meraniami viac či menej líšiť. Veľkosť tejto zmeny sa posudzuje McNemarovým testom tak, že uvažujeme, koľko prvkov súboru pri prvom meraní prejde pri druhom meraní do inej kategórie uvažovanej premennej. Vzťah výsledkov oboch meraní zobrazujeme tabuľkou 2×2 . V každom riadku klasifikujeme výsledky prvého merania a v každom stĺpci výsledky druhého merania (Hendl, 2006).

Dáta je potrebné mať usporiadané tak, ako je v uvedené v tabuľke 1. Ide o štvorcovú tabuľku, v ktorej sú na hlavnej diagonále uvedené počty prvkov súboru, u ktorých nedošlo k zmene, a mimo hlavnej diagonály počet prvkov, u ktorých došlo k zmene (Pecáková, 2008). Predpokladá sa pritom, že prvky súboru sú výberom z multinomického rozdelenia s parametrom n a s pravdepodobnosťami uvedenými v tabuľke 2.

Tab. 1: Početnosti v McNemarovom teste

		Po zásahu		Spolu
		0	1	
Pred zásahom	0	n_{11}	n_{12}	$n_{1.}$
	1	n_{21}	n_{22}	$n_{2.}$
Spolu		$n_{.1}$	$n_{.2}$	n

Tabuľka pre McNemarov test pripomína kontingenčnú tabuľku, ale je zásadne inak postavená. Každá jednotka vo vyjadrenej početnosti zastupuje dve merania alebo dva spárované prípady. Početnosť v každej bunke je početnosť párov. Súčet n jednotiek v sebe nesie $2n$ meraní (Volín, 2007).

Tab. 2: Pravdepodobnosti v McNemarovom teste

		Po zásahu		Spolu
		0	1	
Pred zásahom	0	p_{11}	p_{12}	$p_{1.}$
	1	p_{21}	p_{22}	$p_{2.}$
Spolu		$p_{.1}$	$p_{.2}$	1

Testujeme hypotézu $H_0: p_{1.} = p_{.1}$, ktorá je ekvivalentná hypotéze $H'_0: p_{12} = p_{21}$. Všetky pravdepodobnosti uvedené v tabuľke 2 sú určené neznámymi parametrami p_{11} a p_{12} . Za platnosti hypotézy H'_0 platí $p_{21} = p_{12}$, ďalej $p_{22} = 1 - p_{11} - 2p_{12}$, $\frac{n_{11}}{p_{11}} - \frac{n_{22}}{p_{22}} = 0$, $\frac{n_{12}}{p_{12}} + \frac{n_{21}}{p_{21}} - 2\frac{n_{22}}{p_{22}} = 0$. Odtiaľ dostávame $n_{11} = \frac{n_{22}}{p_{22}}p_{11}$, $n_{12} + n_{21} = 2\frac{n_{22}}{p_{22}}p_{12}$. Použijeme rovnosť $n_{22} = \frac{n_{22}}{p_{22}}p_{22}$. Sčítaním dostaneme $n = \frac{n_{22}}{p_{22}}$. Odhadom pre p_{22} je $\hat{p}_{22} = \frac{n_{22}}{n}$. Potom $\hat{p}_{11} = \frac{n_{11}}{n}$, $\hat{p}_{12} = \frac{n_{12} + n_{21}}{2n}$, $\hat{p}_{21} = \hat{p}_{12}$.

Veličina

$$\chi^2 = \sum_{i=1}^2 \frac{(n_{ii} - n\hat{p}_{ii})^2}{n\hat{p}_{ii}} + \frac{(n_{12} - n\hat{p}_{12})^2}{n\hat{p}_{12}} + \frac{(n_{21} - n\hat{p}_{21})^2}{n\hat{p}_{21}}$$

sa po úprave rovná

$$\chi^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}}$$

a má asymptoticky χ^2 – rozdelenie s jedným stupňom voľnosti. Nulovú hypotézu H_0 zamietame na hladine významnosti α , ak $\chi^2 \geq \chi^2_{\alpha}(1)$. Asymptotický test je možné používať už pri $n_{12} + n_{21} \geq 8$ (Anděl, 2011).

Vzorec pre výpočet testovacej štatistiky je ešte niekedy upravovaný podľa vzoru Yatesovej korekcie, aby sa zohľadnila skutočnosť, že χ^2 – rozdelenie je spojité, kým početnosti dosadzované do vzorca sú prirodzené čísla:

$$\chi^2 = \frac{(|n_{12} - n_{21}| - 1)^2}{n_{12} + n_{21}}$$

Existujú však názory, podľa ktorých je táto korekcia zbytočne prísna. Môžeme však povedať, že výsledky získané za prísnejších podmienok zvyknú byť viac cenené.

McNemarov test je špeciálnym prípadom znamienkového testu. Rozdiel spočíva v tom, že pri znamienkovom teste sa používa štatistika, ktorá má za predpokladu nulovej hypotézy normálne normované rozdelenie, kým pri McNemarovom teste je použitá jej druhá mocnina, ktorá, ako sme už uviedli, má χ^2 – rozdelenie s jedným stupňom voľnosti. Získaná hladina

významnosti je však rovnaká (Řezanková, 2007). Pri malých početnostiach sa vychádza z toho, že za platnosti nulovej hypotézy H_0 je podmienené rozdelenie početností n_{12} pri daných n_{11} a n_{22} binomické $Bi\left(n_{12} + n_{21}, \frac{1}{2}\right)$. Položme $N = n_{12} + n_{21}$. Hypotézu H_0 zamietame, ak je $n_{12} \leq k_1$ alebo $n_{12} \geq k_2$, kde k_1, k_2 sú kritické hodnoty binomického rozdelenia $Bi\left(N, \frac{1}{2}\right)$. Tento test je obojstranný. Využitím variantu založeného na binomickom rozdelení sa dá prevádzať McNemarov test aj proti jednostranným alternatívam (Anděl, 2011).

3. Aplikácia McNemarovho testu v ošetrovatel'stve

Uvedený test aplikujeme na dáta získané z experimentu v oblasti ošetrovatel'stva. V Zariadení sociálnych služieb „Viničky“ v Nitre bolo pozorovaných 32 klientov zariadenia vo veku 65 až 92 rokov. Pozorovanie sa týkalo okrem iného aj toho, či klient mal alebo nemal nejaký úraz. Boli navrhnuté preventívne opatrenia s cieľom znížiť úrazovosť v zariadení. Návrhy preventívnych opatrení pre klientov spočívali v úprave obytných jednotiek: výmena ošúchanej podložky, narovnanie pokrčených kobercov, odstránenie nerovnosti dlážky, odstránenie prahov, správne rozloženie nábytku, odstránenie bariér na obytnej jednotke, odstránenie ostrých častí nábytku, montáž držiadiel a rukovätí v kúpeľni a WC, zabezpečenie správneho osvetlenia. Najčastejším problémom bolo bariérové usporiadanie nábytku, neprítomnosť rukovätí a držiadiel v kúpeľni a WC. Opatrenia a úpravy obytných jednotiek boli zabezpečené v spolupráci s technickým úsekom v mesiacoch júl – november 2010. Edukáciu klientov o používaní správnej pevnej, nešmykľavej obuvi a výmenu nevhodnej obuvi za správnu obuv a edukáciu o nevhodnosti chodiť naboso po mokrej podlahe u klientov vykonala spoluautorka tohto článku v spolupráci s rodinou klientov individuálnou formou komunikácie. Edukácia klientov po uskutočnení edukačných stretnutí pokračovala formou individuálnych rozhovorov počas stretnutí s klientmi.

Po aplikácii preventívnych opatrení sme uskutočnili pozorovania klientov zamerané na výskyt najčastejších rizík u klientov za sledované obdobie mesiacov január – február 2011. Pozorovania boli použité na porovnanie najčastejších rizík u klientov pred a po aplikácii preventívnych opatrení.

V tabuľke 3 sú zaznamenané počty klientov zariadenia s úrazom a bez úrazu pred vykonaním opatrení a po vykonaní opatrení.

Tab. 3: Úrazy pred intervenciou a po intervencii

		úrazy po intervencii		Spolu	
		bez úrazu	s úrazom		
úrazy pred intervenciou	bez úrazu	Počet	10	1	11
		%	31,2%	3,1%	34,4%
	s úrazom	Počet	8	13	21
		%	25,0%	40,6%	65,6%
Spolu	Počet	18	14	32	
	%	56,2%	43,8%	100,0%	

Pred vykonaním opatrení malo nejaký úraz 65,6% sledovaných klientov zariadenia, bez úrazu bolo 34,4% klientov zariadenia. Po vykonaní preventívnych opatrení prišlo k úrazu u 43,8% sledovaných obyvateľov zariadenia, bez úrazu bolo 56,2% obyvateľov zariadenia.

Bez úrazu pred vykonaním opatrení aj po vykonaní opatrení bolo 31,2% klientov zariadenia. Naopak, 40,6% klientov malo úraz pred vykonaním aj po vykonaní preventívnych opatrení. 25% klientov zariadenia pred preventívnymi opatreniami malo úraz, ale po vykonaní opatrení už boli bez úrazu. Len jeden obyvateľ zariadenia bol pred intervenciou bez úrazu a po vykonaní prevencie sa mu stal úraz.

Z tabuľky 4 vidíme, že zo všetkých klientov, ktorí pred intervenciou nemali žiadny úraz, ho 90,9% obyvateľov nemalo ani po vykonaní preventívnych opatrení. Uvažujme teraz len tých klientov, ktorí pred intervenciou mali nejaký úraz. 61,9% z nich ho malo aj po vykonaní prevencie. 38,1% z nich si po vykonaní opatrení polepšilo a nemalo v sledovanom období žiadny úraz.

Tab. 4: Úrazy pred intervenciou a po intervencii – riadkové percentá

		úrazy po intervencii		Total
		bez úrazu	s úrazom	
úrazy pred intervenciou	bez úrazu	90,9%	9,1%	100,0%
	s úrazom	38,1%	61,9%	100,0%
Total		56,2%	43,8%	100,0%

Zistili sme teda, že výskyt úrazov a pádov u klientov sa po aplikácii preventívnych opatrení znížil. Chceme zistiť, či bol vplyv preventívnych opatrení na úrazovosť u klientov zariadenia štatisticky významný. Budeme testovať nulovú hypotézu H_0 : „Vykonanie preventívnych opatrení nemá vplyv na úrazovosť u klientov zariadenia sociálnych služieb“. Na testovanie použijeme McNemarov test. Zvolená hladina významnosti $\alpha = 0,05$. Vypočítame hodnotu testovacej štatistiky

$$\chi^2 = \frac{(n_{12} - n_{21})^2}{n_{12} + n_{21}} = \frac{(1 - 8)^2}{1 + 8} = \frac{49}{9} = 5,4.$$

Hodnotu testovacej štatistiky χ^2 porovnáme s kritickou hodnotou $\chi_{0,05}^2(1) = 3,84$. Keďže hodnota testovacej štatistiky je vyššia ako uvedená kritická hodnota, testovanú hypotézu zamietame na hladine významnosti $\alpha = 0,05$.

Hodnota testovacej štatistiky pri korekcii podľa Yatesa je

$$\chi^2 = \frac{(|n_{12} - n_{21}| - 1)^2}{n_{12} + n_{21}} = \frac{(|1 - 8| - 1)^2}{1 + 8} = \frac{36}{9} = 4.$$

Aj v tomto prípade je hodnota testovacej štatistiky χ^2 vyššia ako kritická hodnota $\chi_{0,05}^2(1) = 3,84$ a testovanú hypotézu zamietame na hladine významnosti $\alpha = 0,05$.

To znamená, že vykonanie opatrení na zníženie úrazovosti u klientov zariadenia sociálnych služieb bolo efektívne.

4. Záver

V príspevku sme sa venovali McNemarovmu testu, ktorý sa používa v prípadoch, keď na skúmaných prvkoch súboru sledujeme dichotomickú premennú dvakrát – pred pokusom (zásahom) a po ňom, a máme zistiť, či prišlo k štatisticky významnej zmene v rozdelení tejto premennej. Teoretické východiská testovania štatistických hypotéz pomocou McNemarovho testu sme aplikovali na zisťovanie efektívnosti protiúrazových opatrení, ktoré boli implementované v zariadení sociálnych služieb. Výskumnú vzorku tvorilo 32 klientov zariadenia vo veku 65 až 92 rokov. Testovali sme nulovú hypotézu „Vykonanie

preventívnych opatrení nemá vplyv na úrazovosť u klientov zariadenia sociálnych služieb“. McNemarovým testom bolo preukázané, že vykonané preventívne opatrenia mali pozitívny vplyv na zníženie menších či väčších úrazov v zariadení sociálnych služieb.

Literatúra

ANDĚL, J. 2011. *Základy matematické statistiky*. Praha: Matfyzpress.

HENDL, J. 2006. *Přehled statistických metod zpracování dat*. Praha: Portál.

MARKECHOVÁ, D. – STEHLÍKOVÁ, B. – TIRPÁKOVÁ, A. 2011. *Základy statistiky pre pedagógov*. Nitra: Univerzita Konštantína Filozofa v Nitre.

PECÁKOVÁ, I. 1998. *Statistika v terenných průzkumech*. Praha: Professional Publishing.

ŘEZANKOVÁ, H. 1997. *Analýza dat z dotazníkových šetření*. Praha: Professional Publishing.

VOLÍN, J. 2007. *Statistické metody ve fonetickém výzkumu*. Praha: Univerzita Karlova v Praze.

Adresy autorov):

Edita Szabová, PaedDr.
Univerzita Konštantína Filozofa v Nitre
Tr. A. Hlinku 1, 949 74 Nitra
edita.szabova@ukf.sk

Edita Tatarová, Mgr.
Zariadenie sociálnych služieb „Viničky“
Považská 14, 949 11 Nitra
edita.tatarova@gmail.com

Shluková analýza a metoda hlavních komponent v aplikaci vyhledávání odlehlých pozorování věkově-specifických měř úmrtnosti v České republice a na Slovensku¹

Cluster analysis and principal component method in the application of searching outliers of age-specific mortality rates in the Czech Republic and Slovakia

Ondřej Šimpach

Abstract: The aim of this paper is to analyse and identify the outliers' observation of age-specific mortality rates for the Czech and Slovak population of males and females. We use the cluster analysis with the nearest neighbour clustering method, which is commonly used in analysing the similarities of objects in the base datasets and the principal component method, which was specially prepared in R software package to identify the outliers in demographic data. Results from both approaches are compared with each other, because it does not provide the same conclusions. Information about the existence of outliers is important for the analyst because the significantly skewed and outlying values of age-specific mortality rates must be smoothed for the continuing of mortality analysis.

Abstrakt: Cílem předkládaného článku je analyzovat a identifikovat odlehlé hodnoty pozorování věkově-specifických měř úmrtnosti u české a slovenské populace mužů a žen. Využito bude shlukové analýzy se shlukovací metodou nejbližšího souseda, která se běžně používá při analýze podobnosti objektů v základním souboru a dále metody hlavních komponent, která byla speciálně připravena do balíčku software R na identifikaci odlehlých hodnot v demografických datech. Výsledky z obou přístupů jsou vzájemně porovnány, neboť neposkytují zcela stejné závěry. Informace o existenci odlehlých pozorování je pro analytika důležitá proto, že významně vychýlené hodnoty věkově-specifických měř úmrtnosti se musí vyrovnávat pro další pokračování v úmrtnostní analýze.

Key words: principal component method, cluster analysis, age-specific mortality rates

Klíčová slova: metoda hlavních komponent, shluková analýza, specifické míry úmrtnosti

JEL classification: C38, C55, J11

1. Úvod

Odlehlá pozorování tvoří většinou vždy komplikace při jakékoliv statistické analýze a brání jejímu korektnímu průběhu. Úkolem každého analytika je nejdříve odlehlá pozorování identifikovat, rozhodnout, zdali jsou významná nebo nejsou a v případě, že by jejich přítomnost mohla ovlivnit korektní výsledek, pak rozhodnout, jakým způsobem se s nimi vypořádat. V našem článku zaměříme svou pozornost na odlehlosti v průbězích věkově-specifických měř úmrtnosti v čase a jako datové matice poslouží základna České republiky a Slovenska v letech 1950–2009. Budou aplikovány dva přístupy, přičemž jeden je založen na identifikaci odlehlých shluků z vypočtených dendrogramů, ten druhý pak na identifikaci z „funkčních“ a „balíčkových“ grafů (Shang, Hyndman, 2010), vytvořených za využití metody hlavních komponent. Hodnoty, které jsou v článku identifikovány jako odlehlé, by musely být pro pokračování v analýze korigovány, aby neohrozily korektní výsledek. Vzhledem

¹ Článek byl zpracován v rámci mezifakultního vědecko-výzkumného projektu Vysoké školy ekonomické v Praze, vedený pod číslem IGS MF/F4/6/2013: Hodnocení výsledků metod shlukové analýzy v ekonomických úlohách.

k tomu, že výsledky z obou přístupů nejsou shodné, zhodnotíme je a provedeme diskusi o jejich vhodnosti.

2. Metodika a data

Pro potřeby analýzy odlehých hodnot budou využita demografická data z „Human Mortality Database“ (HMD) za Českou republiku a Slovensko s roční periodicitou pozorování. Přestože za Českou republiku jsou k dispozici data za roky 1950–2011 v jednoletých věkových skupinách, zkrátíme časové řady pouze do roku 2009, protože slovenská časová řada stále končí rokem 2009 a novější odhady ze strany HMD publikovány nejsou. S využitím počtů zemřelých x -letých v roce t ($M_{x,t}$), expozici k riziku ve věku x a roce t , která je odhadována jako střední stav obyvatelstva x -letého v roce t ($\bar{S}_{x,t}$), budou vypočteny věkově-specifické míry úmrtnosti

$$m_{x,t} = \frac{M_{x,t}}{\bar{S}_{x,t}} \quad (1)$$

pro každou zemi a pohlaví, kde $x = 0, 1, \dots, 100+$ a $t = 1950, 1951, \dots, 2009$. Věkově-specifická matice za každou zemi a pohlaví bude datovým vstupem pro shlukování metodou nejbližšího souseda. Vzhledem k tomu, že se míra úmrtnosti v čase vyvíjí a obecně její úroveň ve většině populací pozvolna klesá, měly by podle principů zmíněné metody být jednotlivé dílčí matice o rozměru $x \times 1$ pozvolna přidávány do shluků a v případě, že je některý průběh těchto měř v daném roce vychýlený, měl by být vyhodnocen jako odlehlý a mezi shluky osamocený. S využitím software Statgraphics Centurion XVI bude aplikována metoda nejbližšího souseda za použití metriky čtvercových euklidovských vzdáleností (viz např. Gordon, 1999)

$$\|a - b\|_2^2 = \sum_i (a_i - b_i)^2, \quad (2)$$

kde a a b jsou libovolné dvě vzdálenosti. Tato metrika má široké uplatnění v různých shlukovacích metodách, přičemž v některých metodách se její použití doporučuje, v jiných je její využití nezbytné (např. Wardova metoda, viz Pivoňka, Löster, 2013, nebo Löster, 2014). Metoda nejbližšího souseda je v článku použita proto, že její výsledky jsou snadno interpretovatelné. Jak uvádí např. Murtagh (1983, s. 354), „algoritmus nejbližšího souseda opakovaně sleduje řetězce shluků $A \rightarrow B \rightarrow C \rightarrow \dots$, kde každý shluk je nejbližší soused toho předchozího, až do dosažení dvojic shluků, které jsou vzájemně nejbližší sousedé“. Vzdálenost dvou shluků je vždy dána vzdáleností dvou nejbližších objektů z nich (Bruynooghe, 1977) a platí, že vzdálenostní funkce ($d = \text{distance function}$)

$$d(A \cup B, C) \geq \min(d(A, C), d(B, C)). \quad (3)$$

Přejdeme nyní k metodě hlavních komponent, kterou se pro potřeby analýzy demografických dat zabýval např. Shang, Hyndman (2010), nebo Hyndman (2012). S využitím počtů zemřelých x -letých v roce t a zemi c , expozici k riziku ve věku x , roce t a zemi c , bude vytvořena tzv. datová „báze“ v software RStudio, naplněna logaritmy věkově-specifických měř úmrtnosti pro každou analyzovanou zemi a každé pohlaví jako

$$BÁZE \leftarrow \ln(m_{x,t}^c) = \ln\left(\frac{M_{x,t}^c}{\bar{S}_{x,t}^c}\right). \quad (4)$$

Tato „báze“ logaritmů je tvořena z důvodu, že využitá struktura se společně s balíčkem „demography“ a „rainbow“ (Hyndman, 2012) využívá pro stochastické modelování úmrtnosti modelem Lee-Cartera (Lee, Carter, 1992)

$$\ln(m_{x,t}^c) = a_x^c + b_x^c k_t^c + e_{x,t}^c, \quad (5)$$

kde a_x^c jsou věkově-specifické profily úmrtnosti nezávislé na čase, b_x^c je doplňková věkově-specifická komponenta určující, jak moc se každá věková skupina mění, když se mění k_t^c a nakonec k_t^c je časově proměnlivý parametr, tzv. „index úmrtnosti“. S využitím metody hlavních komponent, kterou Hyndman (2012) aplikoval do balíčků „demography“ a „rainbow“ jsme schopni identifikovat odlehlé roky jednotlivých zemí a populací. Ty budou graficky zobrazeny na základě výsledků skór hlavních komponent. Autoři Shang, Hyndman (2010) představili dvojrozměrný „funkční“ a dvojrozměrný „balíčkový“ graf, určený pro přehlednou vizualizaci odlehlých pozorování v demografických datech. Tyto funkční a balíčkové grafy vždy obsahují dvě pole. Jedno pole je tmavě šedé, druhé světle šedé. Tmavě šedé pole obsahuje 50 % všech pozorování, přičemž ve funkčním grafu je i černá mediánová křivka. V okolí této mediánové křivky se nacházejí další 2 přerušované, které znázorňují 95% oboustranný interval spolehlivosti. Shang, Hyndman (2010) označují všechny křivky, které leží mimo hranice obou polí jako odlehlé. Tyto křivky jsou ve funkčním grafu zobrazeny barevně s legendou. Dvojrozměrný balíčkový graf neobsahuje žádnou mediánovou křivku, ale Tukeyho hloubkový medián (viz Tukey, 1975). Barevné body s popiskem příslušného pozorování, ležící mimo hranice polí jsou odlehlá.

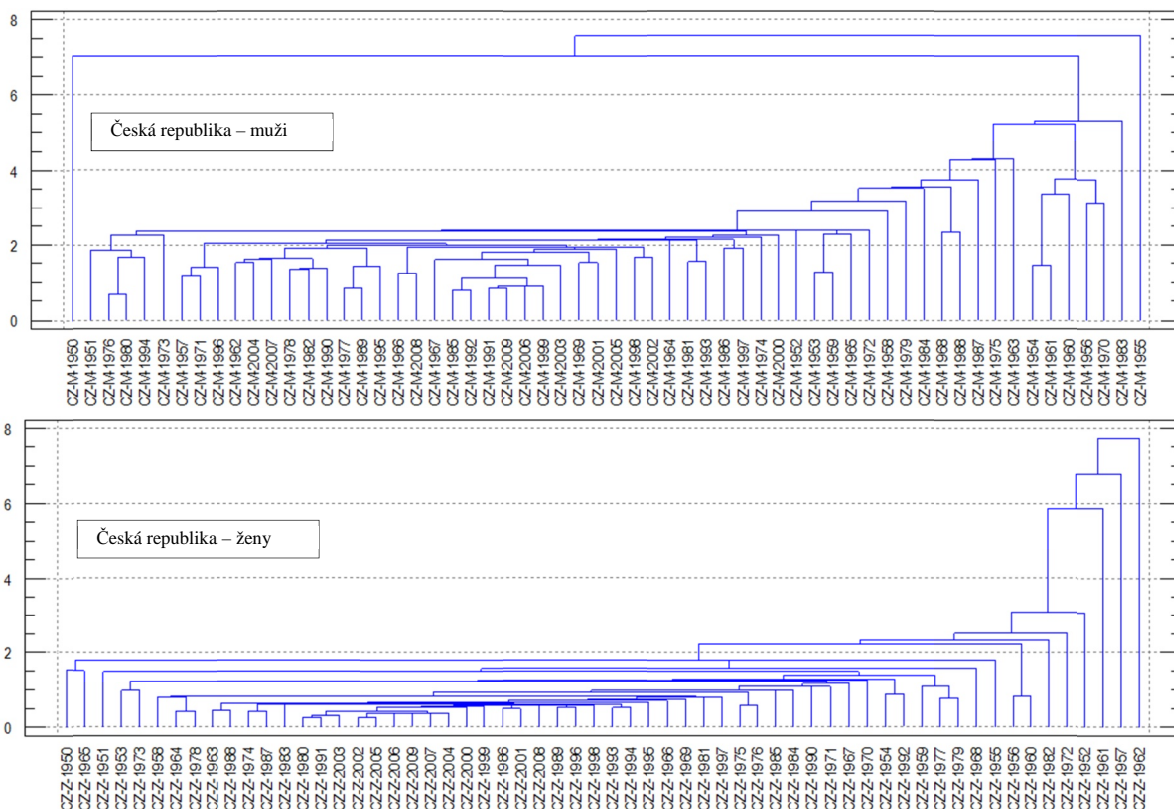
3. Výsledky

V software Statgraphics Centurion XVI byly na datech české a slovenské úmrtnosti mužů a žen vypočteny dendrogramy, ze kterých lze vyčíst podobnost vývoje úmrtnosti mezi jednotlivými lety. Je-li pozorování odlehlé, pak je v dendrogramu většinou osamocené. V případě populace českých mužů (viz Obr. 1 nahoře) se jedná o roky 1950 a 1955. Ženy (Obr. 1 dole) mají podle této metody odlehlé hodnoty v letech 1957, 1961 a 1962. V případě slovenské populace je v datech patrná mnohem vyšší variabilita, protože podobnost nastává až u vyšších hodnot vzdálenosti. Co se týče viditelných osamocených pozorování, jedná se o roky 1950, 1951, 1953, 1970 a 1972 v případě mužů a roky 1950, 1957 a 1961 v případě žen.

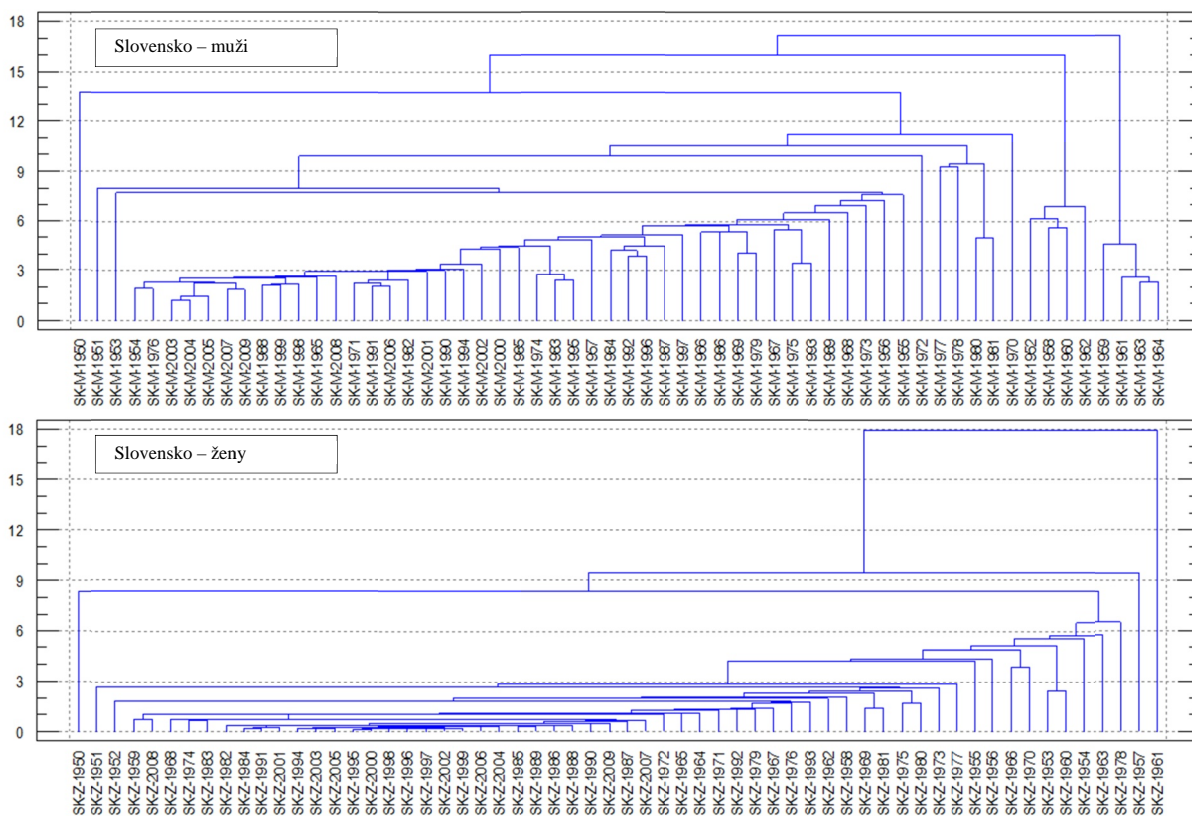
U ženské populace je situace podobná. Odlehlé roky uvažujeme 1950, 1951, 1952. Vyšší hodnoty věkově-specifických měr úmrtnosti jsou zřetelné zejména ve věkových skupinách 5–45 let a dále v nejvyšších věkových skupinách. Na Slovensku jsou detekované odlehlé hodnoty mnohem více viditelné. V případě mužské populace se jedná o roky 1950 a 1951, přičemž výrazně vyšší hodnoty věkově-specifických měr jsou ve věkových skupinách 5–35 let. Ženy mají odlehlých pozorování mnohem více, jedná se o roky 1950, 1951 a 1952 a dále 1964, 1966 a 1969. První skupina odlehlých let je charakterizována vyššími hodnotami věkově-specifických měr úmrtnosti ve věkových skupinách 5–65 let.

Druhá skupina má vysokou variabilitu dat v nejvyšších věkových skupinách a proto byly také pravděpodobně klasifikovány jako odlehlé. Právě odlehlost hodnot a vysoká variabilita v nejvyšších věkových skupinách je v demografii často diskutovaným tématem, neboť hodnoty věkově-specifických měr úmrtnosti je třeba vyrovnávat pomocí vyrovnávacích funkcí.

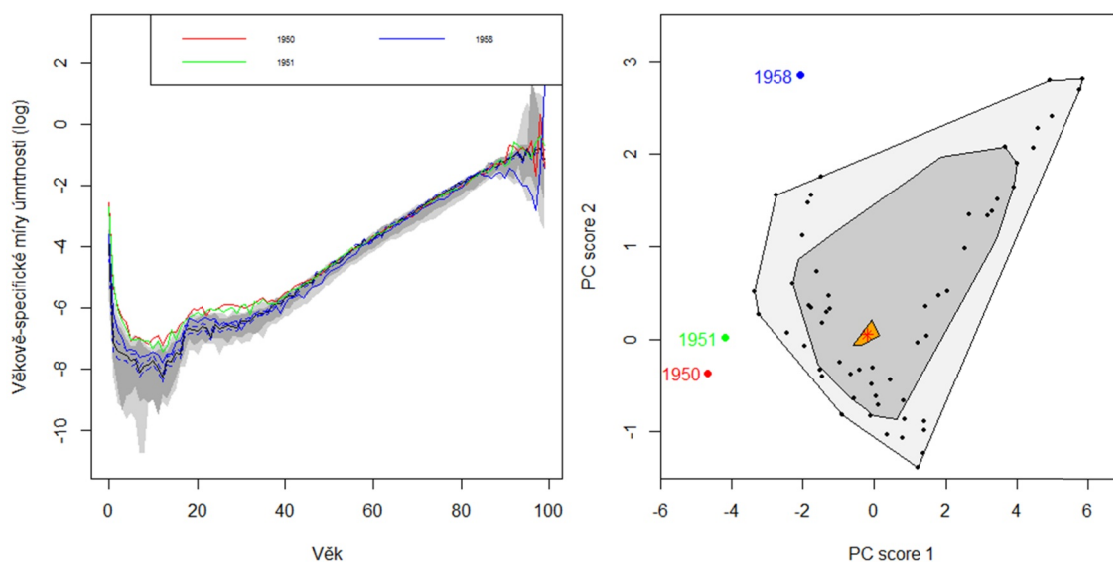
Výsledky poskytnuté balíčkem „demography“ a „rainbow“ (Hyndman, 2012) ze software RStudio, jsou většinou odlišné. Mužská populace České republiky má odlehlé roky 1950, 1951 a 1958. Je to sice poválečné období a počátek komunistického režimu na území tehdejší ČSR, ale tato odlehlost je spíše dána rozkolísaností úmrtnosti v nejvyšších věkových skupinách.



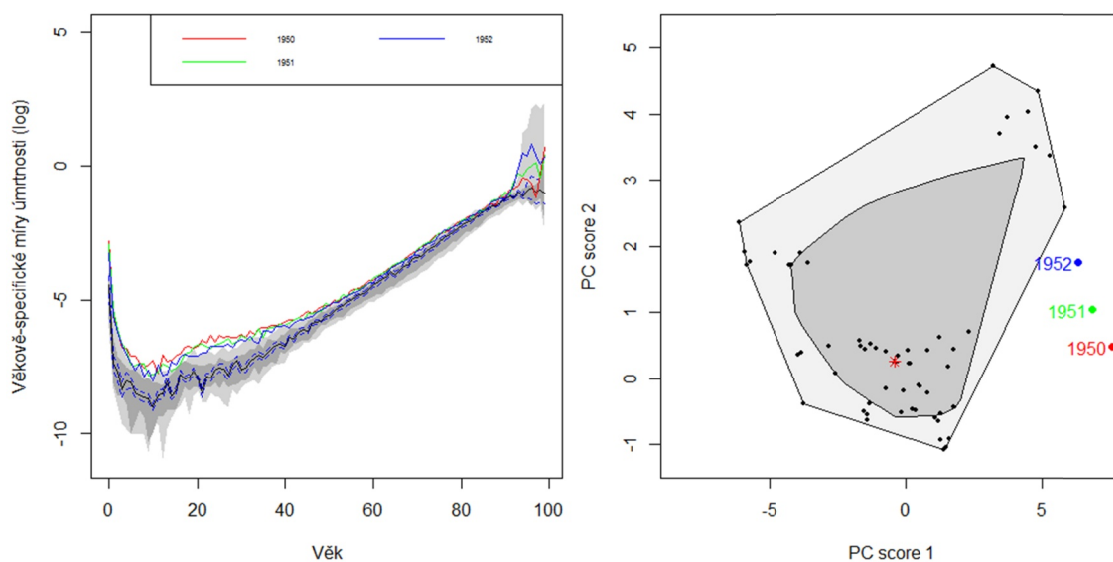
Obr. 1: Dendrogramy pro shluky věkově-specifických měř úmrtnosti pro Českou republiku.
Zdroj: vlastní výpočet v software Statgraphics Centurion XVI



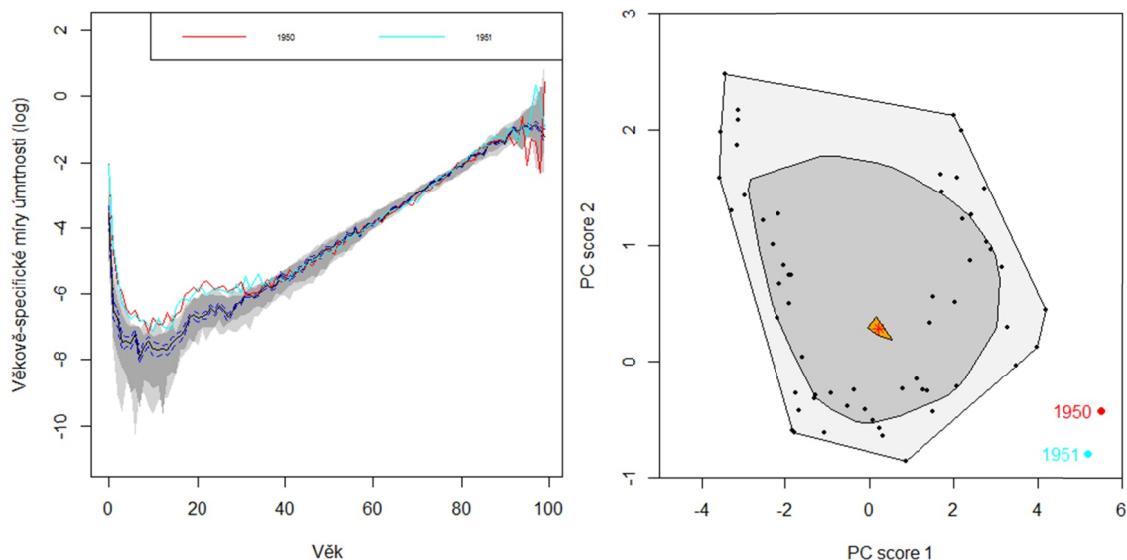
Obr. 2: Dendrogramy pro shluky věkově-specifických měř úmrtnosti pro Slovensko.
Zdroj: vlastní výpočet v software Statgraphics Centurion XVI



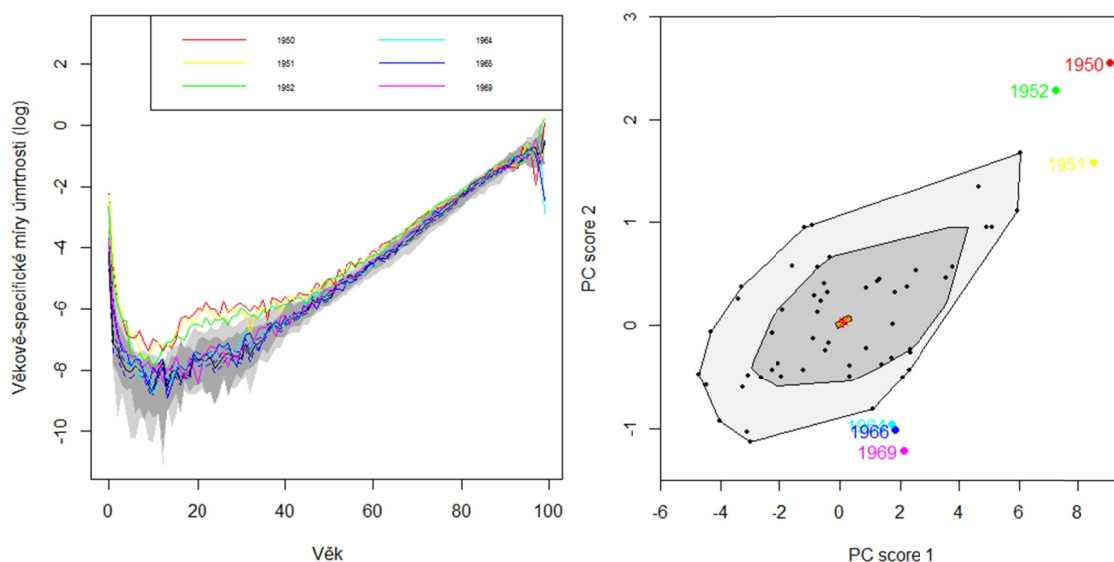
Obr. 3: Odlehlá pozorování v ČR identifikovaná metodou hlavních komponent – muži
Zdroj: vlastní výpočet v software R



Obr. 4: Odlehlá pozorování v ČR identifikovaná metodou hlavních komponent – ženy



Obr. 5: Odlehlá pozorování v SR identifikovaná metodou hlavních komponent – muži



Obr. 6: Odlehlá pozorování v SR identifikovaná metodou hlavních komponent – ženy.

Zdroj: vlastní výpočet v software R

4. Závěr

V předkládaném článku jsme se zabývali srovnáním dvou přístupů v identifikaci odlehlých pozorování na případu věkově-specifických měr úmrtnosti české a slovenské populace, přičemž byla využita shluková analýza s metodou nejbližšího souseda a prezentace výsledků byla za pomoci dendrogramů, a dále metoda hlavních komponent, kde výsledky byly prezentovány formou funkčních a balíčkových grafů. V případě, že bychom použili jinou shlukovací metodu než nejbližšího souseda, dostali bychom pravděpodobně výsledky odlišné, ale nejbližší soused se pro tuto problematiku hodí nejvíc. Z uvedených poznatků tedy můžeme učinit závěr, že vyrovnání věkově-specifických měr úmrtnosti je třeba provádět minimálně u těch populací v konkrétním roce, které byly jak shlukovou analýzou, tak metodou hlavních komponent, identifikovány jako odlehlé, tj. v případě shlukové analýzy a použité metody nejbližšího souseda tvoří osamocené shluky a v případě metody hlavních komponent jsou zobrazeny mimo hranice polí „funkčního“ a „balíčkového“ grafu.

Literatura

- BRUYNOOGHE, M. 1977. Méthodes nouvelles en classification automatique de données taxinomiques nombreuses. *Statistique et Analyse des Données*, vol. 3, pp. 24–42.
- GORDON, A. D. 1999. *Classification, 2nd Edition*. Chapman and Hall, Boca Raton.
- HYNDMAN, R. J. 2012. *demography: Forecasting mortality, fertility, migration and population data*. R package v. 1.16. URL: <<http://robjhyndman.com/software/demography/>>
- LEE, R. D., CARTER, L. R. 1992. Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association*, vol. 87, pp. 659-675.
- LÖSTER, T. 2014. *Metody shlukové analýzy a jejich hodnocení*. 1. vyd. Slaný: Melandrium. 132 s. ISBN 978-80-86175-88-1.
- MURTAGH, F. 1983. A survey of recent advances in hierarchical clustering algorithms. *The Computer Journal*, vol. 26, no. 4, pp. 354–359.
- PIVOŇKA, T. – LÖSTER, T. 2013. Cluster Analysis as a Tool of Evaluating Clusters of the EU Countries before and during Global Financial Crisis from the Perspective of the Labor Market. *Intellectual Economics*, roč. 7, č. 4, s. 411–425. ISSN 1822-8011.
- SHANG, H. L. – HYNDMAN, R. J. 2010. Exploratory graphics for functional data. *Working paper of the Department of Econometrics and Business Statistics*. Monash University, Clayton, Australia, August 3, 2010, pp. 1-9.
- TUKEY, J. W. 1975. Mathematics and the picturing of data. In: *R. D. James, ed., "Proceedings of the International Congress of Mathematicians", vol. 2, Canadian mathematical congress*, Vancouver, pp. 523-531.

Adresa autora:

Ondřej Šimpach, Ing.
Fakulta informatiky a statistiky VŠE v Praze
Nám. W. Churchilla 4, 130 67 Praha 3
ondrej.simpach@vse.cz

Faktory vplývajúce na subjektívne pociťované šťastie

Factors affecting the subjective well-being

Alena Tartalová, Vladan Fejlek

Abstract: In this work we dealt with factors that can affect subjective well-being. The analysis used the latest available data from the database of the World Values Surveys. Overall, three models were estimated. The resulting model has shown the factors that significantly influence subjective well-being are marital status, satisfaction with financial situation, job satisfaction, subjectively perceived health, employment status, socio-economic status and size of community in which the respondent lives.

Abstrakt: V práci sme sa venovali faktorom, ktoré ovplyvňujú subjektívne pociťované šťastie. K analýze boli použité najnovšie dostupné údaje z databázy World Values Surveys. Celkovo boli odhadované tri modely. Výsledný model ukázal, že faktory, ktoré významne vplývajú na subjektívne pociťované šťastie sú rodinný stav, spokojnosť s finančnou situáciou, spokojnosť s pracovným stavom, subjektívne pociťované zdravie, pracovný stav, sociálnoekonomický status a veľkosť sídla, v ktorom respondent žije.

Key words: happiness economics, subjective well-being, econometric model

Kľúčové slová: Ekonomia šťastia, subjektívne pociťované šťastie, ekonometrický model

1. Úvod

Vo viacerých vedeckých článkoch sa skúma bohatstvo jednotlivcov, či spoločnosti, pričom ako ukazovatele pomocou ktorých sa blahobyť meria sú príjem alebo spotreba. Ale napríklad aj celkové zdravie, úroveň vzdelania, vek, pohlavie, či nezamestnanosť môžu mať na celkové bohatstvo vplyv. V tomto príspevku, by sme chceli poukázať na ekonomiu šťastia, ktorá sa zaoberá subjektívne pociťovaním šťastím komplexnejšie ako iba jednosmerným pohľadom na výšku príjmu.

Ekonomia šťastia je prístupom, ktorý vychádza z ekonomických teórií, spojených so sociologickými poznatkami a napriek tomu tento koncept bol prítomný v ekonomických vedách už pri začiatkoch. Ekonomiu šťastia možno brať ako obohatenie bežnej ekonomickej teórie, ale aj ako samostatnú alternatívu (pozri Bruni, 2006 a Bruni, 2007).

Keďže v práci používame rozsiahlo pojem šťastie, alebo pocit šťastia je vhodné definovať si jeho význam, no keďže ide o pojem abstraktný definovanie nie je jednoduché. Šťastie má viacero významov (podľa Diener, 2000).

Šťastie ako zhoda okolností, ktorej výsledok je úspech. Takéto šťastie je z pravidla nezávislé na snažení jednotlivca a je často zamieňané za pojem náhoda.

Šťastie ako krátkodobý pocit, ktorý je výsledkom snahy o dosiahnutie tohto pocitu, je spojený so spotrebou, blahobytom v psychológii hovoríme o hedonisticko-subjektívnom šťastí. Ak dochádza k definícii úžitku, hovoríme o tomto zmysle slova šťastie. Odlišnosti vo vnímaní užitočnosti pramenili so subjektívnej zložky užitočnosti a očakávaní z outputov, pri spotrebe statkov.

Šťastie ako dlhodobý pocit respektíve pretrvávajúci stav. Hovoríme o eudaimonii, ktorú sme rozobrali vyššie.

Šťastie ako pocit trvalý, ide o celkovú spokojnosť s vlastným životom. Ide o najabstraktnejšiu zo všetkých definícií. Takto definované šťastie nemôže byť odhadnuté podľa spotreby, alebo pozície a vlastníctva. Takto definované šťastie môže byť

kvantifikované len jednotlivcom. Táto definícia šťastia je známa pod pojmom subjective well-being a v určitom množstve začleňuje v sebe viac alebo menej predchádzajúce 2 definície. (Diener, 2000)

Táto práca bude pracovať najmä so štvrtou definíciou.

Prvým autorom, ktorý skúmal šťastie a vplyv výšky príjmu na pociťované šťastie, bol Richard Easterlin (1974), Easterlin (2005). Známy je tzv. Easterlinov paradox: Napriek rastúcemu príjmu, ľudia nezaznamenali výrazný nárast v pociťovanom šťastí v dlhodobom horizonte. Tento paradox vytvoril základ k skúmaniu vzájomného vzťahu týchto dvoch veličín. V súčasnosti je najznámejším ekonómom, ktorý sa problematike ekonomie šťastia venuje, švajčiarsky vedec Frey (1995, 2000, 2001). Na Slovensku je táto problematika zatiaľ málo známa a preskúmaná.

2. Ekonometrický model

Pre vytvorenie modelu a analýze vzťahu medzi subjektívne pociťovaným šťastím je potrebné poznať štruktúru dát, na ktorých tieto analýzy postavíme. World Value Survey (WVS) je globálnym projektom na sledovanie a mapovanie ľudských presvedčení a hodnôt. Dáta WVS obsahujú viac ako 95 000 odpovedí zaznamenaných počas 5 vln výskumu, v takmer 100 krajinách sveta. Je jednou z mála databáz, ktorá sa pýta na úroveň pociťovaného šťastia v tak rozsiahlej miere. WVS sleduje ďalšie, pre našu analýzu potrebné premenné, čo je hlavný dôvodom, prečo budeme pracovať práve s touto databázou hodnôt.

WVS zhromažďuje dáta získané prieskumami, čo znamená, že získané dáta možno zaradiť medzi kategoriálne kvalitatívne premenné.

V konštrukcii modelu šťastia budeme používať metódu POLR. Ordered Logit model (OLM), alebo známeho pod označením „proportional odds logistic regression (POLR)“, je logistický model, používaný pre potreby, keď postupná zmena premennej znamená postupnú kvalitatívnu zmenu (Hatrak, 2007). Vysvetľovaná premenná z databázy WVS, úroveň pociťovaného šťastia, ktorá dosahuje hodnoty 1-veľmi šťastný, 2 –celkom šťastný, 3- skôr nešťastný, 4-nešťastný. Vidíme, že s rastúcou hodnotou, dochádza k zmene kvality pociťovaného šťastia. OLM umožňuje odhadnúť zmenu vysvetľovanej premennej z jednej skupiny do inej na základe zmien vo vysvetľujúcich premenných. Problém pri modelovaní podobných dát je, že nevieme aká veľké sú vzdialenosti medzi jednotlivými skupinami, alebo úrovňami kvalitatívnej premennej.

Riešením tohto problému je predpoklad, že medzi úrovňami sú proporcionálne rozdiely, alebo tiež zavedením skrytej premennej y^* , pre ktorú platí :

$$y = 1 \text{ ak } u_0 \leq y^* \leq u_1$$

$$y = 2 \text{ ak } u_1 \leq y^* \leq u_2$$

$$y = 3 \text{ ak } u_3 \leq y^* \leq u_3$$

...

$$y = n \text{ ak } u_{n-1} \leq y^* \leq u_n$$

Kde y je vysvetľovaná premenná, u je vektor nepoznaných hraníc parametrov odhadovaných pomocou hodnôt koeficientov β .

Zostrojíme model 1, u ktorého budeme brať do úvahy premenné o ktorých hovorí teória šťastia ako o významných (Becker (1974), Di Tella a spol. (2001), Veenhoven (1993), White (2007)). V modeli 2 budú zastúpené premenné, ktorých vplyv na šťastie je preštudovanej

literatúry, sporný. P-hodnoty vysvetľujúcich premenných následne porovnáme s hladinou významnosti. Významné premenné následne združíme do finálneho modelu a významnosti overíme. Až model bude obsahovať iba významné premenné pristúpime k vypočítavaniu koeficientov v prostredí jazyka R.

V modeli 1 budeme uvažovať tieto premenné: najvyššie dosiahnuté vzdelanie (VZ), rodinný stav (RS), spokojnosť s finančnou situáciou (SFS), spokojnosť s pracovným stavom (SPS), subjektívne pociťované zdravie (SPZ).

V modeli 2 budeme uvažovať tieto premenné: pohlavie respondenta (P), pracovný stav (PS), vek (V), socioekonomický status (SES), veľkosť sídla v ktorom respondent žije (DS).

V práci uvádzame iba výsledný model 3. Hodnoty koeficientov získané metódou POLR finálneho modelu 3, hovoria o tom, ako a koľko sa musí zmeniť stav vysvetľujúcej premennej, aby došlo k zmene stavu vysvetľovanej premennej o 1 triedu, za predpokladu ceteris paribus. V našom prípade pôjde o zmenu šťastia. Keďže ide o logistický model, koeficienty bude potrebné pred ich interpretáciou potrebné exponenciálne upraviť.

$$H = \beta_0 + \beta_1RS + \beta_2SFS + \beta_3SPS + \beta_4SPZ + \beta_5PS + \beta_6SES + \beta_7DS \quad (1)$$

Pristúpme k overeniu významnosti vysvetľujúcich premenných logistického modelu.

Tab. 14: Hodnoty odhadnutých koeficientov, výsledky testovania ich významnosti

	Value	Std, Error	t value	p value	p value < alfa
RS	0,0313	0,0128	2,4372	0,0148	TRUE
SFS	-0,1994	0,0120	-16,5783	0,0000	TRUE
SPS	-0,1375	0,0131	-10,4747	0,0000	TRUE
SPZ	0,6769	0,0347	19,4849	0,0000	TRUE
PS	-0,0522	0,0243	-2,1468	0,0318	TRUE
SES	0,0420	0,0204	2,0584	0,0395	TRUE
DS	0,2454	0,0652	3,7651	0,0002	TRUE

Zdroj: vlastné spracovanie

Potvrdili sa nám významnosti všetkých premenných. Opäť nemôžeme interpretovať hodnoty koeficientov, z dôvodu toho, že tento postup nám udáva hodnoty vo forme logaritmov. Pre jednoduchšie interpretovanie transformujme hodnoty koeficientov.

Ak by sme chceli v tomto prípade rozhodovať o významnosti premenných na základe t-hodnoty musíme brať v úvahu zmenu stupňov voľnosti, pretože počet premenných vzrástol z 5 na 7, čo hodnotu kvantilu Studentovho rozdelenia posunulo na 1,895.

Tab. 215: Exponenciálne upravené koeficienty modelu 3

RS	SFS	SSP	SPZ	PS	SES	DS
1,0318	0,8192	0,8716	1,9678	0,9492	1,0429	1,2782

Zdroj: vlastné spracovanie

Pôvodné koeficienty sme upravili exponenciálnou transformáciou a dostali sme koeficienty, ktoré môžeme reálne interpretovať. Hodnoty koeficientov hovoria o koľko sa musí zmeniť práve jedna vysvetľujúca premenná, aby došlo k zmene vysvetľovanej premennej o 1 jednotku.

Aby došlo k zmene šťastia o jednu jednotku, v našom prípade k prechodu z jednej skupiny do inej, musí sa za predpokladu *ceteris paribus* zmeniť premenná Rodinný stav o 1,0318 a to znamená, že sa tento stav musí zmeniť napr. z partnerstva na manželstvo, alebo z rozvedeného na partnerstvo.

V prípade skupín SFS a SSP je problematický, pretože po transformácií dát sme stratili negatívne hodnoty a všetky hodnoty nadobudli kladný charakter, ak by sme teda čerpali iba z upravených koeficientov došli by sme k tomu, že zhoršenie spokojnosti v oblasti finančnej, alebo pracovnej o 0,8192, aby došlo k nárastu šťastia o 1 jednotku.

To vyplýva z upravených dát a charakteru premenných. Preto pri faktoroch SFS a SSP budeme ďalej považovať za dôležitú hodnotu koeficientu hodnotu z upravených údajov, no so znamienkom z pôvodne vypočítaného koeficientu.

SPZ nadobúda hodnotu 1,9678, to znamená, že sa musí zmeniť zdravie jednotlivca o skoro 2 úrovne, aby to bolo zaznamenané aj na úrovni šťastia.

Pracovný stav jednotlivca je podobne problematická ako premenné SFS a SSP, tiež nadobúdala v pôvodnej analýze záporné hodnoty, čo by nemuselo hneď znamenať konflikt, keďže rast hodnoty PS nie je vyjadrením niečoho čo by prinášalo konflikt.

Socioekonomická skupiny nadobúda 4 úrovne a hodnota jej koeficientu znamená, že sa musí zmeniť o 1,0429 aby došlo k nárastu šťastia o 1 jednotku. Môžeme povedať, že ak respondent zaznamená zmenu posun v rámci spoločenských tried dochádza u neho k nárastu pocitu šťastia.

Veľkosť koeficientu pri veľkosti sídla hovorí, že ak sa zmení veľkosť sídla o 1,2782 zmení sa subjektívne šťastie o jednu jednotku, v našom prípade to hovorí, že presun jednotlivca z malého až stredného sídla do veľkého jeho šťastie poklesne.

Pre väčšiu prehľadnosť dosiahnutých výsledkov si zhrnieme výsledky do dvoch skupín, do skupiny podľa ich dopadu na subjektívne šťastie. Prvou skupinou nech bude skupina tzv. "efektívnych faktorov", skupina premenných pre ktoré platí :

$$\frac{\Delta x}{\Delta y} \geq 1 \quad (2)$$

A teda, že premenná nemusí prejsť z jednej úrovne do inej aby vyvolala zmenu úroveň pocitovaného šťastia. Do tejto skupiny môžeme zaradiť Spokojnosť s finančnou situáciou, Spokojnosť s prácou a Úroveň pracovného stavu. Tieto premenné plnia dôležitú úlohu v určovaní šťastia jednotlivca.

Druhou skupinou nech bude skupina tzv. „neefektívnych faktorov“, skupina premenných pre ktoré platí:

$$\frac{\Delta x}{\Delta y} < 1 \quad (3)$$

A teda, že premenná musí prejsť z jednej úrovne do inej, aby vyvolala zmenu úrovne pociťovaného šťastia. Do tejto skupiny zaraďujeme Rodinný stav, Subjektívne pociťované zdravie, Socioekonomický status, Veľkosť sídla.

3. Záver

V práci sme sa venovali ekonómií šťastia a hľadaním odpovede na otázku, ktoré faktory šťastia ovplyvňujú. Ekonómia šťastia je teoretickou alternatívou súčasnej ekonómie, ktorá berie za jeden z najvýznamnejších ukazovateľov úroveň subjektívne pociťovaného šťastia. Úloha šťastia bola od začiatkov zakomponovaná do ekonomickej teórie no zavedením merateľnejších veličín a komplikovanosťou merania šťastia sa jeho úloha postupne minimalizovala až nakoniec zanikla. V poslednej časti práce sme sa venovali konštrukcii ekonometrického modelu šťastia v závislosti od faktorov ako vek, pohlavie, vzdelania a pod.. Pri konštrukcii sme najprv zostrojili model 1 s premennými od ktorých sme čakali vysoký vplyv na úroveň šťastia. Model 2 tvorili premenné, ktorých vplyv nebol jasný. Významné premenné z oboch modelov sme skombinovali do modelu 3, kde sme významnosť premenných overili a vypočítali sme hodnoty koeficientov vysvetľujúcich premenných.

Vysvetľujúce premenné sme roztriedili do skupín „efektívne faktory“ a „neefektívne faktory“, podľa ich vplyvu ich zmien na subjektívne šťastie. Modelom sme chceli ukázať, že šťastie môže plniť funkciu ekonomicky merateľnej veličiny.

Príspevok bol vytvorený s podporou vedeckovýskumného projektu VEGA1/0127/11 Priestorová distribúcia chudoby v Európskej únii.

Literatúra

- BECKER, G. 1974. A Theory of Social Interactions. Journal of Political Economy. December 1974, vol. 82, s. 1063 - 1093, ISSN 1558-0970
- BRUNI, L. 2006. Civil Happiness - Economics and human flourishing in historical perspective. New York: Routledge. 2006. ISBN 978-0-415-32628-5.
- BRUNI, L., PORTA, P. (Eds.). 2007. Handbook on the Economics of Happiness. Cheltenham, Northampton: Edward Elgar. ISBN 978-1-84376-826-5
- DIENER, E. 2000 Subjective Well-Being: The Science of Happiness and Proposal for a National Index. In: American Psychologist, Vol. 55, pp. 34-43
- DI TELLA, R., MACCULLOCH, R., OSWALD, J. 2001. Preferences over Inflation and Unemployment: Evidence from Surveys of Happiness. American Economic Review, Vol. 91, s. 335 - 341. ISSN 0002-8282
- EASTERLIN, R. (1995): Will raising the incomes of all increase the happiness of all?. In: Journal of Economic Behaviour & Organization, Volume 27, Issue 1, June 1995, pp.35-47.
- EASTERLIN, R. 2005. Building a Better Theory of Well-Being. V BRUNI, L., PORTA, P. (Eds.) Economics & Happiness - Framing the Analysis. New York : Oxford University Press. 2005, s 29-64. ISBN 978-0-19-928628-7
- FREY, B. S., STUTZER, A. (2000): Happiness, Economy and Institutions. In: The Economic Journal, Vol.110, No. 446, pp. 918-938 ISSN 1424-0459
- FREY, B. S., STUTZER, A. (2001): What Can Economists Learn From Happiness Research?. In: Journal of Economic Literature, 40, 2002, pp. 402-435
- FREY, B.S.: Happiness - Happiness: A Revolution in Economics MIT Press, Cambridge, MA and London, 2008

HATRAK, M. 2007. EKONOMETRIA, Bratislava, IURA Publishing, 2007, 502s., ISBN 978-80-8078-150-7

MCGHEE, P.E. 1999. Health, Healing and the Amuse System. 2. vydanie, 256 str., 1999 ISBN 0787228478

VEENHOVEN, R. (1993), Happiness in Nations: Subjective Appreciation of Life in 56 Nations 1946±1992, Erasmus University Press, Rotterdam

WHITE, A. G. (2007): A global projection of subjective well-being: A challenge to positive psychology. In: Psych-Talk, Issue 56, March 2007, pp.17-20

WORLD VALUES SURVEY 1981-2008 OFFICIAL AGGREGATE v.20090901, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid (<http://www.wvsevsdb.com/wvs/WVSDData.jsp?Idioma=I>)

Adresy autorov:

Vladan Fejlek, Ing.
Ekonomická fakulta, TU Košice
Nemcovej 32, 040 01 Košice

Alena Tartal'ová, Mgr. PhD.
Ekonomická fakulta, TU Košice
Nemcovej 32, 040 01 Košice
alena.tartalova@tuke.sk

Zmena štruktúry spotrebného koša v závislosti od sadzby dane z pridanej hodnoty po prijatí eura

Change of consumer basket structure depending on value added tax rate after adopting euro

Veronika Zlaczka, Viliam Kováč, Jozef Heteš

Abstrakt: Prijatie nového zákonného platidla – meny euro – prinieslo viacero zmien v hospodárskej politike Slovenskej republiky. Tento príspevok má za cieľ ukázať, ako sa táto zmena prejavila v zmene štruktúry spotrebného koša z pohľadu rozličných sadzieb dane z pridanej hodnoty. Analýza poukazuje na zmenu celkových výdavkov domácnosti v závislosti od aplikácie jednotlivých sadzieb tejto dane, ktoré sa na nich uplatňujú. Rozbor výdavkov domácností prezentuje aj zmenu daňového zaťaženia na základe danej dane.

Abstract: The adoption of the new legal currency – euro – has brought a number of changes in the economic policy of the Slovak Republic. This contribution aims to show how this change is reflected in the change of the structure of the consumer basket in terms of the different value added tax rates. The analysis indicates also a change in the total household expenditures, depending on the application of the individual rates of the mentioned tax that are applicable to them. Study of the household expenditures represents a change in the tax burden on the basis of this tax too.

Kľúčové slová: spotrebný koš, daň z pridanej hodnoty, sadzba dane z pridanej hodnoty, daňové zaťaženie.

Key words: consumption basket, value added tax, value added tax rate, tax burden.

JEL classification: E21, E62.

1. Úvod

Vstup Slovenskej republiky do Hospodárskej a menovej únie – Economic and Monetary Union – znamenal výraznú zmenu v hospodárskej politike krajiny. Prijatie novej meny euro dňa 1. 1. 2009 značne ovplyvnilo pohľad na domácu cenotvorbu.

Daňový systém Slovenskej republiky má veľmi bohatú históriu zmien. Viaceré vykonané reformy neobišli ani daň z pridanej hodnoty. Od vzniku štátu prešiel niekoľkými modifikáciami, ktoré boli dôsledkom každoročných právnych úprav a novelizácii. Tieto reformy a novelizácie ovplyvnili vývoj sadzby jednej z najdôležitejších daní, akou daň z pridanej hodnoty nepochybne je. Legislatívna úprava tejto dane sa nachádza v Zákone 222 z roku 2004 o dani z pridanej hodnoty ([8]). Jeho úplné znenie je potrebné vnímať spolu so znením neskoršie prijatých právnych noriem, ktoré ho modifikovali.

2. Vývoj sadzby dane z pridanej hodnoty

Daň z pridanej hodnoty je typickým reprezentantom nepriamej dane. V minulosti existovali a aj v súčasnosti sa uplatňujú viaceré úrovne sadzby dane z pridanej hodnoty. Obvykle sa označujú ako základná sadzba, znížená sadzba a znížená sadzba pre vybrané produkty. Najnižšia sadzba bola v platnosti len obmedzenú dobu. V rokoch 2009 až 2010 sa aplikovala na takzvaný predaj z dvora, pod ktorým sa rozumel predaj výlučne doma vytvorených poľnohospodárskych produktov.

V roku 2004 bola prijatá daňová reforma, ktorá stanovila jednotnú sadzbu dane z pridanej hodnoty vo výške 19 %. Tým došlo k zrušeniu zníženej sadzby dane z pridanej hodnoty, ale

zároveň sa v značnej miere zjednodušila daňová sústava Slovenskej republiky. Takúto podobu si však slovenský daňový systém nenechal dlhú dobu, pretože už od dňa 1. 1. 2007 bola opätovne zavedená nižšia sadzba dane z pridanej hodnoty vo výške 10 %. V poslednom roku obehu slovenskej koruny ako zákonného platidla na území Slovenskej republiky, v roku 2008, bola teda základná sadzba dane z pridanej hodnoty stanovená na úrovni 19 % a znížená sadzba na úrovni 10 %. Dňom 1. 1. 2009, teda dňom zavedenia eura ako zákonného platidla na území Slovenskej republiky vznikla aj druhá znížená sadzba dane z pridanej hodnoty, a to vo výške 6 %, ktorá sa uplatňovala len na takzvaný predaj z dvora. Uvedené sadzby platili do dňa 31. 12. 2010, kedy bola od nasledujúceho roku – od 1. 1. 2011 – základná sadzba zvýšená o 1 percentuálny bod na hodnotu 20 % a zároveň bola zrušená najnižšia sadzba pre predaj domácich poľnohospodárskych produktov. Od tejto zmeny sa už sadzby dane z pridanej hodnoty doposiaľ nemenili.

Popri tovaroch a službách, ktoré podliehajú dani z pridanej hodnoty, sa nachádzajú na trhu aj komodity, ktoré sú oslobodené od tejto dane. Prakticky to znamená, že sa uplatňuje sadzba vo výške 0 %.

3. Spotrebný kôš

Spotrebný kôš predstavuje množinu reprezentatívnych statkov, na základe ktorej sa zisťuje miera zmeny cenovej hladiny v ekonomike. Obsahuje podiely jednotlivých druhov tovarov a služieb, ktoré sú v daných pomeroch obstarávané konečnými zákazníkmi.

Všetky komodity možno zatriediť na základe klasifikácie individuálnej spotreby podľa účelu použitia do dvanástich skupín – skupina 1 – potraviny a nealkoholické nápoje, skupina 2 – alkoholické nápoje a tabak, skupina 3 – odevy a obuv, skupina 4 – bývanie, voda, elektrina, plyn a iné palivá, skupina 5 – nábytok, bytové zariadenie a bežná údržba domu, skupina 6 – zdravotníctvo, skupina 7 – doprava, skupina 8 – pošta a spoje, skupina 9 – rekreácie a kultúra, skupina 10 – vzdelávanie, skupina 11 – hotely a reštaurácie, skupina 12 – rozličné tovary a služby.

Štandard klasifikácie – v origináli nazvaný Classification of Individual Consumption according to Purpose – známy pod akronymom COICOP obsahuje aj skupinu označenú 0, ktorá predstavuje celkový súhrn, teda zahŕňa všetkých dvanásť skupín. Pre úplnosť je potrebné dodať, že existujú aj skupiny 13 a 14, ktoré vyjadrujú individuálnu konečnú spotrebu neziskových organizácií slúžiacich domácnostiam a individuálnu konečnú spotrebu vlády. Všetky uvedené skupiny sa ďalej rozčleňujú na menšie kategórie vo viacerých úrovniach ([1]). Každá krajina si môže túto klasifikáciu adaptovať na vlastnú ekonomiku.

Štatistický úrad Slovenskej republiky na základe údajov získaných od slovenských obchodníkov klasifikáciu upravil tak, ako je to uvedené v tabuľke 1 Štruktúra spotrebného koša na základe klasifikácie individuálnej spotreby podľa účelu použitia.

Ako je možné vidieť v tabuľke 1 Štruktúra spotrebného koša na základe klasifikácie individuálnej spotreby podľa účelu použitia, Štatistický úrad od roku 2008 každoročne upravuje reprezentatívny spotrebný kôš. V žiadnej skupine sa v sledovanom období v medziročnom ponímaní nestalo, aby si ponechala svoju pôvodnú váhu z predchádzajúceho roku. Najvyššia váha je priradená skupine 4, ktorá predstavuje výdavky na bývanie a médiá, kam patria voda, elektrina, plyn a iné palivá. Na druhej strane rebríčka stojí skupina 10, ktorú tvoria výdavky na vzdelávanie.

Tabuľka 16: Štruktúra spotrebného koša na základe klasifikácie individuálnej spotreby podľa účelu použitia

skupina	rok					
	2008	2009	2010	2011	2012	2013
1	16,055 %	16,068 %	18,445 %	17,573 %	16,345 %	16,698 %
2	4,249 %	4,345 %	4,375 %	4,973 %	4,924 %	4,959 %
3	4,509 %	4,482 %	4,623 %	4,422 %	4,395 %	4,308 %
4	28,175 %	28,237 %	25,383 %	27,377 %	27,723 %	27,940 %
5	5,786 %	5,833 %	7,055 %	6,504 %	6,537 %	6,397 %
6	2,632 %	2,928 %	3,711 %	4,333 %	4,342 %	4,509 %
7	9,125 %	9,087 %	7,872 %	7,913 %	8,489 %	7,936 %
8	3,589 %	3,578 %	3,761 %	3,759 %	3,659 %	3,554 %
9	7,823 %	7,895 %	8,463 %	8,467 %	8,296 %	8,317 %
10	1,625 %	1,563 %	1,456 %	1,697 %	1,681 %	1,777 %
11	7,224 %	6,874 %	6,861 %	6,318 %	6,248 %	6,083 %
12	9,208 %	9,112 %	7,995 %	6,664 %	7,359 %	7,522 %
0	100 %	100 %	100 %	100 %	100 %	100 %

Zdroj: vlastné spracovanie podľa údajov Štatistického úradu Slovenskej republiky ([2], [3], [4], [5], [6], [7])

Pri pohľade na rozvrstvenie skupín klasifikácie individuálnej spotreby podľa účelu použitia možno konštatovať, že z daného rozdelenia nie sú jednoznačne stanoviteľné priradené sadzby dane z pridanej hodnoty. Na tovary a služby zahrnuté v skupinách 1, 2, 3, 5, 7 a 11 je uplatňovaná základná sadzba dane z pridanej hodnoty, teda na úrovni 19 %, respektíve 20 %. Skupiny 4, 8 a 9 obsahujú okrem komodít so základnou sadzbou aj tovary a služby, ktoré sú oslobodené od platby dane z pridanej hodnoty. V skupine 6 sa popri základnej sadzbe dane z pridanej hodnoty aplikuje aj znížená sadzba dane z pridanej hodnoty vo výške 10 %. Napokon skupiny 10 a 12 predstavujú súbory tovarov a služieb, ktoré podliehajú všetkým trom spôsobom uplatňovania dane z pridanej hodnoty vo forme základnej sadzby, zníženej sadzby a aj oslobodenia od dane.

4. Analýza daňového zaťaženia domácností daňou z pridanej hodnoty

V našej analýze sa budeme venovať rozboru daňového zaťaženia domácností daňou z pridanej hodnoty. Z výdavkov domácností sme identifikovali podiel výdavkov, ktorý pripadá na hodnotu odvedenej dane z pridanej hodnoty a rozdelili sme túto sumu vzhľadom k aplikácii oboch úrovní daňovej sadzby. Pre observáciu sme si zvolili obdobie začínajúce rokom 2008 z dôvodu, že tento rok predstavoval posledný rok platnosti slovenskej koruny v normálnom obeh.

Pre vykonané kalkulácie sme existenciu druhej zníženej sadzby aplikovali tak, že sme ju priradili do vyššej zníženej sadzby. Túto malú modifikáciu sme vykonali z dôvodu, že predaj vlastných poľnohospodárskych produktov z dvora predstavuje len veľmi malý zlomok celkového predaja, a teda z celkového pohľadu ide o zanedbateľný podiel.

Tabuľka 17: Výdavky domácností na komodity podľa úrovne zaťaženia daňou z pridanej hodnoty

rok	výdavky			
	základná sadzba	znížená sadzba	oslobodené	celkom
2008	3669,05 EUR	84,48 EUR	758,76 EUR	4512,29 EUR
2009	3496,21 EUR	62,95 EUR	784,63 EUR	4343,79 EUR
2010	3548,49 EUR	69,52 EUR	709,12 EUR	4327,13 EUR
2011	3745,08 EUR	59,41 EUR	612,10 EUR	4416,59 EUR
2012	3736,31 EUR	93,83 EUR	650,63 EUR	4480,77 EUR

Zdroj: vlastné spracovanie podľa údajov Štatistického úradu Slovenskej republiky

Tabuľka 2 Výdavky domácností na komodity podľa úrovne zaťaženia daňou z pridanej hodnoty prezentuje sumy výdavkov, ktoré boli vynaložené domácnosťami na obstaranie tovarov a služieb s príslušnou sadzbou dane z pridanej hodnoty, prípadne od dane oslobodených tovarov a služieb. V poslednom roku platnosti slovenskej koruny – v roku 2008 – celkové priemerné výdavky domácnosti predstavovali 4512,29 EUR, z čoho 3669,05 EUR minuli na tovary a služby, pre ktoré platí základná sadzba dane z pridanej hodnoty, výdavky vo výške 84,48 EUR pripadali na komodity zaťažené zníženou sadzbou a 758,76 EUR tvorili výdavky na tovary a služby oslobodené od dane. Po prijatí eura sa celkové výdavky znížili na 4343,79 EUR. Taktiež sa znížili výdavky na tovary a služby so základnou sadzbou na 3496,21 EUR a výdavky na tovary a služby so zníženou sadzbou na 62,95 EUR. Opačný trend možno vidieť vo výdavkoch na tovary a služby oslobodené od dane z pridanej hodnoty – suma výdavkov vzrástla na 784,63 EUR. Podobné smerovanie pokračovalo aj v ďalšom roku 2010 len s rozdielom, že celkové výdavky klesli na 4327,13 EUR. Výdavky v závislosti od jednotlivých sadziieb mali rovnakú tendenciu vývoja. Pozoruhodný skok nastal v roku 2012, kedy sa medziročne zmenili výdavky na komodity s aplikovanou zníženou sadzbou dane z pridanej hodnoty zo sumy 59,41 EUR na 93,83 EUR, čo predstavuje nárast o 57,94 %. Hoci výdavky zviazané so štandardnou sadzbou klesli z 3745,08 EUR na 3736,31 EUR, z pohľadu porovnania celkových výdavkov pri uvedenej medziročnej zmene došlo k nárastu ich hodnoty z 4416,59 EUR na 4480,77 EUR. Tento rozdiel sa pretavil práve do spomínaných rastúcich výdavkov a taktiež do zvýšeného nákupu oslobodených komodít, kde sa minutá suma zdvihla z 612,10 EUR na 650,63 EUR. Z toho vyplýva, že zvýšená medziročná spotreba sa približne rovnomerne prejavila v oboch nárastoch.

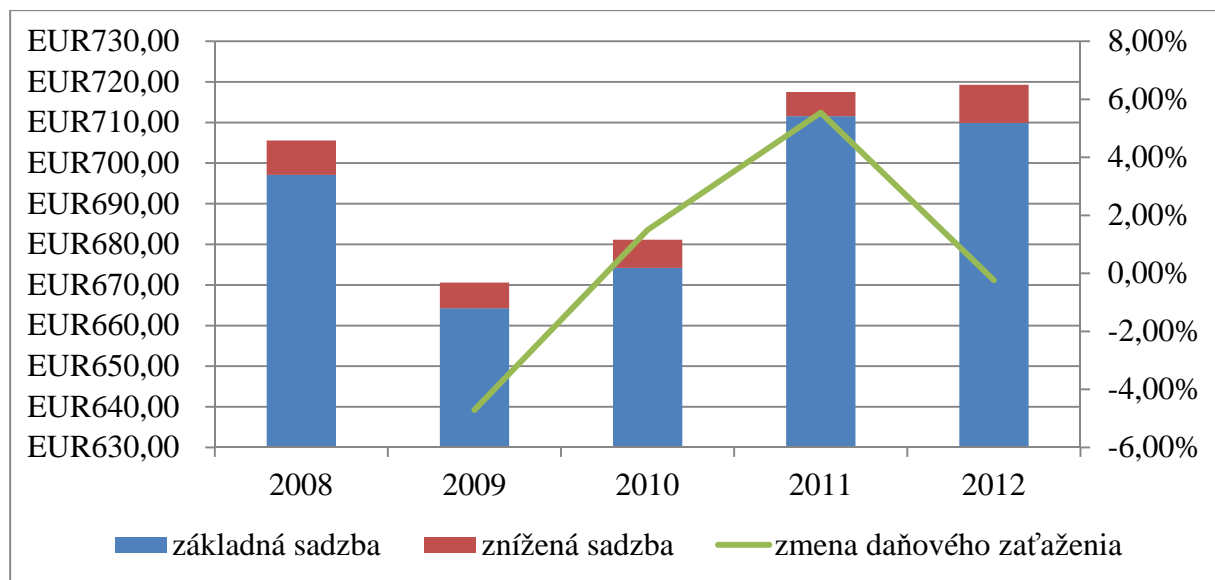
Tabuľka 18: Výdavky domácností na daň z pridanej hodnoty podľa úrovne daňového zaťaženia komodít

rok	daňová povinnosť			podiel daňovej povinnosti na celkových výdavkoch
	základná sadzba	znížená sadzba	celkom	
2008	697,12 EUR	8,45 EUR	705,57 EUR	15,64 %
2009	664,28 EUR	6,30 EUR	670,57 EUR	15,44 %
2010	674,21 EUR	6,95 EUR	681,17 EUR	15,74 %
2011	711,57 EUR	5,94 EUR	717,51 EUR	16,25 %
2012	709,90 EUR	9,38 EUR	719,28 EUR	16,05 %

Zdroj: vlastné spracovanie podľa údajov Štatistického úradu Slovenskej republiky

Tabuľka 3 Výdavky domácností na daň z pridanej hodnoty podľa úrovne daňového zaťaženia komodít zobrazuje sumy výdavkov, ktorú obsahuje množina celkových výdavkov domácností pripadajúcich na odvedenú daň z pridanej hodnoty predávajúcou stranou. Pred prijatím eura za zákonné platidlo domácnosti v roku 2008 v priemere minuli na daň z pridanej hodnoty 705,57 EUR, z čoho 697,12 EUR pripadalo na sumu odvedenú na základe aplikácie základnej sadzby a 8,45 EUR na sumu odvedenú na základe aplikácie zníženej sadzby. V prvom roku obehu eura – v roku 2009 – sa znížila prenesená daňová povinnosť domácností na hodnotu 670,57 EUR, pričom 664,28 EUR pripadalo na základnú sadzbu a 6,30 EUR na zníženú sadzbu. V nasledujúcom roku došlo k zvýšeniu daňového zaťaženia domácností na 681,17 EUR. Z tejto sumy 674,21 EUR predstavovalo odvod vo forme základnej sadzby a 6,95 EUR odvod vo forme zníženej sadzby. Približne rovnaký rastúci trend nasledoval aj v ďalších rokoch sledovaného obdobia.

Obrázok 23: Vývoj výdavkov domácností na daň z pridanej hodnoty podľa úrovne daňového zaťaženia komodít



Zdroj: vlastné spracovanie podľa údajov Štatistického úradu Slovenskej republiky

Ako vidíme na uvedenom grafe Vývoj výdavkov domácností na daň z pridanej hodnoty podľa úrovne daňového zaťaženia komodít, daňový výber štátu zo zníženej sadzby dane z pridanej hodnoty predstavuje len veľmi malý podiel celkového daňového výberu za daň z pridanej hodnoty.

Zaujímavo sa javí krivka zmeny daňového zaťaženia. Po prijatí eura v roku 2009 došlo k medziročnej zmene celkového daňového výberu o 4,71 % smerom nadol. Teda štát zinkasoval z tejto dane menej ako za rok predtým – teda za rok 2008, kedy sa platilo na území Slovenskej republiky ešte slovenskou korunou. Od zmeny meny však daňové zaťaženie nabralo rastúcu tendenciu a v roku 2010 sa zvýšilo o 1,50 % a v roku 2011 o 5,54 %. Až v roku 2012 došlo k jeho stagnácii, kedy sa medziročne znížilo o 0,23 %.

V prípade, ak chceme skúmať na priemerné daňové zaťaženie vzhľadom na rôzne sadzby dane z pridanej hodnoty, mali by sme zamerať svoju pozornosť na hodnotu podielu daňovej povinnosti na celkových výdavkoch domácností. V roku 2008 tvorila celková daňová povinnosť podiel na výdavkoch 15,64 %. Po prijatí eura sa v roku 2009 nepatrne znížila na 15,44 %, v roku 2010 pre zmenu veľmi ľahko zvýšila na 15,74 %, v roku 2011 opäť rast na 16,25 % a napokon v roku 2012 zanedbateľný pokles na 16,05 %. Z uvedených čísel je

zrejme, že zmena meny nemala žiadny vplyv na zmenu priemernej daňovej povinnosti z dani z pridanej hodnoty.

5. Záver

Výsledky našej analýzy ukazujú, že zmena zákonného platidla v Slovenskej republike neprinesla výraznejšie zmeny štruktúry spotrebného koša. Modifikácie jednotlivých tried klasifikácie individuálnej spotreby podľa účelu použitia sa udiali nanajvyš v sume niekoľkých percentuálnych bodov, pričom v roku po zmene meny a v nasledujúcom roku sa medziročné úpravy nevymykali hodnotám bežných medziročných úprav z iných období. Na základe nášho výskumu možno konštatovať, že prijatie eura nemalo vplyv na zmeny štruktúry spotrebného koša prostredníctvom aplikácie rozličných sadzieb dane z pridanej hodnoty. Dokonca sa nijako štatisticky výrazne nezmenili ani celkové výdavky domácnosti po zavedení eura.

Literatúra

- [1] United Nations Publications Board: COICOP; New York: United Nations Publications Board; 2014; <http://unstats.un.org/unsd/cr/registry/regcst.asp?CI=5>
- [2] Spotrebný kôš 2008; Bratislava: Štatistický úrad Slovenskej republiky
- [3] Spotrebný kôš 2019; Bratislava: Štatistický úrad Slovenskej republiky
- [4] Spotrebný kôš 2010; Bratislava: Štatistický úrad Slovenskej republiky
- [5] Spotrebný kôš 2011; Bratislava: Štatistický úrad Slovenskej republiky
- [6] Spotrebný kôš 2012; Bratislava: Štatistický úrad Slovenskej republiky
- [7] Spotrebný kôš 2013; Bratislava: Štatistický úrad Slovenskej republiky; <http://portalkb.statistics.sk/html/showdoc.dodocid=61889.html>
- [8] Zákon 222 o dani z pridanej hodnoty; http://jaspi.justice.gov.sk/jaspiw1/htm_zak/jaspiw_mini_zak_zobraz_clanok1.asp?kotva=k1&skupina=1

Adresy autorov:

Veronika Zlaczka, Ing.
Katedra financií, Ekonomická fakulta, Technická
univerzita v Košiciach
Němcovej 32, Košice, Slovensko
veronika.zlaczka@tuke.sk

Viliam Kováč, Ing.
Katedra financií, Ekonomická fakulta,
Technická univerzita v Košiciach
Němcovej 32, Košice, Slovensko
viliam.kovac@tuke.sk

Jozef Heteš, Ing.
Katedra bankovníctva a investovania, Ekonomická
fakulta, Technická univerzita v Košiciach
Němcovej 32, Košice, Slovensko
jozef.hetes@tuke.sk

Faktory pracovných a podnikateľských príjmov v krajinách EÚ Factors of Labour and Entrepreneurial Income in the EU Countries

Tomáš Želinský

Abstract: The aim of this paper is to identify potential factors of labour (entrepreneurial) income level in the European Union countries. Analysis is based on a model of quantile regression using EU SILC 2012 microdata. According to the results all considered factors (gender, years spent in paid work, degree of urbanisation, education and status in employment) have statistically significant effect on income level.

Abstrakt: Cieľom príspevku je identifikovať potenciálne faktory výšky pracovného (podnikateľského) príjmu v krajinách Európskej únie. Analýza je založená na modeli kvantilovej regresie s použitím mikroúdajov zisťovania EU SILC 2012. Podľa výsledkov majú všetky uvažované faktory (pohlavie, roky praxe, stupeň urbanizácie, vzdelanie a postavenie v zamestnaní) štatisticky významný vplyv na výšku príjmu.

Key words: labour/entrepreneurial income, quantile regression, EU SILC.

Kľúčové slová: pracovný/podnikateľský príjem, kvantilová regresia, EU SILC.

JEL classification: D31, I31.

1. Úvod

Na modelovanie vzťahu medzi príjmom a jeho faktormi je možné použiť napr. tzv. Mincerovu rovnicu, v ktorej sa na strane vysvetľujúcich premenných nachádza ľudský kapitál charakterizovaný počtom rokov, počas ktorých sa jednotliviec vzdelával a potenciálnym počtom rokov pracovných skúseností. Pôvodné Mincerove (1958; 1974) myšlienky boli rozpracované a analyzované mnohými autormi (Lemieux, 2006), pričom za základný tvar Mincerovej rovnice je považovaný nasledovný model (Chiswick, 2006):

$$\ln E_i = b_0 + b_1 S_i + b_2 T_i + b_3 T_i^2 + u_i, \quad (1)$$

kde

E_i je príjem i -tej osoby,

S_i je počet rokov, počas ktorých sa i -tá osoba vzdelávala,

T_i je potenciálny počet rokov pracovných skúseností po škole,

b_k sú odhadované koeficienty, $k = 1, 2, 3$,

u_i je náhodná zložka.

Detailným skúmaním teoretických a empirických zdrojov dochádza Lemieux (2006) k záveru, že jednoduchá forma Mincerovej rovnice (1) je po zohľadnení určitých špecifik aj v súčasnosti vhodným nástrojom na analýzu vzťahu medzi ľudským kapitálom a príjmom.

Cieľom príspevku je s použitím kvantilovej regresie identifikovať faktory potenciálne vplyvajúce na výšku pracovných (a/alebo podnikateľských) príjmov v krajinách Európskej únie a stručne charakterizovať smer a silu štatistického vplyvu vybraných ukazovateľov na výšku príjmu pozdĺž vybraných kvantilov príjmového rozdelenia.

2. Model

Podobne ako napr. Su a Heshmati (2013) vychádzame z rozšírenej Mincerovej rovnice, t. j. pôvodná rovnica (1) je doplnená o vybrané charakteristiky jednotlivca:

$$\ln y_i = a + \mathbf{x}_i \cdot \boldsymbol{\beta} + u_i, \quad (2)$$

kde

$\ln y_i$ je prirodzený logaritmus ročného príjmu i -tej osoby v EUR,

a, b sú odhadované parametre,

x_i je i -tý riadok matice vysvetľujúcich premenných X .

V modeli (2) uvažujeme s nasledovnými vysvetľujúcimi premennými¹:

$poohl$ umelá (dummy) premenná s kategóriami: 1: muž (ref.), 2: žena,

$roky$ počet rokov strávených v platenej práci ako zamestnanec alebo zárobkovo činná osoba²,

$roky^2$ kvadrát premennej roky,

$urban$ stupeň urbanizácie, umelá premenná s kategóriami: 1: územie s hustým osídlením, 2: územie s priemerne hustým osídlením (ref.), 3: územie s riedkym osídlením,

vzd vzdelanie, pričom pôvodné kategórie umelej premennej „najvyššia dosiahnutá úroveň vzdelania podľa ISCED“ boli zlúčené do nových: ZŠ: primárne vzdelanie, SŠ: nižšie, (vyššie) a post-sekundárne vzdelanie (ref.), VŠ: 1. a 2. stupeň terciárneho vzdelania,

$post$ ekonomické postavenie, umelá premenná s kategóriami: 1: podnikateľ / SZČO so zamestnancami (ref.), 2: podnikateľ/SZČO bez zamestnancov, 3: zamestnanec pracujúci za mzdu.

3. Odhad modelu

Model je odhadnutý použitím mikroúdajov zisťovania EU SILC 2012 (Eurostat, 2014) za všetky krajiny EÚ, pričom celkový pracovný/podnikateľský príjem je určený ako súčet nasledovných typov príjmu: peňažný príjem zo zamestnania alebo jemu blízky príjem (PY010G), nepeňažný príjem zo zamestnania (PY020G) a peňažné zisky alebo straty zo samostatnej zárobkovej činnosti (vrátane honorárov) (PY050G).

Vzhľadom na skutočnosť, že vysvetľovaná premenná obsahuje extrémne hodnoty (outliers), metóda najmenších štvorcov (OLS, z angl. *ordinary least squares*) nie je najvhodnejšou metódou na analyzovanie determinantov príjmov. Jednou z vhodných metód je kvantilová regresia, ktorá je robustná voči extrémnym hodnotám. Navyše, pomocou kvantilovej regresie dokážeme zhodnotiť vzťah medzi výškou príjmu a jeho determinantmi pozdĺž rozdelenia príjmu (Koenker a Bassett, 1978; Koenker a Hallock, 2001), resp. zjednodušene povedané, dokážeme určiť, ako pôsobia vysvetľujúce premenné na nízke príjmy a ako na vysoké príjmy.

Vzťah medzi príjmom a jeho determinantmi je skúmaný pre 25., 50., 75. a 90. percentil rozdelenia príjmov. Odhad štandardných chýb koeficientov a testy ich významnosti sú založené na bootstrappingu (ParzenWei a Ying, 1994). Diagnostika odhadnutých modelov kvantilovej regresie je založená na miere $R^1(\tau)$ (analógia k tradičnému koeficientu determinácie). Táto diagnostická metóda bola navrhnutá Koenkerom a Machadom (1999). Koeficient $R^1(\tau)$ meria relatívny „úspech“ modelu kvantilovej regresie na príslušnom kvantile, $R^1(\tau)$ je teda lokálnou mierou vhodnosti použitia modelu pre príslušný kvantil. Tvrdíme, že úplný model (5.18) je na kvantile τ lepší ako model obsahujúci len úrovňový

¹ Pri každej z umelých premenných je referenčná kategória indikovaná ako „(ref.)“ za príslušnou referenčnou kategóriou. Výsledky týkajúce sa ostatných kategórií príslušnej premennej sa interpretujú porovnaním vo vzťahu k referenčnej kategórii.

² Kým Mincerova rovnica predpokladá potenciálny počet rokov praxe, v našom modeli použijeme skutočný počet rokov, nakoľko táto premenná je súčasťou zisťovania EU SILC.

konštantu, ak vplyvom regresorov došlo k významnej zmene i -tej podmienenej kvantilovej funkcie. Koenker a Machado (1999) skúmali správanie navrhnutých mier pomocou simulovaných údajov a podľa ich zistení $R^1(\tau)$ nedáva dobré výsledky (v zmysle, že vykazuje nízke hodnoty) v prípade veľkej variability údajov (týka sa aj údajov o príjme).

Všetky odhady v práci sú uskutočnené v prostredí softvéru R (R Core Team, 2014) s použitím knižnice `quantreg` (Koenker, 2013).

4. Výsledky a diskusia

Odhadnuté modely nám umožňujú identifikovať faktory (vrátane smeru a veľkosti ich vplyvu) pôsobiace na príjmy pozdĺž jednotlivých kvantilov príjmového rozdelenia (Želinský, 2014). Výsledky sú prezentované v tabuľke 1, v ktorej z dôvodu úspory priestoru uvádzame len odhadnuté regresné koeficienty s indikovaním ich štatistickej významnosti (hladina významnosti 0,001: „***“; hladina významnosti 0,01: „**“; hladina významnosti 0,05: „*“ a hladina významnosti 0,1: „.“), kvôli úspore neuvádzame ani odhady úrovňových konštánt.

Odhady regresných koeficientov na príslušnom kvantile indikujú vplyv na príjem pri jednotkovej zmene hodnoty vysvetľujúcej premennej za predpokladu, že hodnoty ostatných premenných sa nemenia.

Pohlavie

Premenná pohlavie je štatisticky významná vo všetkých kvantiloch vo všetkých analyzovaných krajinách a odhadnuté koeficienty nadobúdajú výlučne záporné hodnoty. Vo všeobecnosti to teda znamená, že muži majú vyššie príjmy ako ženy.

Z výsledkov je ďalej zrejmé, že veľkosť vplyvu pohlavia pozdĺž rozdelenia príjmov sa znižuje vo viac ako polovici krajín. Najväčšie rozdiely medzi príjmami mužov a žien na prvom kvantile sa prejavili v krajinách ako Nemecko, Holandsko, Rakúsko, Spojené kráľovstvo a Cyprus, naopak najmenšie v Dánsku, Lotyšsku, Slovinsku, Bulharsku a Litve. Najväčšie rozdiely na treťom kvantile sú v Estónsku, Islande, Nórsku, Spojenom kráľovstve a Nemecku a najmenšie v Maďarsku, na Slovinsku, v Luxembursku, na Malte a v Estónsku.

Podobne ako množstvo iných štúdií zaoberajúcich sa skúmaním vzťahu medzi príjmom a pohlavím, aj výsledky tejto analýzy podporujú predpoklad diskriminácie žien v odmeňovaní na trhu práce, nakoľko príjmy žien sú nižšie ako príjmy mužov vo všetkých kvantiloch.

Roky praxe

Teória (Mincerova rovnica) predpokladá kvadratický vzťah medzi počtom rokov praxe¹ a príjmom. To samozrejme nemusí nutne znamenať, že po dosiahnutí bodu obratu začínajú príjmy pracujúcich klesať (nad touto úvahou by malo väčší zmysel zamýšľať sa, keby súčasťou súboru boli aj dôchodcovia, keďže po ukončení pracovného života príjmy spravidla klesajú), ale že rast príjmov sa bude v čase spomaľovať, šlo by teda o konkávny rast. Odhadnutý regresný koeficient pri kvadratickom člene premennej roky je pozdĺž všetkých kvantilov vo všetkých krajinách záporný a lineárny člen kladný, čo naznačuje, že rast príjmov v závislosti od počtu rokov praxe (a teda aj veku) je skutočne konkávny.

V prípade všetkých krajín platí, že vplyv rokov praxe pozdĺž kvantilov klesá (s výnimkou Maďarska, kde odhadnuté koeficienty neboli štatisticky významné).

¹ V prípade krajín Dánsko, Fínsko, Francúzsko, Island, Litva, Nórsko a Švédsko nebola premenná počet rokov praxe dostupná, a tak bola nahradená vekom respondenta.

Stupeň urbanizácie

Zahrnutím premennej stupeň urbanizácie sledujeme, do akej miery môže hustota osídlenia súvisieť s výškou príjmu ekonomicky aktívnej osoby. Referenčnou kategóriou v tomto prípade bola kategória „územie s priemerne hustým osídlením“¹.

Výsledky naznačujú, že vo väčšine krajín majú osoby žijúce na územiach s hustým osídlením spravidla vyššie príjmy pozdĺž celého príjmového rozdelenia. Výnimku predstavujú Rakúsko, Dánsko, Grécko a Švédsko, kde toto tvrdenie neplatí. U osôb žijúcich na územiach s riedkym osídlením platí, že majú spravidla nižšie príjmy pozdĺž celého príjmového rozdelenia (s výnimkou Spojeného kráľovstva a Luxemburska, kde majú spravidla vyššie príjmy) v porovnaní s osobami žijúcimi na územiach s priemerne hustým osídlením.

Vzdelanie

Medzi príjmom a vzdelaním sa prejavil očakávaný vzťah – ľudia so základným (príp. žiadnym) vzdelaním (ako najvyšším dosiahnutým) dosahujú na jednotlivých kvantiloch nižšie príjmy ako ľudia so stredoškolským vzdelaním a s rovnakými ostatnými charakteristikami.

Osoby so základným a nižším vzdelaním majú na prvom kvartile v priemere o tretinu a na treťom kvartile o štvrtinu nižší príjem ako osoby so stredoškolským vzdelaním a naopak osoby s vysokoškolským vzdelaním majú pozdĺž všetkých kvartilov v priemere o 40 % vyšší príjem ako osoby so stredoškolským vzdelaním.

Postavenie v zamestnaní

Referenčnou kategóriou v rámci premennej postavenie v zamestnaní je „podnikateľ so zamestnancami“. Príjmy podnikateľov bez zamestnancov sú pozdĺž celého príjmového rozdelenia nižšie a štatistický vplyv tejto kategórie na úroveň príjmov pozdĺž rozdelenia vo väčšine krajín slabne. Osoby pracujúce za mzdu, dosahujú vo viacerých krajinách v nízkych kvantiloch v priemere vyššie príjmy ako podnikatelia so zamestnancami. Následne dochádza k obratu a vo vyšších kvantiloch sú príjmy osôb pracujúcich za mzdu výrazne nižšie ako príjmy podnikateľov so zamestnancami.

Takéto zistenia však nemožno považovať za prekvapujúce, nakoľko v prípade podnikateľov môžu príjmy značne kolísať a najvyššie hodnoty príjmu (a teda patriace do vyšších kvartilov) možno očakávať práve u podnikateľov so zamestnancami.

5. Záver

Cieľom príspevku bolo identifikovať faktory potenciálne vplyvajúce na výšku pracovných (a/alebo podnikateľských) príjmov v krajinách Európskej únie a stručne charakterizovať smer a silu štatistického vplyvu vybraných ukazovateľov na výšku príjmu pozdĺž vybraných kvartilov príjmového rozdelenia.

Podobne ako iné výskumy, aj táto štúdia dospela k zisteniam, že existuje významná diferencovanosť medzi príjmami mužov a žien, s rastúcim počtom rokov praxe (a v konečnom dôsledku aj vekom) je spojený konkávny rast príjmov, ľudia žijúci na územiach s hustým osídlením majú spravidla vyššie príjmy ako ostatní a výška príjmu je v pozitívnom vzťahu s najvyšším dosiahnutým vzdelaním.

PodĎakovanie

Napísanie príspevku bolo podporené Vedeckou grantovou agentúrou MŠ SR a SAV v rámci riešenia vedecko-výskumného projektu VEGA 1/0127/11 Priestorová distribúcia

¹ V prípade Islandu a Lotyšska je referenčnou kategóriou „územie s hustým osídlením“, za krajiny Malta, Holandsko a Slovinsko táto premenná nebola dostupná.

chudoby v EÚ. V príspevku boli použité anonymizované z výberového zisťovania EU SILC poskytnuté na výskumné účely na základe kontraktu č. EU-SILC/2011/33 podpísaného medzi Európskou komisiou, Eurostatom a Technickou univerzitou v Košiciach.

Literatúra

- EUROSTAT. 2014. EU SILC 2014 UDB ver 2012-1 from 01-03-14, Cross-sectional data. Luxembourg: European Commission, Eurostat.
- CHISWICK, B. R. 2006. Jacob Mincer, Experience and the Distribution of Earnings. In: GROSSBARD, S. (ed.): *Jacob Mincer: A Pioneer of Modern Labor Economics*, s. 109-126. New York: Springer.
- KOENKER, R. 2013. *quantreg: Quantile Regression*. R package version 5.05.
- KOENKER, R., BASSETT, G. 1978. Regression Quantiles. In: *Econometrica*, Vol. 46, No. 1, pp. 33-50.
- KOENKER, R., HALLOCK, K. 2001. Quantile Regression: An Introduction. In: *Journal of Economic Perspectives*, Vol. 15, No. 4, pp.143-156.
- KOENKER, R., MACHADO, J. A. F. 1999. Goodness of Fit and Related Inference Processes for Quantile Regression. In: *Journal of the American Statistical Association*, Vol. 94, No. 448, pp. 1296-1310.
- LEMIEUX, T. 2006. The „Mincer Equation“ Thirty Years After Schooling, Experience, and Earnings. In: GROSSBARD, S. (ed.): *Jacob Mincer: A Pioneer of Modern Labor Economics*, pp. 127-145. New York: Springer.
- MINCER, J. A. 1958. Investment in Human Capital and Personal Income Distribution. In: *Journal of Political Economy*, Vol. 66, No. 4, pp. 281-302.
- MINCER, J. A. 1974. *Schooling, Experience, and Earnings*. New York: Columbia University Press and National Bureau of Economic Research.
- PARZEN, M. I., WEI, L., YING, Z. 1994. A resampling method based on pivotal estimating functions. In: *Biometrika*, Vol. 81, No. 2, pp. 341-350.
- R CORE TEAM. 2014. *R: A language and environment for statistical computing*. Viedeň: R Foundation for Statistical Computing.
- SU, B., HESHMATI, A. 2013. *Analysis of the Determinants of Income and Income Gap between Urban and Rural China*. IZA Discussion Paper No. 7162.

Adresa autora:

Tomáš Želinský, doc. Ing. PhD.
Ekonomická fakulta, TU Košice
Němcovej 32, 040 01 Košice
tomas.zelinsky@tuke.sk

Tabuľka 1: Odhad parametrov modelu

q	pohl	Ž	roky	roky ²	urban_1	urban_3	vzd_ZŠ	vzd_VŠ	post_2	post_3	R ¹
AT	0,25	-0,637***	0,113***	-0,002***	-0,195**	-0,040	-0,529***	0,352***	-0,210***	0,158*	0,168
	0,5	-0,411***	0,077***	-0,001***	-0,036	-0,077***	-0,364***	0,462***	-0,348***	0,070	0,171
	0,75	-0,327***	0,055***	-0,001***	0,017	-0,069***	-0,432***	0,441***	-0,409***	-0,241***	0,172
	0,9	-0,341***	0,049***	-0,001***	0,048	-0,070***	-0,412***	0,454***	-0,458***	-0,334***	0,189
	OLS	-0,524***	0,097***	-0,002***	-0,035	-0,070*	-0,456***	0,437***	-0,299***	-0,007	0,221
BG	0,25	-0,232***	0,050***	-0,001***	0,098***	-0,203***	-0,743***	0,350***	-0,150	0,079	0,157
	0,5	-0,281***	0,035***	-0,001***	0,172***	-0,075***	-0,318***	0,331***	-0,238***	-0,144*	0,148
	0,75	-0,298***	0,024***	-0,001***	0,221***	-0,022	-0,251***	0,366***	-0,237***	-0,260***	0,152
	0,9	-0,321***	0,023***	-0,001***	0,274***	0,023	-0,234***	0,460***	-0,221*	-0,340***	0,148
	OLS	-0,222***	0,039***	-0,001***	0,134***	-0,154***	-0,448***	0,409***	-0,300***	-0,231***	0,159
CY	0,25	-0,516***	0,092***	-0,002***	0,061*	0,090*	-0,324***	0,523***	-0,567***	-0,147	0,208
	0,5	-0,429***	0,078***	-0,001***	0,050*	0,039	-0,265***	0,533***	-0,405***	-0,076	0,228
	0,75	-0,332***	0,073***	-0,001***	0,065***	0,024	-0,280***	0,571***	-0,314***	-0,031	0,266
	0,9	-0,275***	0,071***	-0,001***	0,046*	0,015	-0,303***	0,578***	-0,279***	-0,104	0,290
	OLS	-0,422***	0,090***	-0,002***	0,057*	0,014	-0,310***	0,560***	-0,391***	-0,118*	0,306
CZ	0,25	-0,473***	0,074***	-0,002***	0,053*	-0,056***	-0,500***	0,471***	-0,340***	0,105	0,142
	0,5	-0,352***	0,050***	-0,001***	0,104***	-0,056***	-0,306***	0,446***	-0,222***	0,034	0,142
	0,75	-0,330***	0,041***	-0,001***	0,082***	-0,082***	-0,271***	0,462***	-0,256***	-0,148*	0,151
	0,9	-0,320***	0,040***	-0,001***	0,113***	-0,071***	-0,315***	0,572***	-0,420***	-0,390***	0,173
	OLS	-0,415***	0,064***	-0,001***	0,082***	-0,056***	-0,364***	0,461***	-0,323***	-0,012	0,224
DE	0,25	-0,784***	0,106***	-0,002***	0,004	-0,118***	-0,607***	0,586***	-0,900***	0,183*	0,248
	0,5	-0,513***	0,075***	-0,001***	0,033*	-0,122***	-0,530***	0,495***	-0,822***	-0,079	0,231
	0,75	-0,375***	0,059***	-0,001***	0,028*	-0,092***	-0,451***	0,428***	-0,805***	-0,443***	0,205
	0,9	-0,346***	0,049***	-0,001***	0,009	-0,097***	-0,395***	0,433***	-0,701***	-0,626***	0,205
	OLS	-0,604***	0,084***	-0,001***	0,038*	-0,112***	-0,459***	0,490***	-0,829***	-0,173***	0,287
DK	0,25	-0,166***	0,327***	-0,004***	-0,054	-0,011	-0,180***	0,248***	-0,765	0,788	0,321
	0,5	-0,200***	0,184***	-0,002***	-0,008	-0,042*	-0,116***	0,240***	-0,446	0,932*	0,228
	0,75	-0,254***	0,127***	-0,001***	-0,009	-0,064***	-0,136***	0,244***	-0,677	0,604	0,200
	0,9	-0,274***	0,106***	-0,001***	0,006	-0,086***	-0,156***	0,283***	-0,615	-0,008	0,213
	OLS	-0,281***	0,264***	-0,003***	-0,025	-0,055*	-0,155***	0,270***	-0,628***	0,224	0,373
EE	0,25	-0,437***	0,059***	-0,001***		-0,049	-0,453***	0,405***	-1,348***	0,036	0,134
	0,5	-0,466***	0,042***	-0,001***		-0,047*	-0,243***	0,433***	-0,938***	-0,054	0,141
	0,75	-0,494***	0,035***	-0,001***		-0,059*	-0,263***	0,442***	-0,784***	-0,344*	0,147
	0,9	-0,508***	0,030***	-0,001***		-0,074*	-0,142*	0,425***	-0,767***	-0,641***	0,140
	OLS	-0,449***	0,050***	-0,001***		-0,089***	-0,287***	0,444***	-0,967***	-0,108	0,182
EL	0,25	-0,323***	0,071***	-0,001***	0,101	-0,026	-0,309***	0,432***	-0,357***	0,282***	0,213
	0,5	-0,232***	0,050***	-0,001***	-0,009	-0,109*	-0,312***	0,367***	-0,440***	-0,003	0,196
	0,75	-0,257***	0,047***	-0,001***	-0,009	-0,094*	-0,245***	0,366***	-0,399***	-0,214***	0,185
	0,9	-0,295***	0,040***	-0,001***	-0,029	-0,075	-0,257***	0,518***	-0,582***	-0,570***	0,190
	OLS	-0,375***	0,061***	-0,001***	0,052	-0,071*	-0,292***	0,481***	-0,481***	0,003	0,274
ES	0,25	-0,317***	0,089***	-0,001***	0,075*	-0,184***	-0,311***	0,397***	-0,175*	0,293***	0,191
	0,5	-0,263***	0,056***	-0,001***	0,065***	-0,077***	-0,195***	0,413***	-0,188***	0,197***	0,187
	0,75	-0,243***	0,043***	-0,001***	0,034*	-0,069***	-0,202***	0,407***	-0,365***	-0,187*	0,199
	0,9	-0,256***	0,039***	-0,001***	0,050***	-0,077***	-0,228***	0,412***	-0,267*	-0,546***	0,193
	OLS	-0,309***	0,076***	-0,001***	0,083***	-0,100***	-0,298***	0,411***	-0,264***	0,011	0,269

Zdroj údajov: EU SILC 2012 (Eurostat, 2014), vlastné spracovanie

Pokračovanie tabuľky 1

	q	pohl_Ž	roky	roky ²	urban_1	urban_3	vzd_ZŠ	vzd_VŠ	post_2	post_3	R ¹
FI	0,25	-0,279***	0,273***	-0,003***	0,045	-0,033	-0,220***	0,279***	-0,665***	0,126*	0,229
	0,5	-0,282***	0,170***	-0,002***	0,021	-0,075***	-0,116***	0,300***	-0,465***	0,003	0,182
	0,75	-0,296***	0,119***	-0,001***	0,072***	-0,055***	-0,094***	0,321***	-0,420***	-0,128***	0,175
	0,9	-0,341***	0,092***	-0,001***	0,086***	-0,066***	-0,088***	0,355***	-0,382***	-0,256***	0,183
	OLS	-0,322***	0,224***	-0,003***	0,044	-0,105***	-0,101***	0,375***	-0,419***	-0,111***	0,295
FR	0,25	-0,380***	0,208***	-0,002***	0,044	-0,026	-0,361***	0,376***	-0,519*	0,139	0,165
	0,5	-0,296***	0,117***	-0,001***	0,047***	-0,044*	-0,190***	0,370***	-0,614***	-0,361***	0,140
	0,75	-0,313***	0,077***	-0,001***	0,061***	-0,048***	-0,179***	0,398***	-0,596***	-0,617***	0,163
	0,9	-0,372***	0,062***	-0,001***	0,062***	-0,060***	-0,153***	0,460***	-0,423***	-0,672***	0,188
	OLS	-0,430***	0,163***	-0,002***	0,059*	-0,011	-0,284***	0,475***	-0,603***	-0,301***	0,213
HU	0,25	-0,313***	-0,014	0,000	0,346*	-0,186*	-0,438***	0,369*	0,053	-0,217	0,814
	0,5	-0,191*	0,005	0,000	0,352***	-0,147	-0,491***	0,485***	-0,026	0,027	0,786
	0,75	-0,148*	0,009	0,000	0,184*	-0,126	-0,470***	0,576***	0,127	0,135	0,785
	0,9	-0,207***	0,016	0,000	0,110	-0,126*	-0,452***	0,576***	0,205	0,256	0,799
	OLS	-0,244***	-0,013*	0,000*	0,252***	-0,137*	-0,443***	0,463***	0,014	-0,036	0,119
IS	0,25	-0,456***	0,193***	-0,002***		-0,062*	-0,084*	0,363***	-0,733***	0,266***	0,241
	0,5	-0,439***	0,142***	-0,001***		-0,065***	-0,151***	0,309***	-0,565***	0,251***	0,270
	0,75	-0,460***	0,117***	-0,001***		-0,067***	-0,128***	0,295***	-0,403***	0,135*	0,254
	0,9	-0,488***	0,121***	-0,001***		-0,054*	-0,113***	0,298***	-0,408***	0,029	0,245
	OLS	-0,482***	0,165***	-0,002***		-0,064*	-0,141***	0,295***	-0,578***	0,202***	0,314
IT	0,25	-0,462***	0,100***	-0,002***	0,040*	-0,118***	-0,429***	0,421***	-0,494***	-0,085*	0,159
	0,5	-0,345***	0,064***	-0,001***	0,027*	-0,107***	-0,334***	0,364***	-0,428***	-0,172***	0,152
	0,75	-0,303***	0,047***	-0,001***	0,061***	-0,065***	-0,311***	0,364***	-0,375***	-0,333***	0,163
	0,9	-0,316***	0,046***	-0,001***	0,089***	-0,055***	-0,307***	0,478***	-0,396***	-0,515***	0,196
	OLS	-0,399***	0,085***	-0,002***	0,039***	-0,088***	-0,437***	0,415***	-0,529***	-0,358***	0,199
LT	0,25	-0,244***	0,148***	-0,002***	0,052	-0,198***	-0,537***	0,702***	-0,258	0,074	0,111
	0,5	-0,254***	0,104***	-0,001***	0,080*	-0,139***	-0,304***	0,594***	-0,340*	-0,188	0,136
	0,75	-0,301***	0,092***	-0,001***	0,169***	-0,023	-0,221***	0,582***	-0,281*	-0,212*	0,134
	0,9	-0,282***	0,089***	-0,001***	0,241***	0,012	-0,185***	0,607***	-0,140	-0,248*	0,127
	OLS	-0,241***	0,125***	-0,001***	0,117***	-0,125***	-0,283***	0,603***	-0,226*	-0,258***	0,201
LU	0,25	-0,457***	0,087***	-0,002***	0,049	0,106***	-0,374***	0,691***	-0,418*	0,068	0,225
	0,5	-0,305***	0,058***	-0,001***	0,059*	0,103***	-0,446***	0,641***	-0,585***	-0,258***	0,262
	0,75	-0,215***	0,043***	-0,001***	0,017	0,078***	-0,486***	0,542***	-0,428***	-0,298***	0,274
	0,9	-0,199***	0,037***	0,000	0,046	0,067***	-0,371***	0,508***	-0,407***	-0,598***	0,282
	OLS	-0,363***	0,070***	-0,001***	0,076***	0,086***	-0,449***	0,590***	-0,421***	-0,180***	0,371
LV	0,25	-0,223***	0,056***	-0,001***		-0,128***	-0,557***	0,651***	-0,388*	0,123	0,104
	0,5	-0,318***	0,034***	-0,001***		-0,109***	-0,275***	0,631***	-0,517***	-0,073	0,113
	0,75	-0,375***	0,033***	-0,001***		-0,126***	-0,277***	0,645***	-0,409***	-0,053	0,135
	0,9	-0,367***	0,030***	-0,001***		-0,154***	-0,281***	0,692***	-0,450***	-0,238***	0,147
	OLS	-0,290***	0,042***	-0,001***		-0,170***	-0,390***	0,704***	-0,587***	-0,186*	0,176
MT	0,25	-0,402***	0,075***	-0,002***			-0,292***	0,428***	-0,26***	-0,052	0,171
	0,5	-0,245***	0,046***	-0,001***			-0,287***	0,374***	-0,297***	-0,131	0,180
	0,75	-0,233***	0,043***	-0,001***			-0,264***	0,382***	-0,521***	-0,368***	0,197
	0,9	-0,219***	0,044***	-0,001***			-0,295***	0,435***	-0,764***	-0,597***	0,223
	OLS	-0,443***	0,076***	-0,002***			-0,318***	0,431***	-0,496***	-0,325***	0,205

Dokončenie tabuľky 1

	q	pohl_Ž	roky	roky ²	urban_1	urban_3	vzd_ZŠ	vzd_VŠ	post_2	post_3	R ¹
NL	0,25	-0,673***	0,112***	-0,002***			-0,212***	0,461***	-0,850***	0,083	0,639
	0,5	-0,479***	0,059***	-0,001***			-0,280***	0,405***	-0,721***	-0,151*	0,636
	0,75	-0,371***	0,047***	-0,001***			-0,220***	0,399***	-0,409***	-0,318***	0,640
	0,9	-0,333***	0,041***	-0,001***			-0,167***	0,424***	-0,440***	-0,492***	0,641
	OLS	-0,567***	0,084***	-0,002***			-0,262***	0,478***	-0,697***	-0,236***	0,250
NO	0,25	-0,421***	0,316***	-0,004***	0,047	-0,061*	-0,360***	0,209***	-0,479***	-0,049	0,297
	0,5	-0,380***	0,210***	-0,002***	0,094***	-0,007	-0,238***	0,206***	-0,380***	-0,121*	0,239
	0,75	-0,397***	0,139***	-0,001***	0,142***	0,010	-0,204***	0,207***	-0,363***	-0,222***	0,204
	0,9	-0,397***	0,139***	-0,001***	0,142***	0,010	-0,204***	0,207***	-0,363***	-0,222***	0,205
	OLS	-0,430***	0,116***	-0,001***	0,176***	0,016	-0,091***	0,235***	-0,456***	-0,437***	0,363
PL	0,25	-0,270***	0,073***	-0,002***	0,020	-0,093***	-0,557***	0,611***	-0,633***	-0,170***	0,143
	0,5	-0,268***	0,048***	-0,001***	0,023	-0,089***	-0,374***	0,551***	-0,413***	-0,260***	0,142
	0,75	-0,292***	0,042***	-0,001***	0,078***	-0,077***	-0,301***	0,564***	-0,275***	-0,297***	0,148
	0,9	-0,348***	0,041***	-0,001***	0,099***	-0,071***	-0,316***	0,597***	-0,235***	-0,315***	0,156
	OLS	-0,317***	0,064***	-0,001***	0,074***	-0,087***	-0,596***	0,599***	-0,483***	-0,265***	0,229
PT	0,25	-0,281***	0,063***	-0,001***	0,019	-0,086***	-0,334***	0,650***	-0,482***	0,126	0,196
	0,5	-0,292***	0,047***	-0,001***	0,042*	-0,077***	-0,363***	0,660***	-0,607***	-0,180*	0,246
	0,75	-0,326***	0,047***	-0,001***	0,069*	-0,102***	-0,392***	0,654***	-0,653***	-0,355***	0,305
	0,9	-0,373***	0,052***	-0,001***	0,109***	-0,140***	-0,444***	0,619***	-0,749*	-0,539*	0,309
	OLS	-0,307***	0,060***	-0,001***	0,029	-0,058*	-0,390***	0,657***	-0,598***	-0,187***	0,293
RO	0,25	-0,250***	0,026***	-0,001***	0,116***	-0,113***	-0,175***	0,445***	-1,495***	0,153	0,517
	0,5	-0,241***	0,020***	0,000***	0,072***	-0,138***	-0,245***	0,490***	-1,185***	-0,017	0,395
	0,75	-0,286***	0,019***	0,000***	0,094***	-0,153***	-0,307***	0,475***	-0,728***	0,010	0,313
	0,9	-0,265***	0,017***	0,000***	0,137***	-0,138***	-0,380***	0,436***	-0,419***	-0,021	0,255
	OLS	-0,290***	0,030***	-0,001***	0,102***	-0,214***	-0,347***	0,454***	-1,079***	0,079	0,605
SE	0,25	-0,299***	0,300***	-0,003***	-0,021	-0,052	-0,161***	0,137***	-0,626***	-0,045	0,247
	0,5	-0,315***	0,173***	-0,002***	0,045*	-0,038*	-0,115***	0,183***	-0,328***	-0,042	0,168
	0,75	-0,295***	0,121***	-0,001***	0,056*	-0,077***	-0,126***	0,173***	-0,284***	-0,072*	0,148
	0,9	-0,314***	0,106***	-0,001***	0,054	-0,086***	-0,129***	0,224***	-0,176***	-0,079	0,151
	OLS	-0,364***	0,251***	-0,003***	-0,005	-0,103***	-0,193***	0,183***	-0,460***	-0,081	0,324
SI	0,25	-0,226***	0,151***	-0,004***			-0,298***	0,627***	-0,552***	0,512***	0,664
	0,5	-0,187***	0,073***	-0,002***			-0,253***	0,610***	-0,594***	0,528***	0,659
	0,75	-0,185***	0,049***	-0,001***			-0,264***	0,595***	-0,441***	0,266*	0,674
	0,9	-0,247***	0,038***	-0,001***			-0,294***	0,645***	-0,370*	0,166	0,682
	OLS	-0,229***	0,109***	-0,003***			-0,379***	0,652***	-0,495***	0,326***	0,231
SK	0,25	-0,292***	0,055***	-0,001***	0,086***	-0,077***	-0,464***	0,338***	-0,298***	0,155*	0,115
	0,5	-0,261***	0,031***	-0,001***	0,078***	-0,060***	-0,328***	0,306***	-0,126	0,029	0,118
	0,75	-0,265***	0,023***	-0,001***	0,083***	-0,051***	-0,298***	0,287***	-0,142*	-0,119*	0,121
	0,9	-0,282***	0,021***	-0,001***	0,089***	-0,021	-0,295***	0,314***	-0,144	-0,217*	0,126
	OLS	-0,306***	0,049***	-0,001***	0,064***	-0,071***	-0,404***	0,310***	-0,191***	-0,001	0,140
UK	0,25	-0,564***	0,058***	-0,001***	0,043	0,008	-0,273***	0,467***	-0,535***	0,271*	0,261
	0,5	-0,447***	0,050***	-0,001***	0,069***	0,078***	-0,272***	0,489***	-0,462***	0,070	0,254
	0,75	-0,395***	0,044***	-0,001***	0,051***	0,086***	-0,216***	0,551***	-0,477***	-0,174	0,246
	0,9	-0,472***	0,053***	-0,001***	0,081***	0,123***	-0,284***	0,572***	-0,611***	-0,474*	0,234
	OLS	-0,523***	0,054***	-0,001***	0,069***	0,038	-0,271***	0,476***	-0,480***	0,108*	0,244

Pracovná migrácia a mobilita v okresoch Slovenskej republiky

Labour migration and mobility in the districts of the Slovak Republic

Martina Žudelová

Abstract: The article deals with migration and population mobility in the Slovak Republic and explores the migration relationships among its districts. The various types of indicators have been used for mapping the inner migration development and mobility and the paper is focused on the internal working migration. The political, economic and structural changes in Slovakia since 1990 have resulted in significant regional disparities in terms of unemployment, infrastructure, brain-drain, increased number of people living below poverty line etc.

Abstrakt: Článok sa venuje problematike pracovnej migrácie a mobility obyvateľov Slovenskej republiky, skúma migračné vzťahy v rámci jej okresov. Hlavná pozornosť je venovaná vnútornej pracovnej migrácii. Prelomové politické, ekonomické a štrukturálne zmeny prebiehajúce na území Slovenska od roku 1990 vyústili do podstatných regionálnych rozdielov v nezamestnanosti, infraštruktúre, brain-drain, počte osôb žijúcich pod hranicou chudoby a pod..

Key words: labour mobility, labour migration, demographic indicators, factor analysis, cluster analysis.

Kľúčové slová: pracovná mobilita, pracovná migrácia, demografické ukazovatele, faktorová analýza, klastrová analýza.

1. Úvod

21. storočie je charakteristické pracovnou migráciou a mobilitou ekonomicky aktívneho obyvateľstva (v texte ďalej aj ako EAO), ktorému sa venuje nasledujúca analýza. Cieľom príspevku je analyzovať vývoj a súčasný stav pracovnej migrácie a mobility so zohľadnením priestorovej polarizácie skúmaných premenných. Využitie sú demografické ukazovatele, faktorová a klastrová analýza a Moranov koeficient priestorovej autokorelácie.

Mobilita (lat. *mobilitas*) v sebe ako základný význam zahŕňa pohyblivosť. Spája sa s pohyblivosťou výrobných faktorov, a to – pôdy, kapitálu a práce. *Migrácia* je prirodzený a nezastaviteľný jav. Je to forma priestorovej mobility, v pravom slova zmysle len taká, v rámci ktorej dochádza k zmene trvalého bydliska. Každý migračný tok a odliv pracovnej sily má svoj smer a príčiny. Mobilitu môžeme definovať a rozlišovať ako *funkčnú* (pracovná, profesijná mobilita) a *geografickú* (priestorová mobilita) (Known, 2009; Lindstorm, Scott 2010).

Z makroekonomického pohľadu sú za migrantov považovaní spravidla tí, ktorí sa rozhodnú migrovať, z dôvodu vysokej nezamestnanosti a nemožnosti pracovného uplatnenia. Dopyt po práci je vysoký, zatiaľ čo ponuka z hľadiska možnosti pracovného uplatnenia a mzdového ohodnotenia je nedostatočná. Rozhodujúcim faktorom, prečo sa niekto rozhodne migrovať, sú rozdiely v rámci regiónov a najmä v mzdách (Massey, 2011; Prochádzková, 2010).

Mikroekonomická neoklasická teória vníma jedinca ako racionálnu bytosť, ktorá porovná náklady spojené s migráciou s možnými výnosmi plynúcimi z migrácie. Migrovať sa rozhodne, ak môže svojím rozhodnutím očakávať výnos. Musí vynaložiť potrebné náklady na cestovanie, nové bývanie, i náklady nefinančného charakteru – vytváranie nových sociálnych väzieb, prispôbovanie sa a pod. (Fias, 2000; Tassinopoulos a Werner, 2010).

V roku 1989 začali na Slovensku veľké zmeny, pri ktorých sa ekonomika transformovala z centrálne plánovanej na trhovo orientovanú. V dôsledku zmien nastal nový trend – sťahovania a dochádzania obyvateľov za prácou mimo okresu svojho trvalého pobytu, dokonca aj do zahraničia.

2. Použité údaje a metodika práce

a. Vstupné údaje

V rámci príspevku pracujeme s tromi databázami:

- prehľad sťahovania obyvateľstva v časovom období rokov 1996 – 2012,
- sčítanie obyvateľov domov a bytov 2011,
- regdat – tematická oblasť zamestnanosť a mzdy.

Pre potreby vyčíslenia demografických ukazovateľov sa primárne využívala databáza prehľadu sťahovania obyvateľstva v časovom období rokov 1996-2012. Konkrétne sekcia vnútorného sťahovania obyvateľstva z pracovných dôvodov. Klastrová a faktorová analýza je postavená na oboch vyššie spomenutých databázach. Vstupnými premennými pre potreby týchto analýz sú: počet nezamestnaných v okrese, prisťahovaní z pracovných dôvodov do okresu, vystťahovaní z pracovných dôvodov z okresu, EAO nedochádzajúce za prácou do iných okresov, EAO dochádzajúce za prácou do iných okresov, EAO v rámci okresu pracujúce v zahraničí, EAO dochádzajúce z iných okresov, počet EAO v okrese, výška nominálnej mzdy. Pri výbere vhodných premenných pre faktorovú a klastrovú analýzu sme boli obmedzení dostupnosťou údajov a štatistík týkajúcich sa slovenských okresov.

b. Použité metódy analýzy vývoja a súčasného stavu pracovnej migrácie a mobility v okresoch Slovenskej republiky

Vyššie uvedené teoretické východiská pre oblasť pracovnej migrácie a mobility sú využité pre potreby hlbšej analýzy, ktorá je primárne zameraná na vnútornú migráciu a mobilitu obyvateľstva z pracovných dôvodov. Pre výpočet *demografických ukazovateľov* boli využité údaje z prehľadu sťahovania obyvateľstva. Z demografických ukazovateľov je v analýze venovaná pozornosť migračnému saldu v období rokov 1996-1999 a 2009-2012. Na účely vyčíslenia *migračného salda* vychádzame zo vzťahu:

$$MS_t = \frac{I - E}{\bar{S}} \times 1000, \quad (1)$$

kde: MS_t je hodnota migračného salda v príslušnom roku, I je počet imigrantov za sledované časové obdobie, E je počet emigrantov za sledované časové obdobie a \bar{S} je stredný stav obyvateľstva.

Hodnotenie efektívnosti migrácie je v príspevku založené na *indexe účinnosti migrácie*, jeho hodnota približujúca sa k 1 je dôkazom priestorovej koncentrácie obyvateľstva:

$$I_{ms_t} = \frac{MS_t}{MO} = \frac{I - E}{I + E} \quad (2)$$

Súčasný stav pracovnej migrácie a mobility je analyzovaný s využitím faktorovej a klastrovej analýzy. *Faktorová analýza* umožnila vytvoriť nové premenné (tzv. faktory) zo súboru pôvodných, zredukovaním počtu vstupných premenných pri zachovaní maximálnej informácie. Predpokladá sa, že vstupné premenné sú dôsledkom spoločných (skrytých) príčin, nie sú známe väzby medzi faktormi, ani ich samotná štruktúra (Král, 2009). Pre ďalšiu analýzu si vyberáme prvé štyri komponenty vysvetľujúce 83% variability premenných. Hodnoty komunální sú pri vstupných premenných taktiež na dostatočne vysokej úrovni. Rotáciou faktorov sme zabránili prípadnej saturácii medzi vstupnými premennými navzájom. Takto určené „nové“ premenné (faktory) už nie sú lineárne závislými a budú pre nás vhodným vstupom pre zhlukovú analýzu (Král, 2009).

Na vytvorenie typológie slovenských okresov na základe ich vzájomnej podobnosti v rámci jednotlivých klastrov (skupín) bola použitá nehierarchická zhlukovacia metóda kmeans.

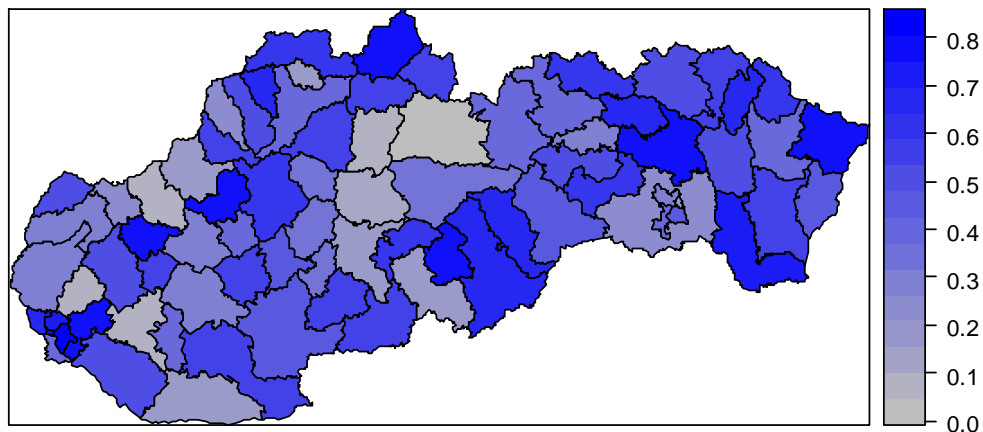
Táto metóda je najvhodnejšia z dôvodu vytvárania malého počtu zhlukov z veľkého počtu objektov. Zhluky vytvorené zo slovenských okresov dodržia podmienku, čím väčšej podobnosti v rámci skupiny a na druhej strane väčší rozdiel medzi zhlukmi. Výsledný počet zhlukov bol stanovený na počet 5.

Prítomnosť a charakter priestorovej polarizácie v slovenských okresoch je prezentovaná prostredníctvom Moranovho I koeficientu priestorovej autokorelácie. Jeho výsledné hodnoty sa môžu pohybovať v intervale hodnôt od (-1,1). Hraničné výsledné hodnoty indikujú prítomnosť priestorovej autokorelácie. Moranov koeficient priestorovej autokorelácie bol aplikovaný na výšku a vývoj nominálnej mzdy v rokoch 2001, 2006 a 2011 v rámci okresov. Pre potreby spomínanej analýzy boli spracované údaje z databázy RegDat, konkrétne tematická oblasť Zamestnanosť a mzdy. Prítomnosť a charakter priestorovej autokorelácie v slovenských okresoch, je sledovaná aplikovaním tohto koeficientu. Pre potreby Moranovho koeficientu priestorovej autokorelácie boli za susedné priestorové jednotky považované tie okresy, ktoré mali spoločnú hranicu.

Všetky analýzy a výpočty prebiehali v prostredí softvéru R (odcítovať), s využitím týchto konkrétnych balíkov `maptools` (Bivand a Lewin-Koh, 2014), `rgdal` (Bivand, Keitt a Rowlingson, 2014), `shape` (Soetaer 2014) a `spdep` (Bivand, 2014).

3. Výsledky a diskusia

Prebiehajúce migračné trendy sa potvrdzujú aj prostredníctvom uvedených základných demografických ukazovateľov (obr. 24, obr. 25). Urbanizácia, vyludňovanie, s tým spojené prehĺbovanie rozdielov medzi okresmi a značná priestorová koncentrácia obyvateľstva sú zreteľné počas celého časového obdobia rokov 1996-2012. Tieto skutočnosti sú spojené s push a pull faktormi pôsobiacimi v okresoch. Geografické podmienky typické pre niektoré oblasti taktiež prispeli k vyprofilovaniu prosperujúcich, napredujúcich a marginálnych regiónov. Miera nezamestnanosti, regionálne disparity neboli tak výrazné pred rokom 1989. Dôsledky poklesu hospodárskeho rastu a úpadku priemyslu sa vo väčšine slovenských okresoch v priebehu 25 rokov zintenzívnili a prehĺbili.

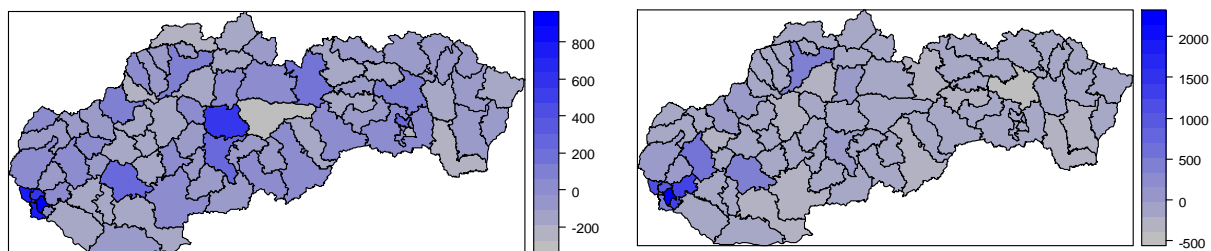


Obr. 26: Index efektívnosti migrácie v okresoch v rokoch 1996-2012.

Zdroj údajov: Prehľad sťahovania obyvateľstva v časovom období rokov 1996 – 2012.
(vlastné spracovanie)

Saldo pracovnej migrácie (Obr. 27) v časovom období rokov 1996-2012 potvrdzuje tieto závery. Mnohé okresy sa vyludňovali počas celého časového obdobia rokov 1996-2012 a dodnes ich obyvatelia opúšťajú. Výrazné postavenie medzi slovenskými okresmi má okres

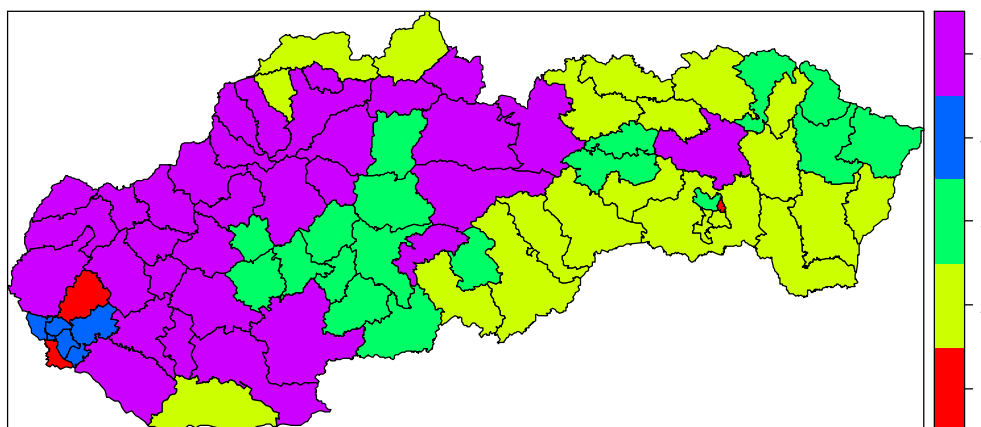
Senec, ktorý je možné porovnať s hlavným mestom Bratislavou, v počte prisťahovaných osôb z pracovných dôvodov



Obr.2: Saldo pracovnej migrácie z pracovných dôvodov v rámci okresov Slovenskej republiky v časovom období rokov 1996-2012.

Zdroj údajov: Prehľad sťahovania obyvateľstva v časovom období rokov 1996 – 2012. (vlastné spracovanie)

Konečné výsledky typológie okresov založené na faktorovej a zhlukovej analýze sú prezentované na Obr. 3. Priemerné hodnoty premenných v rámci jednotlivých klastrov sú uvedené v Tab. 19.



Obr.3: Typológia slovenských okresov na základe faktorovej a klastrovej analýzy.

Zdroj údajov: Prehľad sťahovania obyvateľstva v časovom období rokov 1996 – 2012, SODB 2011. (vlastné spracovanie)

Okresy s najvyšším počtom nezamestnaných osôb majú najvyššiu úroveň dochádzania obyvateľov za prácou. Tieto okresy majú najvyšší počet obyvateľov pracujúcich v zahraničí. Okresy poskytujú menej pracovných príležitostí pre svojich obyvateľov, a tí dochádzajú za prácou do iných okresov. Ide o skupinu okresov nachádzajúcich sa na hranici Slovenskej republiky s našimi zahraničnými susedmi. Obyvatelia týchto okresov uprednostňujú dochádzanie za prácou do zahraničia. Okresy s najvyššou úrovňou nominálnej mzdy prijímajú aj najvyšší počet ekonomicky aktívnych osôb dochádzajúcich za prácou do príslušného okresu a najviac osôb prisťahovaných do okresu z pracovných dôvodov. Okresy s najvyššou úrovňou nominálnej mzdy sú zastúpené len piatimi okresmi - v okolí hlavného mesta. Tieto oblasti majú najvyššie úrovne dochádzky za prácou z dôvodu poskytovania najväčšieho množstva pracovných možností, no ekonomicky aktívne osoby uprednostňujú každodenné dochádzanie za prácou z iných oblastí z dôvodu vysokých životných nákladov. Okresy v rámci zhluku 1 majú najvyšší počet EAO s trvalým pobytom na svojom území a zároveň najvyšší počet EAO dochádzajúcich za svojou prácou do iných okresov. Môžeme teda predpokladať, že ide o oblasti, ktoré sú preferované viac ako tie s nižšími životnými nákladmi v porovnaní so skupinou číslo 4. Túto skutočnosť predpokladáme, z dôvodu neexistencie

údajov poskytujúcich informácie týkajúce sa výšky a štruktúry životných nákladov v rámci slovenských okresov.

Tab. 20: Priemerné hodnoty premenných v rámci vytvorených skupín okresov. (vlastné spracovanie)

Skupina	1	2	3	4	5
Premenná					
Počet nezamestnaných v okrese	102,59	246,58	202,03	83,3	137,85
Počet vystáňovaných z pracovných dôvodov	0,71	0,26	0,66	5,34	0,52
Počet prístáňovaných z pracovných dôvodov	0,72	0,74	1,46	0,71	0,89
EAO nedochádzajúce za prácou	202,62	371,01	434,86	233,41	513,44
EAO dochádzajúce za prácou do okresu z iných okresov v rámci kraja	483,68	166,05	153,48	431,10	152,97
EAO pracujúce v zahraničí	28,21	70,49	62,03	23,10	42,74
EAO dochádzajúce za prácou do okresu z iných krajov	174,20	103,64	110,55	1195,16	99,02
Počet EAO v rámci okresu	548,75	462,02	470,60	508,09	494,59
Výška nominálnej mzdy	797,00	709,00	723,11	1099,00	778,38

Na základe predchádzajúcich výsledkov možno tvrdiť, že v čase dochádzalo k prehľbovaniu regionálnych disparít. Na zhodnotenie stupňa priestorovej polarizácie miezd bol použitý Moranov I koeficient priestorovej autokorelácie.

Aplikovaním Moranovho I koeficientu priestorovej autokorelácie sa potvrdila štatisticky významná prítomnosť priestorovej autokorelácie. Výsledné hodnoty Moranovho koeficientu sú uvedené v tabuľke 2.

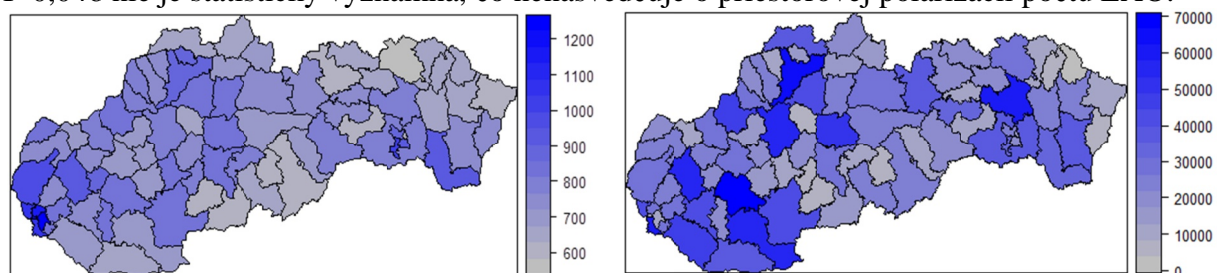
Tab. 2: Výsledné hodnoty Moranovho koeficientu priestorovej autokorelácie

Ukazovateľ	Mzda 2001	Mzda 2006	Mzda 2011
Výsledná hodnota			
p-hodnota	6.804e-14	3.941e-14	4.856e-14
hodnota I	0.51296472	0.515424136	0.524516869

Zdroj údajov: RegDat – tematická oblasť zamestnanosť a mzdy (vlastné spracovanie)

Prítomnosť priestorovej autokorelácie bola sledovaná v troch časových obdobiach. V každom roku je p-hodnota testu významnosti Moranovho koeficientu menšia ako hladina významnosti, čo indikuje prítomnosť štatisticky významnej priestorovej autokorelácie. Hodnota samotného koeficientu v rokoch 2001, 2006 a 2011 je nad hodnotou 0,50, a teda ide o pomerne silnú pozitívnu priestorovú autokoreláciu. Výška nominálnej mzdy bola a je priestorovo autokorelovaná v priestore. Prítomnosť priestorovej autokorelácie potvrdzuje aj obr. 4, z ktorého je zrejmé, že podobné hodnoty nominálnej mzdy sú usporiadané blízko seba.

Na obr. 4 vpravo je znázornená priestorová autokorelácia nominálnej mzdy, vľavo na obr. 4 je znázornená priestorová autokorelácia počtu EAO dochádzajúcich za prácou do zamestnania. Výsledná hodnota koeficientu priestorovej autokorelácie v prípade premennej počet EAO $I=0,046$ nie je štatisticky významná, čo nenasvedčuje o priestorovej polarizácii počtu EAO.



Obr.4: Moranov koeficient priestorovej autokorelácie za rok 2011. (vlastné spracovanie)

4. Záver

Cieľom tohto príspevku bola analýza vývoja pracovnej mobility a migrácie s využitím základných demografických ukazovateľov, faktorovej, klastrovej analýzy a Moranovho koeficientu priestorovej autokorelácie.

Na základe vyššie uvedenej analýzy sú na území Slovenskej republiky od roka 1996 prítomné a zároveň neustále sa prehľbujúce regionálne disparity. Možnosti pracovného uplatnenia, výška nominálnej mzdy, životná úroveň a stupeň rozvoja infraštruktúry sa výrazne líši v rámci okresov.

Otázka pracovnej migrácie a mobility obyvateľstva je rozsiahlym, dokonca až multidisciplinárnym fenoménom, ktorý je možné sledovať z viacerých pohľadov. Samotná analýza bola do značnej miery obmedzená dostupnosťou údajov a ich relevantnosťou.

Pod'akovanie

Napísanie príspevku bolo podporené Vedeckou grantovou agentúrou MŠ SR a SAV v rámci riešenia vedecko-výskumného projektu VEGA 1/0127/11 Priestorová distribúcia chudoby v EÚ.

Literatúra

BIVAND, R., KEITT T., ROWLINGSON B. 2014. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.8-16.

BIVAND, R., LEWIN-KOH N. 2014. maptools: Tools for reading and handling spatial objects. R package version 0.8-29.

BIVAND, R. 2014. spdep: Spatial dependence: weighting schemes, statistics and models. R package version

FIAST, T. 2000. *A Review of Dominant Theories of International Migration*. Oxford: University Press.

KNOWN, D.-B. 2009. *Human Capital and its measurement*. Busan, Korea: The 3rd OECD World Forum on „Statistics, Knowledge and Policy“.

KRÁL, P. et al. 2009. *Viacrozmerné štatistické metódy*. Banská Bystrica: Ekonomická fakulta Univerzity Mateja Bela v Banskej Bystrici.

LINDSTORM, M., SCOTT, T. 2010. *Mobility and Migration*. New York: United Nations Development Programme.

MASSEY, D. S. et al. 1993. Theories of International Migration: A Review and Appraisal. In: *Population and Development Review*. Vol. 19, No. 3, pp 431-466.

PROCHÁDZKOVÁ ILINITCHI, C. 2010. Vybrané teorie migrace a jejich význam při vytváření migračních politik. In: *Acta Oeconomica Pragensia*. Vol. 2010, No. 6, pp. 3-26.

R CORE TEAM (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

SOETAER K.,(2014). shape: Functions for plotting graphical shapes, colors. R package version 1.4.1.

TASSINOPOULOS, A., WERNER, H., KRISTENSEN, S. 2010. *Mobility and Migration of labour in European Union and their specific implications for young people*. European Centre for the Development of Vocational Training.

Adresa autora:

Martina Žudelová, Ing.
Ekonomická fakulta TU v Košiciach
Němcovej 32, 040 01 Košice
martina.zudelova@tuke.sk

Analyzá nezaměstnanosti osob v závislosti na věku, vzdělání a dalších faktorech v mezinárodním srovnání

The analysis of the unemployment depending on age, education and other factors in international comparison

Jitka Bartošová, Klára Siegelová

Abstract: Present article deals with estimation of logistic regression model of the probability of unemployment depending on several factors like age structure of the population, education, health, regions and other quantitative factors. The article is based on the results of the survey EU-SILC for households from Czech Republic, Slovakia and Austria in 2012. Logistic regression model were used for data analysis and the model parameters were estimated in statistical program called SAS.

Abstrakt: Předložený článek se zabývá odhadem logistického regresního modelu pravděpodobnosti nezaměstnanosti v závislosti na některých faktorech, zejména z pohledu věkové struktury obyvatelstva a jeho dosaženého vzdělání, zdraví, regionech a dalších kvantitativních faktorech. Článek vychází z výsledků statistického šetření EU-SILC v roce 2012 a věnuje se domácnostem České Republiky, Slovenska a Rakouska. Parametry modelu byly odhadnuty v prostředí statistického programu SAS.

Key words: unemployment, education, age structure, logistic regression, health

Klíčová slova: nezaměstnanost, vzdělání, věková struktura, model logistické regrese

JEL classification: C14, D31, D63

1. Úvod

Stárnutí populace a růst nezaměstnanosti, mladých absolventů, vzbuzuje obavu v mnoha vyspělých zemích. V nadcházejícím desetiletí Evropské země zažijí strmý nárůst podílu starších lidí a prudký pokles podílu lidí v produktivním věku. Počet pracovníků, kteří každoročně odchází do důchodu, se bude zvyšovat a pravděpodobně nakonec převýší počet nových subjektů vstupujících na trh práce. Evropská unie a její orgány se problémům zaměstnanosti a sociální soudržnosti věnují v rámci různých projektů financovaných z evropských rozpočtů. Pro Českou republiku byly stanoveny následující národní úkoly v oblasti zaměstnanosti osob různých věkových kategorií:

- zvýšit míru zaměstnanosti osob ve věku 20-64 let na 75% (současnost 65,6%),
- zvýšit míru zaměstnanosti žen 20-64 let na 65% (současnost 57,3%),
- zvýšit míru zaměstnanosti starších osob 55-64 let na 55% (za rok 2010 46,5%),
- snížit míru nezaměstnanosti mladých osob ve věku 15-24 let o třetinu oproti roku 2010 (v roce 2010 činila 18,4%),
- snížit míru nezaměstnanosti osob s nízkou kvalifikací (stupeň ISCED 0-2) o čtvrtinu oproti roku 2010 (současnost 25%).

Tyto úkoly by měly být splněny do roku 2020. Uvedené cíle doprovází ještě další národní cíle týkající se snižování počtu domácností žijících pod hranicí chudoby, problémů se snižující se kupní silou osob v důchodovém věku apod.

Aktuálnost této problematiky potvrzuje řada článků věnovaných analýzám příjmové nerovnosti, rizika monetární chudoby a materiální deprivace, nezaměstnanosti a demografického vývoje, které byly v poslední době publikovány v odborných a vědeckých časopisech. U nás i v zahraničí. Z českých a slovenských publikací můžeme jmenovat např. práce Bartošové (2013), Bartošové a Želinského (2013), Bílkové (2012a,b), Fialy a

Langhamrové (2013 a 2014), Michálka a Veselovské (2012), Pauhofové (2012), Stankovičové, Vlačuhy a Ivančíkové (2013), Večerníka (2013), Želinského (2013), Želinského a Pauhofové (2013). Další četné publikace, které s touto problematikou úzce souvisí, se zabývají modelováním příjmové, resp. výdajové distribuce a diagnostikou faktorů, které ji ovlivňují. V posledních letech byly publikovány např. práce Malé (2013), Marka (2013), Marka a Vrabce (2013), Pacákové, Lindy a Sipkové (2012), Řezankové a Löstera (2013), Řezankové a Želinského (2014), Sipkové a Sipka (2012), Šimpacha (2012 a 2013), Želinského a Stankovičové (2012) a další.

Datová základna je tvořena výsledky výběrového šetření příjmů a životních podmínek domácností EU-SILC z roku 2012, které je povinné pro všechny členské státy Evropské Unie.¹ Základní jednotkou členění je hospodářící domácnost a její osoby, které mají ve vybraném bytě jediné či hlavní bydliště.

2. Logistická regrese

K modelování podmíněné pravděpodobnosti kategoriální (dichotomická) proměnné v závislosti na vektoru vysvětlujících proměnných spojitého či kategoriálního typu slouží logistická regrese. Logistická regrese je součástí zobecněného lineárního modelu. Jedná se o nelineární model, k jehož linearizaci se nejčastěji využívá logitová nebo probitová transformace vysvětlované proměnné. Logitová transformace spočívá v převedení podmíněné pravděpodobnosti p na tzv. *logit* (p), tj. logaritmus šance (*odds*), takže $\text{logit}(p) = \ln(\text{odds})$, resp. $\text{logit}(p) = \log(\text{odds})$. Šance (*odds*) je podílem pravděpodobnosti výskytu očekávané události (úspěchu) k výskytu neočekávané události (neúspěchu), takže $\text{odds} = \frac{p}{1-p}$. Vztah mezi logitem, tj. hodnotou $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$, resp. $\log\left(\frac{p}{1-p}\right)$, a vektorem vysvětlujících proměnných (x_1, x_2, \dots, x_k) je lineární a lze ho vyjádřit vztahem

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1)$$

Zpětnou transformací dostaneme vztah mezi pravděpodobností p a vektorem vysvětlujících proměnných. Platí

$$p = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}} \quad (2)$$

Závislost má nelineární charakter a jejím grafem je sigmoidní křivka, tzv. S-křivka. Pro více informací viz např. Stankovičová a Vojtková (2007).

3. Modelování pravděpodobnosti nezaměstnanosti osob podmíněné vybranými faktory

3.1. Volba proměnných modelu

Analýza vlivu vybraných faktorů na riziko nezaměstnanosti, tj. pravděpodobnosti, že daná osoba bude nezaměstnaná, byla provedena ve třech sousedících zemích – v České republice, na Slovensku a v Rakousku. Do modelu byly zařazeny ty faktory, u nichž lze očekávat, že budou mít na riziko nezaměstnanosti osob statisticky významný vliv. Výběr proměnných uvádí tabulka 1.

Ekonomický status osoby v čele domácnosti byl očištěn od osob v důchodovém věku a osob pracovně neaktivních a ve výběru zůstaly pouze osoby v produktivním věku, které byly rozděleny na osoby

- *zaměstnané*

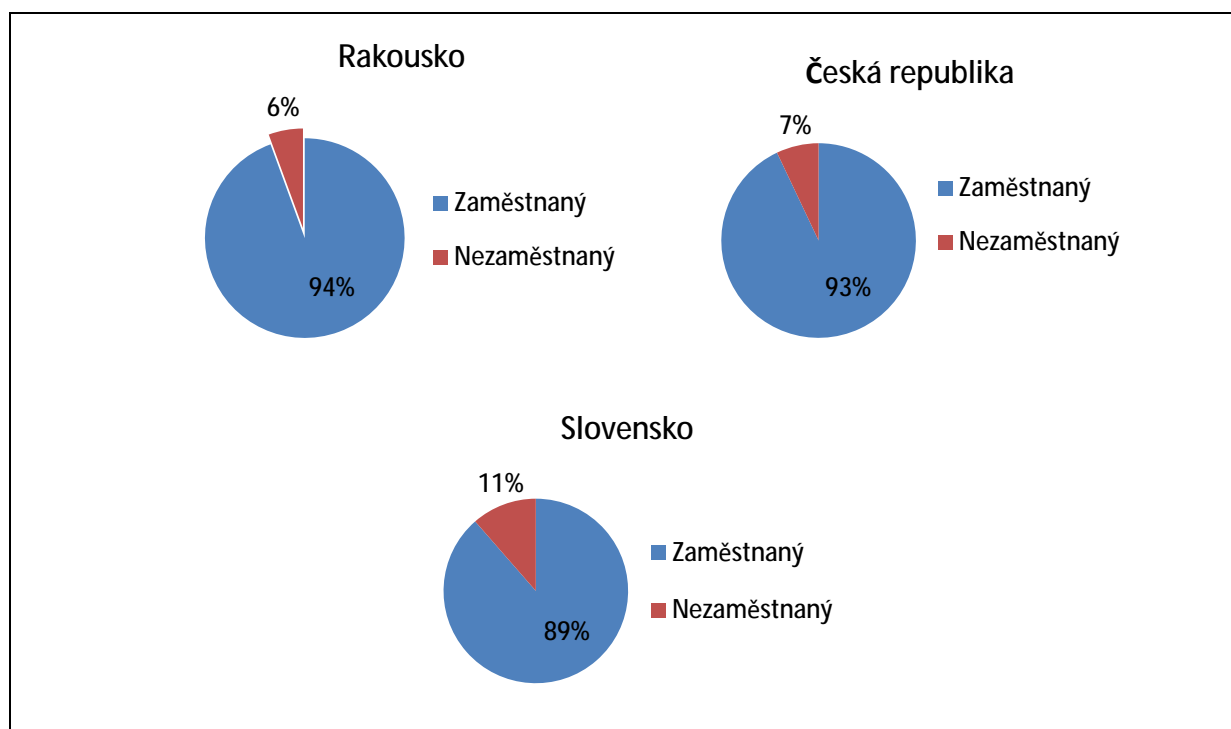
¹ Poprvé bylo toto šetření provedeno Českým statistickým úřadem v roce 2005 pod názvem Životní podmínky v roce 2005.

- nezaměstnané

Procentuální zastoupení zaměstnaných a nezaměstnaných osob v Čechách, na Slovensku a v Rakousku je vidět na obrázku 1.

Tabulka 1 Seznam proměnných a jejich obměn (variant)

Kategoriální proměnné	Obměny
<i>ekonomický status</i>	2 varianty
<i>vzdělání</i>	6 variant – dle nomenklatury ISCED
<i>věkové skupiny</i>	4 varianty
<i>pohlaví</i>	2 varianty
<i>zdravotní stav</i>	5 varianty
<i>regiony v České republice</i>	8 variant
<i>regiony na Slovensku</i>	1 varianta
<i>regiony v Rakousku</i>	3 varianty
Numerické proměnné	Typ
<i>dávky v nezaměstnanosti</i>	spojitá nezáporná
<i>příjmy z majetku</i>	spojitá nezáporná



Obrázek 1 Procentní zastoupení zaměstnaných a nezaměstnaných osob

Z uvedených grafů vyplývá, že podíl nezaměstnaných osob byl v roce 2011 největší na Slovensku (11%), následovalo Česko (7%) a nakonec Rakousko (6%). Aktuálnější procentuální nezaměstnanost, naměřená v srpnu 2014, činila v Rakousku 4,7%, v Čechách 6,3% a na Slovensku 13,3%.

Dalším faktorem zařazeným do analýzy byl *region*. V České republice se jedná o osm regionů

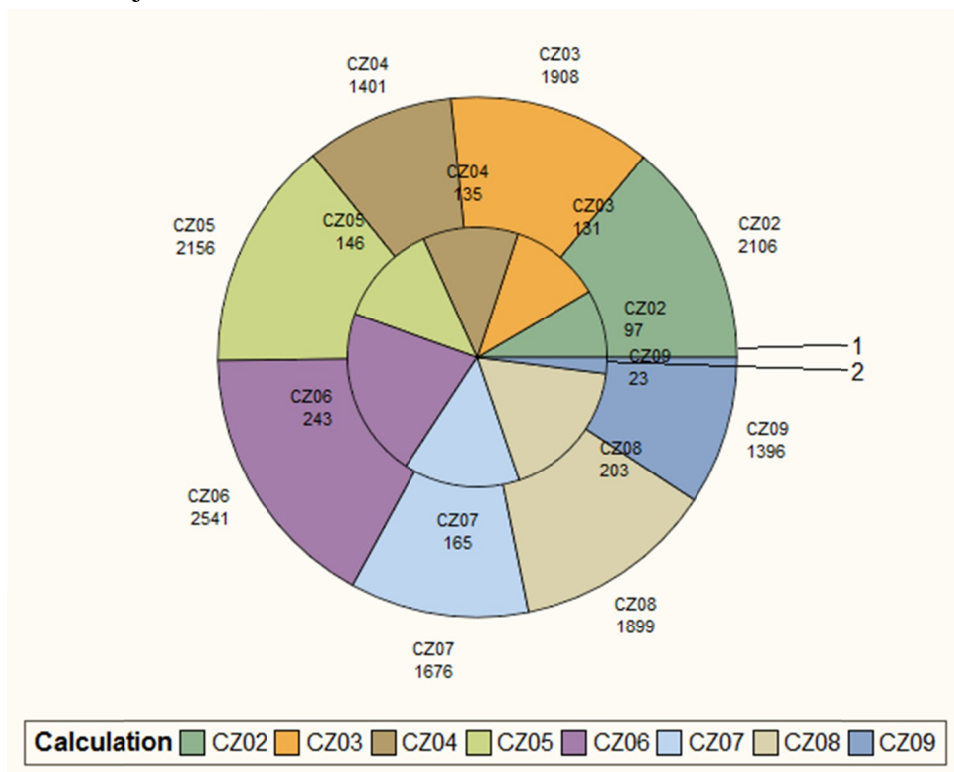
- CZ02 – Středočeský kraj
- CZ03 – Plzeňský a Jihočeský kraj
- CZ04 – Karlovarský a Ústecký kraj
- CZ05 – Liberecký, Královehradecký a Pardubický kraj

- CZ06 – Vysočina a Jihomoravský kraj
- CZ07 – Olomoucký a Zlínský kraj
- CZ08 – Moravskoslezský kraj
- CZ09 – Praha

Slovenský soubor není bohužel do regionů vůbec rozdělen a v Rakousku jsou tři regiony

- AT1 –
- AT2 –
- AT3 –

Zaměstnanost (hodnota 1) a nezaměstnanost (hodnota 2) osob žijících v jednotlivých regionech České republiky je zachycena na obrázku 2. Je vidět, že tradičně nejnižší nezaměstnanost je v Praze a Středočeském kraji, nejvyšší pak na Vysočině a v Jihomoravském kraji.



Obrázek 2 Zaměstnanost v regionech ČR

Proměnná **věková kategorie** byla vytvořena kategorizací osob podle věku do těchto skupin:

- 1 0-29 let
- 2 30-44 let
- 3 45-64 let
- 4 65 a více

Procentuální zastoupení nezaměstnaných osob podle uvedených věkových kategorií byla pro všechny státy v celku vyrovnána, to znamená, že žádná z věkových kategorií nebyla zastoupena ve větším poměru než jiné věkové kategorie. Zajímavé je, že v České republice se například mezi nezaměstnanými osobami nevyskytují lidé nad 65 let, ale v Rakousku a Slovensku ano, avšak v nepatrném počtu. A na druhé straně Česká republika jediným státem z této trojice, kde je výraznější zastoupení osob ve věku nad 65 let, které pracují.

Proměnná **vzdělání** byla v datovém souboru charakterizována 7 variantami podle stupnice definované ISCED, a to

- 0 *predprimární (mateřské školy)*
- 1 *primární (1. stupeň základní školy)*
- 2 *nižší sekundární (2. st. základní, 1. – 4. roč. osmiletých středních)*
- 3 *vyšší sekundární (středoškolské, maturita či vyučení)*
- 4 *postsekundární (pomaturitní studium)*
- 5 *bakalářské*
- 6 *magisterské nebo doktorské*

Vzhledem k nevýznamnému vlivu na modelovanou proměnnou, byla v dalším kroku provedena agregace vzdělání do tří kategorií

- 2 *primární (kategorie 0 – 2)*
- 3 *sekundární (kategorie 3 – 4)*
- 5 *terciární (kategorie 5 – 6)*

Jako další vysvětlující proměnná, která by mohla pravděpodobnost nezaměstnanosti ovlivnit, bylo vybráno **pohlaví**

- *muž*
- *žena*

a do modelu byl zařazen rovněž **zdravotní stav**

- 1 *velmi dobrý*
- 2 *dobrá*
- 3 *uspokojivý*
- 4 *špatný*
- 5 *velmi špatný*

a dvě numerické (peněžní) proměnné – **příjmy z majetku** a **dávky v nezaměstnanosti**, které by mohly ovlivňovat motivaci k práci.

3.2 Konstrukce logistického modelu

Modelovaná binární proměnná – **ekonomický status** – reprezentuje *zaměstnanost*, resp. *nezaměstnanost* osob. K modelování byla použita logistická regrese s logitovou transformací vysvětlované proměnné. Modely byl zkonstruován v programu SAS EG, který k odhadu parametrů využívá metodu maximální věrohodnosti. Pro hodnocení výstižnosti modelu je použit zobecněný koeficient determinace a ke zjišťování významnosti proměnných (resp. jejich obměn) v modelu slouží Waldův test.

V průběhu konstrukce modelu došlo k postupnému vyřazení některých proměnných, protože do modelu vstupovaly jako statisticky nevýznamné. V tabulkách 2 – 4 jsou uvedeny výsledky odhadu parametrů redukovaných modelů.

Referenční skupinou je ve všech třech modelech tvořena *mužem ve věku 30-45 let s dobrým zdravotním stavem a středním stupněm vzdělání* (v modelu pro Českou republiku se jedná o člověka z Prahy, tj. z **regionu CZ09**).

- **Česká republika (CZ)**

Z tohoto modelu byly postupně vyřazeny proměnná **pohlaví**, **příjem z majetku** a **věkové kategorie** jako statisticky nevýznamné. Ostatní proměnné byly významné, kromě několika obměn proměnné **region** – konkrétně se jedná o region CZ03 zahrnující kraj *Plzeňský* a *Jihočeský* a CZ05 zahrnující kraj *Liberecký*, *Královéhradecký* a *Pardubický*.

Výstižnost modelu, měřená zobecněným koeficientem determinace, je velmi nízká – pouze 4,3%. Použitá množina proměnných tedy dostatečně nevysvětluje změny pravděpodobnosti

nezaměstnanosti. Bude potřeba hledat ještě další vysvětlující proměnné, které by kvalitu modelu zvýšily.

abulka 2 Odhad parametrů modelu pro Česko

Parametr	Obměna	Bodový odhad	Chyba odhadu	Waldův test	p-hod.
<i>konstantní člen</i>		-2.0306	0.1694	143.7315	<.0001
<i>dávky v nezam. (tis. EUR)</i>		1.1900	0.0890	177.3319	<.0001
<i>region</i>	CZ02	-0.3171	0.0999	10.0683	0.0015
<i>region</i>	CZ03	0.0004	0.0902	0.0000	0.9963
<i>region</i>	CZ04	0.2711	0.0903	9.0031	0.0027
<i>region</i>	CZ05	0.0229	0.0852	0.0722	0.7882
<i>region</i>	CZ06	0.366	0.0717	26.0430	<.0001
<i>region</i>	CZ07	0.3684	0.0829	19.7700	<.0001
<i>region</i>	CZ08	0.4759	0.0769	38.3198	<.0001
<i>zdravotní stav</i>	<i>velmi dobrý</i>	-0.6669	0.0755	78.1070	<.0001
<i>zdravotní stav</i>	<i>uspokojivý</i>	-0.2125	0.0714	8.8585	0.0029
<i>zdravotní stav</i>	<i>špatný</i>	0.336	0.0898	13.9911	0.0002
<i>zdravotní stav</i>	<i>velmi špatný</i>	0.9812	0.1606	37.3439	<.0001
<i>vzdělání</i>	<i>primární</i>	2.2083	0.5747	14.7631	0.0001
<i>vzdělání</i>	<i>nižší sekundární</i>	0.4314	0.1737	6.1676	0.0130
<i>vzdělání</i>	<i>postsekundární</i>	-0.981	0.3431	8.1764	0.0042
<i>vzdělání</i>	<i>bakalářské</i>	-1.204	0.1895	40.3507	<.0001
<i>vzdělání</i>	<i>magist. nebo dokt.</i>	0.1662	0.0318	27.3827	<.0001

S využitím odhadnutých parametrů uvedených v tabulce 2 můžeme zapsat linearizovaný logistický model pro Českou republiku. Po jeho odlogaritmování dostaneme nelineární (exponenciální) model pro šance ve tvaru

$$\begin{aligned}
 odds(CZ) &= \frac{p}{1-p} \\
 &= 0,1313 \cdot 1,0012^{dávky\ v\ nezam. (v\ tis. EUR)} \cdot 0,7283^{CZ02} \cdot 1,0004^{CZ03} \\
 &\cdot 1,3114^{CZ04} \cdot 1,0232^{CZ05} \cdot 1,4416^{CZ06} \cdot 1,4420^{CZ07} \cdot 1,6095^{CZ08} \\
 &\cdot 0,5127^{velmi\ dobrý\ zdr.\ stav} \cdot 0,8086^{uspokojivý\ zdr.\ stav} \cdot 1,3993^{zdr.\ stav\ špatný} \\
 &\cdot 2,6677^{zdr.\ stav\ velmi\ špatný} \cdot 9,1002^{primár.\ vzdělání} \\
 &\cdot 1,5394^{nižší\ sekund.\ vzdělání} \cdot 0,3749^{postsekund.\ vzdělání} \cdot 0,3000^{bakalář.\ vzdělání} \\
 &\cdot 1,1808^{magist.\ nebo\ dokt.\ vzdělání}
 \end{aligned}$$

odhadnutých hodnot parametrů modelu pro Českou republiku vyplývá, že šance být nezaměstnaný s vyšší **dávek v nezaměstnanosti** prakticky nemění. Nevýznamné jsou v Čechách také **příjmy z majetku** a **věkové kategorie**.

Ze získaných výsledků je dále vidět, že oproti Praze (region CZ09) je šance na nezaměstnanost významně vyšší ve všech **regionech** České republiky kromě CZ03 a CZ05. Nejvyšší je v regionu CZ02 (o 72,83%). Na dalších místech se postupně umístily regiony CZ08 (60,95%), CZ07 (44,20%), CZ06 (44,16%) a CZ04 (31,14%).

Co se týče **zdravotního stavu**, je zřejmé, že v porovnání s dobrým zdravím šanci být nezaměstnaný snižuje stav velmi dobrý (o 48,73%) a uspokojivý (o 19,14%). Naproti tomu špatný zdravotní stav tuto šanci zvyšuje a o 39,93% a velmi špatný dokonce o 166,77%.

Dalším vlivným faktorem je v Čechách také stupeň dosaženého **vzdělání**. Oproti referenční variantě – střednímu stupni – šance na nezaměstnanost vzroste nejvíce u osob s primárním vzděláním, a to dokonce na devítinásobek (o 810%). Ke zvýšení vede také

vzdělání nižší sekundární (o 54,94%) a překvapivě i magisterské nebo doktorské (o 18,08%). U osob se vzděláním postsekundárním a bakalářským dochází naopak ke snížení o 63,51% a o 70,00%.

- **Slovensko (SK)**

Při konstrukci modelu pro Slovensko byla proměnná *vzdělání* agregována do dvou skupin *základní a střední a vysokoškolské* (včetně doktorského), protože při podrobnějším dělení neměly vzdělanostní kategorie statisticky významný vliv. Statisticky nevýznamné se jevily opět některé *věkové kategorie*, konkrétně se jednalo o osoby *nad 65 let*, a také *zdravotní stav velmi špatný*. Zajímavé je to, že proměnná *pohlaví* je zde na rozdíl od Čech významná.

Zobecněný koeficient determinace je i v tomto případě velice nízký – činí pouze 4,75%, takže ani na Slovensku do modelu vybrané vysvětlující proměnné dostatečně nevystihují změny rizika nezaměstnanosti.

Tabulka 3 Model odhadu parametrů pro Slovensko

Parametr	Obměna	Bodový odhad	Chyba odhadu	Waldův test	p-hod.
<i>konstantní člen</i>		-1.2405	0.0934	176.3729	<.0001
<i>příjmy z majetku (tis. EUR)</i>		-0.3201	0.0230	190.7164	<.0001
<i>dávky v nezam. (tis. EUR)</i>		0.6982	0.0510	185.1534	<.0001
<i>zdravotní stav</i>	<i>velmi dobrý</i>	-0.2451	0.0469	27.3556	<.0001
<i>zdravotní stav</i>	<i>uspokojivý</i>	0.0831	0.0503	2.7311	0.0984
<i>zdravotní stav</i>	<i>špatný</i>	0.1761	0.0682	6.6615	0.0099
<i>zdravotní stav</i>	<i>velmi špatný</i>	0.1708	0.1188	2.0672	0.1505
<i>pohlaví</i>	<i>žena</i>	0.0671	0.0200	11.2300	0.0008
<i>věková kategorie</i>	<i>0-29 let</i>	0.7247	0.0852	72.3285	<.0001
<i>věková kategorie</i>	<i>45-64 let</i>	-0.3583	0.0853	17.6383	<.0001
<i>věková kategorie</i>	<i>65 a více</i>	-0.0628	0.2420	0.0674	0.7951
<i>vzdělání</i>	<i>základní</i>	0.3677	0.0253	211.3861	<.0001

P o jeho odlogaritmování lineárního modelu s koeficienty uvedenými v tabulce 3 dostaneme pro Slovensko následující nelineární model pro šance:

$$odds(SK) = \frac{p}{1-p} = 0,2892 \cdot 0,7261^{\text{příjmy z majetku (tis.EUR)}} \cdot 2,0097^{\text{dávky v nezam.(tis.EUR)}} \cdot 0,7826^{\text{velmi dobrý zdr. stav}} \cdot 1,0867^{\text{uspokojivý zdr. stav}} \cdot 1,1926^{\text{zdr. stav špatný}} \cdot 1,1863^{\text{zdr. stav velmi špatný}} \cdot 1,0694^{\text{žena}} \cdot 2,0641^{0-29 \text{ let}} \cdot 0,6989^{45-64 \text{ let}} \cdot 0,9391^{65 \text{ a více let}} \cdot 1,4444^{\text{zákl. vzdělání}}$$

Z odhadu parametrů modelu vyplývá, že na rozdíl od Čech na Slovensku vzroste šance být nezaměstnaný s navýšením *dávek v nezaměstnanosti* (o tisíc EUR) na dvojnásobek, přesněji o 100,97%. Významnou roli, ale opačnou, hrají také *příjmy z majetku*, díky nimž klesne šance o 27,39%. Protože zde chybí informace o situaci v jednotlivých *regionech*, nelze vliv této proměnné na Slovensku sledovat.

Oproti *dobrému zdravotnímu stavu* se na obvyklé 5%-ní hladině významnosti liší šance být nezaměstnaný pouze v případě *velmi dobrého* zdraví, kdy dojde ke zvýšení o 78,26%, a v případě *špatného* zdraví, který působí naopak snížení o 19,26%. *Uspokojivý* zdravotní stav, který zvyšuje šanci o 8,67%, je statisticky významný pouze na 10%-ní hladině a *špatný* je v modelu statisticky nevýznamný.

Na rozdíl od Čech hraje na Slovensku významnou roli také *pohlaví* a *věk* osoby, s výjimkou málo početné skupiny pracujících ve věku *65 a více let*. V porovnání s *muži* mají *ženy* šanci být nezaměstnané o něco vyšší (o 6,94%) a ve srovnání s věkovou kategorií *30 – 44 let* se šance u osob *do 30 let* zvyšuje na dvojnásobek (o 106,41%) a u osob ve věku *45 – 64 let* se naopak snižuje o 30,11%.

Jako problematický se v modelu jevil stupeň dosaženého **vzdělání**, který zaujal v modelu statisticky významnou pozici až po agregaci do dvou variant. Lze tedy konstatovat, že oproti osobám se středním a vysokoškolským vzděláním vzroste u osob se základním vzděláním šance na nezaměstnanost o 44,44%.

- **Rakousko (AT)**

V modelu pro Rakousko bylo nutné proměnnou *vzdělání* opět agregovat do dvou skupin, stejně jako v modelu pro Slovensko. Z p-hodnot Waldova testu uvedených v tabulce 3 je také vidět, že obměny *zdravotní stav uspokojivý* a *věkové kategorie 65 a více let a do 30 let* nemají na modelovanou proměnnou statisticky významný vliv.

Tabulka 4 Model odhadu parametrů pro Rakousko

Parametr	Obměna	Bodový odhad	Chyba odhadu	Waldův test	p-hod.
<i>konstantní člen</i>		-2.2850	0.1217	352.3319	<.0001
<i>příjmy z majetku</i> (tis. EUR)		0.0132	0.0034	13.8802	0.0002
<i>dávky v nezam.</i> (tis. EUR)		0.2550	0.0130	402.4186	<.0001
<i>region</i>	AT1	0.3723	0.0510	53.3500	<.0001
<i>region</i>	AT2	-0.1952	0.0669	8.5173	0.0035
<i>zdravotní stav</i>	<i>velmi dobrý</i>	-0.9381	0.0766	149.8374	<.0001
<i>zdravotní stav</i>	<i>uspokojivý</i>	0.1003	0.0775	1.6754	0.1955
<i>zdravotní stav</i>	<i>špatný</i>	0.5451	0.1080	25.4522	<.0001
<i>zdravotní stav</i>	<i>velmi špatný</i>	0.9136	0.1675	29.7602	<.0001
<i>pohlaví</i>	<i>žena</i>	0.0736	0.0373	3.8971	0.0484
<i>věková kategorie</i>	<i>0-29 let</i>	0.1937	0.1230	2.4774	0.1155
<i>věková kategorie</i>	<i>45-64 let</i>	-0.4594	0.1214	14.3281	0.0002
<i>věková kategorie</i>	<i>65 a více</i>	0.4285	0.3308	1.6778	0.1952
<i>vzdělání</i>	<i>základní</i>	0.4417	0.0393	126.0762	<.0001

Hodnota koeficientu determinace (5,99%) opět ukazuje na velmi nízkou výstižnost zkonstruovaného modelu. Znovu se tedy potvrzuje, že zvolené proměnné nevystihují hlavní příčiny vedoucí ke změnám modelované proměnné.

Po transformaci lineárního modelu pro logity jsme dostali nelineární model pro šance ve tvaru

$$\begin{aligned}
 odds(AT) &= \frac{p}{1-p} \\
 &= 0,1018 \cdot 1,0133^{\text{příjmy z majetku (tis.EUR)}} \cdot 1,2905^{\text{dávky v nezam.(tis.EUR)}} \\
 &\cdot 1,4511^{AT1} \cdot 0,8227^{AT2} \cdot 0,3914^{\text{velmi dobrý zdr. stav}} \\
 &\cdot 1,1055^{\text{uspokojivý zdr. stav}} \cdot 1,7248^{\text{zdr. stav špatný}} \cdot 2,4933^{\text{zdr. stav velmi špatný}} \\
 &\cdot 1,0764^{\text{žena}} \cdot 1,2137^{0-29 \text{ let}} \cdot 0,6317^{45-64 \text{ let}} \cdot 1,5350^{65 \text{ a více let}} \\
 &\cdot 1,5553^{\text{zákl. vzdělání}}
 \end{aligned}$$

Z odhadu parametrů modelu pro Rakousko vyplývá, že v této zemi vzroste šance být nezaměstnaný při zvýšení *dávek v nezaměstnanosti* (o tisíc EUR) o 29,05%, zatímco při zvýšení *příjmů z majetku* (o tisíc EUR) vzroste pouze o 1,33%. Co se týče vlivu *regionů*,

můžeme konstatovat, že oproti hodnotě v regionu AT3 je šance v regionu AT1 vyšší o 45,11%, zatímco v regionu AT2 je naopak nižší o 17,73%.

Na rozdíl od osob s *dobrym zdravotním stavem* je šance být nezaměstnaný u osob s *velmi dobrým* zdravím o 60,86% nižší, u osob se *špatným* zdravím je naopak o 72,48% vyšší a v případě osoby s *velmi špatným* zdravotním stavem je v Rakousku tato šance vyšší dokonce o 149,33%. Změna způsobená vlivem *uspokojivého* zdravotního stavu je v modelu statisticky nevýznamná.

Podobně jako na Slovensku i zde hraje významnou roli *pohlaví* a *věk* osoby, ovšem s výjimkou obou krajních věkových kategorií, do nichž náleží osoby *do 30 let* a osoby *ve věku 65 a více let*. V porovnání s *muži* mají *ženy* šanci být nezaměstnané vyšší o 7,64% a ve srovnání s věkovou kategorií *30 – 44 let* se šance u osob ve věku *45 – 64 let* snižuje o 36,83%.

Obměny proměnné vzdělání bylo nutno opět agregovat do dvou variant, podobně jako na Slovensku. Oproti osobám se středním a vysokoškolským vzděláním vzroste u osob s pouze základním vzděláním šance na nezaměstnanost o 55,53%.

4. Závěr

Prezentovaná analýza byla věnována míře ovlivnění nezaměstnanosti osob ve třech sousedících státech (České republice, na Slovensku a v Rakousku) výší příjmu z majetku, úrovni dávek v nezaměstnanosti, pohlavím, vzděláním a zdravotním stavem osoby a regionem, v němž daná osoba žije. K tomuto účelu byl zkonstruován logistický regresní model s binární vysvětlující proměnnou ekonomický status osoby s variantami zaměstnaná/nezaměstnaná. Koeficienty determinace modelů byly ve všech třech zemích velmi nízké (mezi 4% a 6%), proto je potřeba hledat další vysvětlující proměnné, které přispějí ke zvýšení výstižnosti modelů.

Z odhadu parametrů logistických modelů vyplynuly následující skutečnosti:

K faktorům, které jsou statisticky významné ve všech třech státech, patří především zdravotní stav osoby, dále pak výše dávek v nezaměstnanosti, stupeň dosaženého vzdělání (na Slovensku a v Rakousku agregovaný do dvou kategorií) a region, ovšem pouze v případě, že existuje příslušná data podle regionů, tj. v Čechách a Rakousku.

V Čechách není statisticky významnou proměnnou příjem z majetku, pohlaví a věk osoby. Na Slovensku a v Rakousku jsou statisticky významné všechny zvolené vysvětlující proměnné, ovšem ve slovenských datových souborech zcela chybí dělení na regiony.

Ve všech státech šanci být nezaměstnaný ovlivňuje především zdravotní stav, v Čechách ho však svým vlivem daleko předčí primární stupeň vzdělání. Překvapivé je, že v Čechách nemá na tuto šanci vliv ani pohlaví ani věková kategorie osoby.

Poděkování

Článek vznikl s podporou interního grantového projektu Vysoké školy ekonomické v Praze č. F6/30/2014.

Literatura

BARTOŠOVÁ, J. 2013. *Finanční potenciál domácností – kvantitativní metody a analýzy*. Praha: Professional Publishing.

BARTOŠOVÁ, J. – ŽELINSKÝ, T. 2013. Extent of poverty in the Czech and Slovak Republics fifteen years after split. In: *Post-Communist Economies*, Vol. 25, No. 1. pp. 119–131.

- BÍLKOVÁ, D. 2012a. Recent Development of the Wage and Income Distribution in the Czech Republic. In: *Prague Economic Papers*. Vol. 21, No. 2. pp. 233–250.
- BÍLKOVÁ, D. 2012b. Development of Wage Distribution of the Czech Republic in Recent Years by Highest Education Attainment and Forecasts for 2011 nad 2012. In: Löster, T., Pavelka, T. (eds.): *6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 162–182.
- FIALA, T. – LANGHAMROVÁ, J. 2013. Vývoj ekonomického a sociálního zatížení a stárnutí populace. In: *Politická ekonomie*, Vol. 61, No. 3. pp. 338–355.
- FIALA, T. – LANGHAMROVÁ, J. 2014. Modelling of the Future Development of the Total Amount of Premium Paid and Total Amount of Old-Age Pensions in the Czech Republic. In: *Politická ekonomie*, Vol. 62, No. 2. pp. 232–248.
- LABUDOVÁ, V. – VOJTKOVÁ, M. – LINDA, B. 2010. Application of multidimensional methods to measure poverty. In: *E+M Ekonomie a management*. Vol. 13, No. 1, pp. 6–21.
- LÖSTER, T. – LANGHAMROVÁ, J. 2012. Disparities between Regions of the Czech Republic for Non-business Aspects of Labour Market. In: Löster, T., Pavelka, T. (eds.): *6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 689–702.
- MALÁ, I. 2013. Použití konečných směsí logaritmicke-normálních rozdělení pro modelování příjmů českých domácností. In: *Politická ekonomie*, Vol. 61, No. 3. pp. 356–372.
- MAREK, L. 2013. Some Aspects of Average Wage Evolution in the Czech Republic. In: Löster, T., Pavelka, T. (eds.): *7th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 947–958.
- MAREK, L. – VRABEC, M. 2013. Probability Models of Wage Distribution. In: Vojáčková, H. (ed.): *Proceedings of 31th International Conference Mathematical Methods in Economics 2013*. Jihlava: Vysoká škola polytechnická. pp. 575–581.
- MICHÁLEK, A. – VESELOVSKÁ, Z. 2012. Vývoj a komparácia vybraných charakteristík nerovnosti a chudoby v krajinách EU. In: Pauhofová, I., Želinský, T. (eds.): *Nerovnosť a chudoba v Európskej únii a na Slovensku*. Košice: TU Košice. s. 11–22.
- PACÁKOVÁ, V. – LINDA, B. – SIPKOVÁ, L'. 2012. Distribution and Factors of the Highest Wages in the Slovak Republic. In: *Ekonomický časopis*, Vol. 60, No. 9. pp. 918–934..
- PAUHOFOVÁ, I. 2012. Generovanie chudoby vo vidieckych regiónoch Slovenska v krízovom období. In: Pauhofová, I., Želinský, T. (eds.): *Nerovnosť a chudoba v Európskej únii a na Slovensku*. Košice: TU Košice. s. 49–56.
- ŘEZANKOVÁ, H. – LÖSTER, T. 2013. Shluková analýza domácností charakterizovaných kategoriálními ukazateli. In: *E+M. Ekonomie a Management*, Vol. 16, No. 3. pp. 139–147.
- ŘEZANKOVÁ, H. – ŽELINSKÝ, T. 2014. Factors of Material Deprivation Rate in the Czech Republic by Household Type. In: *Ekonomický časopis*, Vol. 62, No. 4. pp. 394–410.
- SIPKOVÁ, L'. – SIPKO, J. 2012. Analysis of Income Inequality of Employees in the Slovak Republic. In: Löster, T., Pavelka, T. (eds.): *6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1032–1042.
- STANKOVIČOVÁ, I. – VLAČUHA, R. – IVANČÍKOVÁ, L. 2013. Trend Analysis of Monetary Poverty Measures in the Slovak and Czech Republic. In: Löster, T., Pavelka, T. (eds.): *7th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1334–1343.
- STANKOVIČOVÁ, I. – VOJTKOVÁ, M. 2007. Viacrozmerné štatistické metódy s aplikáciami. Iura Edition.

ŠIMPACH, O. 2012. Logit and Probit Models in the Probability Analysis: Change in the Probability of Death of Celiac Disease Patients. In: *Statistika*, Vol. 49, No. 4. pp. 67–80.

ŠIMPACH, O. 2013. Application of Cluster Analysis on Demographic Development of Municipalities in the Districts of Liberecký Region. In: Löster, T., Pavelka, T. (eds.): *7th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1390–1399.

VEČERNÍK, J. 2013. The changing role of education in the distribution of earnings and household income: the Czech Republic in 1988–2009. In: *Economics of Transition*, Vol. 21, No.1. pp. 111–133.

ŽELINSKÝ, T. 2013. Estimation of Subjective (Income) Poverty Lines :Sensitivity-Specificity Curves Approach. In: Hu, J. (ed.): *Social Sciences and Society. Book Series:Advances in Education Research*, Vol. 6. pp. 42–47.

ŽELINSKÝ, T. – PAUHOFVÁ, I. 2013. High Income Households in Eastern EU Countries. In: Hu, J. (ed.): *3rd Conference on Applied Social Science (ICASS 2013)*, Vol. 1. pp. 73–78.

ŽELINSKÝ, T. – STANKOVIČOVÁ, I. 2012. Spatial Aspects of Poverty in Slovakia. In: Löster, T., Pavelka, T. (eds.): *6th International Days of Statistics and Economics, Conference Proceedings*. Slaný: Melandrium. pp. 1228–1235.

Evropská agentura pro bezpečnost a ochranu zdraví při práci: Články. [online]. [cit. 2014-11-8]. Dostupné z: <http://osha.europa.eu/cs/publications/articles/promoting-active-ageing-in-the-workplace>

Adresa autora

Klára Siegelová
Vysoká škola ekonomická v Praze,
Fakulta managementu,
Jarošovská 1117/2
Jindřichův Hradec 377 01
klara_siegelova@hotmail.com

Jitka Bartošová
Vysoká škola ekonomická v Praze,
Fakulta managementu
Jarošovská 1117/2
Jindřichův Hradec 377 01
bartosov@fm.vse.cz

Analyza štruktúry poskytovateľov dlhodobej zdravotnej starostlivosti v Banskobystrickom kraji

Analysis of the long-term care providers structure in Banska Bystrica region

Beáta Gavurová, Samuel Koróny

Abstract: The article reflects an actual issue of long-term health care. The Banska Bystrica Region was chosen to analyze the structure of social service providers in order to evaluate the actual situation and specifications in existing social model of a given region. We focused on aging index in a longer period of time within our analysis as the demographic changes represent very significant factor in the process of social programmes' and strategies' formation. The analysis's outputs pointed to an inevitability of approach which focuses on deeper structured data and realization of multi-dimensional analyses that reflect on demographic aspects, morbidity and social dependency of inhabitants in a given region.

Abstrakt: Príspevok reflektuje na aktuálnu problematiku dlhodobej zdravotnej starostlivosti. Na analýzu štruktúry poskytovateľov sociálnych služieb bol zvolený Banskobystrický kraj s cieľom zhodnotenia aktuálnej situácie a špecifik v existujúcom sociálnom modeli daného regiónu. V rámci analýzy sme sa zamerali aj na index starnutia v dlhšom časovom horizonte, nakoľko demografické zmeny sú významným faktorom v procese tvorby sociálnych programov a stratégií. Výstupy analýzy poukázali na nevyhnutnosť prístupu k hlbšie štruktúrovaným údajom a realizácie viacdimeziálnych analýz reflektujúcich na demografické aspekty, morbiditu a sociálnu odkázanosť obyvateľov Banskobystrického regiónu.

Key words: long-term health care, aging index, aging population, social system, health system, social service providers.

Kľúčové slová: dlhodobá zdravotná starostlivosť (DZS), index starnutia, starnutie populácie, sociálny systém, zdravotný systém, poskytovatelia sociálnych služieb.

JEL classification: I11, I15

1. Úvod

Globálne starnutie vyvolá štrukturálne zmeny v ekonomike, zdravotníctve, sociálnej sfére, štruktúre pracovného trhu, školstve, kultúre, vo fungovaní domácností, ako aj v spolužití rodín (Ilmarinen, 2008). Vyrovnanie sa s dvojnásobným počtom seniorov si vyžaduje nastavenie efektívnych procedurálnych aktivít v ekonomike s inštitucionálnymi podporami. Neustále proklamovaná problematika globálneho starnutia je vysoko aktuálna nielen v medzinárodnom, ale aj v národnom meradle a je diskutovaná v mnohých odborných kruhoch v kontexte zdravotnej a sociálnej politiky (Vajda et al. 2009; Vajda, Vravec, 2011). Na základe údajov Eurostatu je zrejmé, že prirodzený prírastok populácie (rozdiel medzi počtom živonarodených detí a zomretých osôb) sa v EÚ zastaví v roku 2015. V roku 2035 na odvrátenie demografickej krízy už nebude dostačujúca ani miera migrácie (Population projections 2008 – 2060). Podľa aktualizovanej Prognózy vývoja obyvateľstva v okresoch SR do roku 2025 (Vaňo, Bleha, 2008) bude Banskobystrický kraj patriť medzi tri kraje, v ktorých sa očakáva úbytok celkového počtu obyvateľstva. Najviac to bude v okrese Brezno o 7 % a v okrese Detva o 6,1 %. Proces populačného starnutia bude pokračovať vo všetkých okresoch Slovenskej republiky. Staršie obyvateľstvo bude na západe a juhu Slovenska a vo veľkých mestách, mladšie obyvateľstvo bude na severe a východe Slovenska.

Uvedené skutočnosti odôvodňujú význam aktívneho prístupu v riešení problematiky starnutia vo všetkých krajinách. Je nevyhnutné na ňu apelovať a ovplyvňovať tvorcov ekonomických,

sociálnych, vzdelanostných a iných politík na národnej, ako i na regionálnej úrovni na podporu procesov aktívneho starnutia populácie, aby sa vyšší, resp. vysoký vek obyvateľov na Slovensku nestal limitujúcim faktorom merania prínosu a užitočnosti jednotlivca pre spoločnosť.

2. Vývoj hodnôt indexu starnutia v Banskobystrickom kraji

Na meranie procesu starnutia sa využívajú mnohé demografické ukazovatele. Vybrali sme si index starnutia a priemerný vek obyvateľov analyzovaného regiónu. Index starnutia je možné vypočítať viacerými metodikami (medzinárodnými podľa OECD, WHO a Eurostatu, ako aj národnou metodikou podľa ŠÚ SR). Z dôvodu obmedzenia rozsahu a obsahu príspevku sme sa zamerali na výpočet indexu starnutia len podľa slovenskej metodiky a selektovali len Banskobystrický kraj. Index starnutia podľa slovenskej metodiky (ŠÚ SR) vyjadruje počet osôb v poproduktívnom veku pripadajúci na osoby v predproduktívnom veku. Podľa metodiky ŠÚ SR sa do poproduktívneho veku započítavajú muži vo veku 60 a viac a ženy 55 a viac (v porovnaní podľa metodiky EÚ sú to všetky osoby vo veku 65 a viac), do predproduktívneho veku osoby od 0 – 14 rokov. Tým dochádza k značným rozdielom vo výsledných hodnotách získaných rôznymi metodikami, čo obmedzuje komparáciu v týchto aspektoch. Tabuľka 1 zobrazuje hodnoty indexu starnutia v SR podľa ŠÚ SR od roku 2001 do roku 2011 (údaje za rok 2012 v čase analýzy ešte neboli dostupné), ako aj indexu starnutia v Banskobystrickom kraji.

Tabuľka 1: Index starnutia v SR a v Banskobystrickom kraji

Lokalita/pohlavie	Index starnutia										
	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Slovensko											
Spolu	96,74	101,44	106,37	111,24	116,34	122,07	127,99	133,6	137,78	141,68	166,9
Muži	63,88	66,49	69,19	71,99	74,6	77,71	81,49	85,44	88,41	91,66	95,8
Ženy	131,17	138,09	145,38	152,45	160,16	168,68	176,9	184,21	189,77	194,26	175,3
Banskobystrický kraj											
Spolu	105,33	110,29	115,33	120,22	125,26	131,13	137,27	143,63	148,39	152,97	156,5
Muži	68,29	71,18	73,85	76,9	79,46	82,59	86,31	90,58	93,92	97,91	101,5
Ženy	144,13	151,33	158,94	165,46	173,03	181,66	190,57	199,22	205,54	210,43	213,9

Ak sa pozrieme na hodnoty v tabuľke 1 vidíme, že v analyzovaných rokoch narastal podiel počtu obyvateľov v poproduktívnom veku k počtu obyvateľov v predproduktívnom veku pri oboch pohlaviach nielen v Banskobystrickom kraji, ale aj v rámci Slovenska. Tento fakt odôvodňuje významnosť riešenia problematiky starnutia obyvateľstva aj v dlhodobej perspektíve a upozorňuje na mnohé závažné súvisiace trendy.

3. Metodológia a údajová základňa

Pre naše analýzy sme využili dostupnú databázu „Poskytovateľov sociálnych služieb“ (ďalej PSS) za celú SR, ktorá obsahovala k 1.6.2014 celkovo 3518 položiek a 20 premenných, z ktorých sú v tomto príspevku vybrané: Forma sociálnej služby, Zriaďovateľ, Druh soc. služby, Okres poskytovania soc. služby a Kapacita zariadenia. Na spracovanie údajov sme použili softvér MS Excel 2010, konkrétne kontingenčné tabuľky a grafy. Kontingenčné tabuľky obsahujú údaje triedené podľa dvoch alebo viacerých štatistických

znakov. Kombinácia triediacich hľadísk umožňuje vystihnúť vzájomné súvislosti a štruktúru PSS. Taktiež je možné zobraziť údaje kontingenčnej tabuľky priamo v grafe, čo je vhodné na aktuálne grafické zobrazenie každej zmeny (t.j. použitého filtra) v tabuľke.

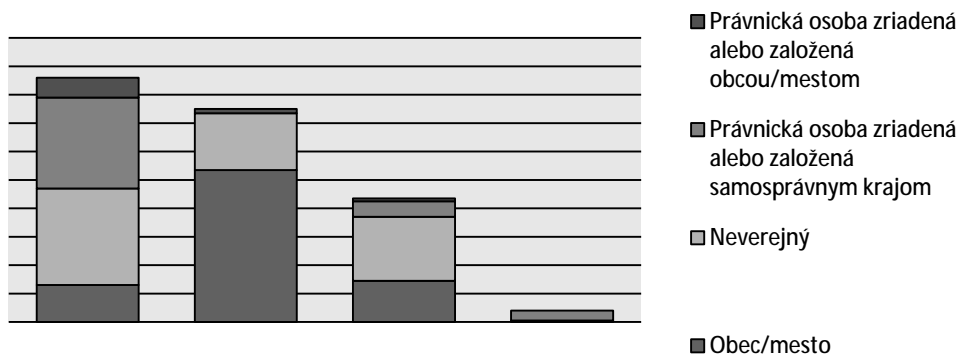
4. Výsledky a diskusia - Analýza PSS v Banskobystrickom kraji

V tejto časti predstavujeme parciálne výstupy analyzovanej štruktúry PSS v Banskobystrickom kraji podľa cielených kritérií. Tabuľka 2 deklaruje štruktúru PSS podľa zriaďovateľa v jednotlivých mestách Banskobystrického kraja.

Tabuľka 2: Počet a podiel PSS podľa poskytovateľa a formy soc. služby v Banskobystrickom samosprávnom kraji

BbSK – počet PSS	Poskytovateľ											Spolu	
	Obec/mesto			Neverejný			P.o. zriadená samosp. Krajom			P.o. zriadená obcou/mestom			
	počet	% r.	% s.	počet	% r.	% s.	počet	% r.	% s.	počet	% r.		% s.
Pobytová - ročná	26	15%	16%	68	40%	44%	64	37%	78%	14	8%	74%	172
Terénna	107	71%	66%	40	27%	26%				3	2%	16%	150
Ambulantná	29	33%	18%	45	52%	29%	11	13%	13%	2	2%	11%	87
Pobytová - týždenná				1	13%	1%	7	88%	9%				8
Spolu:	162			154			82			19			417

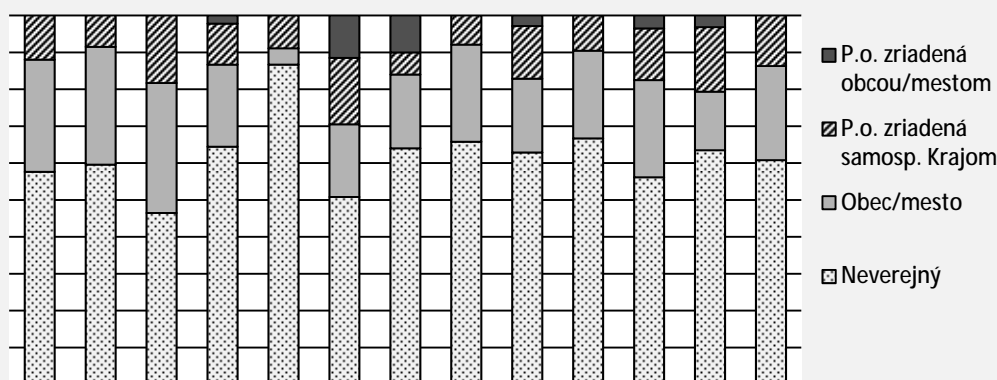
Banskobystrický samosprávny kraj - počet PSS



Obrázok 1: BbSK - počet PSS podľa formy soc. služby

Ako vyplýva z údajov v tabuľke 2 a z obrázku 1, z celkového počtu PSS v okresoch Banskobystrického kraja dominujú PSS zriadené obcou/mestom. Tento trend, ktorý je zreteľný aj v iných krajoch je signálom snahy obce/mesta podieľať sa rozhodujúcou mierou na poskytovaní sociálnych služieb v danom regióne. Najvyšší počet PSS je v terénnej forme, aj z pohľadu štruktúry PSS, ako aj formy (107 PSS). To súvisí aj s nastaveným a využívaným modelom sociálneho systému v danom kraji. Obrázok 2 poskytuje grafické znázornenie miery zastúpenia jednotlivých skupín PSS podľa zriaďovateľa v okresoch Banskobystrického kraja.

Banskobystrický samosprávny kraj - percentuálne zastúpenie PSS



Obrázok 2: BbSK (okresy) - relatívny podiel PSS podľa zriaďovateľa

Aj na obrázku 2 vidíme dominantné zastúpenie PSS zriadených obcou/mestom. V posledných rokoch je zreteľný znížený dopyt po zariadeniach pre seniorov a nárast počtu občanov s vyšším stupňom odkázanosti na pomoc inej fyzickej osoby. Vytvárajú sa v rámci existujúcich zariadení nové druhy inštitucionálnych služieb: špecializované zariadenia, zariadenia podporovaného bývania, domov na pol ceste a rehabilitačné stredisko (Konceptia rozvoja sociálnych služieb v Banskobystrickom samosprávnom kraji na roky 2011 – 2017). Tabuľka 3 poskytuje detailný pohľad na štruktúru PSS podľa formy sociálnych služieb v jednotlivých okresoch Banskobystrického kraja.

Tabuľka 3: BbSK (okresy) - počet PSS podľa formy soc. služby

BbSK Okresy	Terénna	Ambulantná	Pobytová ročná	Pobytová - týždenná	Spolu
Banská Bystrica	33	27	30	2	92
Banská Štiavnica	20	16	11		47
Brezno	37	18	16		71
Detva	23	12	9	1	45
Krupina	18	14	13		45
Lučenec	24	21	15	1	61
Poltár	25	14	11		50
Revúca	22	11	5		38
Rimavská Sobota	25	23	22		70
Veľký Krtíš	23	13	6		42
Zvolen	23	17	16	1	57
Žarnovica	22	18	20	3	63
Žiar nad Hronom	25	12	14		51

Forma poskytovania sociálnej služby v danom regióne napovedá o smere nastavenia sociálnej politiky daného regiónu. Prevažujúce formy sú terénna a ambulatná. Každá z nich sa vyznačuje svojimi špecifikami, ako to aj uvádza Zákon o sociálnych službách (Zákon č. 448/2008 Z.z. o sociálnych službách v znení neskorších predpisov). Ambulantná sociálna

služba sa poskytuje osobe, ktorá dochádza, je sprevádzaná alebo je dopravovaná do miesta poskytovania sociálnej služby. Miestom jej poskytovania môže byť aj zariadenie. Terénna sociálna služba sa poskytuje osobe v jej prirodzenom sociálnom prostredí a možno ju poskytovať aj prostredníctvom terénnych sociálnych programov s cieľom predchádzania sociálnemu vylúčeniu osoby, rodiny a komunity, ktoré sú v nepriaznivej sociálnej situácii. Pobytová sociálna služba sa poskytuje v prípade, ak jej súčasťou je ubytovanie a to ako celoročná alebo týždenná sociálna služba.

Poskytovanie terénnej sociálnej služby alebo ambulantnej sociálnej služby má prednosť pred pobytovou sociálnou službou. Ak terénna alebo ambulantná sociálna služba nie je vhodná, resp. účelná a jej poskytnutím sa dostatočne nevyrieši nepriaznivá sociálna situácia osoby, poskytuje sa pobytová sociálna služba. Z hľadiska preferencií pri pobytovej sociálnej službe má prednosť týždenná pred celoročnou. Občan má právo na výber formy poskytovanej služby a toto právo musí byť dodržané bez ohľadu na uvádzané pravidlá prednosti. Sociálna služba môže byť poskytovaná aj s využitím telekomunikačných technológií v prípadoch, ktoré sa javia ako účelné pre konkrétnu situáciu, resp. riešenie problematiku. Výsledky analýzy štruktúry PSS je vhodné konfrontovať aj s kapacitami zariadení sociálnych služieb v kontexte počtu seniorov v okresoch analyzovaného kraja. Z uvedeného dôvodu sme vybrali zariadenia, ktoré sú priamo určené seniorom (označené ako Zariadenia pre seniorov) aj tie, ktoré sú určené dospelým fyzickým osobám odkázaným na pomoc iných fyzických osôb, teda aj seniorom (označené ako DSS – Domov sociálnych služieb, Denný stacionár, Zariadenie opatrovateľskej služby, Zariadenie pre seniorov). Uvedenú štruktúru v jednotlivých okresoch analyzovaného kraja znázorňuje tabuľka 4.

Tabuľka 4: BbSK – kapacita zariadení PSS podľa druhu a formy soc. služby

BbSK - Kapacita zariadení PSS	Forma soc. služby			Spolu
	Pobytová - ročná	Ambulantná	Pobytová - týždenná	
Druh soc. služby				
Domov sociálnych služieb (DSS)	2623	244	85	2952
Zariadenie pre seniorov	1371			1371
Útulok	339			339
Špecializované zariadenie	184	11	7	202
Zariadenie opatrovateľskej služby	193			193
Zariadenie podporovaného bývania	166			166
Denný stacionár		114		114
Zariadenie núdzového bývania	109			109
Nocľaháreň		76		76
Rehabilitačné stredisko	6	27	3	36
Domov na pol ceste	28			28
Spolu	5019	472	95	5586

Ako uvádza tabuľka 4, dominantnú pozíciu v danej štruktúre PSS z hľadiska uvádzaných kapacít v registri PSS majú Domovy sociálnych služieb a Zariadenia pre seniorov. Uvedený fakt je odôvodniteľný aj tým, že Domovy sociálnych služieb, ako aj Zariadenia pre seniorov nie sú striktné špecializované len na vybrané druhy diagnóz, väčšinou sú prispôbené rôznym a viacerým typom diagnóz, ktoré sú príznačné pre danú vekovú kategóriu. Táto širšia orientácia poskytovania sociálnych služieb predurčuje aj vyšší dopyt po sociálnej a zdravotnej starostlivosti a zároveň aj vyššiu využiteľnosť kapacít PSS, resp. môže signalizovať eventuálne aj ich potenciálny nedostatok.

4. Záver

Situácia ohľadom sociálnej a zdravotnej starostlivosti starších a zdravotne a sociálne odkázaných občanov na Slovensku je neuspokojivá. V posledných rokoch sa rozpadol systém

následnej zdravotnej starostlivosti, rehabilitačných a doliečovacích oddelení, ako aj liečební pre dlhodobu chorých. Nie je zabezpečená ani dostatočná podpora domácej ošetrovateľskej starostlivosti v prirodzenom prostredí pacienta. V domovoch sociálnych služieb absentuje prepojenie so zdravotnou starostlivosťou. Dostupnosť sociálnych služieb je v súčasnosti ovplyvnená rozhodnutiami samosprávnych krajov a obcí. Preto je nevyhnutné zabezpečiť kvalitnú údajovú základňu o celom systéme sociálnych služieb, aby bolo možné realizovať viacdimeziálne analýzy reflektujúce na neustále zmeny v externom prostredí, ako aj na dostupnosť sociálnych a zdravotných služieb, morbiditu obyvateľov regiónov a mnohé ďalšie relevantné fakty, ktorých podchytenie umožní nastaviť správne smery sociálnej politiky daného regiónu, ako aj krajiny.

Príspevok uvádza čiastkové výsledky výskumu podporeného projektom VEGA č. 1/0929/14 „Viacdimeziálne ekonomicko-finančné zhodnotenie procesu zavádzania a využívania systému jednotnej zdravotnej starostlivosti a kvantifikácia jej finančných dopadov na systém zdravotníctva v SR.“

Literatúra

- VAŇO, B., BLEHA, B. 2008. PROGNOZA VÝVOJA OBYVATEĽSTVA V OKRESOCH SR DO ROKU 2025. Infostat, výskumné demografické centrum, 2008.
- ILMARINEN, J. 2008. *Ako si predĺžiť aktívny život. Stárnutie a kvalita pracovného života v Európskej únii*. 1. vyd. Bratislava: 2008, 467 s. ISBN 978-80-07-01658-3
- KONCEPCIA ROZVOJA SOCIÁLNYCH SLUŽIEB V BANSKOBYSSTRICKOM SAMOSPRÁVNOM KRAJI NA ROKY 2011 – 2017. Banskobystrický samosprávny kraj. [cit. 16. novembra 2014]. Dostupné na: <http://www.vucbb.sk/portal/sites/default/files/bbsk/dokumenty/2011/koncepcne-materialy/koncepcia-rozvoja-socialnych-sluzieb-v-banskobystrickom-samospravnom-kraji-na-roky-2011-2017.pdf>
- POPULATION PROJECTIONS 2008 – 2060. [cit. 4. apríla 2014]. Dostupné na: <http://europa.eu/rapid/pressReleasesAction.do?reference=STAT/08/119&format=HTML&aged=0&language=EN&guiLanguage=en>.
- RADA PRO STÁRNUTÍ SPOLEČNOSTI WEF.2014. Globální agenda Rady pro stárnutí společnosti: principy. [cit. 4. apríla 2014]. Dostupné na: <https://sites.google.com/site/brainjogginginfo/rada-pro-starnuti-spolocnosti-wef>
- Zákon č. 447/2008 Z.z. o peňažných príspevkoch na kompenzáciu ťažkého zdravotného postihnutia a o zmene a doplnení niektorých zákonov.
- Zákon č. 448/2008 Z.z. o sociálnych službách a o zmene a doplnení zákona č. 455/1991 Zb. o živnostenskom podnikaní (živnostenský zákon) v znení neskorších predpisov.
- VAJDA, V., VRAVEC, J. 2011. CBA analýza zavedenia ehealth v SRN s aplikáciou na Slovenskú republiku. In: *Journal of Management and Business: Research and Practice*. 2011, 3(2), pp. 148-160.
- VAJDA, V. et al. 2009. Identifikácia súčasných a očakávaných potrieb firiem v oblasti informačných a komunikačných technológií v regióne východného Slovenska. Košice, TU, 2009, 121 s. ISBN 978-80-553-0279-9

Adresa autorov:

doc. Ing. Beáta Gavurová, PhD.
MBAEkonomická fakulta TUKE
Katedra bankovníctva a investovania
Němcovej 32
040 01 Košice
E-mail: beata.gavurova@tuke.sk

RNDr. Samuel Koróny, PhD.
Inštitút ekonomických vied
Ekonomická fakulta UMB
Cesta na amfiteáter 1
974 01 Banská Bystrica
E-mail: samuel.korony@umb.sk

Statistická analýza environmentálních dat na příkladu měření elementární plynné rtuti

Statistical analysis of environmental data measurements on the example of gaseous elemental mercury

Petr Veselík, Alice Dvorská

Abstract: This paper deals with evaluation of parallel measurements of environmental data describing the concentration of gaseous elemental mercury in the atmosphere. The measurements were conducted with two identical instruments in a 10 min time step, between December 2012 and June 2013. The data was graphically visualised and evaluated by using exploratory data analysis. To perform data analysis parametric and nonparametric methods were used. The aim of the study was to compare measurement values from the two instruments.

Abstrakt: Příspěvek se zabývá vyhodnocením paralelních měření environmentálních dat zachycujících koncentrace elementární plynné rtuti v ovzduší. Měření byla prováděna v 10 minutových časových intervalech za období od prosince 2012 do června 2013 dvěma identickými analyzátoři. Data byla graficky znázorněna a vyhodnocena pomocí explorativní datové analýzy. Ke statistické analýze získaných dat byly použity parametrické i neparametrické metody. Cílem této studie bylo porovnání získaných hodnot měření oběma přístroji.

Keywords: Environmental data, Normality tests, Nonparametric methods, Homogeneity of measurements

Klíčová slova: Environmentální data, Testy normality, Neparametrické metody, Homogenita měření

JEL classification: C00, C12

1. Úvod

Statistika hraje ve vyhodnocování a interpretaci environmentálních dat stále významnější úlohu. V případě monitoringu znečištění životního prostředí chemickými látkami spočívá její význam např. v analýze časových řad, časové variability měření, korelací mezi koncentracemi monitorovaných látek, vyhodnocování jejich vzájemných vztahů metodami vícerozměrné analýzy apod. U látek, které se v životním prostředí vyskytují ve stopových množstvích a jejich koncentrace jsou často pod detekčními a kvantifikačními limity příslušných analytických metod, si lze významně vypomoci řešením problematiky cenzorovaných dat.

Příkladem environmentálního monitoringu je v této studii sledování koncentrací rtuti v ovzduší České republiky. V posledních desetiletích totiž koncentrace rtuti v ovzduší vykazují nárůst a z důvodů jejích chemicko-fyzikálních vlastností se na ni začalo nahlížet jako na globální polutant (UNECE 2010). Z hlediska zdraví člověka je zásadní fakt, že po odstranění z atmosféry do jiné složky prostředí je rtuť metylací transformována na vysoce toxickou a bioakumulativní metylrtuť (Clarkson a Magos, 2006).

Častým problémem je u aplikace statistických metod na soubory environmentálních dat malá velikost těchto souborů. Environmentální monitoring v „klasické“ podobě, tj. spočívající v odběru vzorku, jeho uchování a transportu do laboratoře a následné analýze je časově a logisticky náročný. Naopak kontinuální automatický monitoring „in situ“, kdy je studovaná látka odebírána a analyzována jedním přístrojem přímo na místě zájmu poskytuje nesrovnatelně vyšší množství dat. Až donedávna však byl tento typ měření z instrumentálního hlediska velmi složitý.

Rozdíl v získaném množství dat oběma výše zmíněnými metodami je markantní a ilustrován pomocí následujícího srovnání. Na Observatoři Košetice provozované v oblasti Vysočiny Českým hydrometeorologickým ústavem byl v letech 2006-2013 prováděn manuální odběr vzorků vzduchu analyzovaného na obsah rtuti 1x týdně. Takové odběrové schéma je dostačující z hlediska sledování dlouhodobých trendů. Od roku 2012 je na přilehlé Atmosférické stanici (AS) Křešín u Pacova provozované Centrem výzkumu globální změny AV ČR, v. v. i., prováděn automatický kontinuální monitoring atmosférické rtuti s časovým krokem 5 minut. Takovéto rozlišení je nezbytné pro potřeby vysoce časově rozlišených studií chemických procesů, dálkového transportu látky v ovzduší a vyhodnocení epizod s naměřenými vysokými koncentracemi rtuti. V tomto typu studií hraje statistika také významnou roli a její aplikace je podpořena právě velkým množstvím změřených dat.

Na AS Křešín u Pacova jsou v provozu dva automatické analyzátoři rtuti TEKRAN 2537B, jediné v České republice. Stanici tvoří zejména 250 m vysoký atmosférický stožár, který umožňuje dlouhodobé měření vertikálního gradientu koncentrací rtuti v ovzduší ve výškách 3 a 240 m nad zemí, což je experimentální design důležitý pro výše zmíněné vysoce časově rozlišené studie. Je realizován umístěním jednoho přístroje v přízemním kontejneru a druhého v malém kontejneru umístěném přímo na věži. Aby bylo možné vertikální gradient (tj. rozdíl v koncentracích rtuti změřený oběma přístroji způsobený meteorologickými podmínkami, intenzitou zdrojů rtuti a dalšími environmentálními faktory) identifikovat, je třeba nejprve popsat rozdíly v měření obou identických přístrojů. Právě to je cílem této práce.

2. Popis dat

V této studii byla použita data zachycující koncentrace celkové plynné rtuti ve vzorcích ovzduší získaná ze dvou identických přístrojů TEKRAN 2537B umístěných paralelně vedle sebe od prosince 2012 do června 2013 v přízemním kontejneru AS Křešín u Pacova. Měřicí cyklus přístroje TEKRAN 2537B je desetiminutový: 5 minut probíhá odběr vzorku vzduchu a dalších 5 minut analýza. Aby nedocházelo k časovým prodlevám, jsou v jednom přístroji umístěny dvě vzorkovací a analytické cesty časově jdoucí proti sobě: zatímco na jedné cestě probíhá odběr vzorků, na druhé probíhá analýza vzorku předchozího a naopak. V konečném důsledku tedy přístroj TEKRAN 2537B poskytuje koncentrační data každých 5 minut. AS Křešín u Pacova byla až do dubna 2014 závislá na přenosu dat přes linky blízké Observatoře Košetice, která umožňuje maximální časové rozlišení 10 minut. Proto jsou v této studii použita data 10minutová, tj. průměr vždy z jednoho odběru a analýzy na každé cestě přístroje.

Z datového souboru byla odstraněna neplatná měření, tedy hodnoty nižší než $0,1 \text{ ng/m}^3$, což je detekční limit přístroje. Dále pak byly odstraněny záznamy z automatické kalibrace přístrojů, které probíhají každý den po 19. hodině a vedou k tomu, že hodnoty poskočí cca o řád nahoru a další hodnoty, které byly dle technických záznamů o provozu přístroje označeny za neplatné. Za účelem vytvoření více homogenních datových souborů byla naměřená data roztríděna do menších časových období - přibližně týdenních intervalů. Pro každý interval byl spočítán vektor rozdílů hodnot naměřených prvním a druhým přístrojem. Daný vektor byl následně podroben statistické analýze. Na základě grafického znázornění pomocí boxplotů (viz obrázek 1) byly dále odstraněny hodnoty označené jako extrémní.

3. Použité statistické testy

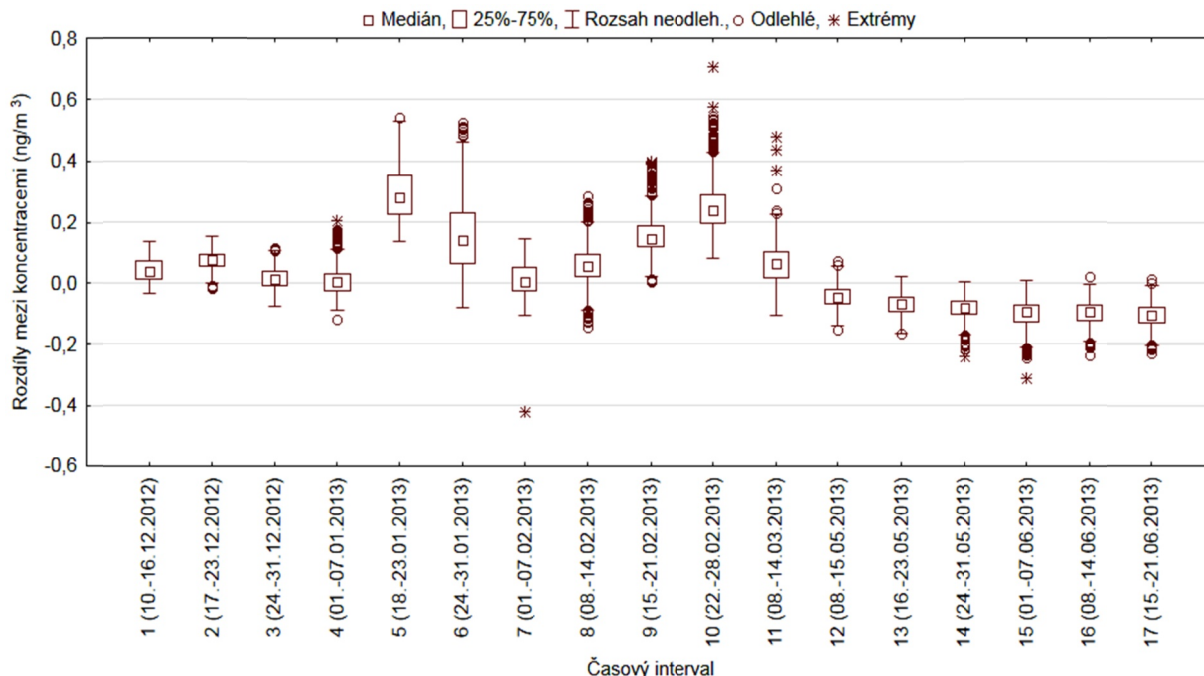
Pro srovnání funkce obou přístrojů byl použit parametrický i neparametrický přístup. Všechny použité statistické testy byly prováděny na 5% hladině významnosti. Před aplikací parametrických testů jsme ověřovali normalitu dat, a to pomocí X kvadrát testu dobré shody (Anděl, 2007, s. 271), Kolmogorovova-Smirnovova (Anděl, 2007, s. 240) a Andersonova-Darlingova testu (Cullen a Frey, 1999, s. 155). Vzhledem k tomu, že zjištěné rozdíly mezi měřeními oběma přístroji nelze ve většině případů považovat za výběry z normálního

rozdělení, byl kladen důraz na testy neparametrické. Pro test shody mediánů časových období byl použit Kruskalův-Wallisův test (Anděl, 2007, s. 243) a pro testování shody rozptylů byl proveden Flignerův-Killeenův test (Gardner, 2012, s. 253). Wilcoxonův párový test (Anděl, 2007, s. 233) byl použit pro posouzení odlišnosti měření obou přístrojů, testovala se tedy hypotéza, že medián pro vektor rozdílů měření je roven nule.

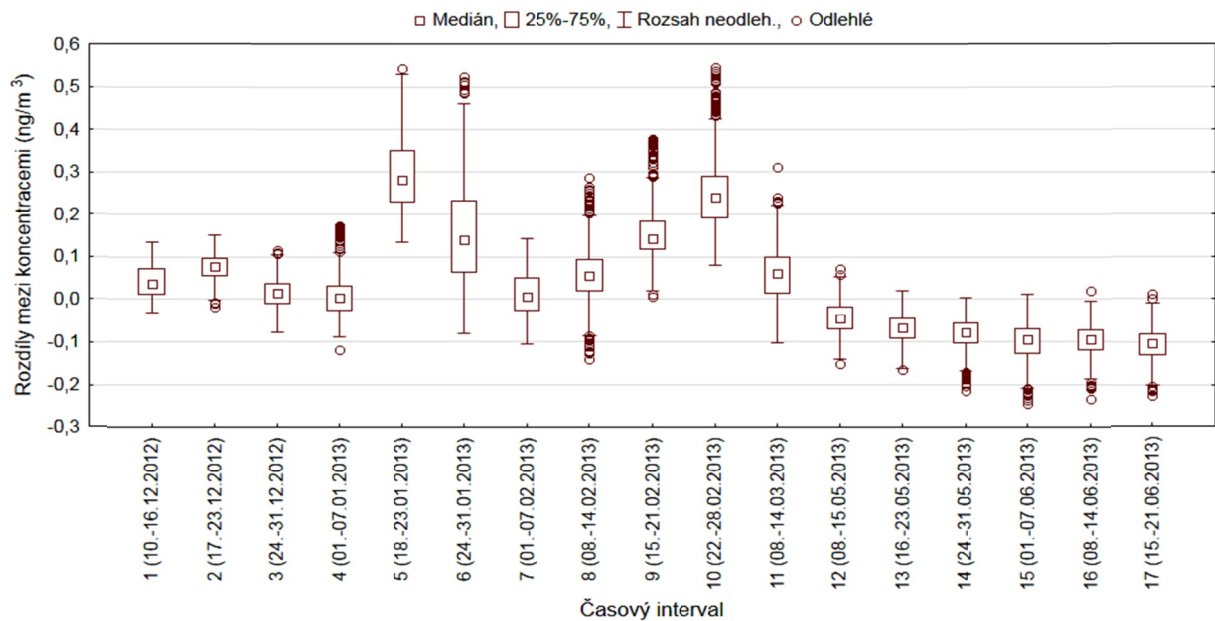
Nicméně data jsou velkého rozsahu, a proto bylo možné vyjít z asymptotické normality a použít i testy parametrické. Test rovnosti středních hodnot byl proveden pomocí jednofaktorové analýzy rozptylu (Anděl, 2007, s. 210) a Bartlettovým testem (Anděl, 2007, s. 214) byla testována homogenita rozptylů. Hypotéza o shodě středních hodnot vektoru rozdílů měření získaných prvním a druhým přístrojem (tedy pro vektor rozdílů) byla testována t-testem (Anděl, 2007, s. 75).

4. Grafická analýza dat

Grafické srovnání rozdílů párových měření získaných oběma přístroji ve všech 17 týdenních intervalech bylo provedeno pomocí krabicového grafu (viz obrázek 1). Z tohoto grafu je patrné, že hodnoty mediánů se v čase mění. Lze například vypožorovat, že mezi 4. a 5. časovým intervalem došlo k výraznému nárůstu, zatímco mezi 10. a 11. intervalem ke znatelnému poklesu. V obou případech se jedná o meziobdobí, během kterých došlo k technickým zásahům v podobě uvedení vždy jednoho z přístrojů mimo provoz a servisu (např. výměně tlakové lahve s nosným plynem argonem). Tato skutečnost by mohla částečně objasnit tyto skokové změny. Vzhledem k tomu, že detekční limit přístroje TEKRAN 2537B je $0,1 \text{ ng/m}^3$, lze mediánové hodnoty rozdílů měření v týdnech 1 až 4, 7 až 8 a 11 až 17 považovat za malé. Dále je zřejmé, že v období od 4. do 7. a od 8. do 11. časového intervalu je variabilita, stejně jako výskyt extrémních hodnot, v jednotlivých týdnech největší. Z obrázku 2, který znázorňuje párová měření bez extrémních hodnot je vidět, že rozdělení rozdílů týdenních měření nevykazuje výraznou asymetrii kolem mediánu.

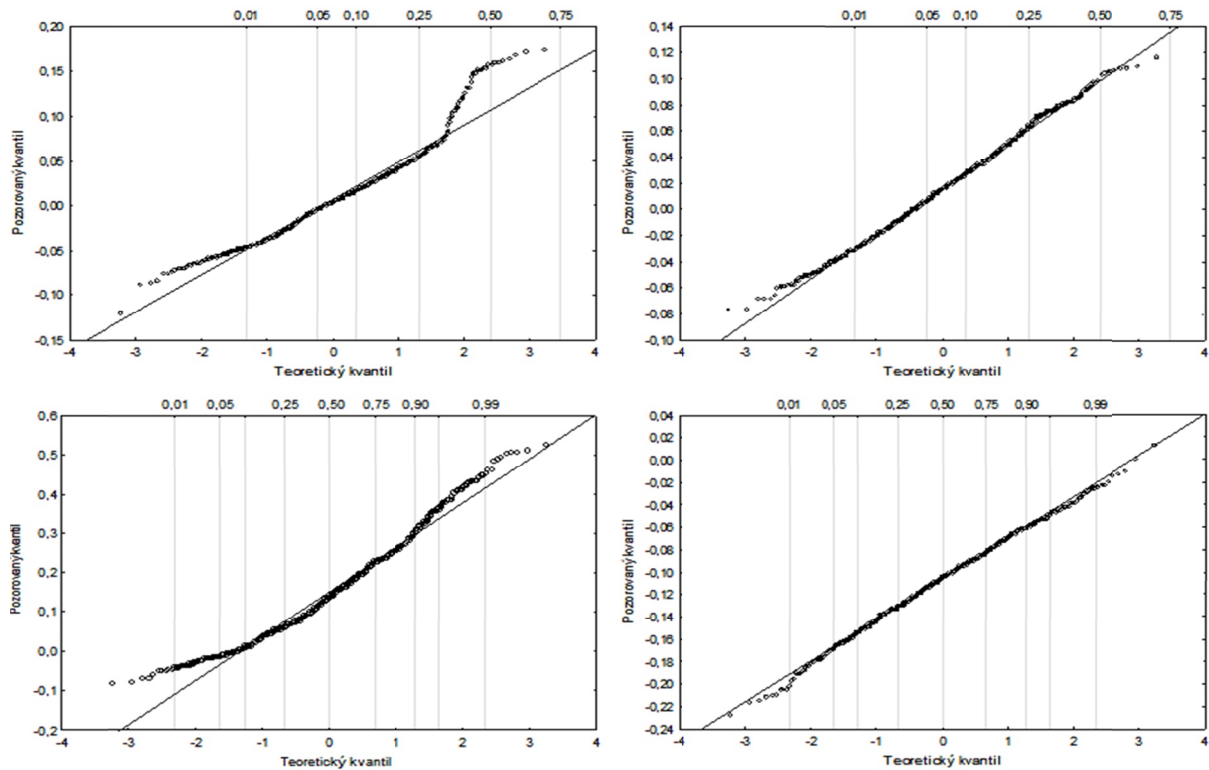


Obr.28: Krabicový graf s extrémními hodnotami pro všechny časové intervaly



Obr.2: Krabicový graf bez extrémních hodnot pro všechny časové intervaly

Dále budeme zkoumat, zda lze rozdíly považovat za výběry z normálního rozdělení. Za tímto účelem byly pro jednotlivá časová období konstruovány Q-Q ploty. Pokud dané body leží přímo na proložené přímce nebo v její těsné blízkosti, můžeme předpokládat, že data pocházejí z normálního rozdělení. Na obrázku 3 jsou zkonstruovány čtyři Q-Q ploty pro vybraná týdenní období, kde v levém sloupci jsou znázorněny Q-Q ploty, které vykazují značné odchýlení od normality a naopak v pravém sloupci jsou uvedeny Q-Q ploty, kde se vektor hodnot zdá být výběrem z normálního rozdělení.



Obr. 3: Q-Q ploty pro čtyři časová období (v levém sloupci 4. interval (nahore) a 6. interval (dole), v pravém sloupci 3. interval (nahore) a 17. interval (dole)). Pořadová čísla intervalů viz tabulka 1

5. Výsledky statistických testů

V předchozí části byly zmíněny grafické prostředky sloužící k posouzení středních hodnot (resp. mediánů) a variability rozdílů měření. Pro ověření normality rozdílů byly zkonstruovány Q-Q ploty (viz obrázek 3). V této části uvedeme v tabulce 1 výsledky statistických testů pro shodu středních hodnot (resp. mediánů), shodu rozptylů a ověření normality.

Tab. 1: Výsledky testů normality, Wilcoxonova párového testu a t-testu pro všechny časové intervaly.

INTERVAL	N	A-D	χ^2	K-S	WPT	T-T
1. 10. - 16. 12. 2012	737	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
2. 17. - 23. 12. 2012	919	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$1,31e^{-06*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
3. 24. - 31. 12. 2012	1088	0,0501	0,0533	0,2005	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
4. 01. - 07. 01. 2013	972	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	0,0008*	$2 \times 10^{-6*}$
5. 18. - 23. 01. 2013	834	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
6. 24. - 31. 01. 2013	1067	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
7. 01. - 07. 02. 2013	954	$<1 \times 10^{-6*}$	$3,57e^{-06*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
8. 08. - 14. 02. 2013	951	$<1 \times 10^{-6*}$	0,0142*	$2,42e^{-05*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
9. 15. - 21. 02. 2013	955	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
10. 22. - 28. 02. 2013	957	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
11. 08. - 14. 03. 2013	949	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$1,38e^{-06*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
12. 08. - 15. 05. 2013	1107	0,0022*	0,0414*	0,0063*	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
13. 16. - 23. 05. 2013	1107	0,0180*	0,0003*	0,0110*	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
14. 24. - 31. 05. 2013	1113	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$6,55e^{-06*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
15. 01. - 07. 06. 2013	970	$<1 \times 10^{-6*}$	$6,40e^{-06*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
16. 08. - 14. 06. 2013	971	0,0040*	0,0060*	0,0107*	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$
17. 15. - 21. 06. 2013	973	0,1570	0,1764	0,0774	$<1 \times 10^{-6*}$	$<1 \times 10^{-6*}$

Symbolem * jsou označeny hodnoty příslušné statistiky, kdy se na 5% hladině významnosti zamítá odpovídající nulová hypotéza, první sloupec udává pořadové číslo období (týdne) a druhý sloupec uvádí konkrétní časové rozmezí, pro které bylo měření vyhodnocováno, N udává počet měření v daném časovém intervalu, A-D, χ^2 a K-S jsou p-hodnoty provedených testů normality, WPT je p-hodnota Wilcoxonova párového testu a T-T je p-hodnota t-testu.

Z provedených testů dobré shody je zřejmé, že na 5 % hladině významnosti nezamítáme nulovou hypotézu o normalitě dat ve 3. a 17. týdenním intervalu. To je také patrné z již dříve zkonstruovaných Q-Q plotů na obrázku 3 v grafech v pravém sloupci (horní graf: 3. interval, spodní graf: 17. interval). Z provedené analýzy dále vyplývá, že existují statisticky významné rozdíly mezi měřeními obou přístrojů ve všech uvažovaných obdobích, což je zřejmé z výsledků Wilcoxonova párového testu a t-testu. Tedy přesněji řečeno, na 5 % hladině významnosti zamítáme nulovou hypotézu, že střední hodnota (resp. medián) pro rozdíly měření je rovna nule.

Cílem další aplikace těchto testů na vybrané časové bloky složené z na sebe navazujících intervalů bylo nalezení takových bloků, které by byly homogenní vzhledem ke střední hodnotě (mediánu) nebo rozptylu. Výsledky pro takové homogenní bloky jsou zachyceny v tabulce 2. Zmíněny jsou pouze časové bloky, ve kterých jsme nezamítli hypotézu o shodě rozptylů nebo středních hodnot.

Tab. 2: Výsledky dílčích testů na ověření homogenity časových bloků.

INTERVAL		PARAMETRICKÉ		NEPARAMETRICKÉ		IS	
		BA-T	ANOVA	FL-K	KR-W	DM	HM
13+14	16. - 31. 05. 2013	0,4345	$<1 \times 10^{-6*}$	0,3383	$<1 \times 10^{-6*}$	-0,076	-0,074
15+16	01. - 14. 06. 2013	$1,81e^{-06*}$	0,0714	0,0002*	0,4856	-0,100	-0,097
16+17	08. - 21. 06. 2013	0,6827	$<1 \times 10^{-6*}$	0,4948	$<1 \times 10^{-6*}$	-0,103	-0,100

Symbolem * jsou označeny hodnoty příslušné statistiky, kdy se na 5% hladině významnosti zamítá odpovídající nulová hypotéza, v prvním a druhém sloupci tabulky je uvedeno sledované časové období, BA-T a ANOVA jsou p hodnoty Bartlettova testu a ANOVY, FL-K a KR-W jsou p-hodnoty Fligner-Killeenova a Kruskal-Wallisova testu, DM a HM jsou dolní a horní meze intervalu spolehlivosti pro střední hodnotu rozdílu v jednotlivých spojených obdobích.

Hypotéza o shodě středních hodnot, resp. mediánů nebyla zamítnuta na 5% hladině významnosti pro 15. a 16. interval (01. - 14. 06. 2013). Hypotéza o shodě rozptylů nebyla zamítnuta na 5% hladině významnosti pro dva časové bloky: 13. a 14. interval (16. - 31. 05. 2013); 16 a 17 interval (08. - 21. 06. 2013). Tyto výsledky jsou potvrzeny také na obrázku 2, kde lze homogenní úseky dobře rozeznat.

6. Závěr

Statistické zpracování datových souborů z paralelního měření dvěma identickými přístroji vedlo k zjištění, že oba přístupy, jak parametrický tak i neparametrický udávají srovnatelné výsledky testů. Protože datové soubory jsou velkého rozsahu, což u environmentálních dat nebývá časté, lze spoléhat na asymptotickou normalitu. V takovém případě můžeme tedy použít i metody pro normální data a spoléhat na robustnost těchto metod.

Z hlediska monitoringu obsahu rtuti v ovzduší dvěma automatickými přístroji umožňujícími kontinuální měření je důležité zjištění, že rozdíly v získaných údajích jsou většinou malé (medián rozdílu měření u 13 ze 17 zkoumaných časových intervalů nepřesahoval detekční limit přístroje, tj. $0,1 \text{ ng/m}^{-3}$), avšak mohou kolísat. Toto zjištění bylo umožněno déle trávající (sedmiměsíční) paralelní odběrovou kampaní, což je v protikladu k podobným studiím, které jsou obvykle navrženy pro časové období řádu týdnů nezahrnujících běžné provozní problémy přístrojů (např. Gustin et al., 2013). Variabilita zjištěných rozdílů v měřeních oběma přístroji je důležité zjištění a je třeba ji mít na paměti zejména v době před a po servisním zásahu na přístrojích.

Poděkování

Data zpracovaná v tomto článku byla získána z přístrojů pořízených v rámci projektu OP VaVpl z fondů EU a státního rozpočtu České republiky (Projekt: CzechGlobe Centre for Global Climate Change Impact Studies, Reg. No. CZ.1.05/1.1.00/02.0073). Článek vznikl za podpory projektu specifického výzkumu SV14-FEM-K101-01-MICH.

Literatura

- [1] ANDĚL, Jiří. *Základy matematické statistiky*. Druhé opravené vydání. Praha: MATFYZPRESS, 2007, 358 s. ISBN 80-7378-001-1.
- [2] CLARKSON T.W. a L. MAGOS, 2006. *The toxicology of mercury and its chemical compounds*. Critical Reviews in Toxicology 36, 609-662.
- [3] CULLEN, A. C. a H. C. FREY. *Probabilistic techniques in exposure assessment: a handbook for dealing with variability and uncertainty in models and inputs*. New York: Plenum, 1999, 336 s. ISBN 0-306-45957-4.

Dostupné z: <http://stat.ethz.ch/R-manual/Rpatched/library/stats/html/fligner.test.html>

- [4] GARDENER, M. *The Essential R Reference*. První. Indianapolis: Wiley, 2012, 576 s. ISBN 978-1-118-39141-9.
- [5] Gustin M.S., Huang J., Miller M.B., Peterson C., Jaffe D.A., Ambrose J., Finley B.D., Lyman S.N., Call K., Talbot R., Feddersen D., Mao H., Lindberg S.E., 2013. Do we understand what the mercury speciation instruments are actually measuring? Results of RAMIX. *Environmental Science & Technology* 47, 7295-7306
- [6] UNECE 2010. *Hemispheric Transport of Air Pollution 2010: Part. B: Mercury. Air Pollution Studies No.18*. UN New York and Geneva, 192 s.

Adresy autorů:

Petr Veselík, Ing.
Univerzita obrany, katedra ekonometrie
Šumavská 4, 662 10 Brno
petr.veselik@unob.cz

Alice Dvorská, RNDr. Ph.D.
Centrum výzkumu globální změny AV ČR,
v. v. i.
Bělidla 4a, 603 00 Brno
dvorska.a@czechglobe.cz

Zo života SŠDS

From live of SSDS

Z histórie seminárov Výpočtová štatistika

From the history of seminars Computational Statistics

Pri príležitosti 23. ročníka seminára Výpočtová štatistika uvádzame stručnú chronológiu predošlých ročníkov.

Prvý seminár sa uskutočnil 9. - 10. 12. 1986 z iniciatívy zamestnancov Katedry štatistiky VŠE v Bratislave a Katedry štatistiky VŠE v Prahe zaoberajúcimi sa problematikou využitia výpočtovej techniky v riešení štatistických úloh. Príspevky účastníkov boli uverejnené v Informáciách SDŠS č. 3 a č. 4 v roku 1986.

Miestom konania Seminárov bola do roku 2011, budova Infostat-u, v roku 2012 kongresová sála ŠÚ SR na Hanulovej 5/c v Bratislave. Väčšina seminárov sa organizovala v spolupráci so Štatistickým úradom SR (resp. SŠU v Bratislave) a Infostat-om Bratislava (resp. VUSEIaR Bratislava), v roku 2012 je spoluorganizátorom aj Prírodovedecká fakulta UK Bratislava. V aktuálnom 22. ročníku seminára bolo miesto konanie Aula Prírodovedeckej fakulty UK v Bratislave a druhá časť akcie pre mladých: Pohľady do analytiky - Analytika očami profesionálov - pásmo prednášok v zasadačej miestnosti Fakulty managementu UK, Odbojárrov 10, Bratislava.

Druhý seminár prebehol 8. 12. 1987, tretí seminár 11. - 12. 12. 1990. Potom nastala prestávka v organizácii seminárov Výpočtovej štatistiky a 4. seminár sa uskutočnil 7. - 8. 12. 1994.

Od 5. seminára uskutočneného 5. - 6. 12. 1996 sa už realizuje každoročne ako medzinárodný seminár.

6. medzinárodný seminár Výpočtová štatistika sa uskutočnil 4.- 5. 12. 1997,
7. medzinárodný seminár Výpočtová štatistika sa uskutočnil 3. - 4. 12. 1998,
8. medzinárodný seminár Výpočtová štatistika sa uskutočnil 2. - 3. 12. 1999,
9. medzinárodný seminár Výpočtová štatistika sa uskutočnil 7. - 8. 12. 2000,
10. medzinárodný seminár Výpočtová štatistika sa uskutočnil 6. - 7. 12. 2001,
11. medzinárodný seminár Výpočtová štatistika sa uskutočnil 5. - 6. 12. 2002,
12. medzinárodný seminár Výpočtová štatistika sa uskutočnil 4. - 5. 12. 2003,
13. medzinárodný seminár Výpočtová štatistika sa uskutočnil 2. - 3. 12. 2004,
14. medzinárodný seminár Výpočtová štatistika sa uskutočnil 1. - 2. 12. 2005,
15. medzinárodný seminár Výpočtová štatistika sa uskutočnil 7. - 8. 12. 2006,
16. medzinárodný seminár Výpočtová štatistika sa uskutočnil 6. - 7. 12. 2007,
17. medzinárodný seminár Výpočtová štatistika sa uskutočnil 4. - 5. 12. 2008,
18. medzinárodný seminár Výpočtová štatistika sa uskutočnil 3. - 4. 12. 2009,
19. medzinárodný seminár Výpočtová štatistika sa uskutočnil 2. - 3. 12. 2010,
20. medzinárodný seminár Výpočtová štatistika sa uskutočnil 1. - 2. 12. 2011,
21. medzinárodný seminár Výpočtová štatistika sa uskutočnil 6. - 7. 12. 2012,
22. medzinárodný seminár Výpočtová štatistika sa uskutočnil 5. - 6. 12. 2013 a
23. medzinárodný seminár Výpočtová štatistika sa uskutočnil 4. - 5. 12. 2014.

Príspevky 2. seminára boli opublikované v Informáciách SDŠS č. 1/1989 a od 3. seminára sa publikujú v samostatnom Zborníku príspevkov príslušného seminára. Od 14. seminára sú

príspevky publikované vo vedeckom časopise SŠDS FORUM STATISTICUM SLOVACUM (FSS). Príspevky mladých v rámci prehliadky prác mladých štatistikov a demografov boli publikované spolu s príspevkami účastníkov seminára Výpočtová štatistika. Počnúc 19-tým ročníkom seminára Výpočtová štatistika je vydávaný, z príspevkov zaslaných do Prehliadky prác mladých štatistikov a demografov, samostatný zborník, v spolupráci s Klubom Dispersus

Zameraním seminára je problematika na rozhraní počítačových vied a štatistiky.

Tematické okruhy posledných seminárov sa nemenia:

- praktické využitie paketov štatistických programov,
- práca s rozsiahlymi súbormi údajov,
- vyučovanie výpočtovej štatistiky a príbuzných predmetov,
- praktické aplikácie výpočtovej štatistiky,
- iné.

V čase konania seminára Výpočtová štatistika sa uskutočňuje aj **prehliadka prác mladých štatistikov a demografov**. Táto akcia prebieha od 7. seminára. Na 8. medzinárodnom seminári prezentovalo svoje práce 5 mladých štatistikov a demografov, na 9. medzinárodnom seminári už bolo 20 prác mladých štatistikov a demografov, na 10. bolo prihlásených 26 prác a na 11. bolo prihlásených 18 prác, ale vzhľadom na niekoľko prác vypracovaných skupinou autorov bol počet účastníkov vyšší než predošlý rok. Na 12. seminári bolo prihlásených 19 prác, pričom niektoré sú prácou viacerých autorov. Na ďalšom 13. seminári bolo prihlásených 9 prác od 12 autorov. V rámci 14. seminára bolo prihlásených 15 sólových prác mladých autorov. Na 15. seminári bolo prihlásených 20 prác mladých autorov. V rámci 16. seminára bolo prihlásených 17 sólových prác mladých autorov. V rámci 17. seminára bolo prihlásených 15 sólových prác mladých autorov. V 18. ročníku bolo prihlásených 12 sólových prác mladých autorov. V 19. ročníku bolo prihlásených 15 prác autorov. V 20. ročníku seminára bolo prihlásených 15 prác mladých autorov. V 21. ročníku seminára bolo prihlásených 19 prác mladých autorov, z toho jedna práca bola napísaná dvoma autormi. V 22. ročníku seminára bolo prihlásených 9 prác mladých autorov, z toho dve práce mali po dvoch autorov. V aktuálnom ročníku 23. seminára bolo prihlásených 9 prác mladých autorov.

Prípadní záujemcovia z radov mladých štatistikov a demografov (za mladých považujeme štatistikov a demografov pred ukončením vysokej školy) môžu získať informácie na www.ssds.sk, blok akcie a na e-mailových adresách:

chajdiak@statis.biz ; jan.luha@fmed.uniba.sk ; iveta.stankovicova@fm.uniba.sk.

Informácie o najbližšom seminári získate na webovskej stránke SŠDS www.ssds.sk.

doc. Ing. Jozef Chajdiak, CSc.
STU Bratislava
predseda SŠDS

RNDr. Ján Luha, CSc.
LFUK Bratislava
vedecký tajomník SŠDS

doc. Ing. Iveta Stankovičová, PhD.
FM UK Bratislava
predsedníčka programového a
organizačného výboru seminára
Výpočtová štatistika

OBSAH CONTENTS

	Foreword	1
	Predhovor	2
Ľudmila Bartóková	Vonkajšia rovnováha po roku 2000 v podmienkach SR a ČR – čistý export a čisté úspory External balance after 2000 in Slovakia and the Czech Republic - net exports and net savings	3
Jitka Bartošová, Jakub Stejskal	Vývoj prostorové distribúcie príjmovej nerovnosti v Európe v letech 2007-2012 Development of spatial distribution of income inequality in Europe	9
Martin Bod'a	Poznámka k váženému dynamickému rozkladu finančného ukazovateľa na báze multiplikatívnej väzby A note on weighted dynamic decomposition of a financial indicator based on a multiplicative relationship	18
Kornélia Cséfalvaiová, Jana Langhamrová, Jitka Langhamrová	Civilizačné choroby a stredná dĺžka života Civilization diseases and life expectancy	24
Tomáš Fiala, Zdeněk Pavlík	Zvyšovanie dôchodového veku – jedna z možných ciest k finančnej udržiteľnosti dôchodových systémů v ČR, SR i Európe Rising of Retirement Age – A Possible Way to Sustainability of Pension Systems in Czech Republic, Slovakia and Europe	29
Lukáš Friga, Vladimír Gazda	Režim vývoja burzového indexu z hľadiska klastrovej analýzy Stock Exchange Index Development Regime as an Application of the Cluster Analysis	35
Simona Fučíková, Michaela Antovová	Zdraví a ľudský kapitál v Českej republike a na Slovensku Health and human capital in the Czech Republic and the Slovakia	40
Marek Gróf	Ukazovatele centrality pri hodnotení polohy predajní Centrality measures in retail unit location evaluation	45
Jozef Chajdiak	Miery inovácie – prehľad príkladov Degree of innovation - summary of examples	50
Jozef Chajdiak, Lucia Coskun	Rozbor súboru vinárskych podnikov pomocou všeobecného lineárneho regresného modelu Analysis of a wine companies file using GLM	54
Štefan Kováč	Vzťah investícií a hrubej pridanej hodnoty v IKT sektore Relation between investment and gross value added in ICT sector	59
Jana Kubanová, Michal Kuban	Hodnocení úrovně kompetencí získaných studiem Evaluation of the level of competence gained by studying	65
Nikolay Kulbakov	Segmentace států EU podle sektorových indikátorů důvěry Segmentation of EU according sectorial confidence indicators	69
Viera Labudová	Príjmová nerovnosť na Slovensku, 2004 – 2010 Income inequality in Slovakia, 2004-2010	75
Bohdan Linda, Michal Kuban	Úspěšnost studentů v bakalářském stupni studia The success rate of students in the bachelor study	80
Radka Lipovská, Peter Tóth	Modelovanie pozitívnych vzťahov na báze stochastického agentovo orientovaného modelu Positive relationships modelling based on stochastic agent-based model	84

Elena Makhalova, Kornélia Cséfalvaiová, Iva Pecáková	Odhad správného počtu shluků pro překrývající se data Estimating the Correct Number of Clusters for Overlapping Data	90
Tomáš Marcinko	Dôsledky porušenia predpokladu nezávislosti v prípade jednovýberového Studentovho t-testu Consequences of Independence Assumption Violation Regarding One-Sample Student's t-test	97
Igor Melicherčík, Gábor Szűcs	Vplyv vybraných faktorov na zisk životnej poisťovne pri doživotných dôchodkoch vyplácaných z úspor v II. dôchodkovom pilieri na Slovensku The influence of selected factors on the profit of life insurance company in the case of life annuities paid from the second pillar in Slovakia	103
Rajmund Mirdala, Aneta Svrčeková	International Financial Integration and Volatility of Financial Flows Medzinárodná finančná integrácia a volatilita finančných tokov	115
Tomáš Pavelka, Tomáš Löster	Využití shlukové analýzy ke srovnání dopadů celosvětové finanční krize na trh práce zemí Evropské unie The use of cluster analysis for a comparison of the global financial crisis effects on the labour market of the European Union	125
David Pejčoch	Paradox imputace Imputation Paradox	131
Tomáš Pivoňka, Tomáš Löster	Shluky nefinančních podniků evropských zemí dle metodiky národních úctů Clusters of the non-financial organizations of European union countries by the national account methodology	137
Ivana Polakevičová, Edita Szabová, Tatiana Kamenská	Prieskum závislosti medzi finančnými možnosťami študentov generácie Y a nakupovaním na zľavových portáloch Survey on the Dependence between the Financial Options of Generation Y Students and Shopping at Discount Portals	143
Petra Širůčková	Řešení problému polynomické kalibrace metodou maximální věrohodnosti Solving the problem of polynomial calibration by the maximum likelihood method	148
Patrik Sleziak	Analyza dat pomocou štatistického softvéru R Analysis of data using R statistical software	159
Patrik Sleziak, Miroslav Sabo	Rodové rozdiely v prevalencii špecifických fóbií Gender differences in the prevalence of specific phobias	165
Iveta Stankovičová, Róbert Vlačuha	Vývoj miery rizika monetárnej chudoby v regiónoch Slovenska Trend of monetary poverty measures in Slovak regions	171
Edita Szabová, Edita Tatarová	McNemarov test a jeho aplikácie v ošetrovatel'stve McNemar's Test and Its Application in Nursing	178
Ondřej Šimpach	Shluková analýza a metoda hlavních komponent v aplikaci vyhledávání odlehlých pozorování věkově-specifických měř úmrtnosti v České republice a na Slovensku Cluster analysis and principal component method in the application of searching outliers of age-specific mortality rates in the Czech Republic and Slovakia	183
Alena Tartal'ová, Vladan Fejlek	Faktory vplývajúce na subjektívne pociťované šťastie Factors affecting the subjective well-being	190

Veronika Zlaczká, Viliam Kováč, Jozef Heteš	Zmena štruktúry spotrebného koša v závislosti od sadzby dane z pridanej hodnoty po prijatí eura Change of consumer basket structure depending on value added tax rate after adopting euro	196
Tomáš Želinský	Faktory pracovných a podnikateľských príjmov v krajinách EÚ Factors of Labour and Entrepreneurial Income in the EU Countries	202
Martina Žudelová	Pracovná migrácia a mobilita v okresoch Slovenskej republiky Labour migration and mobility in the districts of the Slovak Republic	210
Jitka Bartošová, Klára Siegelová	Analýza nezaměstnanosti osob v závislosti na veku, vzdelání a dalších faktorech v mezinárodním srovnání The analysis of the unemployment depending on age, education and other factors in international comparison	217
Beáta Gavurová, Samuel Koróny	Analýza štruktúry poskytovateľov dlhodobej zdravotnej starostlivosti v Banskobystrickom kraji Analysis of the long-term care providers structure in Banska Bystrica region	228
Petr Veselík, Alice Dvorská	Statistická analýza environmentálných dat na príkladu měření elementární plynné rtuti Statistical analysis of environmental data measurements on the example of gaseous elemental mercury	234
	Zo života SSDS From live of SSDS	
	Z histórie seminárov Výpočtová štatistika From the history of seminars Computational Statistics	242
	OBSAH CONTENTS	244

SAS® onDemand for Academics

Pre akademické účely sprístupňujeme SAS aj online prostredníctvom cloudu. Spracovanie analýz prebieha na vzdialenom hositeľskom serveri, čím každému odpadávajú náklady na hardvér, správu softvéru, inštaláciu a diskový priestor. Študenti aj pedagógovia tak pre potreby výučby môžu získať non-stop prístup k najnovšej verzii softvéru SAS nielen v škole, ale aj na ich domácich počítačoch – všade tam, kde majú prístup k vysokorychlostnému internetu.

Dostupný softvér:

SAS® Enterprise Guide® (Base SAS)
 SAS® Enterprise Miner™
 SAS® Forecast Server
 SAS® Studio

Viac informácií nájdete na:

www.sas.com/slovakia/academic/soda

E-learning

Študentom a pedagógom zapojeným do nášho akademického programu ponúkame zadarmo aj obľúbené SAS e-learningové kurzy. K dispozícii máme viacero e-kurzov od základných po veľmi pokročilé. Prístupové heslá k jednotlivým kurzom zasielame na požiadanie mailom. Ak o ne máte záujem, neváhajte nás kontaktovať e-mailom.

Dostupné e-learningové kurzy pre Academic:

- SAS® Enterprise Guide® 1: Querying and Reporting (EG4.3)
- SAS® Programming Introduction: Basic Concepts
- SAS® Programming 1: Essentials
- SAS® Programming 2: Data Manipulation Techniques
- SAS® SQL 1: Essentials
- SAS® Programming 3: Advanced Techniques and Efficiencies
- SAS® Macro Language 1: Essentials
- SAS® Enterprise Guide® 1: Querying and Reporting (EG4.2)
- Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression
- Rapid Predictive Modeling for Business Analysts (EM 6.2)
- SAS® Enterprise Guide®: ANOVA, Regression, and Logistic Regression
- Applied Analytics Using SAS® Enterprise Miner™
- Rapid Predictive Modeling for Business Analysts (EM 7.1)
- SAS® Enterprise Guide® 1: Querying and Reporting (EG 5.1)
- SAS® Enterprise Guide®: ANOVA, Regression, and Logistic Regression [4.3 and 5.1]
- SAS® Enterprise Guide® 2: Advanced Tasks and Querying (EG 5.1)
- Applied Analytics Using SAS® Enterprise Miner™
- Introduction to Statistical Concepts
- Querying, Reporting, and Analyzing Data Using SAS® Enterprise Guide®
- Applied Analytics Using SAS® Enterprise Miner™
- Applied Analytics Using SAS® Enterprise Miner™



The Power to

KNOW

“

„Vďaka tomu, že študenti využívajú vysoko profesionálny analytický softvér, získavajú predstavu o tom, v akých situáciách sa dajú analýzy dát využiť. Učia sa, ako v praxi používať rôzne analytické metódy a interpretovať výsledky, ktoré zo softvéru vziať. Rozhodovať iba na základe intuície sa dnes nedá. Analýzy dát a získavanie užitočných informácií zohrávajú v biznisových rozhodnutiach čoraz väčšiu úlohu.“

doc. Ing. Iveta Štankovičová, PhD.
vysokoškolský pedagóg
Fakulta managementu
Univerzita Komenského v Bratislave

“

„Po zoznámení sa so softvérom SAS® sa pre mňa stala štatistika po množstve teoretických poznatkov oveľa atraktívnejšou. Zaujímavé grafické spracovanie i široká škála možností využitia boli hlavné dôvody, prečo som sa rozhodol venovať väčšiu pozornosť práve tomuto softvéru.“

Ondrej Dúžik
doktorand
Fakulta hospodárskej informatiky
Ekonomická univerzita v Bratislave

“

„Študenti sú neraz milo prekvapení, že môžu pracovať s aktuálnym softvérom, ktorý využíva toľko firiem po celom svete. Veľmi pozitívne vnímali aj prednášky SAS konzultantov u nás na fakulte, ktorí im priblížili využitie SAS-u v praxi. Som presvedčená, že vedieť aspoň niečo zo SAS-u, prináša študentom na pohovoroch značnú konkurenčnú výhodu oproti ostatným a nadobudnuté vedomosti im budú nápomocné v ich budúcom zamestnaní.“

Ing. Renáta Prokeiová PhD.
vysokoškolský pedagóg
Fakulta ekonomiky a manažmentu
Slovenská poľnohospodárska univerzita v Nitre



THE
POWER
TO KNOW.

SAS je popredným svetovým poskytovateľom riešení v oblasti biznis analytiky a odborných služieb. Celosvetovo Softvér SAS® využíva viac ako 70 000 organizácií pre zlepšenie svojej výkonnosti pomocou spracovania obrovských objemov dát, čo vedie napr. k rýchlejšim a presnejším strategickým podnikateľským rozhodnutiam, získavajúci lepší vzťah so zákazníkmi a dodávateľmi a dodržiavaniu regulačných požiadaviek. Spoločnosť SAS má viac ako 13 600 zamestnancov vo viac ako 400 pobočkách v 55 štátoch sveta. Na Slovensku má svoje zastúpenie od roku 1995 a špecializuje sa hlavne na finančný sektor, telekomunikácie a energetiku. Viac informácií nájdete na: www.sas.com/slovakia.



Zaujímavé pozície s nadnárodnými zodpovednosťami

Finančné centrum zdieľaných služieb (Finance & Controlling shared service center)
Tvoja skvelá budúcnosť

- reportovanie
- plánovanie a rozpočtovanie
- správa pohľadávok
- spracovanie zákazníckych zmlúv
- výkazy peňažných tokov
- interné platby
- účasť na mesačných finančných uzávierkach

Pridaj sa k tým najlepším!

T · · Systems ·

Kontakt:

petra.judicakova@t-systems.sk

Pokyny pre autorov

Jednotlivé čísla vedeckého recenzovaného časopisu FORUM STATISTICUM SLOVACUM sú prevažne tematicky zamerané zhodne s tematickým zameraním akcií SŠDS. Príspevky v elektronickej podobe prijíma zástupca redakčnej rady na elektronickej adrese uvedenej v pozvánke na konkrétne odborné podujatie Slovenskej štatistickej a demografickej spoločnosti. Akceptujeme príspevky v slovenčine, češtine, angličtine, nemčine, ruštine a výnimočne po schválení redakčnou radou aj inom jazyku. Názov word-súboru uvádzajte a posielajte v tvare: **priezvisko_nazovakcie.doc resp. docx**

Forma: Príspevky písané výlučne len v textovom editore MS WORD, verzia 6 a vyššia, písmo Times New Roman CE 12, riadkovanie jednoduché (1), formát strany A4, všetky okraje 2,5 cm, strany nečíslovať. Tabuľky a grafy v čierno-bielom prevedení zaradiť priamo do textu článku a označiť podľa šablóny. Bibliografické odkazy uvádzať v súlade s normou STN ISO 690 a v súlade s medzinárodnými štandardami. Citácie s poradovým číslom z bibliografického zoznamu uvádzať priamo v texte.

Rozsah: Maximálny rozsah príspevku je 6 strán.

Príspevky sú recenzované. Redakčná rada zabezpečí posúdenie príspevku oponentom.

Príspevky nie sú honorované, poplatok za uverejnenie akceptovaného príspevku je minimálne 30 € Za každú stranu navyše je poplatok 5 €

Štruktúra príspevku: (Pri písaní príspevku využite elektronicкую šablónu: <http://www.ssds.sk/> v časti Vedecký časopis, Pokyny pre autorov.). **Časti v angličtine sú povinné!**

Názov príspevku v slovenskom jazyku (štýl Názov: Time New Roman 14, Bold, centrovať)

Názov príspevku v anglickom jazyku (štýl Názov: Time New Roman 14, Bold, centrovať)

Vynechať riadok

Meno1 Priezvisko1, Meno2 Priezvisko2 (štýl normálny: Time New Roman 12, centrovať)

Vynechať riadok

Abstrakt: Text abstraktu v slovenskom jazyku, max. 10 riadkov (štýl normálny: Time New Roman 12).

Abstract: Text abstraktu v anglickom jazyku, max. 10 riadkov (štýl normálny: Time New Roman 12).

Kľúčové slová: Kľúčové slová v slovenskom jazyku, max. 2 riadky (štýl normálny: Time New Roman 12).

Key words: Kľúčové slová v anglickom jazyku, max. 2 riadky (štýl normálny: Time New Roman 12).

JEL classification: Uviesť kódy klasifikácie podľa pokynov v:

http://www.aeaweb.org/journal/jel_class_system.php

Vynechať riadok a nastaviť si medzery odseku pre nadpisy takto: medzera pred 12 pt a po 3 pt. Nasleduje vlastný text príspevku v členení:

1. **Úvod** (štýl Nadpis 1: Time New Roman 12, bold, zarovnať vľavo, číslovať,)
2. **Názov časti 1** (štýl Nadpis 1: Time New Roman 12, bold, zarovnať vľavo, číslovať)
3. **Názov časti 1. . .**
4. **Záver** (štýl Nadpis 1: Time New Roman 12, bold, zarovnať vľavo, číslovať)

Vlastný text jednotlivých častí je písaný štýlom Normal: písmo Time New Roman 12, prvý riadok odseku je odsadený vždy na 1 cm, odsek je zarovnaný s pevným okrajom. Riadky medzi časťami a odsekmi nevynechávajú. Nastavte si medzi odsekmi medzeru pred 0 pt a po 3 pt.

5. **Literatúra** (štýl Nadpis 1: Time New Roman 12, bold, zarovnať vľavo, číslovať)

[1] Písať podľa normy STN ISO 690

[2] GRANGER, C.W. – NEWBOLD, P. 1974. Spurious Regression in Econometrics. In: Journal of Econometrics, č. 2, 1974, s. 111 – 120.

Adresa autora (-ov): Uved'te svoju pracovnú adresu!!! (štýl Nadpis 1: Time New Roman 12, bold, zarovnať vľavo, adresy vpísať do tabuľky bez orámovania s potrebným počtom stĺpcov a s 1 riadkom):

Meno1 Priezvisko1, tituly1 (študenti ročník)

Pracovisko1 (študenti škola1)

Ulica1, 970 00 Mesto1

meno1.priezvisko1@mail.sk

Meno2 Priezvisko2, tituly2 (študenti ročník)

Pracovisko2 (študenti škola2)

Ulica2, 970 00 Mesto2

meno2.priezvisko2@mail.sk

FORUM STATISTICUM SLOVACUM

vedecký recenzovaný časopis Slovenskej štatistickej a demografickej spoločnosti

Vydavateľ:

Slovenská štatistická a demografická
spoločnosť
Miletičova 3
824 67 Bratislava 24
Slovenská republika

Redakcia:

Miletičova 3
824 67 Bratislava 24
Slovenská republika

Fax: 02/39004009

e-mail:

chajdiak@statis.biz
jan.luha@fmed.uniba.sk

Dátum vydania: december 2014

Registráciu vykonalo:

Ministerstvo kultúry Slovenskej republiky

Dátum registrácie: 22. 7. 2005

Evidenčné číslo: EV 3287/09

Tematická skupina: B1

Periodicita vydávania:

minimálne 2 krát ročne

Objednávky:

Slovenská štatistická a demografická
spoločnosť
Miletičova 3, 824 67 Bratislava 24
Slovenská republika

IČO: 178764

DÍČ: 2021504276

Číslo účtu:

IBAN: SK70 0900 0000 0000 1146 9672

ISSN 1336-7420

Redakčná rada:

RNDr. Peter Mach – *predseda*

Doc. Ing. Jozef Chajdiak, CSc. – *šéfredaktor*

RNDr. Ján Luha, CSc. – *vedecký tajomník*

členovia:

Prof. RNDr. Jaromír Antoch, CSc.

Ing. František Bernadič

Doc. RNDr. Branislav Bleha, PhD.

Ing. Mikuláš Cár, CSc.

Ing. Ján Cuper

Prof. RNDr. Gejza Dohnal, CSc.

Ing. Anna Janusová

Doc. RNDr. PaedDr. Stanislav Katina, PhD.

Prof. RNDr. Jozef Komorník, DrSc.

RNDr. Samuel Koróny, PhD.

Doc. Dr. Jana Kubanová, CSc.

Doc. RNDr. Bohdan Linda, CSc.

Prof. RNDr. Jozef Mládek, DrSc.

Doc. RNDr. Oľga Nánásiová, CSc.

Doc. RNDr. Karol Pastor, CSc.

Mgr. Michaela Potančoková, PhD.

Prof. RNDr. Rastislav Potocký, CSc.

Doc. RNDr. Viliam Páleník, PhD.

Ing. Marek Radvanský, PhD.

Prof. Ing. Hana Řezanková, CSc.

Doc. Ing. Iveta Stankovičová, PhD.

Prof. RNDr. Beata Stehlíková, CSc.

Prof. RNDr. Anna Tirpáková, CSc.

Prof. RNDr. Michal Tkáč, CSc.

Doc. Ing. Vladimír Úradníček, PhD.

Ing. Boris Vaňo

Doc. Ing. Mária Vojtková, PhD.

Prof. RNDr. Gejza Wimmer, DrSc.

Ročník: X.

Číslo: 6/2014

Cena výtlačku: 30 EUR

Ročné predplatné: 120 EUR