Výpočtová Štatistika



Computational Statistics

2014

Prehliadka prác mladých štatistikov a demografov

(Zborník)

SLOVENSKÁ ŠTATISTICKÁ A DEMOGRAFICKÁ SPOLOČNOSŤ

Recenzenti
Michal ILLOVSKÝ (Katedra matematiky a deskriptívnej geometrie Stavebná fakulta STU)
Lukáš PASTOREK (Fakulta informatiky a statistiky, Vysoká škola ekonomická v Praze)
Tomáš ŽELINSKÝ (Ekonomická fakulta, Technická univerzita v Košiciach)

Vydala Slovenská štatistická a demografická spoločnosť
Bratislava, 2014

ISBN 978-80-88946-66-3



K organizovaniu seminára a vydaniu zborníka významne prispeli hlavní sponzori podujatia spoločnosti ***T-Systems Slovakia, s. r. o.*** *a* ***SAS Slovakia, s. r. o.***

Obsah

PREDSLOV 3

INFORMÁCIE O SEMINÁRI 4

Paula BANASOVÁ
Analýza fertility metódou indexu premenlivého zloženia
Analyze of fertility using Index variable composition method 5

Jakub DANKO
Hľadanie štruktúry kapitálových trhov
Search the Structure of Capital Markets 10

Lukáš KALINA
Analýza singulárneho spektra aplikovaná na časový rad prichádzajúcich hovorov
Singular spectrum analysis aplied on call center arrivals time series 18

Martin KAMENÍK
Simulácia evolúcie v opakovanej dileme samaritána
The Simulation of Evolution in the Iterated Samaritan's Dilemma 27

Michal MEŠŤAN
Komparácia aktívneho a pasívneho prístupu k úsporám v dôchodkových schémach
Comparison of active and passive saving approach in pension schemes 37

Lukáš RIES
Vzájomná súvislosť a predikcia vývoja miery inflácie a nezamestnanosti v podmienkach Slovenskej republiky
Mutual relation and prediction of the inflation rate and the unemployment rate in the conditions of Slovak Republic 45

Zuzana ROŠŤÁKOVÁ
Rozšírenie pravdepodobnostného spánkového modelu
Innovation of probabilistic sleep model 55

Ľubica ŠTIBLÁROVÁ
Finančná integrácia vs. vonkajšia nerovnováha? Aplikácia medzinárodnej investičnej pozície
Financial integration vs. external imbalances? Application of the international investment position 65

Martin ZORIČÁK
Modelovanie daňovej konkurencie/kooperácie v systéme dvoch krajín
Tax competition/cooperation modelling in two countries system 75

Predslov

Vážení kolegovia a kolegyne, milí študenti, drahí priatelia

Minulý rok sa niesol v duchu osláv Medzinárodného roka štatistiky, ktorý mal za cieľ zvýšiť povedomie o štatistike a zviditeľniť štatistiku ako nástroj na riešenie pálčivých problémov dnešného sveta. Do tejto osvety sa v priebehu minulého roka zapojilo viac ako 2300 organizácií zo 128 krajín sveta.

Nezostalo iba pri popularizácii. V Londýne sa stretlo sto najprominentnejších zástupcov štatistického sveta, aby sformulovalo víziu budúceho vývoja štatistickej disciplíny (Statistics and Science: A Report of the London Workshop on the Future of the Statistical Sciences). Okrem vyzdvihnutia prínosu znovuzrodenej Bayesovskej štatistiky, Markových reťazcov, nových prístupov v klinickom a experimentálnom biologickom výskume, sa vášnivo diskutovalo o úlohe štatistiky a štatistikov pri analýze tzv. „Big Data“.

Malé štatistické súbory s jasne definovanými problémami sa v priebehu 20. storočia zmenili na obrovské dátové súbory s enormným „n“, veľkým „p“ a komplexnými otázkami. Domnelá priepasť medzi štatistikmi a počítačovými vedcami sa nezadržateľne zmenšuje. Zvládnutie výpočtovej náročnosti už nie je exkluzívny problém IT špecialistov.

Na tento trend nadväzuje i iná téma, ktorá zarezonovala na seminári v Londýne. Má štatistik budúcnosť? Nebude nutné preformulovať jeho schopnosti v prospech čoraz častejšie sa vyskytujúcej pozície dátového vedca?

Našou snahou je ukázať, že napriek meniacemu sa svetu, názvosloviu alebo marketingu, štatistické metódy zostávajú nevyhnutným nástrojom, ktorý nemá náhradníka. Som veľmi rád, že sa i tento rok našla skupina mladých a nadaných študentov, ktorí sa rozhodli preukázať, že aktívne využívajú štatistiku ako jednu zo svojich kľúčových schopností. Nakoniec to budú absolventi, ktorí ukážu odpoveď na otázku, s akou zo svojich schopností sa najviac identifikujú.

Lukáš Pastorek
za organizačný výbor

SLOVENSKÁ ŠTATISTICKÁ A DEMOGRAFICKÁ SPOLOČNOSŤ

KLUB DISPERSUS
STEP OUT OF RANGE

Informácie o seminári

Výpočtová štatistika 2014 – Computational Statistics 2014

Prehliadka prác mladých štatistikov a demografov

Čas konania: 4. – 5. decembra 2014

Miesto konania: Prezentačné centrum AMOS, pavilón B1, Prírodovedecká fakulta UK, Mlynská dolina, 842 15 Bratislava 4.

Slovenská štatistická a demografická spoločnosť (SŠDS) a *Klub Dispersus* zorganizovali celoslovenské stretnutie študentov so záujmom o aplikáciu matematicko-štatistických výskumných metód v rámci 23. medzinárodného seminára *Výpočtová štatistika 2014* a jeho programovej časti *Prehliadka prác mladých štatistikov a demografov*.

Program pre študentov:

4. 12. 2014 (štvrtok): Fórum analytických prác študentov

16:00 – 18:15 Prehliadka prác mladých štatistikov a demografov

18:15 – 22:00 Panelová diskusia, stretnutie členov klubu

5. 12. 2014 (piatok): Analytika očami profesionálov

09:00 – 09:40 Behaviorálna segmentácia klientov banky a jej použitie
Peter Kertys, *VÚB banka*

09:45 – 10:25 Štatistické metódy v oblasti trhu práce
Michal Páleník, *Inštitút zamestnanosti*

10:30 – 11:10 Budúcnosť rodín a „nerodín“ na Slovensku
Branislav Bleha, *Prírodovedecká fakulta UK*

11:15 – 11:55 Vizualizácia štatistických dát v prostredí Internetu – nástroje a formáty
Martin Řezáč, *SAS Slovakia*

12:00 Vyhlásenie výsledkov študentskej súťaže, ocenenie víťazov
Lukáš Pastorek, *Klub Dispersus*

Analýza fertility metódou indexu premenlivého zloženia
Analyze of fertility using Index variable composition method

Paula Banasová

Abstract: This article analyzes the fertility Slovak women using the method of the index variable composition, comparing the years 1996 and 2011. Z výsledkov vyplýva, že vyšší podiel na plodnosti majú ženy žijúce v mestách.

Abstrakt: Príspevok analyzuje fertility slovenských žien pomocou metódy indexu premenlivého zloženia, porovnáva roky 1996 a 2011. The results indicate that a higher proportion of fertility had women living in cities.

Key words: fertility, individual index, index variable composition, urban and rural population

Kľúčové slová: fertilita, individuálne indexy, index premenlivého zloženia, mestská a vidiecka populácia

1. Úvod

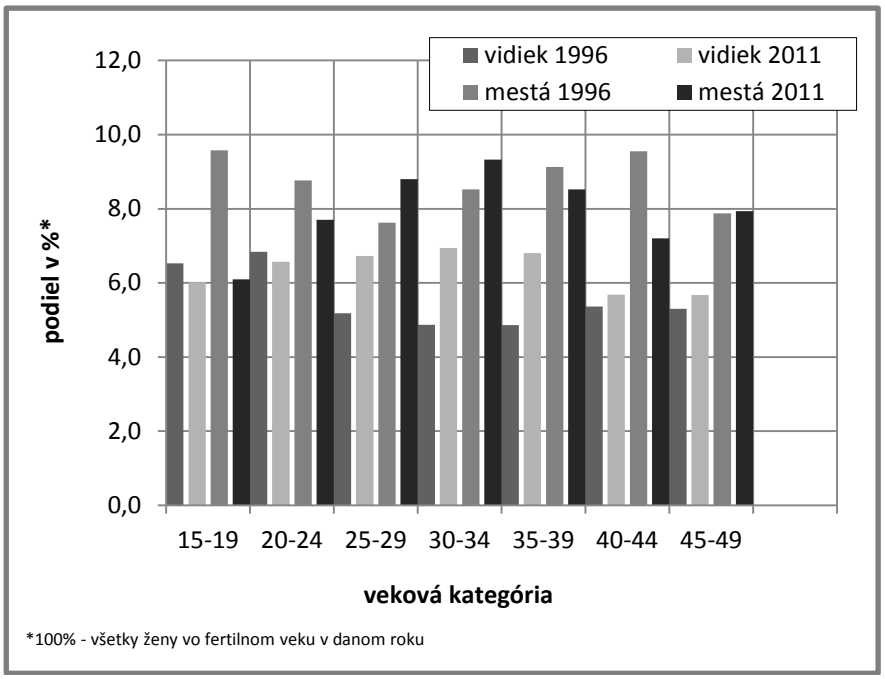
Príspevok sa venuje zmenám v plodnosti žien vidieka a miest Slovenska v horizonte rokov 1996 – 2011. Fertilita bola analyzovaná metódou rozkladu zložených individuálnych indexov, konkrétne indexu premenlivého zloženia. Môžeme ho vyjadriť ako súčin dvoch indexov, indexu štruktúry a indexu stáleho zloženia, pomocou ktorých môžeme vyjadriť, do akej miery je všeobecná miera plodnosti ovplyvnená zmenou vekovej štruktúry žien a do akej miery zmenou špecifických mier plodnosti žien podľa veku v sledovanom období. Týmto spôsobom boli rozložené individuálne indexy populácie žien miest na jednej strane a populácie žien vidieckych sídel na strane druhej. Vyčlenenie mestskej a vidieckej subpopulácie žien bolo zvolené na základe kritéria štatútu mesta – t. j. za mestá boli považované tie sídla, ktoré disponovali štatútom mesta v roku 2011. Vo vyhodnocovaní výsledkov boli subpopulácie žien komparované. Zdrojové dáta pochádzajú z Pramenných diel 1996 – 2011 ŠÚ SR, tiež z on-line dostupných databáz Infostatu.

2. Urbánny a rurálny priestor

Článok sa venuje komparácii intenzity reprodukcie mestských a vidieckych žien na Slovensku. Rozdelenie populácie je opodstatnené, keďže 54% obyvateľov žije v mestách, 46% na vidieku. Z obr.1 vyplýva, že rozdiely v zastúpení žien vo vekových kategóriách vo fertilnom veku badáme pri roku 1996 rovnako ako pri roku 2011. Podiel žien v reprodukčnom veku v mestách je vyšší ako na vidieku, teda sa ženy v mestách výraznejšie môžu podieľať na celkovom počte narodených detí. V procese rozlišovania urbánneho a rurálneho priestoru ide primárne o problém vyhraničenia, regionalizácie a definície vidieka a mesta, ktoré zahŕňa širokú škálu priestorových a nepriestorových aspektov. [3] Všeobecne uznávanou definíciou je podľa Bednářovej [2] jedine medzinárodná definícia OECD, ktorá vymedzuje vidiek na základe hustoty zaľudnenia, a to 150 ob./km^2, resp. 100 ob./km^2 a menej. Slovensko má priemernú hustotu zaľudnenia 110 ob./km^2, hlavné mesto 1124 ob./ km^2. Túto klasifikáciu v upravenom spôsobe používa aj Eurostat.

V príspevku bola použitá klasifikácia definovaná Zákonom SNR č. 369/1990 Zb. o obecnom zriadení a definícia mestského sídla podľa zákona NR SR č. 453/2001 Z. z. § 22. Na

základe tohto kritéria má Slovenská republika 138 štatutárnych miest vyhlásených NR SR a 2 753 obcí bez štatútu mesta, teda vidieckych obcí[1].



Obr.1: Podiel žien vo fertilnom veku podľa bydliska a vekovej štruktúry v rokoch 1996 a 2011. Zdroj: ŠÚ SR, vlastné výpočty [5]

3. Index premenlivého zloženia

Indexy sú pomernými číslami, slúžia na porovnávanie ukazovateľov. V našom ponímaní ich budeme chápať ako tie, ktoré vyjadrujú zmenu ukazovateľov v čase. Index premenlivého zloženia patrí k skupine individuálnych (rovnorodých) zložených indexov, ktoré sa používajú pri konštrukcii indexov homogénnych (sčítateľných) veličín. Pre tento príspevok bola zvolená metóda porovnávajúca zmeny vzhľadom na bázický rok, teda pôjde o tzv. bázický index premenlivého zloženia. Do súčinu v tomto indexe vstupujú dva čiastkové indexy, a to index stáleho zloženia, ktorý vyjadruje vplyv zmeny extenzitnej zložky (zmeny počtu žien v jednotlivých vekových kategóriách) a druhý, index štruktúry, vyjadrujúci vplyv zmeny intenzitnej zložky (zmeny intenzity fertilít jednotlivých vekových kategórií žien).

Index stáleho zloženia dostaneme štandardizáciou váh, keď „obidva priemery (priemernú intenzitnú veličinu v obidvoch obdobiach) stanovíme pri použití len jednej nezmenenej štruktúry extenzitných veličín. [1] Diferenciáciu vplyvu intenzitného indexu zistíme tak, že stanovíme priemernú intenzitnú veličinu v oboch obdobiach na základe rovnakých údajov. [1]. Skonštruovaný index premenlivého zloženia je potom nasledovný:

$$I^p = \frac{p^1}{p_0} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} \qquad (1)$$

V konštrukcii indexu rozlišujeme nasledujúce premenné: q0 – stredný stav žien v tisícoch v danej 5-ročnej vekovej kategórii pre počiatočný sledovaný rok, q1 - stredný stav žien v tisícoch v danej 5-ročnej vekovej kategórii pre záverečný sledovaný rok, p0 - špecifické miery

[1] Údaj platný k okamihu sčítania k 21. máju 2011.

fertility pre 5-ročné vekové kategórie za počiatočný sledovaný rok, p1 - špecifické miery fertility pre 5-ročné vekové kategórie za záverečný sledovaný rok.

Index premenlivého zloženia (IPZ) je rozložený dvoma spôsobmi, aby boli porovnané zmeny a vplyvy jednotlivých komponentov vstupujúcich do konštrukcie. V prvom spôsobe považujeme pri indexe štruktúry (ISTR) vekovú štruktúru z počiatočného roku (1996) za stabilnú a menia sa individuálne fertility pre počiatočný a záverečný rok. Pri indexe stáleho zloženia (ISZ) sa štruktúra žien podľa veku mení, špecifické plodnosti ostávajú statické, a to z posledného sledovaného roku (2011). Prvý spôsob rozkladu indexu premenlivého zloženia:

$$IPZ = ISZ \;.\; ISTR$$

$$\frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_1}{\sum q_1}} \cdot \frac{\frac{\sum p_0 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} \qquad (2)$$

Prepólovaním stabilných hodnôt z predchádzajúceho spôsobu vznikne druhý rozklad, kde v indexe štruktúry zostáva veková štruktúra z koncového roku stabilnou (z roku 2011), meníme však individuálne hodnoty fertilít. V indexe stáleho zloženia zostane hodnota špecifickej plodnosti z počiatočného roku nemenná, veková štruktúra sa obmieňa z oboch porovnávaných rokov. Druhý spôsob rozkladu indexu premenlivého zloženia:

$$IPZ = ISZ \;.\; ISTR$$

$$\frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} = \frac{\frac{\sum p_1 q_0}{\sum q_0}}{\frac{\sum p_0 q_0}{\sum q_0}} \cdot \frac{\frac{\sum p_1 q_1}{\sum q_1}}{\frac{\sum p_0 q_0}{\sum q_0}} \qquad (3)$$

Z konštrukcie indexu premenlivého zloženia vyplýva, že odráža zmenu dvoch priemerných úrovní, na ktorej sa podieľali jednak zmeny intenzitných veličín a jednak zmeny ich nositeľov. [1] Metóda bola aplikovaná na údaje z rokov 1996 a 2011 pre Slovensko, mestskú a vidiecku populáciu žien, diferencovanú na základe kritéria o štatúte miest podľa zákona platného k roku 2011.

4. Analýza vplyvu vekovej štruktúry a individuálnych fertilít na plodnosť slovenských žien v mestách a na vidieku

Vplyv vekovej štruktúry a samotnej fertility na celkovú plodnosť na Slovenku za obdobie 1996 - 2011 sumarizuje tab.1. Z výsledkov vyplýva, že index premenlivého zloženia je u miest vyšší ako na vidieku. Znamená to, že ženy v mestách majú v roku 2011 vzhľadom na rok 1996 vyšší vplyv na reprodukciu ako ženy na vidieku. Priemerná fertilita v meste sa zvýšila 1,161-krát, tj. o 16%. Na vidieku došlo k opačnému procesu, počas obdobia 1996 – 2011 sa znížila 0,905-krát, čiže takmer o 10%. V oboch prípadoch, zohrala na zmene celkovej fertility slovenských žien primárnu úlohu zmena vekovej štruktúry.

Pri analýze jednotlivých čiastkových indexov, badáme rozdiely v možnom reprodukčnom správaní sa žien. V prípade, že by sme brali do úvahy len vplyv vekovej štruktúry žien z roku 1996, tak by priemerná fertilita za sledované obdobie (smerom k roku 2011) stúpla v urbánnom priestore o 6,4% (ISTR1[2]=1,064) a v rurálnom o 9,56% (ISTR1=1,956). Ak by

[2] ISTR1, ISTR2 – index štruktúry vypočítaný prvým, resp. druhým spôsobom rozkladu indexu premenlivého zloženia

sme ponechali vekovú štruktúru z roku 2011, stúpla by u mestských žien plodnosť až o 11,6% (ISTR2=1,116). V subpopulácii vidieckych žien naopak, len o 4,8% (ISTR2=1,048).

Na základe porovnania výsledkov len pri zmene vekového zloženia žien vidíme, že ak by ostalo vekové zloženie žien z roku 1996, tak by ženy z rurálneho prostredia mali na plodnosť väčší vplyv ako mestské, paradoxne, z dlhodobého hľadiska by takýto vývoj priniesol Slovensku nižšiu celkovú plodnosť. Z pohľadu na stav indexu štruktúry pre Slovensko môžeme na základe tab.1 konštatovať, že celková fertilita by pri podmienkach vekového zloženia z roku 2011 bola vyššia, na úrovni 8,7% ako keby sme ponechali vekovú štruktúru z roku 1996, keďže tak by bol nárast len o 2,7%.

Ak ponecháme plodnosť na úrovni aká bola v roku 1996, tak by plodnosť žien v mestách stúpli o 1,041-krát a u vidieckych žien o 0,864-krát. Rovnaký záver ukazuje aj stabilizovanie fertility na hladine z roku 2011, keďže potvrdilo, že ženy z urbánneho priestoru ovplyvňujú vývoj plodnosti o 1,091-krát, teda významnejšie ako ženy z vidieka ($ISZ1^3$=0,463).

Tab.1: Indexy premenlivého zloženia fertilít na Slovensku, na vidieku a v mestách v období 1996 – 2011. Zdroj: ŠÚ SR, Infostat, vlastné výpočty [5]

	SR		***Mesto***		***Vidiek***	
Spôsob rozkladu indexu premenlivého zloženia	prvý	druhý	prvý	druhý	prvý	druhý
Index štruktúry	1,027	1,087	1,064	1,116	1,956	1,048
Index stáleho zloženia	1,019	0,963	1,091	1,041	0,463	0,864
Index premenlivého zloženia	1,046	1,046	1,161	1,161	0,905	0,905

V mestskej a vidieckej populácii sa počas celého sledovaného obdobia menia role tzv. nositeliek plodnosti. Kým na začiatku sledovaného obdobia to vďaka vekovej štruktúre a intenzite fertility boli ženy žijúce na vidieku, tak na konci skúmaného obdobia sú to ženy z mestského prostredia. Potvrdzuje aj Šprocha [4] a to tak, že rast intenzity plodnosti je v súčasnosti väčší v mestách ako na vidieku.

Významnými faktormi ovplyvňujúcimi tento vývoj sú okrem prevahy počtu žien z miest aj celkové zmeny spoločenskej klímy, globalizácia, životného štýlu, individualizmu, nárastu propagácie rodovej rovnosti a pod., čo demografia označuje ako druhý demografický prechod. Komplex faktorov a aspektov, ktoré charakterizujú súčasné smerovanie populácie, vytvára nový model rodinného a reprodukčného správania.

5. Záver

Na základe podrobnej analýzy pomocou metódy indexu premenlivého zloženia môžeme zhodnotiť, že intenzita fertility je pod väčším vplyvom subpopulácie žien z miest ako tých z vidieka. Článok ukazuje na zvyšujúci sa potenciál vplyvu na celkovú plodnosť u žien z urbánneho prostredia počas celého sledovaného obdobia. Role nositeliek plodnosti sa za posledných 15 rokov vymenili v dôsledku zmien vekovej štruktúry významnejšie, ako z dôsledku zmien individuálnych fertilít.

[3] ISZ1, ISZ2 – index stáleho zloženia vypočítaný prvým, resp. druhým spôsobom rozkladu indexu premenlivého zloženia

Literatúra

[1] BAKYTOVÁ, H., URGON, M., KONTŠEKOVÁ, O. (1975). Základy štatistiky (druhé vydanie. vyd.). Bratislava: ALFA, 1975. s. 392.

[2] BEDNÁŘOVÁ, H. 2011.: Vybrané problémy vymezení venkova. In: *Forum statisticum Slovacum.* 2011, č.6. s. 14-20. ISSN 1336-7420.

[3] KUSENDOVÁ, D. (2011). Problematika vyhraničenia urbánneho a vidieckeho priestoru – geografický výskum a prax. In: *Forum statisticum Slovakum*. 2011, č. 6. s. 97-104. ISSN 1336-7420

[4] ŠPROCHA, B. 2008.: Reprodukčné správanie mestského a vidieckeho obyvateľstva Slovenska. Bratislava. Výskumné demografické centrum, Infostat. 2008. s.72 [3.11.2014] ISBN 978-8089398-01-0. Dostupné na: http://www.infostat.sk/vdc/pdf/Mesto_vidiek.pdf.

[5] ŠTATISTICKÝ ÚRAD SR, 2011.: Pramenné diela – 1996 – 2011. [10.11.2014] Dostupné na: http://portal.statistics.sk/showdoc.do?docid=6674

Adresa autora:

Paula Banasová, Bc.
Prírodovedecká fakulta
Univerzita Komenského v Bratislave
2. ročník, 2. stupeň
paula.banasova@gmail.com

Hľadanie štruktúry kapitálových trhov
Search the Structure of Capital Markets

Jakub Danko

Abstract: The article presents the method of identification of capital market structure in USA, which is based on mutual relationships between individual sub-indices Dow Jones Industrial Average (DJIA). Structure here is identified by R. Mantengna's (1999) methodology that is known as the methodology of Minimum Spanning Tree. It identifies the dominant correlation relationships between individual DJIA sub-indices in the presented research. Minimum Spanning Trees are assessed on the 20-day rolling windows return of sector sub-indices during the period of March 2001 - September 2014. These were then grouped, using the method of clustering, into groups which include similar graphical representations of capital market.

Abstrakt: V článku je prezentovaný spôsob identifikácie štruktúry kapitálového trhu v USA, ktorý je založený na vzájomných vzťahoch medzi jednotlivými sektorovými podindexmi Dow Jones Industrial Average (DJIA). Štruktúra je tu identifikovaná metodológiou R. Mantegnu (1999) známej ako metodológia minimálnej kostry grafu (Minimum Spanning Tree). Tá v prezentovanom výskume identifikuje dominantné korelačné vzťahy medzi jednotlivými podindexmi DJIA. Minimálne kostry grafu sú tu odhadnuté na 20 dňových neprekrývajúcich oknách výnosností sektorových podindexov z obdobia marec 2001 - september 2014. Tie sme neskôr s pomocou zhlukovej analýzy zoskupili do skupín obsahujúcich podobné grafové reprezentácie kapitálového trhu.

Key words: Minimum Spanning Tree, Graph Theory, Clustering, Dow Jones Industrial Average.

Kľúčové slová: Minimálna kostra grafu, Teória grafov, Zhluková analýza, Dow Jones Index.

1. Úvod

Cieľom tohto článku je pomocou rôznych štatistických metód a metód diskrétnej matematiky nájsť a identifikovať špecifické režimy kapitálového trhu. Režim trhu chápeme ako určitý stav na trhu a jedným z jeho prejavov môžu byť napríklad stabilita alebo nestabilita, teda odchýlenie od nejakého dlhodobého trendu. Našim cieľom je najmä identifikovať tento režim a následne predikovať vývoj trhu v budúcnosti.

Základy metodológie minimálnej kostry grafu s cieľom pochopenia štruktúry kapitálových trhov položil Rosario Mantegna (1999). [2] Onnela a Chakraborti (2003) tiež vychádzajú z výpočtu minimálnej kostry grafu, avšak na rozdiel od predchádzajúceho prístupu má funkcia mapujúca korelácie medzi vrcholmi na vzdialenosti nelineárny priebeh. Túto novú sieťovú štruktúru pomenovali dynamický graf akcií. [7] Bonanno a Caldarelli (2004) sa zaoberali tým, ako vplýva na výnosnosť finančných inštrumentov časový horizont a zistili, že s klesajúcim časovým horizontom sa štruktúra grafu reprezentujúceho trh postupne mení z komplexnej na jednoduchú. Zaoberali sa tiež volatilitou časových radov cien akcií a finančnými trhmi z globálneho pohľadu. [8] Následne začali autori využívať pri analýze pokročilejšie metódy teórie grafov, napríklad reprezentáciu trhu planárnymi grafmi (grafy, ktorých hrany sa nepretínajú). Podobná metodológia sa začala využívať aj pri analyzovaní trhu s menovými pármi, teda pri analýze výmenných kurzov – Mizuno, Takayasu (2005) [9], Naylor, Rose, Moyle (2007). [10] Medzi najnovšie analýzy patria práce autorov Kenett, Tumminello, Mantegna (2010) [11], kde sa autori zaoberajú takzvanými parciálnymi koreláciami. Tieto parciálne korelácie následne využívajú ako nástroj na nájdenie

najrelevantnejších hrán v grafe. Taktiež vychádzajú z planárnych grafov a v práci popisujú aj nový typ závislej siete, u ktorej zahrnutie hrán, respektíve prepojení závisí od stredných hodnôt a štandardných odchýlok jednotlivých uzlov (vrcholov) tejto siete. Vyššie spomenutí autori pri svojich analýzach prevažne využívajú údaje z New York Stock Exchange (NYSE), jedna z analýz vychádzala aj z Tel Aviv Stock Exchange (TASE). [1]

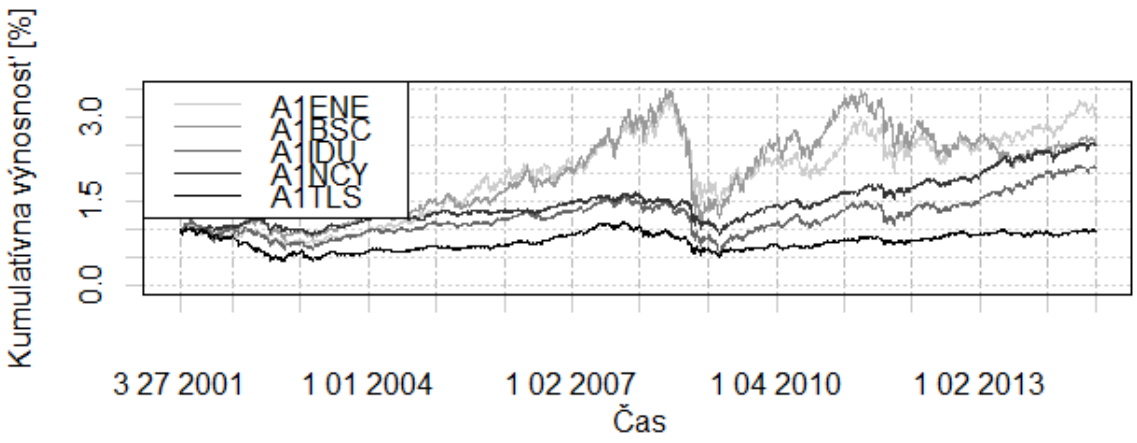
2. Vstupné údaje a úvodné analýzy

Americký burzový index DJIA obsahuje sektorové podindexy (pozri Tab. 1), ktoré sme využili v našej analýze.

Tab. 1: Podindexy ako vybrané sektory americkej ekonomiky, vlastné spracovanie podľa: http://www.djindexes.com/globalfamily/

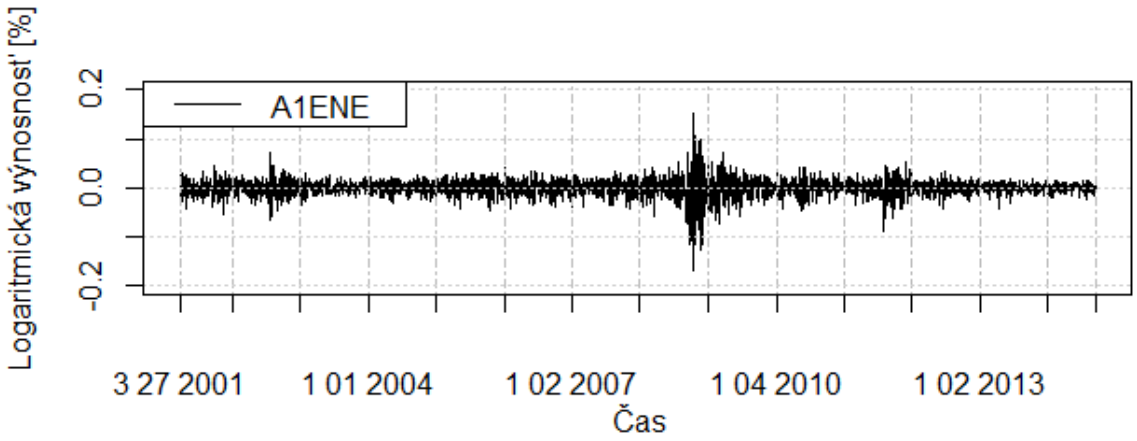
Skratka (označenie) podindexu	Celý názov podindexu
A1ENE	DOW JONES AMERICAS ENERGY STOCK INDEX
A1BSC	DOW JONES AMERICAS BASIC MATERIALS STOCK INDEX
A1IDU	DOW JONES AMERICAS INDUSTRIAL STOCK INDEX
A1NCY	DOW JONES AMERICAS CONSUMER NON-CYCLICAL STOCK INDEX
A1TLS	DOW JONES AMERICAS TELECOMMUNICATIONS INDEX
A1UTI	DOW JONES AMERICAS UTILITIES STOCK INDEX
A1FIN	DOW JONES AMERICAS FINANCIAL STOCK INDEX
A1TEC	DOW JONES AMERICAS TECHNOLOGY STOCK INDEX
A1HCR	DOW JONES AMERICAS HEALTHCARE STOCK INDEX

Časové rady uzatváracích cien týchto deväť podindexov sme analyzovali v období od 27. 3. 2001 do 15. 9. 2014. Počas tohto obdobia sme získali takmer *3 500* údajov o uzatváracích cenách. Pre každý z podindexov sme vypočítali logaritmické denné výnosnosti. Pre ilustráciu kumulatívny tvar prvých piatich podindexov uvádzame na Obr. 1.



Obr. 1: Kumulatívne výnosnosti vybraných podindexov (vlastné spracovanie v programe R)

Pre názornosť uvádzame aj logaritmickú výnosnosť podindexu A1ENE na Obr. 2. Z obrázku je zrejmé, že veľký pokles na Obr. 1 v roku 2008 sa prejavil v danom roku na Obr. 2 vysokou volatilitou výnosnosti tohto burzového podindexu.



Obr. 2: Logaritmická výnosnosť podindexu A1ENE (vlastné spracovanie v programe R)

Na databáze denných výnosností sme vytvárali neprekrývajúce časové okná dĺžky *20*, kde sme sledovali vzájomné krížové korelácie. Vzhľadom k tomu, že sme mali necelých *3 500* denných pozorovaní, vzniklo *174* neprekrývajúcich sa časových okien.

3. Prístup na základe teórie grafov

Ak by sme predpokladali, že naše imaginárne portfólio tvorí práve týchto deväť podindexov, korelačná matica by predstavovala klasický pohľad na dané portfólio. Tento pohľad je relatívne jednoduchý, nedokáže však zachytiť zložitejšie väzby ani štruktúru daného portfólia. Veľmi vhodným reprezentantom vzťahu jednotlivých finančných nástrojov tvoriacich portfólio sa javí graf, ako nástroj diskrétnej matematiky. Graf je definovaný množinou vrcholov, ktoré v našom prípade predstavujú jednotlivé podindexy a množinou hrán, ktoré predstavujú nejaký vzťah medzi vrcholmi. My by sme chceli, aby hrany predstavovali vzdialenosť medzi jednotlivými podindexmi. Ak chceme uvažovať o vzdialenosti, musíme definovať konkrétnu metriku. Tu preberáme prístup viacerých autorov [2], [3], [4], ktorí využívajú vzťah:

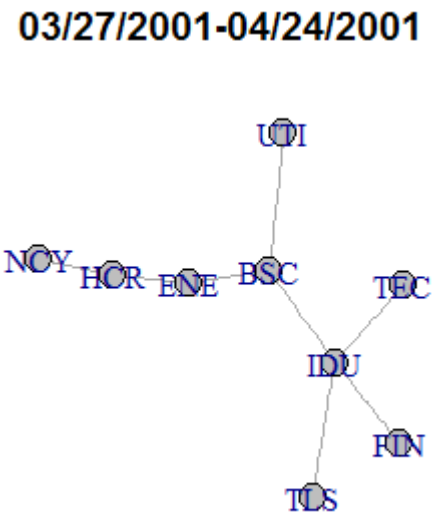
$$d(i,j) = \sqrt{2.(1-\rho_{i,j})} \tag{1}$$

kde $\rho_{i,j}$ je Pearsonov korelačný koeficient *i*-teho a *j*-teho podindexu v rámci vyšetrovaného okna

Z korelačných matíc sme týmto spôsobom vytvorili matice vzdialeností. Získali sme teda *174* matíc vzdialeností. Tieto matice sú symetrické, rozmeru *9x9* (*9* podindexov), na hlavnej diagonále majú nuly a hodnota v *i*-tom riadku a *j*-tom stĺpci predstavuje vzdialenosť *i*-teho a *j*-teho podindexu. Takto definované matice už môžu byť reprezentované kompletnými grafmi s deviatimi vrcholmi a *36* hranami (každý z deviatich vrcholov je spojený s každým, teda máme $\binom{9}{2} = 36$ hrán). Získali sme takto *174* kompletných grafov, ktoré ilustrujú vzdialenosti medzi podindexami za dané časové okno. Celá analýza prebieha v prostredí programu R, ktorý pre prácu s grafmi používa knižnicu igraph.

Ak je graf reprezentovaný maticou vzdialenosti, dá sa v ňom vypočítať minimálna kostra. Vychádzame pri tom z toho, že každý súvislý graf má kostru, pričom táto kostra predstavuje taký podgraf grafu, ktorý obsahuje všetky vrcholy (je faktorom), je súvislý a nemá kružnice (je stromom). Ak graf má nejaké hranové ohodnotenie (v našom prípade dĺžka hrany), potom určite musí v grafe existovať aspoň jedna taká kostra, ktorá má minimálny súčet hranových ohodnotení a takúto kostru budeme nazývať minimálna. Existuje viacero algoritmov na hľadanie minimálnej kostry grafu, reprezentovaného maticou vzdialenosti, napríklad Kruskalov alebo Primov algoritmus. [5] Knižnica igraph pri výpočte minimálnej kostry využíva Primov algoritmus na hranovo vážených grafoch (reprezentovaných maticou vzdialenosti). Algoritmus je nasledovný: v prvom kroku zvolíme ľubovoľný vrchol a do kostry pridáme hranu s minimálnym ohodnotením, ktorá s týmto vrcholom inciduje. Takto nám vznikne strom. Následne v ďalších krokoch k aktuálnemu stromu pridáme hranu, ktorej jeden vrchol leží v strome a druhý leží mimo stromu a to takú, ktorá má minimálne ohodnotenie. Minimálna kostra súvislého grafu s n vrcholmi teda predstavuje graf s n vrcholmi a $n - 1$ hranami. [5]

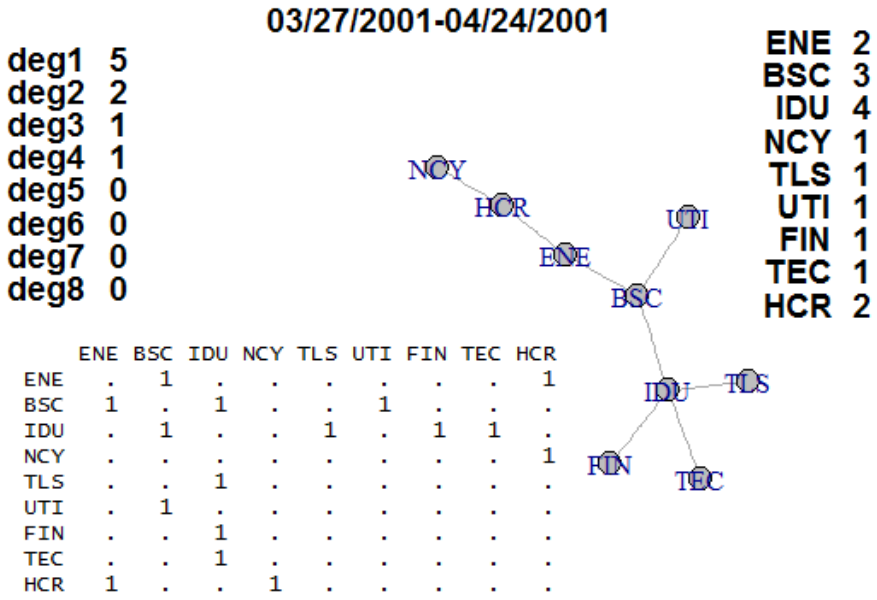
Zo *174* kompletných hranovo ohodnotených grafov sme si vypočítali *174* minimálnych kostier. Príklad minimálnej kostry prvého časového okna je na Obr. 3, pričom ako sme uviedli, ide o graf o *9* vrcholoch (podindexoch) a *8* hranách (minimálnych vzdialenostiach medzi nimi).



Obr. 3: Príklad odhadnutej minimálnej kostry grafu (vlastné spracovanie v programe R)

Tieto kostry sme z dôvodu počítačového spracovania potrebovali nejako reprezentovať. Prvou možnosťou ako reprezentovať grafy je matica susednosti. Túto sme uvažovali dvojako, buď ako hranovo váženú, teda dĺžka hrany predstavuje danú hodnotu v matici susednosti alebo ako neváženú binárnu maticu, ktorá vyjadruje susednosť (dvojica vrcholov sa nachádza v minimálnej kostre) jednotkou a nesusednosť (dvojica vrcholov sa nenachádza v minimálnej kostre) nulou. Takáto reprezentácia je úplne jednoznačná, nemôže sa stať, žeby poukazovala na rozdielne grafy. Pre počet kostier úplného grafu môžeme použiť Cayleho formulu, podľa ktorej pre každé $n \geq 3$ je počet kostier na daných n vrcholoch rovný n^{n-2}. [5] Pri deväťprvkovej množine vrcholov je možné vytvoriť až $9^7 = 4\,782\,969$ minimálnych kostier, čo sa nám zdalo byť pre naše účely pomerne veľa. Preto sme sa v ďalšom postupe snažili zamerať na modelovanie tried ekvivalencie minimálnych kostier, pričom kostry prislúchajúce k jednej triede ekvivalencie by mali mať rovnako veľké početnosti vrcholov prvého stupňa,

druhého stupňa až *n-1*. stupňa. Výhodou týchto tried ekvivalencie je ich jednoduchosť a menšia početnosť v porovnaní s maticou susednosti. Na Obr. 4 môžeme vidieť rôzne reprezentácie minimálnej kostry prvého časového okna:



Obr. 4: Možnosti reprezentácie minimálnej kostry vybranými triedami ekvivalencií a neváženou maticou susednosti (vlastné spracovanie v programe R)

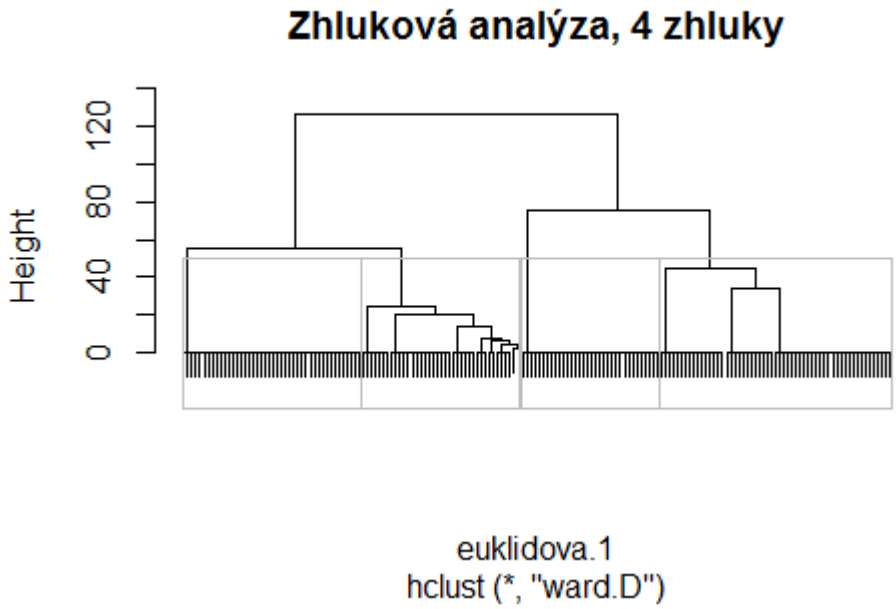
Ako najvhodnejšiu reprezentáciu sme teda považovali reprezentáciu okna *n-1*-prvkovým vektorom, ktorý obsahuje početnosti stupňov vrcholov.

4. Zhluková analýza

Na začiatku zhlukovej analýzy sme mali *174* vektorov dĺžky *8*, každý vektor za jedno časové okno. Cieľom je zhlukovaním identifikovať určitý počet klastrov (zhlukov), pričom predpokladáme, že relatívne podobné minimálne kostry sa budú zhlukovať do spoločných klastrov a budú tak predstavovať nejaký režim trhu (napríklad stabilitu, prudký pokles a podobne). Zhlukovú analýzu sme realizovali taktiež v prostredí programu R a využili sme pritom Wardovu metódu. Ide o metódu minimálneho rozptylu a predstavuje zlučovaciu hierarchickú metódu. Pri tejto metóde sa podobnosť objektov, respektíve zhlukov, meria ako súčet štvorcov medzi objektmi z dvoch zhlukov, sčítaná cez všetky atribúty daných objektov. Pri danej metóde sa snažíme o minimalizáciu súčtu rozptylov cez všetky novovytvorené zhluky. Je preto veľmi vhodná, keďže cieľom je vytvárať zhluky tak, aby sme v rámci konkrétneho zhluku vytvorili čo najväčšiu previazanosť medzi jeho objektmi. [6] Za optimálny počet zhlukov sme zvolili 4. Výsledky zhlukovej analýzy uvádzame na Obr. 5.

Ak sme identifikovali 4 základné zhluky (klastre), vieme o každom jednom časovom okne povedať, v ktorom z týchto zhlukov sa nachádza. Keďže sa do jedného klastra zhlukujú triedy ekvivalencie predstavujúce také minimálne kostry, ktoré majú veľmi podobnú štruktúru, predpokladáme, že konkrétny klaster predstavuje konkrétny režim, ktorý sa nachádza na trhu. Tento predpoklad vychádza z empirických pozorovaní, kedy napríklad v prípade veľkých poklesov na trhu majú výnosnosti akcií tendenciu skorelovať a teda byť pri sebe veľmi blízko, čo by v našej minimálnej kostre malo vytvoriť tvar „hviezdy“ (taký graf, ktorý má jeden vrchol stupňa *n-1* a *n-1* vrcholov stupňa *1*. Na druhej strane empirické pozorovania poukazujú na to, že v prípade stabilného trhu sa výnosnosti akcií vyvíjajú každá vlastným smerom, preto

tam nepredpokladáme, že akcie budú až tak veľmi blízko seba, teda uvažujeme, že v našich minimálnych kostrách sa začnú objavovať dlhšie cesty.



Obr. 5: Výsledky zhlukovej analýzy a identifikácia 4 zhlukov (vlastné spracovanie v programe R)

5. Matica prechodu

Zo zhlukovej analýzy sme pre každú jednu zo *174* tried ekvivalencie určili zhluk, do ktorého by mala minimálna kostra daného časového okna patriť. Identifikovali sme štyri režimy trhu, o ktorých zatiaľ nemáme žiadne informácie, poznáme iba štruktúry trhu reprezentované minimálnymi kostrami, ktoré do nich patria. Ďalšou otázkou je, ako sa príslušnosti k jednotlivým zhlukom vyvíjajú v čase. Ak predpokladáme, že americké burzové indexy vďaka svojej obrovskej likvidite predstavujú trhy blízke efektívnym trhom, potom môžeme prijať predpoklad o neexistencii pamäte tohto trhu a výnosnosti teda považovať za realizácie markovovského procesu. Pri prijatí tohto predpokladu uvažujeme o matici prechodu, ktorá predstavuje štvorcovú maticu rozmeru *4x4* (lebo sme identifikovali *4* zhluky – režimy trhu), ktorá v *i*-tom riadku a *j*-tom stĺpci vyjadruje počet režimov trhu, ktoré boli v nejakom časovom okne v *i*-tom a v nasledujúcom časovom okne v *j*-tom zhluku. Súčet prvkov na hlavnej diagonále tejto matice preto predstavuje počet režimov, ktoré ostali v rovnakom zhluku aj v nasledujúcom období. Túto maticu uvádzame v Tab. 2:

Tab. 2: Matica prechodu 174 časových okien (vlastné spracovanie v programe R)

	row.names	1.zhluk	2.zhluk	3.zhluk	4.zhluk	sum
1	1.zhluk	13	14	5	11	43
2	2.zhluk	13	25	7	12	57
3	3.zhluk	6	10	13	5	34
4	4.zhluk	11	8	9	11	39
5	sum	43	57	34	39	173

Z tejto matice sme následne odhadli maticu pravdepodobnosti prechodu z jedného zhluku do iného. Interpretácia tejto matice je nasledovná: hodnota v *i*-tom riadku a *j*-tom stĺpci predstavuje pravdepodobnosť, že ak trh nachádza v režime *i*, v nasledujúcom období sa bude nachádzať v režime *j*. Danú maticu pravdepodobnosti uvádzame v Tab. 3:

Tab. 3: Matica pravdepodobnosti prechodu 174 časových okien (vlastné spracovanie v programe R)

row.names	1.zhluk	2.zhluk	3.zhluk	4.zhluk
1.zhluk	0.3023256	0.3255814	0.1162791	0.2558140
2.zhluk	0.2280702	0.4385965	0.1228070	0.2105263
3.zhluk	0.1764706	0.2941176	0.3823529	0.1470588
4.zhluk	0.2820513	0.2051282	0.2307692	0.2820513

6. Ďalšie možnosti analýzy

V článku sme popísali doterajšie výsledky, ktoré sme analýzou amerického kapitálového trhu získali. Analýza tým samozrejme nekončí, avšak malý priestor tohto článku autorovi neumožňuje, aby sa venoval všetkým častiam, uvedieme však aké sú plány pre nasledovné analýzy. Nakoľko sme si v prípade jednej triedy ekvivalencie rozdelili trh na základe zhlukovej analýzy na niekoľko režimov, budeme zhlukovú analýzu vykonávať aj na iných triedach ekvivalencie a na matici susednosti, pričom výsledky porovnáme. Okrem toho by sme chceli získať aj nejaké informácie o trhu, ktorý sa nachádza v konkrétnom zhluku, ktorý sme určili zhlukovou analýzou. Plánujeme to realizovať prostredníctvom súhrnného DJIA indexu, kde na daných časových oknách vypočítame základné štatistiky a budeme testovať, či tieto charakteristiky sú rovnaké pre všetky klastre alebo sa nám podarí identifikovať nejaké základné charakteristiky jednotlivých klastrov. Maticu pravdepodobnosti prechodu by sme chceli použiť pri výpočte takzvaných markovovských reťazcov, pri ktorých by sme chceli vedieť vypočítať pravdepodobnosť toho, že ak sa trh, ktorý identifikujeme na základe vybraných charakteristík nachádza v určitom režime, v priebehu nasledujúceho krátkeho obdobia sa dostane do iného konkrétneho režimu. Taktiež chceme zistiť, či trhy v určitých režimoch majú tendenciu po dlhšom čase konvergovať k nejakému konkrétnemu režimu alebo majú striedavý vývoj.

Záver

Článok predstavuje zhrnutie doterajšej analýzy štruktúry kapitálového trhu v USA reprezentovaného podindexmi DJIA indexu. Za použitia štatistických metód a metód teórie grafov sa nám podarilo identifikovať určité trhové režimy, ktoré máme v pláne v budúcnosti ďalej analyzovať.

Literatúra

[1] FINANCIAL NETWORK ANALYTICS (FNA): A Short History of Correlation Networks Research. [online]. [cit. 10. 11. 2014]. Dostupné na: <http://www.fna.fi/blog/2012/07/06/a-short-history-of-correlation-networks-research>.

[2] MANTEGNA, R.N. 1999. Hierarchical structure in financial markets. In: *The European Physical Journal B*, s. 193 – 196.

[3] YINGHUA ZHANG. 2009. Stock Market Network Topology Analysis Based on a Minimum Spanning Tree Approach. Thesis.

[4] MISKIEWICZ, J. 2012. Analysis of Time Series Correlation. The Choice of Distance Metrics and Network Structure. In: *Acta Physica Polonica A*, s. B89 – B94.

[5] CZAP, JÚLIUS. 2013. Teória grafov. *Nepublikované študijné materiály zverejnené v MOODLE*, EkF TU v Košiciach.

[6] WARD, J.H. 1963, Hierarchical Grouping to Optimize an Objective Function. In: *Journal of the American Statistical Association*, s. 236 - 244.

[7] ONNELA, J. P. – CHAKRABORTI, A. – KASKI, K. – KERTÉSZ, J. – KANTO, A. 2003. Asset Trees and Asset Graphs in Financial Markets. In: *Physica Scripta*, s. 48 – 53.

[8] BONANNO, G. – CALDARELLI, G. – LILLO, F. – MICCICHÉ, S. – VANDEWALLE, N. – MANTEGNA, R. N. 2004. Networks of Equities in Financial Markets. In: *The European Physical Journal B.*

[9] MIZUNO, T. – TAKAYASU, H. – TAKAYASU, M. 2005. Correlation Networks Among Currencies.

[10] NAYLOR, M. – ROSE, L. – MOYLE, B. 2007. Topology of Foreign Exchanges Markets Using Hierarchical Structure Methods. In: *Physica A.*

[11] KENETT, D. – TUMMINELLO, M. – MADI, A. – GERSHGOREN, G. – MANTEGNA, R. N. – BEN-JACOB, E. 2010. Dominating Clasp of the Financial Sector Revealed by Partial Correlation Analysis of the Stock Market.

Adresa autora:

Jakub Danko, Bc.
Ekonomická fakulta
Technická univerzita
2. ročník, 2. stupeň
dankojakub@gmail.com

Analýza singulárneho spektra aplikovaná na časový rad prichádzajúcich hovorov

Singular spectrum analysis aplied on call center arrivals time series

Lukáš Kalina

Abstract: We present a model-free approach to analyse high frequency and noisy time series, which lack structure and to decompose the original series into a small number of additive components such as trend, oscillatory components and noise. We apply the singular spectrum analysis method on real call center arrivals, which are known to be noisy and "structureless", using the SAS/ETS software.

Abstrakt: Prezentujeme neparametrický model na analýzu vysokofrekvenčných, nestacionárnych časových radov a na ich dekompozíciu do malého počtu komponentov akými sú trend, harmonický komponent a biely šum. Aplikujeme metódu analýzy singulárneho spektra na reálne prichádzajúce hovory, ktoré bývajú zašumené a "bez štruktúry", použitím SAS/ETS softvéru.

Key words: singular spectrum analysis, time series, call center arrivals, SAS/ETS

Kľúčové slová: analýza singulárneho spektra, časový rad, prichádzajúce hovory, SAS/ETS

1. Úvod

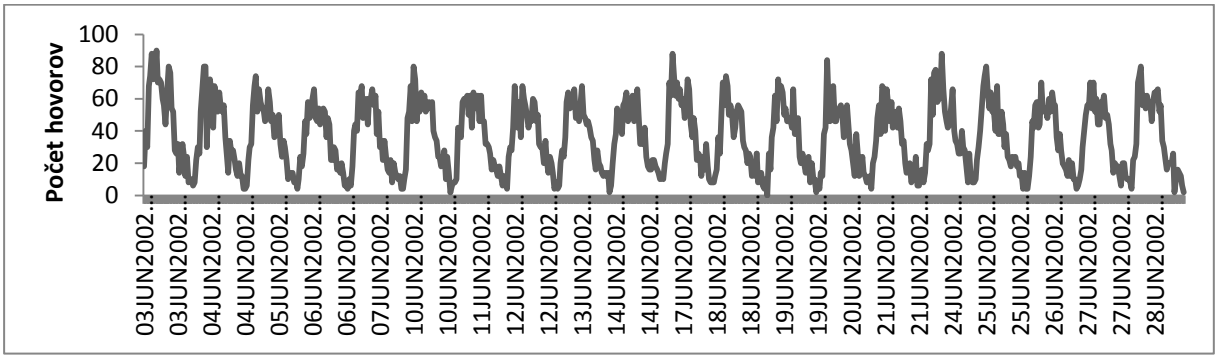
V odbore zákazníckeho servisu je jednou z najťažších výziev práve odhad očakávanej práce, pod ktorou môžeme napríklad chápať prichádzajúce hovory, emaily, listy alebo príchod zákazníkov do predajne ale aj v zdravotníctve napríklad počet pacientov prichádzajúcich na pohotovosť. Historické dáta majú veľkú granularitu, nakoľko je dôležité aby prichádzajúce hovory, zákazníci alebo pacienti boli čo v najkratšom čase obslúžený. Inak môže nastať lievikový efekt, ktorý bude viesť k nespokojnosti, stresu a stratám či už finančným alebo v prípade zdravotníctva stratám na životoch. Tým že sa ale údaje zbierajú s veľmi vysokou frekvenciou za krátke časové úseky sú nestacionárne v strednej hodnote alebo variabilite. Vo väčšine prípadov tradičné metódy časových radov zlyhávajú vo vysvetlení funkčných závislostí a nie sú schopné zachytiť dynamiku systému. V našej práci by sme chceli predstaviť modernú metódu nazývanú analýza singulárneho spektra (ďalej SSA z anglického názvu Singular Spectrum Analysis), ktorá kombinuje prvky analýzy časových radov a metódy viacrozmernej štatistiky. Jej výhodou je, že je neparametrická a nepredpokladá stacionaritu časového radu alebo biely šum reziduí. Podľa [6,7,8] bola prvý krát predstavená v 80. rokoch minulého storočia ale do dnešnej doby je spracovaná iba v niekoľkých ucelených publikáciách z dôvodu jej výpočtovej náročnosti a nízkej popularity. Zatiaľ sa nezaraďuje medzi klasické štatistické metódy, ktoré sa preberajú na základných kurzoch.

2. Informácie o vstupných dátach

Zpracované údaje pochádajú z call centra jednej americkej banky, ktorá má kontaktné centrá v New York, Pennsylvania, Rhode Island a Massachussets. Kontaktné centrá spracujú denne približne 300 tisíc hovorov a zabezpečujú služby pre rozličné segmenty klientov. V našej práci sa zameriame na retail linku, ktorá tvorí ¾ z celkového počtu prichádzajúcich hovorov. Budeme analyzovať prichádzajúce hovory v dňoch od pondelka do piatka, v operačných časoch od 07:00 do 24:00, v kalendárnych dňoch od 01.06.2002 do 30.8.2002. Na obrázku 1 môžeme pozorovať počet prichádzajúcich hovorov za mesiac jún 2002. Zvolili sme kratšie obdobie aby bolo jednoduchšie pozorovať časový rad. V tabuľke 1 zhrnieme úvodné fakty o charaktere údajov:

Tab.4: Sumár úvodných faktov analyzovanej databázy

Fakt	Hodnota
Počet analyzovaných dní:	**65**
Počet sviatkov	**1**
Dĺžka sledovanej periódy počas dňa [min]	**30**
Počet periód počas dňa	**34**



Obr.3: Počet prichádzajúcich hovorov za mesiac jún 2002

3. Algebraický zápis vstupnej databázy

V našej práci budeme rozlišovať dva druhy zápisov vstupných dát:

a) v maticovom tvare:

Nech je počet prichádzajúcich hovorov zapísaný v maticovom tvare označený **A** kde počet prichádzajúcich hovorov v perióde dňa j, dňa i a typu dňa d je označený $a_{ij}^{(d)}$, kde $j = 1, \dots, m$, $i = 1, \dots, n$, $d \in \{1,2,3,4,5\}$ a n_d je počet dní typu d.

Napríklad počet prichádzajúcich hovorov v utorok 11.06.2002 od 07:00 do 07:30 môžeme zapísať $a_{11,1}^{(2)}$.

b) v tvare stĺpcového vektora (časového radu):

Nech je počet prichádzajúcich hovorov zapísaný v tvare vektora $\mathbf{A} = (a_1, \dots, a_i, \dots, a_N)^{\mathrm{T}}$.

4. Imputácia údajov

Ako je z tabuľky 1 zrejmé v sledovanom období máme jeden sviatok, pre ktorý je charakteristický nižší počet hovorov ako je očakávaný počet charakteristický pre daný typ dňa a iný vzor (pattern) prichádzajúcich hovorov počas jednotlivých periód. Z toho dôvodu sme sa rozhodli tento deň imputovať použitím nasledovného postupu:

1. Predpokladáme, že prichádzajúce hovory do retail linky sú z dvojnásobne stochastického modelu Poissonovho pravdepodobnostného rozdelenia s parametrom $\lambda_j^{(d)}$.
2. Určíme neskreslený bodový odhad parametra $\lambda_j^{(d)}$, ktorým je výberová charakteristika

$$\bar{X}^{(d)}{}_j = \frac{1}{n_d} \sum_{i=1}^{n} a_{ij}^{(d)} \qquad (1)$$

Kde $\bar{X}^{(d)}{}_j$ môžeme interpretovať ako priemerný počet hovorov v perióde dňa j a v type dňa d.

3. Počet hovorov v špeciálny deň, v danú periódu dňa nahradíme pomocou náhodného generátora čísel s Poissonovým pravdepodobnostným rozdelením s parametrom $\bar{X}^{(d)}{}_j$.

Dáta spracovávame pomocou softvéru SAS a pokročilých programovacích techník. V prípade imputácie používame vyhľadávaciu tabuľku, ktorá obsahuje nasledovné premenné:

typ dňa, perióda dňa, priemerný počet hovorov pre daný typ dňa a periódu dňa ($\bar{X}^{(d)}{}_j$) so zloženým indexom na type a perióde dňa. Iterujeme cez tabuľku prichádzajúcich hovorov, ktorá obsahuje nasledovné premenné: kalendárny deň, perióda dňa, typ dňa, indikátor špeciálneho dňa a počet prichádzajúcich hovorov. V prípade, že je v programovom dátovom vektore indikátor sviatku nastavený na hodnotu 1 tak pomocou *set key* príkazu vyberieme z vyhľadávacej tabuľky parameter $\bar{X}^{(d)}{}_j$ a nahradíme za počet prichádzajúcich hovorov hodnotu, ktorá sa nám vráti z funkcie *RAND('POISSON',* $\bar{X}^{(d)}{}_j$*)*. Priemerná absolútna relatívna zmena parametra $\bar{X}^{(d)}{}_j$ po imputácii bola v porovnaní s parametrom pred imputáciou iba 0,51 % preto môžeme konštatovať relatívne nízke skreslenie údajov.

5. Použitá metodológia

SSA analýza je neparametrická metóda, ktorá dokáže pracovať s akýmkoľvek procesom: lineárnym, nelineárnym, stacionárnym, nestacionárnym, gaussovým alebo negaussovým. Tradičné metódy analýzy časových radov predpokladajú stacionaritu časového radu a normalitu reziduí. SSA však nemá žiadne predpoklady. Táto metóda dekomponuje časový rad na jej základné prvky a následne ju rekonštruuje [6,7,8].

Nech $N > 2$. Predpokladajme časový rad kladných reálnych čísel $\boldsymbol{A} = (a_1, \dots, a_i, \dots, a_N)^T$ s dĺžkou N. SSA metóda pozostáva z týchto fáz a krokov:

- Prvá fáza: Dekompozícia

1. Krok: Premapovanie

V tomto kroku sa rozdelí pôvodný časový rad na niekoľko viacdimenzionálnych časových radov, ktoré sú závislé od toho istého časového radu, posunuté o L období dozadu.
Nech je L konštanta, reprezentujúca dĺžku oneskorenia, tak že $1 < L < N$. A konštanta K pre ktorú platí $K = N - L + 1$ reprezentuje počet oneskorených vektorov

$$\boldsymbol{A_i} = (a_i, \dots, a_{i+L-1})^T, 1 \le i \le K \tag{2}$$

ktoré majú dimenziu L. Trajekčná matica časového radu $\boldsymbol{A}$:

$$\boldsymbol{A} = [\boldsymbol{A_1}, \dots, \boldsymbol{A_K}] \tag{3}$$

má oneskorené vektory v stĺpcoch. Môžeme ju taktiež zapísať nasledovne:

$$\boldsymbol{A} = a_{ij} = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_K \\ a_2 & a_3 & a_4 & \dots & a_{K+1} \\ \dots & \dots & \dots & \dots & \dots \\ a_L & a_{L+1} & a_{L+2} & \dots & a_N \end{pmatrix} \tag{4}$$

Platí, že $a_{ij} = a_{i+j-1}$ a prvky na diagonále sú konštantné. Projekčná matica sa nazýva taktiež aj Hankerova matica. L by sme mali zvoliť dostatočne veľké aby sme zachytili základné vlastnosti a dynamiku pôvodného časového radu. V prípade, že pracujeme s dátami so sezónym komponentom tak by malo platiť, že $L \ge s$ kde s je počet sezón. Ak zvolíme príliš malé L môžu sa zmiešať komponenty časového radu a tým by sa nedali správne interpretovať.

2. Krok: Singulárna dekompozícia

Na projekčnú maticu teraz aplikujeme singulárnu dekompozíciu (ďalej SVD z anglického názvu Singular Value Decomposition). Nech $\sigma_1, \sigma_2, \dots, \sigma_L$ sú singulárne čísla matice $\boldsymbol{AA^T}$ zoradené vzostupne tak, že $\sigma_1 \ge \sigma_2 \ge \dots \ge \sigma_L \ge 0$ (v niektorej literatúre sa používa názov odmocnené vlastné čísla lebo platí $\sigma_i = \sqrt{\lambda_i}$) a vlastné vektory $\boldsymbol{u_1}, \dots, \boldsymbol{u_L}$ matice $\boldsymbol{AA^T}$ korešpondujú s príslušnými vlastnými číslami. Nech $d = max(i, ktoré\ \sigma_i > 0)$. Ak platí $\boldsymbol{v_i} = \boldsymbol{A^T u_i}/\sigma_i$ môžeme SVD trajekčnej matice $\boldsymbol{A}$ zapísať nasledovne:

$$\boldsymbol{A} = \boldsymbol{a_1} + \dots + \boldsymbol{a_d} \tag{5}$$

Kde $\boldsymbol{a_i} = \sigma_i \boldsymbol{u_i v_i^T}$. Stĺpcové vektory $\boldsymbol{a_i}$ majú hodnosť 1 preto ich môžeme považovať za elementárne matice. Kolekcia $(\sigma_i \boldsymbol{u_i v_i})$ sa nazýva taktiež i-ta vlastná trojička SVD (z anglického názvu eigentriple).

- Druhá fáza: Rekonštrukcia

3. Krok: Zoskupovanie vlastných trojičiek

Po dosiahnutí rozloženia (5), rozdelíme vektory $\{1, \dots, d\}$ na disjunktné skupiny $\boldsymbol{I_1}, \dots, \boldsymbol{I_m}$. Nech je $\boldsymbol{I} = \{\mathrm{i}_1, \dots, \mathrm{i}_p\}$. Potom výsledná matica $\mathbf{A}_{\boldsymbol{I}}$ korešpondujúca so skupinou $\boldsymbol{I}$ je definovaná ako $\boldsymbol{A} = \boldsymbol{a}_{\mathbf{i}_1} + \cdots + \boldsymbol{a}_{\mathbf{i}_p}$. Tieto matice sú vypočítané pre $\boldsymbol{I} = \boldsymbol{I_1}, \dots, \boldsymbol{I_m}$ a rozloženie $\boldsymbol{A} = \boldsymbol{a_1} + \cdots + \boldsymbol{a_d}$ vedie k ďalšej dekompozícii, ktorú môžeme zapísať nasledovne:

$$\boldsymbol{A} = \boldsymbol{a}_{\boldsymbol{I}_1} + \cdots + \boldsymbol{a}_{\boldsymbol{I}_m} \tag{6}$$

Proces hľadania ortogonálnych skupín $\boldsymbol{I_1}, \dots, \boldsymbol{I_m}$ sa nazýva zoskupovanie do vlastných trojičiek. Zoskupovanie slúži na tvorbu aditívnych komponentov časového radu.

4. Krok: Diagonálne priemerovanie

V poslednom kroku SSA metódy transformujeme každú maticu zoskupenej dekompozície do nového časového radu s dĺžkou N.

Nech $\boldsymbol{Y}$ je $L \times K$ matica s prvkami $y_{ij}, 1 \le i \le L, 1 \le j \le K$. Nastavíme $L^* = min(L, K), K^* = max(L, K)$ a $N = L + K - 1$. Nech je $y_{ij}^* = y_{ij}$ ak $L > K$ a $y_{ij}^* = y_{ji}$ ak $L \le K$. Diagonálnym priemerovaním transformujeme maticu $\boldsymbol{Y}$ do časových radov $\boldsymbol{y_1}, \dots, \boldsymbol{y_N}$ pomocou nasledovných vzorcov:

$$\boldsymbol{y_k} = \begin{cases} \frac{1}{k}\sum_{m=1}^{k} y_{m,k-m+1}^* & pre\ 1 \le k < L^*, \\ \frac{1}{L^*}\sum_{m=1}^{L^*} y_{m,k-m+1}^* & pre\ L^* \le k < K^*, \\ \frac{1}{N-k+1}\sum_{m=k-K^*+1}^{N-K^*} y_{m,k-m+1}^* & pre\ K^* \le k < N. \end{cases} \tag{7}$$

Vzorec súvisí s priemerovaním prvkov matice na jej diagonále $i + j = k + 1$. Pri výbere $k = 1$ dostávame $\boldsymbol{y_1} = y_{1,1}$, pre $k = 2$ dostávame $\boldsymbol{y_2} = (y_{1,2} + y_{2,1})/2$ a tak ďalej. Matica $\boldsymbol{Y}$ je podobne ako trajekčná matica - Hankelova matica.

Diagonálne priemerovanie aplikované na maticu $\boldsymbol{A_{I_k}}$, tvorí časový rad $\widetilde{\boldsymbol{A}}^{(\boldsymbol{k})} = (\widetilde{\boldsymbol{a}}_{\boldsymbol{1}}^{(\boldsymbol{k})}, \dots, \widetilde{\boldsymbol{a}}_{\boldsymbol{N}}^{(\boldsymbol{k})})$. Jej inicializačné časové rady $\boldsymbol{a_1}, \dots, \boldsymbol{a_N}$ sú dekomponované do sumy m časových radov:

$$\boldsymbol{a_n} = \sum_{k=1}^{m} \tilde{a}_n^{(k)} \tag{8}$$

6. Softvérové riešenie SSA

Pri riešení používame SAS edukačný analytický balíček a SAS JMP. Skôr ako sme začali analyzovať časový rad pomocou ďalej spomenutých procedúr bolo potrebné najskôr zadefinovať v SAS-e kustomizované intervaly, ktoré zodpovedajú nami použitým dátam. V sledovanom období máme 3 mesiace zachytávajúce pracovné dni od pondelka do piatka a pracovný čas od 07:00 do 24:00 pričom jeden sledovaný interval má 30 minút. Vytvoríme tabuľku s atribútmi *Begin*, *End* a *Season*, obsahujúcu všetky polhodinové intervaly, ktoré ideme analyzovať. Tabuľkou inicializujeme náš vlastný interval *WorkHours* pomocou príkazu

```
options intervalds = (<Nazov_Intervalu> = <Nazov_tabulky>);
```

Pre technický detail použitia nastavenia *intervalds* odporúčame [5].

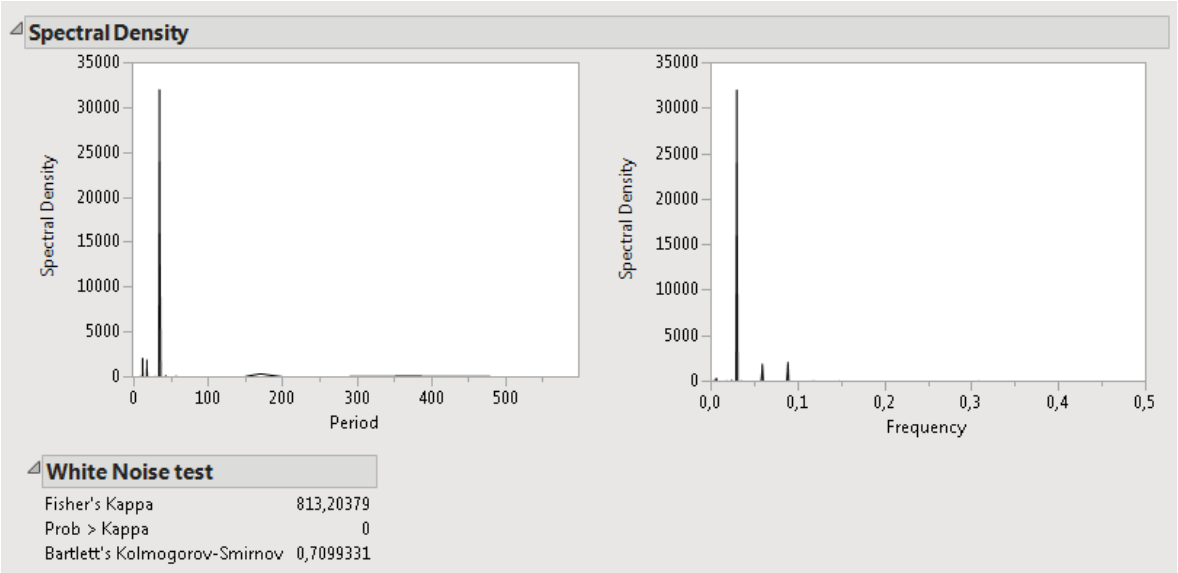
Čo sa týka metódy SSA existuje niekoľko softvérov, ktoré nám poskytujú možnosť aplikácie metódy SSA a ako príklad môžeme uviesť implementáciu SSA do MATLAB alebo R. Zvolili sme si štatistický softvér SAS/ETS 12.1, ktorý obsahuje na túto činnosť špeciálne nastavenie *SSA* v rámci procedúry *PROC TIMESERIES*. Pre technické detaily odporúčame [2, 4]. Implementácia SSA v SAS-e používateľa síce odľahčí od náročných kalkulácií, v metodike uvedených, krokov ale poskytuje podľa nášho názoru relatívne málo výstupov, ktoré by sme mohli využiť pri hlbšej analýze. Ako príklad môžeme uviesť neprítomnosť výstupu, pomocou SVD dekomponovanej, trajekčnej matice $\boldsymbol{a_i}$ na vlastnú trojičku $\sigma_i \boldsymbol{u_i} \boldsymbol{v_i^T}$. Na druhej strane si môžeme byť istý, správnosťou výsledkov nakoľko sa SAS používa hlavne na

komerčné účely. Z toho dôvodu sa pri vývoji a testovaní nových metód zapája široká odborná verejnosť a dôraz sa kladie na kvalitu a presnosť, čo sa pri nekomerčných riešeniach jednoznačne zaručiť nedá.

V rámci procedúry *PROC TIMESERIES* je možné si zvoliť aj grafické zobrazenie použitím *plot=(SSA),* ktorý nám poskytne histogram zobrazujúci singulárne čísla, kumulatívny histogram vysvetlenej variability singulárnych čísel a porovnanie jednotlivých komponentov s originálnym časovým radom. Ďalej je možné presmerovať výstup do SAS tabuľky, ktorý obsahuje názov atribútu, časový interval, cyklus, sezónu, originálny časový rad a jednotlivé aditívne komponenty.

7. Praktická časť

Náš časový rad s dĺžkou $N = 2210$ dekomponujeme SSA metódou na jednotlivé komponenty, ktorými sú: trend, harmonický signál a náhodný prvok. Najskôr ale zvolíme veľkosť oneskorenia časového radu L.



Obr.4: Analýza spektrálnej hustoty

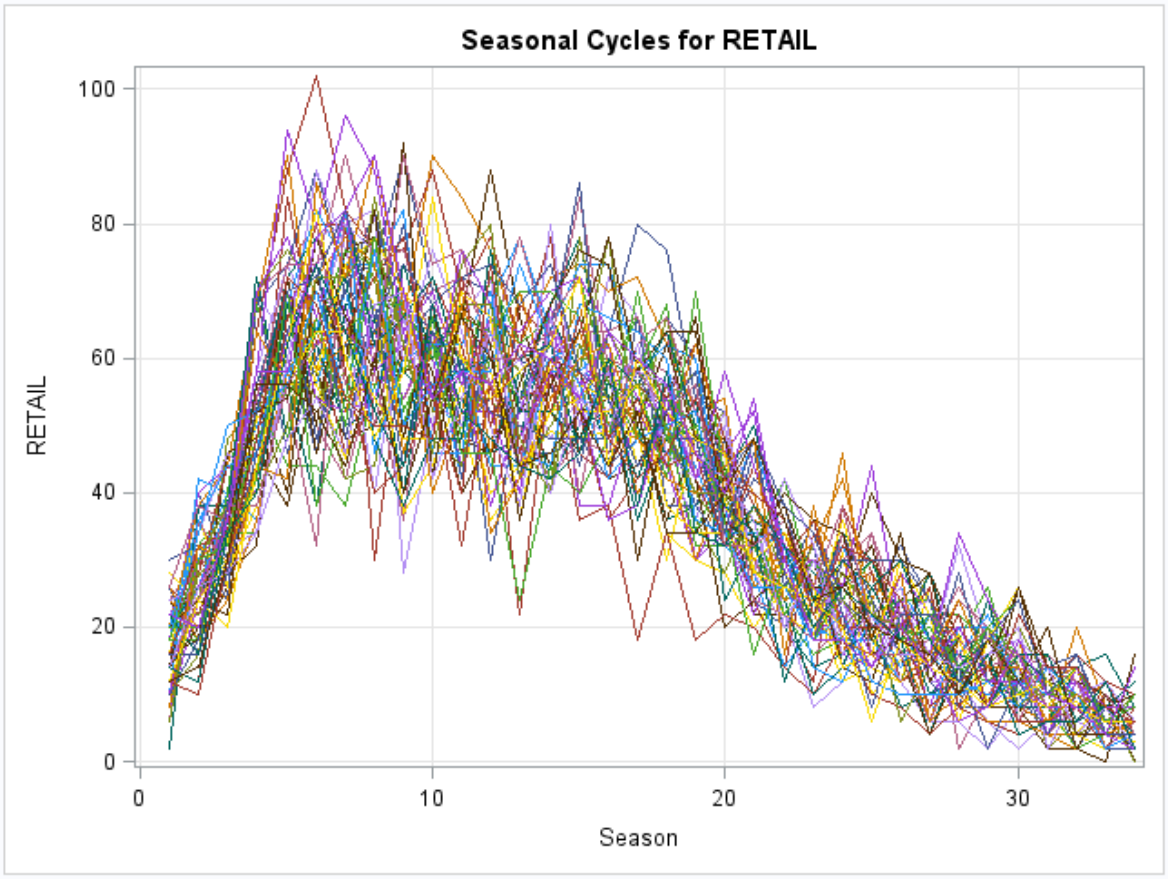
Rozhodneme sa na základe analýzy spektrálnej hustoty, ktorá je na obrázku 2. Jej myšlienkou je rozložiť stacionárny časový rad do kombinácie sinusoíd s náhodnými a nekorelovanými koeficientami. Testovacia štatistika je *Fisher´s Kappa*, ktorá nadobúda hodnotu 813,20. Testujeme hypotézu H_0 podľa, ktorej sú hodnoty časového radu z normálneho rozdelenia s rozptylom 1 proti alternatívnej hypotéze, že časový rad má periodický komponent. Ak je p-hodnota (na obrázku 2 *Prob >Kappa)* väčšia ako hladina významnosti α zamietame nulovú hypotézu. Ako je z obrázku zrejmé môžeme zamietnuť hypotézu H_0, konštatovať prítomnosť periodického komponentu a tým nastaviť veľkosť oneskorenia pôvodného časového radu na $L = 34$.

V dole uvedenom kóde používame $L = LENGTH = 34$. Špecifikovali sme ho tiež v nastavení *SEASONALITY* aby sme mohli zobraziť aj sezónne cykly. Do *SAS* príkazu *GROUPS* sa definujú skupiny, do ktorých je dekomponovaný pôvodný časový rad. Prvá skupina obsahuje prvý lag (lag = oneskorenie), druhá skupina obsahuje druhý lag a tak ďalej.

Namiesto príkazu *GROUPS* môžeme použiť napríklad *THRESHOLDPCT*, ktorým určíme veľkosť poslednej skupiny na základe kumulatívneho percenta variability singulárnych čísel. Odporúčame ale manuálne určiť komponenty v skupinách a až na základe ich grafickej analýzy rozhodnúť o zaradení oneskorení do skupín. Skupiny nemusia nevyhnutne obsahovať lagy, ktoré idú bezprostredne za sebou.

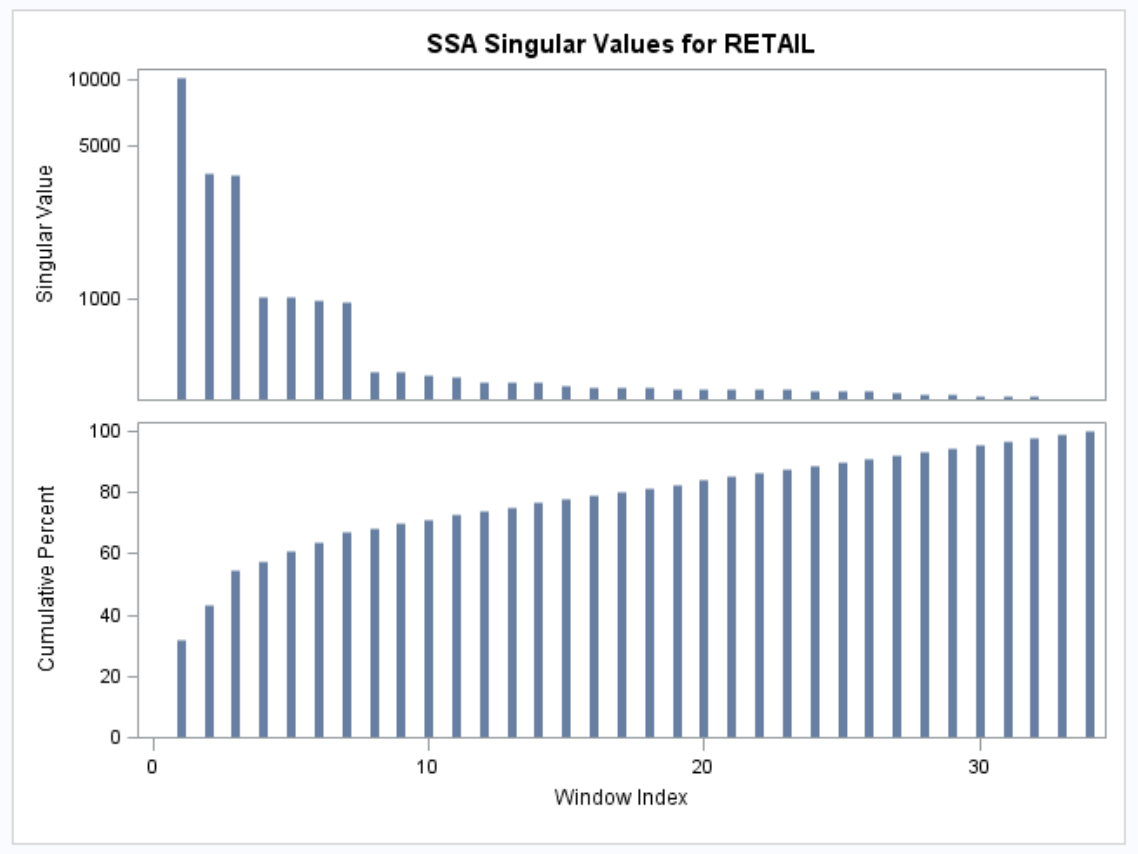
```
proc timeseries data = SE_ARSE_TS_INPUTED outssa = SSA_OUTPUT
plots=(series cycles ssa) SEASONALITY = 34;
SSA / LENGTH = 34 groups = (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)
(12)(13)(14)(15)(16)(17)(18)(19)(20)(21)(22)(23)(24)(25)(26)
(27)(28)(29)(30)(31)(32)(33)(34);
id DATE_TIME interval = WorkHours;
var RETAIL;
run;
```

(*Pozn.: V uvedenom kóde v príkaze id je interval = WorkHours. „WorkHours“ je nami vytvorený interval. Obyčajne sa používajú prednastavené intervaly typu Weekday, Month, Year.)*



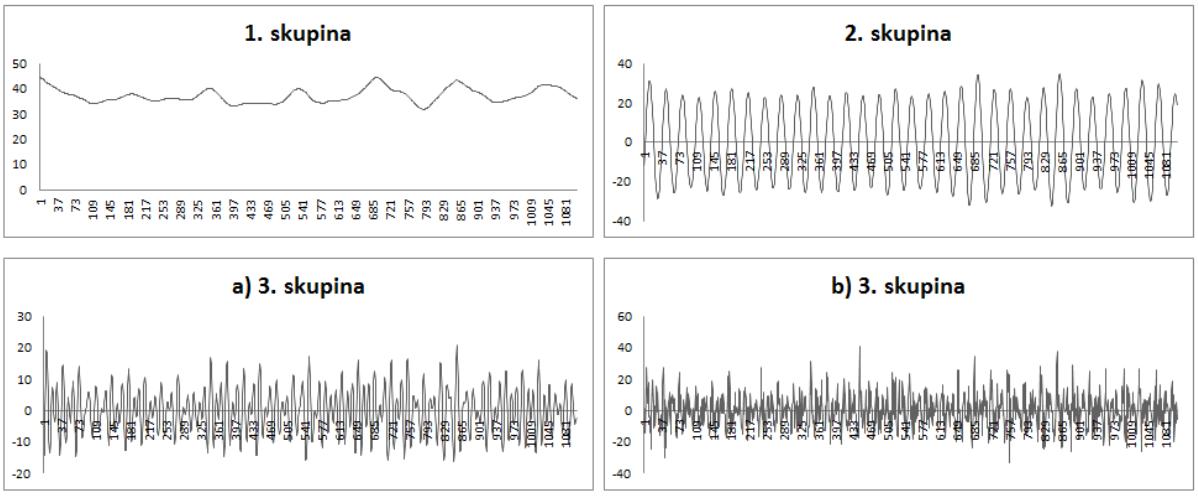
Obr.5: Graf sezónnych cyklov prichádajúcich hovorov počas dňa

Obrázok 3 reprezentuje graf sezónnych vzorov, ktoré sú charakteristické v priebehu dňa. Horizontálna os s názvom *Season* by sa dala substituovať jednotlivými polhodinovými intervalmi počas pracovnej doby. Približne o 9:00 je počet hovorov najvyšší. Pomalým tempom klesá do 16:30 následne počet prichádzajúcich hovorov klesá rýchlejšie až do konca prevádzkovej doby.



Obr. 6: Vlastné spektrum prichádzajúcich hovorov

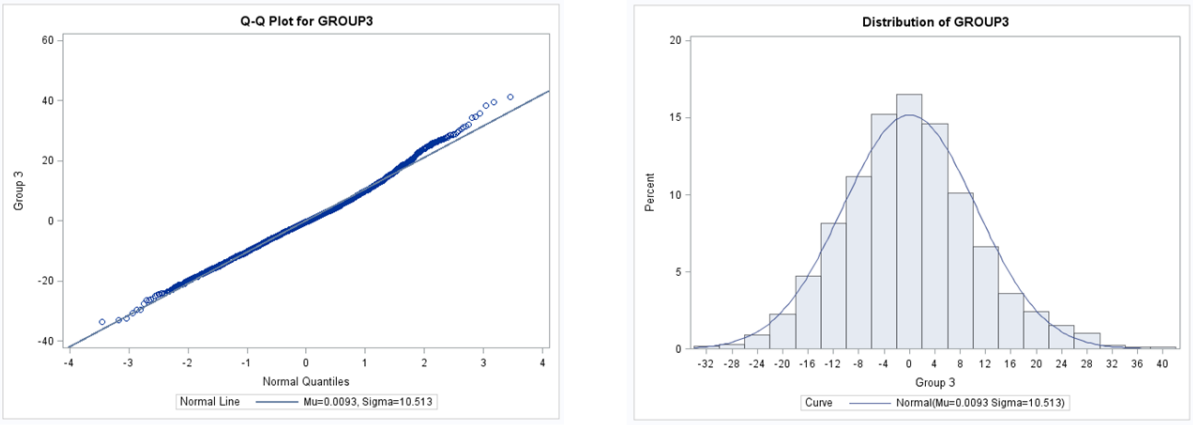
V prvom histograme na obrázku 4 môžeme pozorovať singulárne čísla σ_i a v druhom ich kumulatívne percento vysvetlenej variability trajekčnej matice. Z týchto grafov môžeme odpozorovať prvé náznaky spektrálnych skupín. Prvá skupina zrejme obsahuje 1. vlastnú trojičku $\boldsymbol{a_{I_1}} = \boldsymbol{a_1} = \sigma_1 \boldsymbol{u_1 v_1^T}$, druhá skupina 2. a 3. vlastnú trojičku $\boldsymbol{a_{I_2}} = \boldsymbol{a_2} + \boldsymbol{a_3} = \sigma_2 \boldsymbol{u_2 v_2^T} + \sigma_3 \boldsymbol{u_3 v_3^T}$, tretia skupina 4., 5., 6. a 7. vlastnú trojičku $\boldsymbol{a_{I_3}} = \sum_{i=4}^{7} a_i$ a posledná skupina zachytáva zostávajúcu variabilitu. O tvorbe skupín sa ale rozhodneme až na základe grafickej analýzy jednotlivých elementárnych matíc $\boldsymbol{a_i}$



Obr.7: Spektrálne skupiny

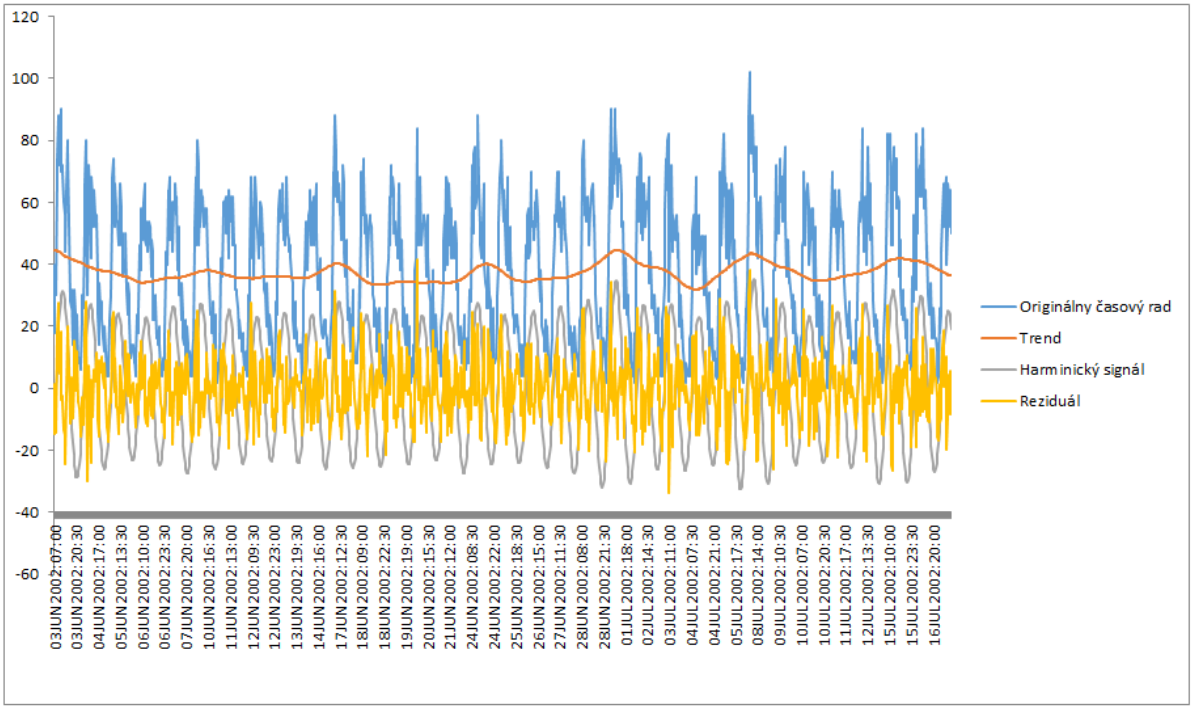
Na obrázku 5 môžeme pozorovať jednotlivé spektrálne skupiny, ktoré sme si predbežne definovali na základe singulárnych čísiel. Prvú skupinu môžeme interpretovať ako trend. Ako je zrejmé z vertikálnej osi zachytáva úroveň časového radu prichádzajúcich hovorov. Druhú

skupinu interpretujeme ako harmonický signál, ktorý pravidelne osciluje okolo horizontálnej osy so strednou hodnotou 0 a konštantným rozptylom. V tretej skupine (označenej *a) 3.skupina*) pozostávajúcej z 4. až 7. vlastnej trojičky nieje jednoznačne zrejmý žiadny funkčný vzťah preto sme sa ju rozhodli spojiť zo zvyšnými vlastnými trojičkami, ktoré môžeme pozorovať na grafe s názvom *b) 3.skupina* a interpretovať ako náhodnú zložku.



Obr.8: Analýza náhodnej zložky

Na obrázku 6 je grafická analýza náhodnej zložky. Na základe Q-Q grafu a histogramu môžeme predpokladať, že je normálne rozdelená so strednou hodnotou približne 0 a rozptylom 10.513.



Obr.9: Graf originálneho časového radu a jeho komponentov

Na obrázku 7 môžeme pozorovať originálny časový rad, ktorý je dekomponovaný na jednotlivé aditívne komponenty.

8. Záver

Ak sú jednotlivé komponenty separované, tak že zachytávajú iba jedinečný a špecifický funkčný vzťah akým je napríklad lineárny (alebo linearizovateľný) trend, sezónne, harmonické prvky, cykly (v prípade ročných časových radov) alebo náhodný prvok môžeme extrapolovať použitím prognostických metód časových radov akými sú napríklad:

- modely s konštantnou úrovňou (Náhodná prechádzka, Brownow model jednoduchého exponenciálneho vyrovnávania, Prognostické kĺzavé priemery)
- modely časových radov s trendom (Modely s konštantnými alebo variabilnými parametrami v čase) alebo
- modely periodických časových radov (CENSUS X-11, Holtovo-Wintersovo exponencionálne vyrovnávanie...).

V [6,7,8] navrhujú okrem tradičných metód prognózovania časových metód použiť lineárne rekurentné relácie. Myšlienkou tejto metódy, je že časový rad pozostáva z lineárnej kombinácie exponenciálnych, polynomických a harmonických časových radov.
Analýza singulárneho spektra je neodmysliteľnou súčasťou analýzy časových radov tam kde tradičné metódy nedokážu zachytiť komplexnú dynamiku systémov, na základe ktorých sa riadia prichádzajúce požiadavky na zákaznícke centrum.

Literatúra

[1]DENG, CH. 2014. Time Series Decomposition Using Singular Spectrum Analysis. In: *Electronic Theses and Dissertations,* Paper 2352. http://dc.etsu.edu/etd/2352. s. 12-64.

[2]LEONARD, M., JOHN, M., ELSHEIMER, B., KESSLER, M. 2010. An introduction to Singular Spectrum Analysis with SAS/ETS® Software. In: *SAS Global Forum 2010*. s. 1-15.

[3]GILLARD, J., KNIGHT, V., ANTHONY, V. 2014. Using singular spectrum analysis to obtain staffing levels requirements in emergency units. In: *Journal of the Operational Research Society,* roč. 65, č.5, s. 735-746.

[4]SAS INSTITUTE INC., 2012. SAS/ETS 12.1 Users Guide. Cary, NC, USA. In:

http://support.sas.com/documentation/cdl/en/etsug/65545/PDF/default/etsug.pdf

[5]SAS INSTITUTE INC., 2014. JMP 11 Online Documentation. Cary, NC, USA. In:

http://www.jmp.com/support/help/Time_Series_Commands.shtml

[6] GOLYANDINA, N., NEKRUTKIN, V., ZHIGLJAVSKY, A., 2001. Analysis of time series structures. SSA and related techniques. *Chapman & Hall/CRC*, s. 4-56.

[7]GOLYANDINA, N., KOROBEYNIKOV, A. 2013. Basic Singular Spectrum Analysis and R Forecasting. In: *Computational Statistics & Data Analysis*, s 1-7

[8]GOLYANDINA, N., ZHIGLJAVSKY, A., 2013. Singular Spectrum Analysis for Time Series. *SpringerBriefs in Statistics Series*. s. 6-178.

Adresa autora:

Lukáš Kalina, Bc.
Fakulta hospodárskej informatiky
Ekonomická univerzita v Bratislave
2. ročník, 2. stupeň
lukas.kalina.89@mail.sk

Simulácia evolúcie v opakovanej dileme samaritána
The Simulation of Evolution in the Iterated Samaritan's Dilemma

Martin Kameník

Abstract: In the following thesis a simulation of co-evolutionary behaviour of agents playing iterated Samaritan's dilemma against each other is presented. Strategies of agents in a population of Parasites and Samaritans are evolved through a genetic algorithm, which includes operators such as selection, crossover and mutation. This algorithm makes them relatively better. Strategies are evaluated by fitness value in every generation. Agents make decisions about next steps according to the history of three previous games. The most common result of presented simulation concerning co-evolution of agents in an environment of Samaritan's dilemma was permanent exploitation of Samaritan by Parasite. Only in one case, agents evolved into a form in which the Parasite was forced to work.

Abstrakt: V článku prezentujeme simuláciu koevolučného vývoja správania agentov, ktorí proti sebe hrajú opakovanú dilemu samaritána. Stratégie agentov v populácii parazitov aj samaritánov podliehajú evolúcii – genetickému algoritmu, ktorého súčasťou sú genetické operátory ako selekcia, kríženie a mutácia. Ten z nich robí relatívne lepšie stratégie, ktoré sú hodnotené v každej generácii hodnotou fitness, pričom agenti sa rozhodujú o voľbe nasledujúceho ťahu na základe histórie troch predchádzajúcich ťahov. Najčastejším výsledkom našej simulácie koevolučného vývoja agentov v prostredí samaritánovej dilemy je permanentné vykorisťovanie samaritána parazitom. Iba v jednom prípade sa agenti vyvinuli do podoby, v ktorej samaritán prinútil parazita pracovať.

Key words: samaritan's dilemma, genetic algorithm.

Kľúčové slová: dilema samaritána, genetický algoritmus.

1. Úvod

Evolúcia je dokázaná forma adaptácie na dynamické a komplexné prostredie. Genetický algoritmus predstavuje vyhľadávací algoritmus, ktorý napodobňuje základné rysy prirodzenej evolúcie. V tejto práci používame tento genetický algoritmus k učeniu agentov hrať samaritánovu dilemu, hru s nenulovým súčtom, ktorú prvý krát predstavil Buchanan a prostredníctvom maticovej hry znázornil negatívne dôsledky altruistického správania.

Inšpiráciou, ako viesť nasledujúcu simuláciu, sa nám stal článok The Evolution of Strategies in the Iterated Prisoner's Dilemma, v ktorom autor, okrem iného, využíva túto nedeterministickú metódu riešenia problému k tvorbe stratégii a k simulovaniu koevolučného vývoja agentov v prostredí opakovanej dilemy väzňa. Našim cieľom je však aplikácia tohto koevolučného prístupu na opakovanú dilemu samaritána. Tento prístup predpokladá existenciu dvoch populácii agentov, ktorí volia ťahy na základe výsledkov 3 predchádzajúcich kôl hry. Po odohraní hry sú prirodzenou selekciou zvolení najlepší agenti a tí podliehajú genetickým operátorom, ktoré slúžia k hľadaniu lepších stratégii.

2. Dilema samaritána

Buchanan predstavil pojem Dilemy Samaritána v ekonómii, aby ukázal, že altruizmus môže vyvolať nežiaduce správanie príjemcov založené na zlyhaní samaritána konať strategicky. Buchanan takisto identifikoval, že Dilema Samaritána odráža určité charakteristiky moderného sociálneho štátu, pretože výhody vo forme dávok sociálneho zabezpečenia sú dôvodom pre oslabenie motivácie pracovať. Dilemu znázornil ako

dvojmaticovú hru v aktívnej a pasívnej podobe, pričom v našej práci sa obmedzíme len na pasívnu Samaritánovu dilemu.

	B_1	B_2
A_1	I 4 ; 2	II 1 ; 1
A_2	III 2 ; 3	IV 3 ; 4

Obrázok 10 Pasívna dilema samaritána

Máme hráča A – Samaritán reprezentovaný stratégiami A_1 – nepomáhať a A_2 – pomáhať a hráča B, ktorý predstavuje Parazita so stratégiami B_1 – pracovať a B_2 – nepracovať. Samaritán v pasívnej forme preferuje situáciu v bunke I, ktorá predstavuje sociálne optimálnu situáciu, zatiaľ čo Parazit preferuje situáciu v bunke IV inak aj sociálne neakceptovateľnú situáciu, ktorej by sa spoločnosť chcela vyhnúť. Tieto bunky sú Nashovou rovnováhou hry.

Dilema Samaritána vlastne predstavuje jeho vykorisťovanie kvôli neschopnosti odmietnuť pomáhať a teda Parazit strategickým chovaním vo forme voľby stratégie B_2 núti Samaritána použiť stratégiu A_2. Základným rozdielom medzi aktívnym a pasívnym samaritánom je, že v prvom prípade mu je určité strategické správanie nariadené jeho dominantnou stratégiou, zatiaľ čo v druhom prípade je mu diktované správanie hernou situáciou, do ktorej ho donútil Parazit.

3. Simulácia

Uvažujme o populácii samaritánov a parazitov v podobne r-tíc $S=(s_1,s_2,...,s_r)$ a $P=(p_1,p_2,...,p_r)$ pozostávajúcich z r agentov so stratégiami definovanými v podobe binárnych l-tíc $s_i=(s_{i,1},s_{i,2},...,s_{i,l})$ a $p_j=(p_{j,1},p_{j,2},...,p_{j,l})$, $i,j\in\{1,2,...,r\}$, pričom agenti hrajú opakovanú dilemu samaritána. Pre tieto binárne l-tice teda platí $s_{i,a},p_{j,b}\in\{0,1\}$, $a,b\in\{1,2,...,l\}$.

Nasledujúcu časť analyzujeme z pohľadu samaritána, ale analogicky môžeme dané postupy aplikovať aj na parazita. Počas existencie jedincov v generácii t každý samaritán s_i odohraje m kôl s každým parazitom p_j z opozitnej populácie. Každému jedincovi z populácie je po odohraní všetkých kôl priradená hodnota fitness, ktorá determinuje efektívnosť jedinca pri riešení problému:

$$f_t(s_i)=\frac{\sum_{k=1}^{m}\sum_{j=1}^{r}\varphi_k(s_i,p_j)}{mr} \qquad (1)$$

kde $\varphi_k(s_i,p_j)$ vyjadruje výplatu i-teho samaritána proti j- temu parazitovi v k-tom kole hry. Cieľom agentov je teda maximalizácia priemernej výplaty.

Jedinci s vyššou priemernou výplatou majú právo prispieť viacerými potomkami do generácie $t+1$. K vhodnému výberu jedincov využívame metódu nazývanú turnajová selekcia (s vrátením) matematicky analyzovanú autormi Blickle a Thiele. Tento selekčný mechanizmus náhodne zvolí dve l- tice s_{i_1} a s_{i_2}, $i_{1,2}\in\{1,2,...,r\}$ z S a za predpokladu $f_t(s_{i_1})>f_t(s_{i_2})$ postupuje do generácie $t+1$ agent s_{i_1}, v opačnom prípade postupuje s_{i_2}. V prípade ak $f_t(s_{i_1})=f_t(s_{i_2})$ postupuje jedinec s_{i_1}. Daný postup opakujeme r- krát.

Očakávaný počet výskytov $p(s_i)$ potomkov jedinca s_i v generácii $t+1$ je potom daný ako:

$$p(s_i) = r\left[\left(\frac{i}{r}\right)^2 - \left(\frac{i-1}{r}\right)^2\right] \tag{2}$$

za predpokladu $f_t(s_1) < f_t(s_2) < ... < f_t(s_r)$, teda že každý agent v populácii dosiahne unikátnu hodnotu fitness. V tom prípade môžeme i interpretovať ako počet agentov s hodnotou fitness $f_t(s_i)$ alebo nižšou.

Nakoľko sa v generácii $t+1$ objavuje potomok jedinca, je potrebné ho vytvoriť. Na to využívame dva genetické operátory.

V prvom rade ide o genetický operátor kríženie. Majme náhodne vybrané l- tice $s_{i_1} = (s_{i_1,1}, s_{i_1,2}, ..., s_{i_1,l})$ a $s_{i_2} = (s_{i_2,1}, s_{i_2,2}, ..., s_{i_2,l})$ z S, ktoré úspešne prešli selekčným mechanizmom. Zvolíme pozíciu $1 < \lambda \leq l$ na základe rovnomerného rozdelenia pravdepodobnosti, pričom vychádzajúc z jednobodového kríženia nám vznikajú dve nové l- tice $(s_{i_1,1}, ..., s_{i_1,\lambda-1}, s_{i_2,\lambda}, ..., s_{i_2,l})$ a $(s_{i_2,1}, ..., s_{i_2,\lambda-1}, s_{i_1,\lambda}, ..., s_{i_1,l})$, ktoré predstavujú dvoch nových agentov s novými stratégiami. Všetky l- tice, ktoré prešli selekčným mechanizmom sú náhodne spárované a preto daný postup opakujeme $r/2$-krát. Literatúra odporúča, aby kríženie bolo aplikované s určitou vysokou pravdepodobnosťou. V našom prípade to je pravdepodobnosť 0,85.

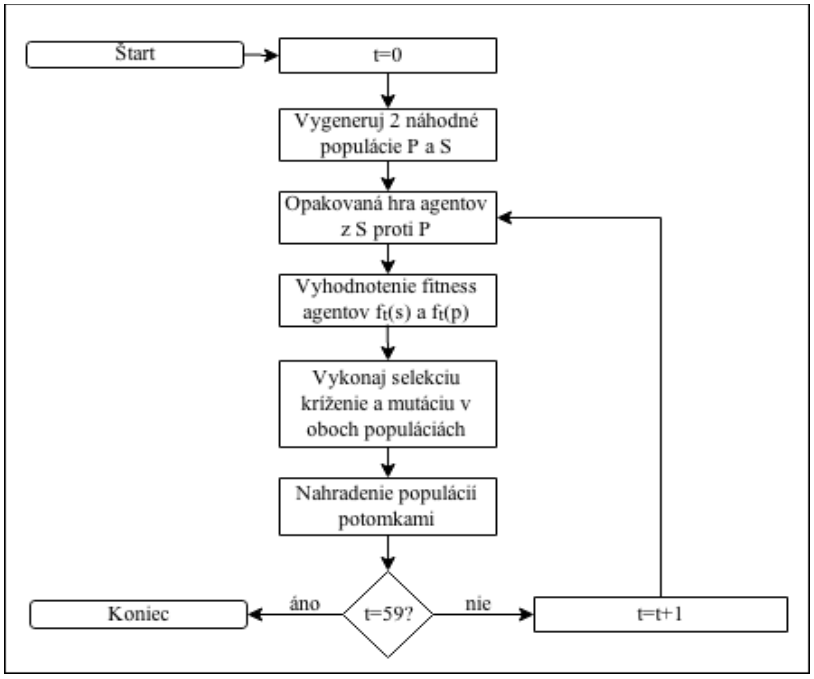
Okrem toho sa pre zachovanie diverzity v populácii využíva genetický operátor mutácia, ktorý je definovaný ako náhodné prehodenie každého bitu l-tice $s_i = (s_{i,1}, s_{i,2}, ..., s_{i,l})$ s pravdepodobnosťou $0{,}005$.

Tento proces vedie k vzniku nových agentov a takisto nových populácií samaritánov a parazitov, pričom stratégiami sa jedinci podobajú ich relatívne úspešnejším rodičom z predchádzajúcej generácie. Následne sa celý tento postup opakuje. V takomto koevolučnom procese prostredie pre prvú populáciu pozostáva z druhej populácie a naopak, pričom začína sa s náhodnými populáciami s nízkymi priemernými výplatami. Prvá populácia sa pokúša adaptovať na prostredie druhej populácie a simultánne druhá populácia sa snaží adaptovať na prostredie pozostávajúce z prvej populácie.

Predpokladáme, že hráči volia svoje ťahy ($A_1 - A_2$, $B_1 - B_2$) v nasledujúcom kole na základe histórie 3 predchádzajúcich hier, teda histórie vlastných ťahov a ťahov protihráča. Hráčom je umožnené voliť iba čisté stratégie. Ďalej označme možnosti $A_1 - A_2$ resp. $B_1 - B_2$ ako 1 – 0. Nakoľko hra v jednom kole môže nadobudnúť 4 rôzne výsledky, tak existuje $4^3 = 64$ možných histórií s pamäťou 3. Napríklad história opakovaných volieb stratégií „Pomáhať - Pracovať“ počas 3 predchádzajúcich kôl je „111111“ (nepárna pozícia vyjadruje ťahy hráča a párna pozícia ťahy protihráča). Reakciu na takúto minulosť sa nachádza v prvku $s_{i,64}$. Takto má každá história, ktorá môže vzniknúť, priradený určitý prvok v l- tici. Avšak v prvých troch kolách hry ešte nebola vytvorená história s požadovanou veľkosťou, preto je kódovanie nasledovné. Prvok $s_{i,65}$ predstavuje počiatočný ťah v prvom kole, $s_{i,66}$ a $s_{i,67}$ vyjadrujú aký ťah bude použitý v druhom kole v závislosti od predchádzajúceho ťahu oponenta a prvky $s_{i,68}$, $s_{i,69}$, $s_{i,70}$, $s_{i,71}$ vyjadrujú, aký ťah bude použitý v treťom kole v závislosti od ťahov oponenta vykonaných počas dvoch predchádzajúcich kôl.

V našom prípade je teda $l = 71$. Množina všetkých l- tíc predstavuje prehľadávaný priestor $\Omega = \{0,1\}^l$, v ktorom ak označíme každú l- ticu celým číslom, tak jej mohutnosť je

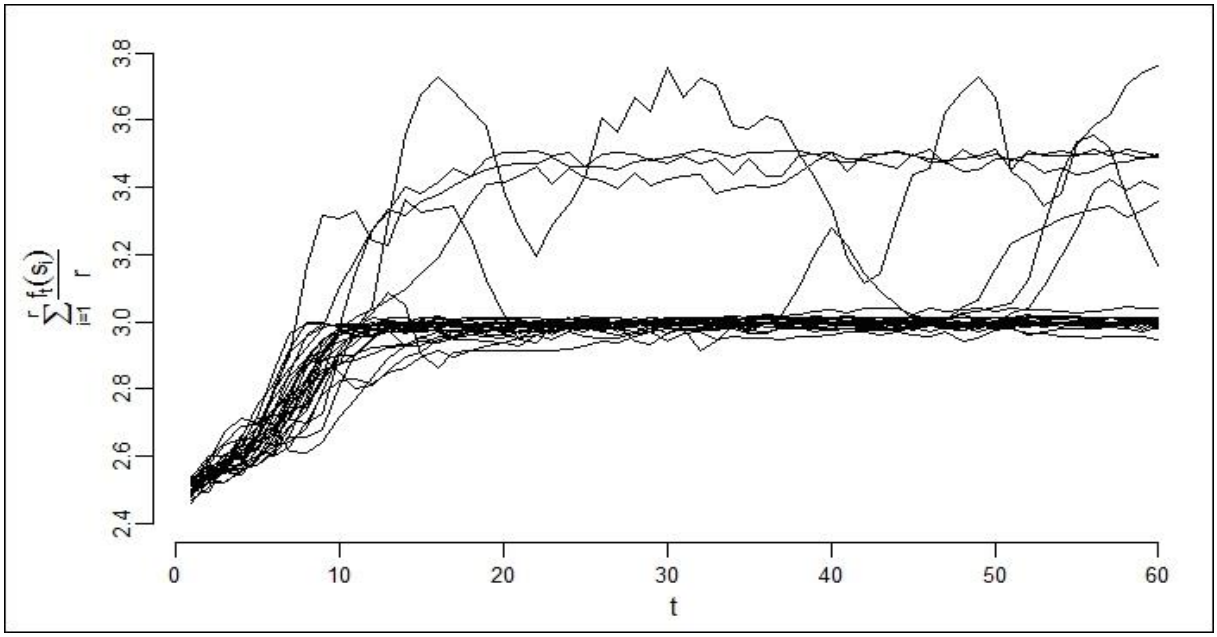
$|\Omega| = 2^{71}$, čo je značne rozsiahly počet dostupných stratégií. V našej práci ďalej uvažujeme s $r = 70$ a $m = 70$.



Obrázok 11 Jednoduchý vývojový diagram simulácie

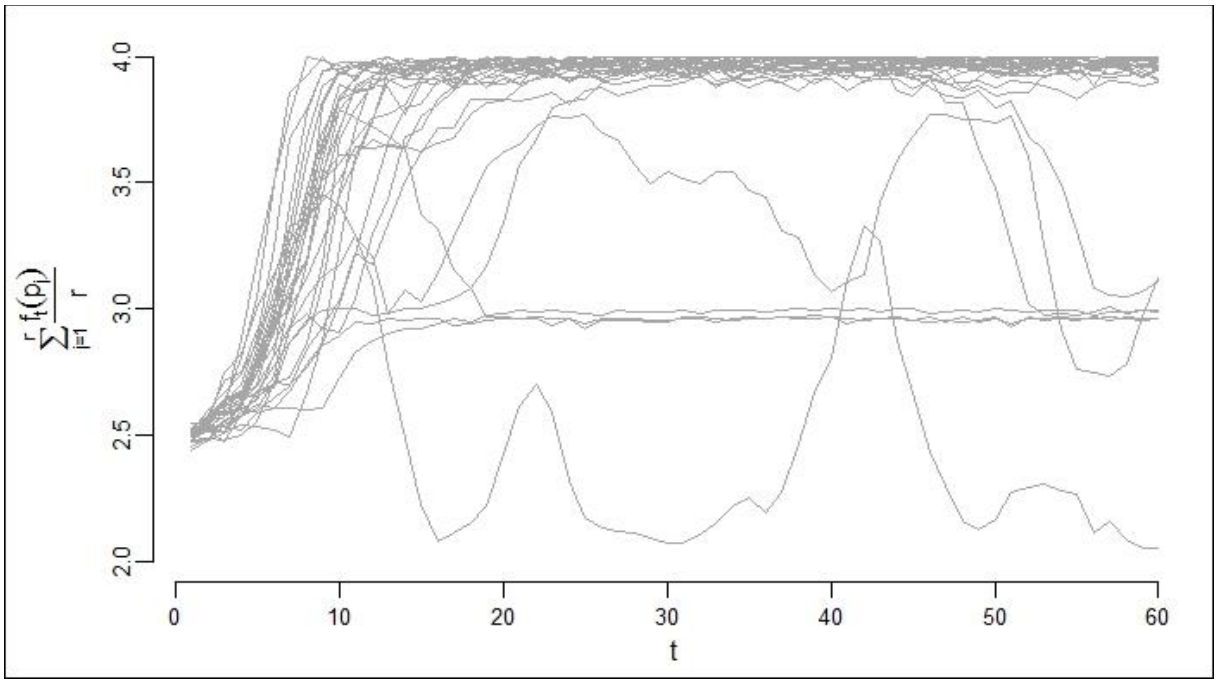
4. Výsledky

Celá simulácia bola vytvorená v programovacom jazyku R s využitím balíkov GA a diagram. Je potrebné zdôrazniť, že genetický algoritmus spadajúci do oblasti heuristických algoritmov môže nadobúdať rôzne výsledky. Preto sme každú simuláciu dilemy samaritána zopakovali 30 krát pri rovnakých nastaveniach, ale s náhodne vygenerovanými počiatočnými populáciami samaritánov a parazitov. Obrázok č. 3 zachytáva vývoj priemerného fitness za populáciu samaritánov počas 60 generácií. Súhrne teda znázorňuje všetkých 30 simulácii. Zároveň pre 4 simulácie sme zvýšili počet generácii, počas ktorých pokračuje ich vývoj, ktorý nebol v priebehu 60 generácií jednoznačne ukončený.



Obr. 12: Priemerná výplata za populáciu Samaritánov v jednotlivých generáciách

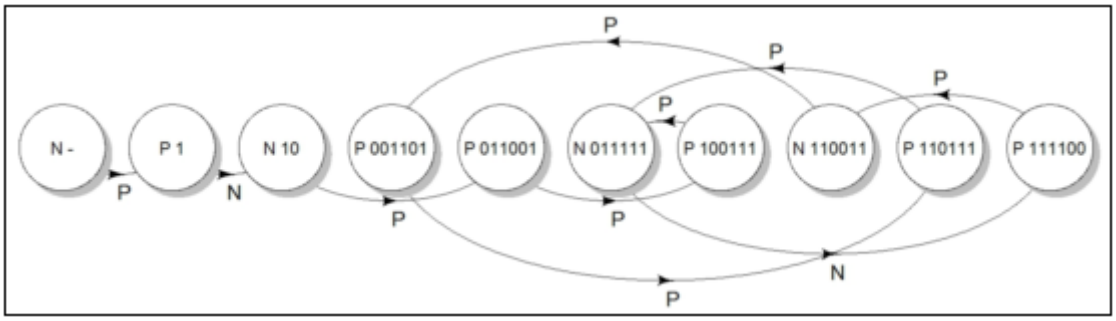
Obdobný graf z pohľadu parazita zachytáva obrázok č. 4.



Obrázok 13: Priemerná výplata za populáciu Parazitov v jednotlivých generáciách

Celkovo v 23 simuláciách skonvergoval vývoj hier počiatočných náhodných agentov pomerne rýchlo do stavu, v ktorom je samaritán v každom kole vykorisťovaný parazitom, teda Nashova rovnováha z bunky I na obrázku č. 1 má najväčšiu početnosť výskytov. Už približne od 18. generácie táto sociálne neakceptovateľná situácia jednoznačne prevažovala vo väčšine simulácií.

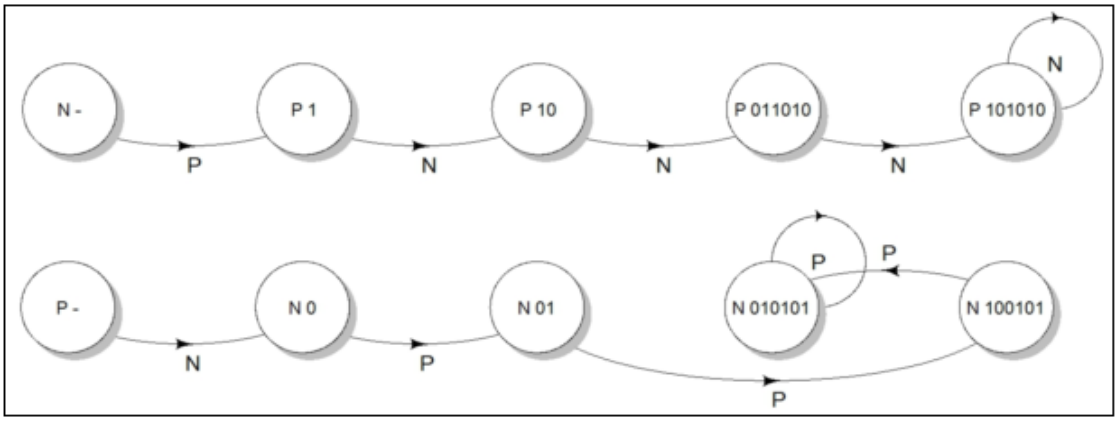
Intuitívnejší popis priebehu opakovanej hry dvoch agentov znázorňuje prechodový diagram. Vrcholy reprezentujú históriu 3 hier, ktoré predchádzali danému stavu spolu s reakciou agenta na danú históriu (v prípade prvého vrcholu história ešte nebola vytvorená, zatiaľ čo ďalšie dva vrcholy zobrazujú iba históriu ťahov protihráča). Orientované hrany a slučky predstavujú ťah protihráča, ktorý taktiež reaguje na históriu predchádzajúcich hier. Diagram grafu na obrázku č. 5 naznačuje ako môže vyzerať priebeh opakovanej dilemy samaritána dvoch náhodne vygenerovaných stratégií.



Obrázok 14 Diagram vývoja hry 2 náhodne vygenerovaných agentov

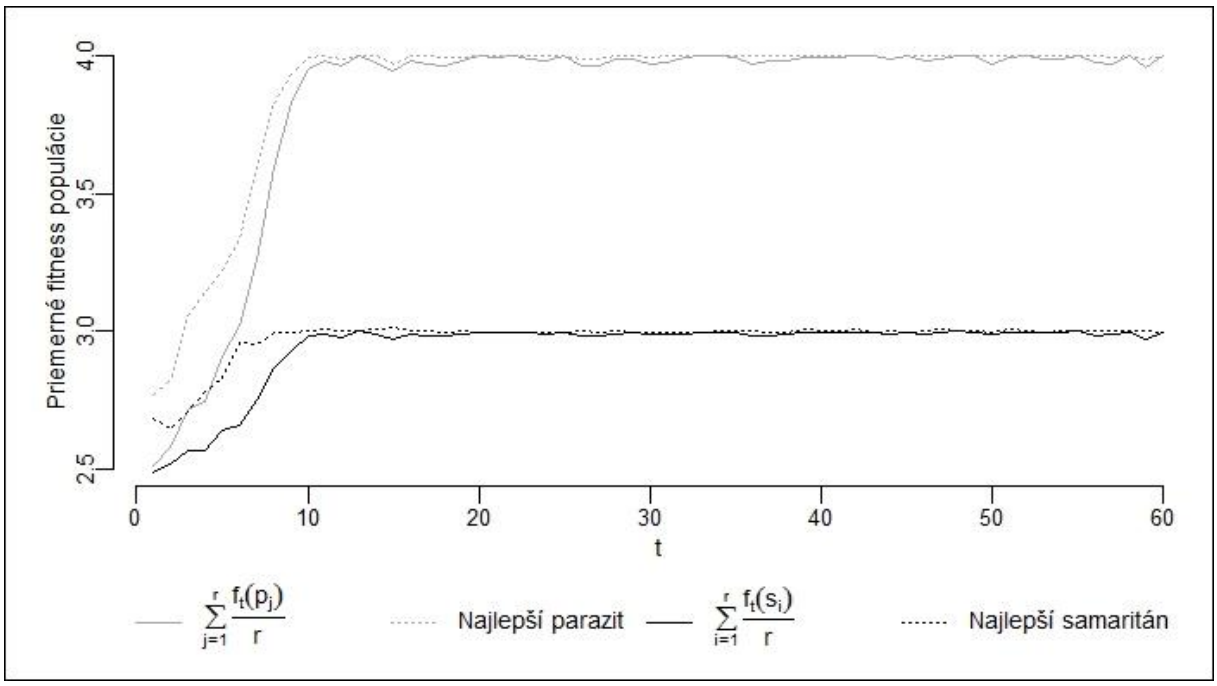
Diagram na obrázku č. 6 hore zobrazuje vývoj hry priemerného samaritána[4] proti priemernému parazitovi z vybranej simulácie, ktorá skonvergovala do stavu vykorisťovania a diagram dole zobrazuje rovnakú situáciu, ale z pohľadu parazita.

[4] Ak konkrétny prvok v l-tici je 0 pre viac ako 50% jedincov v populácii tak ide o priemerný prvok - bit, následne podobným prepočtom všetkých bitov dostaneme priemerného jedinca populácie.



Obrázok 15 Typický priebeh hry hráčov konečných populácií v 23 simuláciách

V prvom kole sa síce hráči voľbou stratégií „Nepomáhať - Pracovať“ ocitnú v bunke I, ale hneď na to je voľba oboch opačná. Po odohraní ďalších 3 kôl je už počiatočný ťah vylúčený z pamäte oboch hráčov a vyžaduje sa od nich reakcia na opakované vykorisťovanie samaritána parazitom. Následná reakcia „Pomáhať - Nepracovať“ ich privedie do stavu, v ktorom je história rovnaká ako v predchádzajúcom kole. V dôsledku toho dochádza k neustálemu opakovaniu rovnakých ťahov hráčov vyjadrenej slučkou v oboch diagramoch.



Obrázok 16 Vývoj priemerných výplat v simulácii č. 23

Väčšina zo spomínaných 23 simulácií mala podobný priebeh ako na obrázku č. 7. Priemerná hodnota výplat všetkých počiatočných populácii samaritánov bola 2,51 so smerodajnou odchýlkou 0,020. Pre populácie parazitov je počiatočná priemerná výplata 2,49 a smerodajná odchýlka 0,023. Následná selekcia agentov s najvyššou priemernou výplatou a vznik nových v podobe potomkov úspešných agentov pôsobí na rast priemernej výplaty oboch populácii v priebehu nasledujúcich generácii. Avšak rast hodnoty priemernej výplaty populácie parazitov je prudší ako v populácii samaritánov a tým prvky $s_{i,43}$, $p_{j,22}$ (reakcia na „101010“ resp. „010101“) sa stávajú dominantnými.

V nasledujúcej tabuľke sú uvedené najvyššie početnosti, ktoré vyjadrujú koľko priemerných samaritánov (parazitov) si osvojilo v rámci konkrétnej histórie reakciu pomáhať

resp. nepomáhať (nepracovať resp. pracovať). Vychádzame z 23 simulácií, ktoré skonvergovali do stavu, kedy je samaritán vykorisťovaný parazitom.

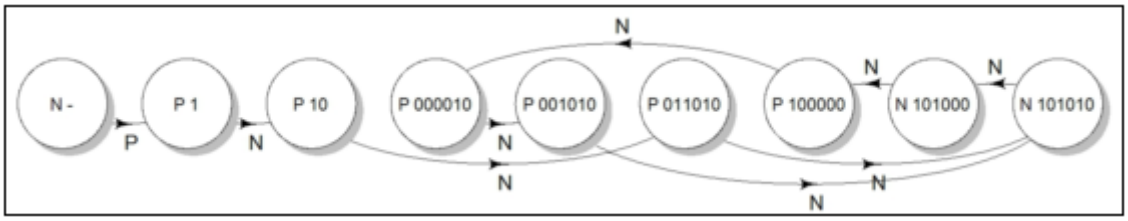
Tab.: 1 Zhoda určitých prvkov stratégií v simuláciách

História	Pomáhať	História	Nepomáhať	História	Nepracovať	História	Pracovať
"000000"	20	**"001011"**	17	**"000001"**	16	**"111000"**	17
"000010"	16	**"010011"**	16	**"000010"**	16		
"001010"	22	**"010101"**	17	**"000101"**	21		
"010000"	18	**"011111"**	16	**"001000"**	18		
"010100"	16	**"111111"**	16	**"001101"**	16		
"011001"	17			**"010000"**	16		
"011010"	21			**"010001"**	20		
"100000"	16			**"010101"**	23		
"101000"	16			**"100001"**	16		
"101010"	23			**"100101"**	19		
"101101"	16			**"101001"**	16		
"110011"	16			**"110101"**	16		
"110110"	18						
"111010"	17						

Všetci priemerní samaritáni z 23 simulácii reagovali stratégiou „Pomáhať“ na vykorisťovanie počas predchádzajúcich troch kôl. Ak si z celej pamäte troch hier vyberieme iba históriu posledného kola, tak vidíme často sa objavujúcu rovnakú reakciu na vyňatú históriu „****10“ ale aj „****00“. Naopak, väčšina Samaritánov zvolí stratégiu „Nepomáhať“ ak tejto voľbe predchádzala opakovaná kombinácia „Nepomáhať - Pracovať“ alebo aj „Pomáhať - Pracovať“.

Analogicky aj všetkých 23 priemerných parazitov reagovalo rovnako na vykorisťovanie samaritána počas 3 prechádzajúcich kôl. Väčšina stratégií parazitov si vyvinulo rovnakú reakciu „Nepracovať“ na históriu posledného kola „****01“.

Tentoraz predpokladajme, že v simuláciách, ktorých výsledkom je permanentné vykorisťovanie dôjde u niektorého zo samaritánov k mutácii prvku $s_{i,43}$, a to zo stratégie „Pomáhať“ na „Nepomáhať“. Môžeme to interpretovať ako situáciu, kedy sa samaritán po 3 kolách vykorisťovania rozhodne nepomáhať, zatiaľ čo l-tice parazitov ostávajú nezmenené. Vývoj takejto hry z pohľadu samaritána zobrazuje nasledujúci diagram.



Obrázok 17 Vývoj hry pri mutácii prvku 43. prvku l-tice samaritána

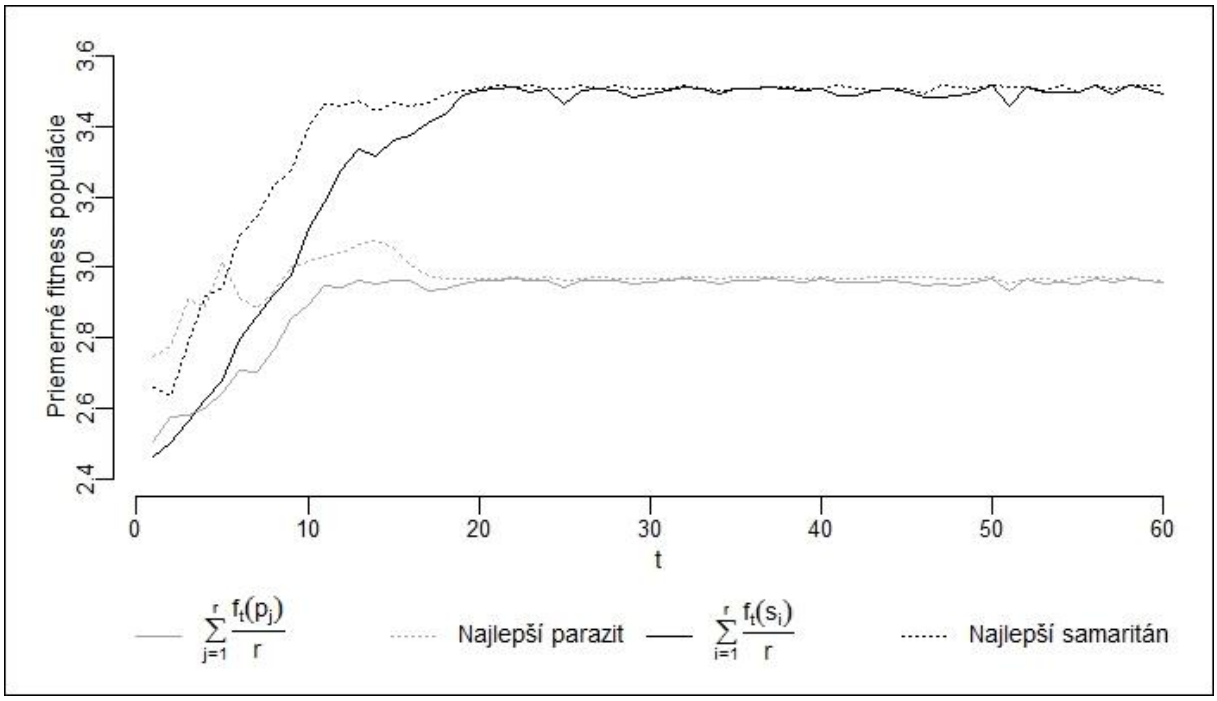
Počiatočný vývoj je zachovaný až do momentu, kedy sa od agentov vyžaduje reakcia zakódovaná v prvkoch $s_{i,43}$, $p_{j,22}$. Vieme, že parazit zareaguje stratégiou „Nepracovať“ a samaritán zmutovanou stratégiou „Nepomáhať“. Pozoruhodným dôsledkom evolučného vývoja stratégii je následný vývoj, kedy sa agenti po ďalších 3 kolách dostanú do pôvodného stavu. Od samaritána sa znova vyžaduje reakcia na históriu predstavujúcu jeho vykorisťovanie počas troch predchádzajúcich kôl. Obdobný jav je zaznamenaný aj v ďalších simuláciách, ktorých výsledkom bolo vykorisťovanie samaritána parazitom. Avšak samaritán s takouto mutáciou kvôli nižšej priemernej výplate v populácii Samaritánov neprežije a bude počas nasledujúcich generácii z populácie odstránený selekčným mechanizmom.

Priebeh hry v tých simuláciách, ktoré neskonvergovali do stavu opakovaného vykorisťovania (simulácie 1, 17, 19, 24 a 25) sa vyvíjal podobne ako na obrázku.



Obrázok 18 Typický priebeh z pohľadu samaritána konečných populácií v 5 simuláciách

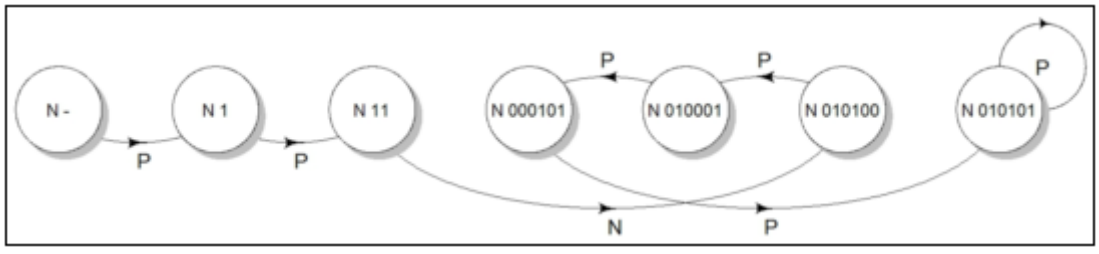
Odlišnosti sa v týchto simuláciách vyskytovali iba počas prvých kôl, zatiaľ čo ich spoločným znakom bol rovnaký priebeh v posledných dvoch vrcholoch diagramu. Parazit nedokázal úplne prinútiť samaritána k permanentnej pomoci a v dôsledku toho dochádza od určitého kola k striedavej výmene reakcií a teda aj výplat hráčov. Hráči sa striedavo pohybujú medzi bunkami I a IV, ktoré predstavujú Nashovu rovnováhu pasívnej dilemy samaritána.



Obrázok 19 Vývoj priemerných výplat v simulácii č. 24

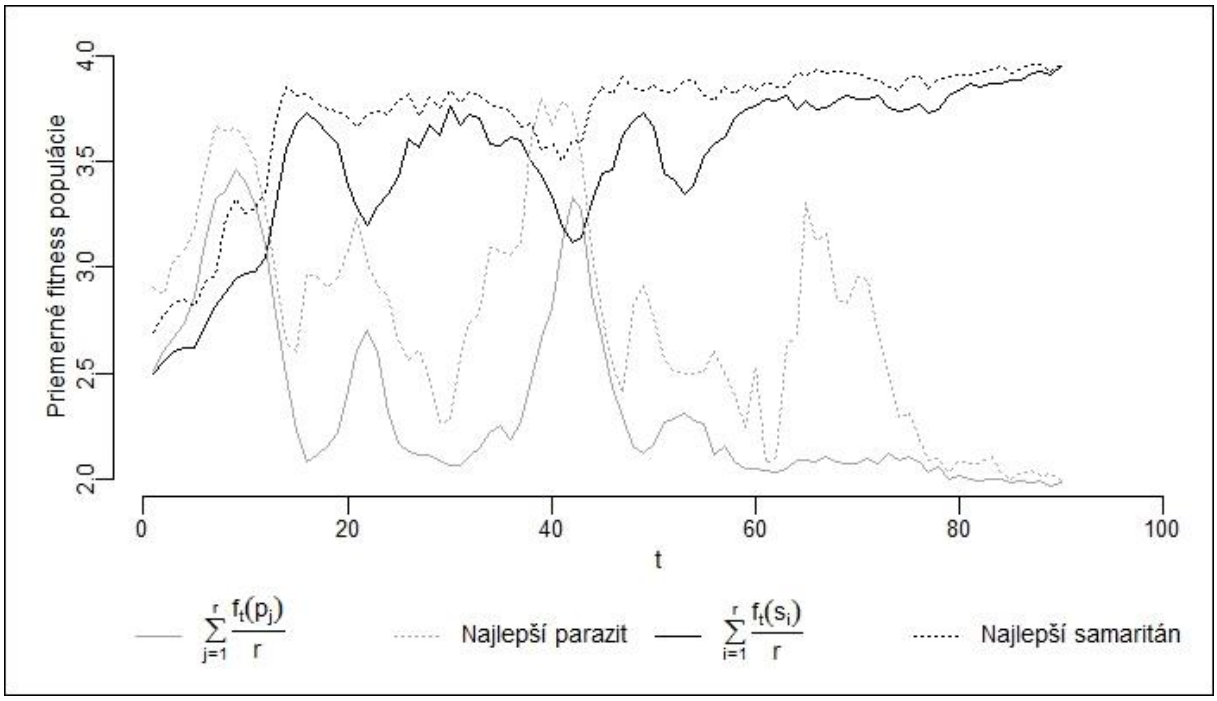
Obdobný vývoj ako na obrázku č. 10 bol aj v simuláciách č. 17 a 25, zatiaľ čo vo zvyšných dvoch sa samaritáni dokázali vyvinúť do podoby zobrazenej na obrázku č. 11 až v konečných generáciách.

Iba v simulácii č. 8 nebol priemerný parazit schopný permanentne a ani striedavo vykorisťovať priemerného samaritána. Naopak, samaritán sa vyvinul do podoby, v ktorej dokázal prinútiť priemerného parazita pravidelne pracovať, čo predstavuje sociálne optimálny stav.



Obrázok 20 Priebeh hry z pohľadu samaritána konečných populácií v simulácii č. 8

Tento vývoj ako naznačuje aj obrázok č. 12 nebol taký jednoznačný ako v predchádzajúcich dvoch prípadoch.



Obrázok 21 Vývoj priemerných výplat v simulácii č. 8

5. Záver

Kombináciou vhodnej reprezentácie stratégii a genetického algoritmu bolo možné simulovať vývoj správania agentov založeného na evolúcii v prostredí opakovanej dilemy samaritána, ktorá bola predmetom našej analýzy. Z náhodných stratégii sa prostredníctvom selekčného mechanizmu, kríženia a mutácie tvoria nové a lepšie stratégie, ktoré sú vhodnejšie prispôsobené na prostredie opozitnej populácie. V dôsledku opakovania tohto procesu vývoj konverguje k určitému výsledku. Je ale potrebné dodať, že stratégie vytvorené koevolúciou môžu byť účinné proti populácií predstavujúcu prostredie, na ktoré sa adaptujú, ale ich výkonnosť môže byť nižšia, ak budú hrať opakovanú dilemu samaritána proti inej populácii.

Aj keď stratégie agentov a tým aj priebeh hry skonverguje do určitého stavu, neznamená to, že v priebehu nasledujúcich generácií nemôže nastať zmena vo vývoji. Vyplýva to z toho, že ani jedna z populácii neukončí svoj vývoj, ale pokúša sa o neustálu adaptáciu na prostredie tvoriace druhú populáciu. Môžeme si však všimnúť, že z náhodných populácií pozostávajúcich z nie veľmi racionálnych agentov, je častým výsledkom evolučného vývoja práve opakované vykorisťovanie samaritána parazitom, čo predstavuje sociálne neakceptovateľnú situáciu. V menšom počte prípadov už vznikali populácie, ktorých agenti sa striedavo nachádzali v bunke I a IV na obrázku č. 1 a iba v jednom prípade bola populácia samaritánov schopná donútiť populáciu parazitov k opakovanej práci.

Literatúra

AXELROD, R., 1987. The Evolution of Strategies in the Iterated Prisoner's Dilemma. In: L. Davis, ed. 1975. *Genetic Algorithms and Simulated Annealing*. London: Pitman, s. 32-41.

BLICKLE, T – THIELE, L.1995. A mathematical analysis of tournament selection. In: L. Eshelman, ed., *Proceedings of the 6th International Conference on Genetic Algorithms*. San Francisco: Morgan Kaufmann Publishers, s. 9-16. URL: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=EF38AF9206ACD5D4A9668154CCA7CFCC?doi=10.1.1.52.9907&rep=rep1&type=pdf

BUCHANAN, J.M., 1972. The Samaritan's Dilemma. In: E. Phelps, ed. 1975. *Altruism, Morality and Economic Theory*. New York: Russell Sage, s. 71-85.

ENGELBRECHT, A. P. 2007. *Compuational Intelligence: An Introduction, Second Edition*. John Wiley & Sons, Ltd
ERRITY, A., 2003. *Evolving Strategies for Prisoner´s dilemma*. Dublin: Dublin City University
FOGEL, D.B., 1993. Evolving Behaviors in the Iterated Prisoner's Dilemma. In: *Evolutionary Computation*, roč. 1, č. 1, s. 77-97
KOZA, J. R., 1992. *Genetic programming: on the programming of computers by means of natural selection*. Cambridge, MA: MIT Press
SCHMITT, L.M., 2001. Theory of genetic algorithms. In: *Theoretical Computer Science*, roč. 259, č. 1, s. 1-61
WEISE, T. 2009. *Global Optimization Algorithms – Theory and Application, Second edition*. Thomas Weise

Adresa autora:

Martin Kameník, Bc.
Ekonomická fakulta
Technická univerzita v Košiciach
2. ročník, 2. stupeň
martinkamenik3@gmail.com

Komparácia aktívneho a pasívneho prístupu k úsporám v dôchodkových schémach

Comparison of active and passive saving approach in pension schemes

Michal Mešťan

Abstract: Main objective of the paper is to assess validity of assumption whether the life-cycle saving strategy based on active fund switching is significantly better performing than passive saving approach based on continuous investments solely into the bond pension fund using dynamic microsimulation model. Active fund switching is based on a comparison of cumulative daily return on pension funds over a certain period and subsequent allocation of savings to the fund by the decision-making mechanism..

Abstrakt: Cieľom príspevku je overiť predpoklad, podľa ktorého sporiteľ aktívnym prístupom k sporeniu založenom na prestupovaní medzi dvoma dôchodkovými fondmi dosiahne vyšší výnos ako pri pasívnom prístupe založenom na sporení si výlučne v jednom dôchodkovom fonde s využitým dynamického mikrosimulačného modelu. Aktívny prístup k sporeniu je založený na porovnávaní kumulatívnej dennej výnosnosti dôchodkových fondov za určité obdobie a následnej alokácie úspor do fondu podľa rozhodovacieho mechanizmu.

Key words: pension funds, saving strategy, investment strategy, active investment strategy, passive investment strategy, microsimulation model.

Kľúčové slová: dôchodkové fondy, sporiaca stratégia, investičná stratégia, aktívne investovanie, pasívne investovanie, mikrosimulačný model.

1. Úvod

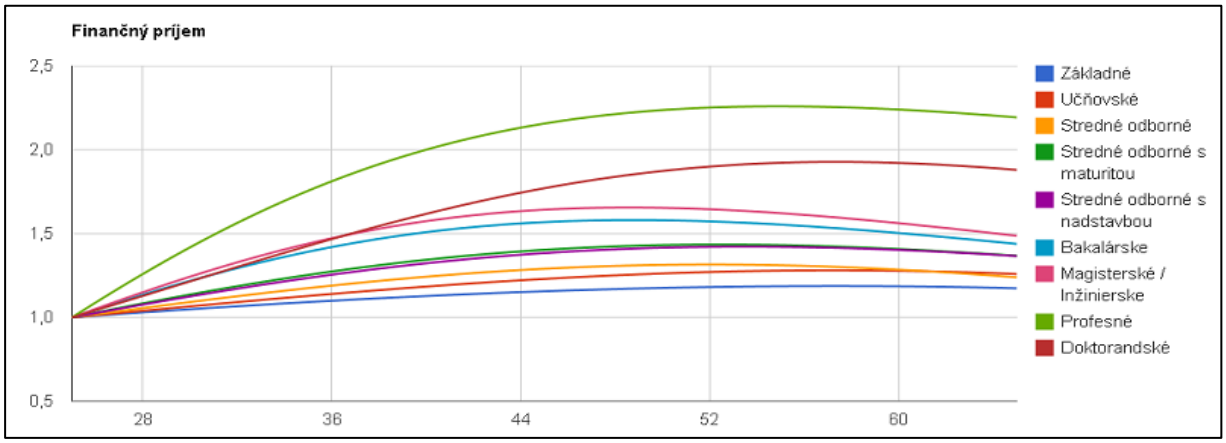
Pri sporení v kapitalizačnom pilieri na Slovensku majú sporitelia možnosť alokácie svojich úspor do dvoch fondov, z ktorých jeden musí byť vždy dlhopisový garantovaný dôchodkový fond. Väčšina sporiteľov v kapitalizačnom pilieri pôsobí konzervatívne, nakoľko sa nachádzajú v dlhopisových dôchodkových fondoch, čo môže byť spôsobené podľa Baláža (2009) nielen ich averziou k riziku, ale aj ich inertným správaním. Vysokú inertnosť sporiteľov potvrdili vo svojom výskume aj Šebo a Virdzek (2013), ktorá je spojená najmä z nedostatočným poznaním fungovania systému a možnosťami, ktoré sporiteľom poskytuje. Prostredníctvom mojej práce chcem sporiteľom prezentovať možnosti alokácie úspor do dvoch dôchodkových fondov a komparáciou aktívneho a pasívneho prístupu k sporeniu demonštrovať, že zvýšenou aktivitou v podobe prestupovania medzi fondmi môžu sporitelia dosiahnuť vyššie zhodnotenie svojich úspor ako pri inertnom prístupe (sporenie výlučne v dlhopisovom dôchodkovom fonde počas celej doby sporenia).

2. Mikrosimulačný model (MikroSIM)

Všetky výpočty a simulácie boli uskutočnené prostredníctvom mikrosimulačného modelu (MikroSIM), ktorý postupne budujeme s kolegami z Inštitútu sporenia a investovania a Ekonomickej fakulty UMB v Banskej Bystrici, v rámci riešenia viacerých výskumných projektov, na ktorých sa v súčasnosti podieľam. Vytvorený dynamický mikrosimulačný model dôchodkového systému na Slovensku obsahuje deväť typov agentov v závislosti od dosiahnutého vzdelania a ich celoživotné príjmové funkcie. Detailnejšie metodologické otázky modelovania celoživotnej príjmovej funkcie v modeli MikroSIM ako aj samotnú podstatu fungovania MikroSIM uvádzajú vo svojich publikáciách Šebo, Virdzek a Šebová (2014) alebo Šebo a Šebová (2013).

Pre potreby realizácie môjho výskumu som mal možnosť využívať funkcie modelu MikroSIM, aj keď jeho funkčná stránka nie je podstatou môjho príspevku. Obrázok 1

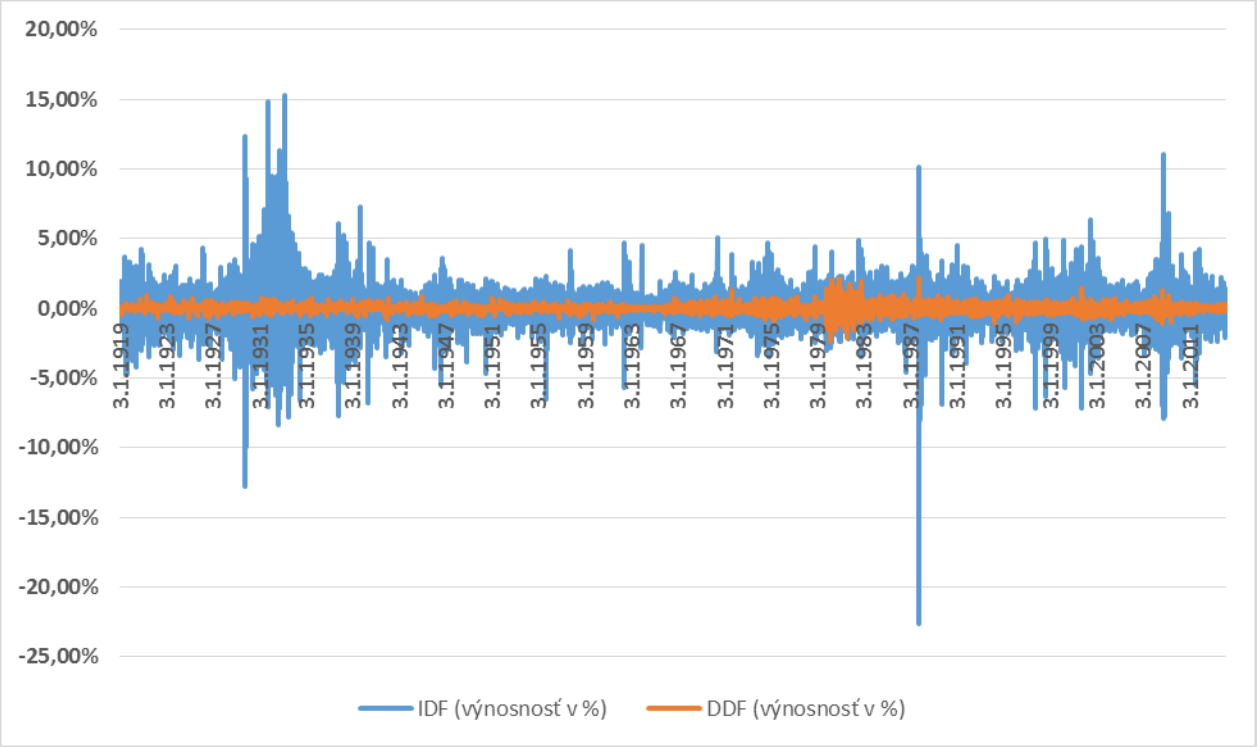
poskytuje grafické znázornenie vývoja celoživotných príjmových funkcií jednotlivých agentov od 25 roku do 62 roku veku života.



Obr. 22: Modelované celoživotné príjmové funkcie vybraných 9 sporiteľov

Virdzek (2012) vo svojej práci navrhol metodiku zostrojenia dlhopisového a indexového dôchodkového fondu a pre každý fond aj aktuálnu hodnotu dôchodkovej jednotky (AHDJ). Pri indexovom dôchodkovom fonde (IDF) ide o kopírovanie indexu Dow Jones Industrial Average (DJIA) a pri dlhopisovom dôchodkovom fonde (DDF) vývoj výnosov 20 ročných dlhopisov USA. Historickú výnosnosť IDF a DDF prezentuje graf na obrázku 2. K tomuto kroku nás viedla skutočnosť neexistencie finančných nástrojov na Slovensku a v ostatných európskych krajinách s dostatočne dlhou históriou a likviditou ako to je pri nami zvolených podkladových nástrojoch. MikroSIM obsahuje aj časový rad inflácie meranej indexom spotrebiteľských cien (CPI) v USA. Hodnoty vyššie uvedených časových radov sú na dennej báze od januára 1919 do apríla 2014 (pri CPI sme mali prístup iba k mesačným dátam, ktoré sme jednoduchou lineárnou extrapoláciou pretransformovali na denné dáta). Z príslušného časového radu sme pre každú simuláciu vybrali denné hodnoty inflácie, o ktoré sme navyšovali hodnotu celoživotnej príjmovej funkcie sporiteľa v čase. Tento krok vyplýva z ekonomických poznatkov, ktoré hovoria o vplyve inflácie na rast mzdy.

Dôležitou súčasťou konštrukcie MikroSIM bolo vytvorenie systémových pravidiel, na ktorých bude fungovať dôchodkový systém. Pravidlá priamo korešpondujú so zákonným nastavením fungovania starobného dôchodkového sporenia na Slovensku v zmysle platnej legislatívy. V našom systéme uvažujeme o 18 % odvodoch do dôchodkového systému z vymeriavacieho základu sporiteľa, z ktorých 6 % bude odvedených do druhého (kapitalizačného) piliera starobného dôchodkového sporenia. Tu predpokladáme zjednodušený stav, ktorý nastane po postupnom navyšovaní príspevkov od roku 2017, a zároveň priemeruje zmeny v odvodoch počas posledných 10 rokov.



Obr. 23: Historická denná výnosnosť (v %) indexového dôchodkového fondu (IDF) a dlhopisového dôchodkového fondu (DDF) od 1919 do 2014

Vstupné podmienky pre modelovanie výkonnosti individuálnych sporivých stratégií v modeli MikroSIM sú nasledovné:

I. Sporiteľ môže presúvať úspory medzi fondmi maximálne raz za mesiac (aj keď legislatíva platná k dnešnému dňu umožňuje sporiteľom meniť pomer sporenia svojich úspor medzi dvoma fondmi každý deň),

II. Realizácia zmeny pomeru sporenia nastáva o T + 3 dní, t.j. 3. deň po zadaní pokynu na zmenu pomeru sporenia,

III. Zákonné obmedzenie prestupu medzi fondmi podľa Zákona č. 43/2004 Z. z. o starobnom dôchodkovom sporení § 92 ods. 1 (znenie účinné od 1.1.2015), ktoré ustanovuje, že sporiteľ musí mať v dlhopisovom garantovanom dôchodkovom fonde v prvý deň kalendárneho mesiaca, v ktorom dovŕši:

a) 52 rokov veku, najmenej 10 % čistej hodnoty svojho majetku,
b) 53 rokov veku, najmenej 20 % čistej hodnoty svojho majetku,
c) 54 rokov veku, najmenej 30 % čistej hodnoty svojho majetku,
d) 55 rokov veku, najmenej 40 % čistej hodnoty svojho majetku,
e) 56 rokov veku, najmenej 50 % čistej hodnoty svojho majetku,
f) 57 rokov veku, najmenej 60 % čistej hodnoty svojho majetku,
g) 58 rokov veku, najmenej 70 % čistej hodnoty svojho majetku,
h) 59 rokov veku, najmenej 80 % čistej hodnoty svojho majetku,
i) 60 rokov veku, najmenej 90 % čistej hodnoty svojho majetku,
j) 61 rokov veku, 100 % čistej hodnoty svojho majetku.

IV. V tomto štádiu sme abstrahovali pri vytváraní modelu od vplyvu poplatkov, pričom v realite majú výrazný vplyv na zníženie výkonnosti akejkoľvek investičnej stratégie Mešárová (2014). S realizáciou aktívnej investičnej stratégie nie sú spojené žiadne priame finančné náklady v podobe poplatkov, pretože prestup medzi fondmi u toho istého správcu nie je zaťažený osobitnými poplatkami.

3. Vytvorenie sporiacej stratégie

Pri vytváraní sporiacich stratégií som rešpektoval diferencované prístupy k investovaniu v penzijných schémach, akými sú napríklad Mertonova-Samuelsonova celoživotná sporiaca stratégia alebo stratégia, ktorú uvádza Sharpe (2011), založená na dynamickej optimalizácii portfólia prostredníctvom dynamicky sa meniacej averzie k riziku. Uvedené prístupy ma viedli k vytvoreniu viacerých sporiacich stratégií, ktoré môže sporiteľ uplatňovať pri správe svojich úspor v penzijných schémach.. V práci budem sledovať uplatnenie stratégie vyvinutej na základe porovnania výnosnosti jednotlivých fondov v krátkom časovom období. Stratégia má názov SONYA.

Stratégia sa a priori rozhoduje o alokácii 100 % (0 %) alebo 0 % (100 %) úspor do IDF (DDF). Rozhodovací mechanizmus je

$$\sum_{t=1}^{n} \frac{(IDF_{AHDJ_t} - IDF_{AHDJ_{t-1}})}{IDF_{AHDJ_{t-1}}} > \sum_{t=1}^{m} \frac{(DDF_{AHDJ_t} - DDF_{AHDJ_{t-1}})}{DDF_{AHDJ_{t-1}}} \quad (1)$$

kde

IDF AHDJ - aktuálny kurz dôchodkovej jednotky(AHDJ) indexového dôchodkového fondu (IDF),
DDF AHDJ - aktuálny kurz dôchodkovej jednotky (AHDJ) dlhopisového fondu (IDF),
n – počet dní, pre ktoré kalkulujem kumulatívny výnos v IDF dozadu od dnešného dňa,
m – počet dní, pre ktoré kalkulujem kumulatívny výnos v DDF dozadu od dnešného dňa,
t – je deň poslednej známej AHDJ (v našom modeli je to AHDJ z pred troch dní).

Ak je ľavá strana nerovnice väčšia ako pravá, sporiteľ bude alokovať 100 % svojich úspor do IDF, ak je ľavá strana menšia ako pravá, sporiteľ bude alokovať 100 % svojich úspor do DDF, pričom budem rešpektovať zákonné obmedzenie prestupu medzi fondmi uvedené v predchádzajúcej časti.

4. Komparácia aktívneho a pasívneho prístupu k riadeniu úspor v MikroSIM

Prostredníctvom MikroSIM modelu som uskutočnil 55 simulácií sporiteľa s dosiahnutým II. stupňom vysokoškolského vzdelania (7. agent v našom MikroSIM). Pre každú simuláciu bude dodržaný 40 ročný sporivý horizont, v ktorom simulujeme príjem sporiteľa vo veku od 22 rokov do 61 rokov. Aktuálna hodnota dôchodkovej jednotky (AHDJ) na začiatku každej simulácie bude vždy nastavená na hodnotu 0,3000 €, pričom vývoj AHDJ pre príslušný fond bude determinovaný historickou výnosnosťou finančného nástroja uvedeného v predchádzajúcej časti príspevku. Pre každú simuláciu sme s historických dát použili 40 rokov dlhý spojitý časový rad pre IDF, DDF a aj infláciu. Sporiteľ do systému pre potreby simulácií vstupuje po roku, t.j. prvá simulácia začína simulovaním výnosnosti fondov z historických dát od roku 1919 do 1958, druhá simulácia začína simulovaním výnosnosti fondov z historických dát od 1920 do 1959, a posledná simulácia od roku 1974 do roku 2014.

Simulácie poskytnú sporiteľovi rôznu výšku naakumulovaných úspor v čase pri aktívnom respektíve pasívnom prístupe k sporeniu.

Pre lepšiu prezentáciu výsledkov dávam do pomeru konečnú hodnotu úspor pri aktívnom, respektíve pasívnom, prístupe k sporeniu k celkovým zaplateným príspevkom (2), v ktorom hodnotím výlučne iba o koľkonásobne väčšie zhodnotenie by sme dosiahli pri aktívnom, respektíve pasívnom, prístupe k sporeniu a následne overujem predpoklad uvedený v úvode práce tým, že dávam do pomeru hodnotu úspor (ODÚ) dosiahnutú pri uplatňovaní stratégie SONYA k hodnote úspor pri pasívnom sporení do dlhopisového dôchodkového fondu (3).

$$K_t = \frac{ODÚ_{DDF_t}}{ODÚ_{\sum príspevky_t}}, t \in < 1919{,}2014 > \quad (2)$$

$$K_t = \frac{ODÚ_{Sonya_t}}{ODÚ_{\sum príspevky_t}}, t \in < 1919{,}2014 > \quad (3)$$

$$K_t = \frac{ODÚ_{Sonya_t}}{ODÚ_{DDF_t}}, t \in < 1919{,}2014 > \quad (4)$$

kde

ODÚ – hodnota úspor dosiahnutá v čase t stratégiou SONYA respektíve hodnotu úspor dosiahnutú sporením v dlhopisovom dôchodkovom fonde,

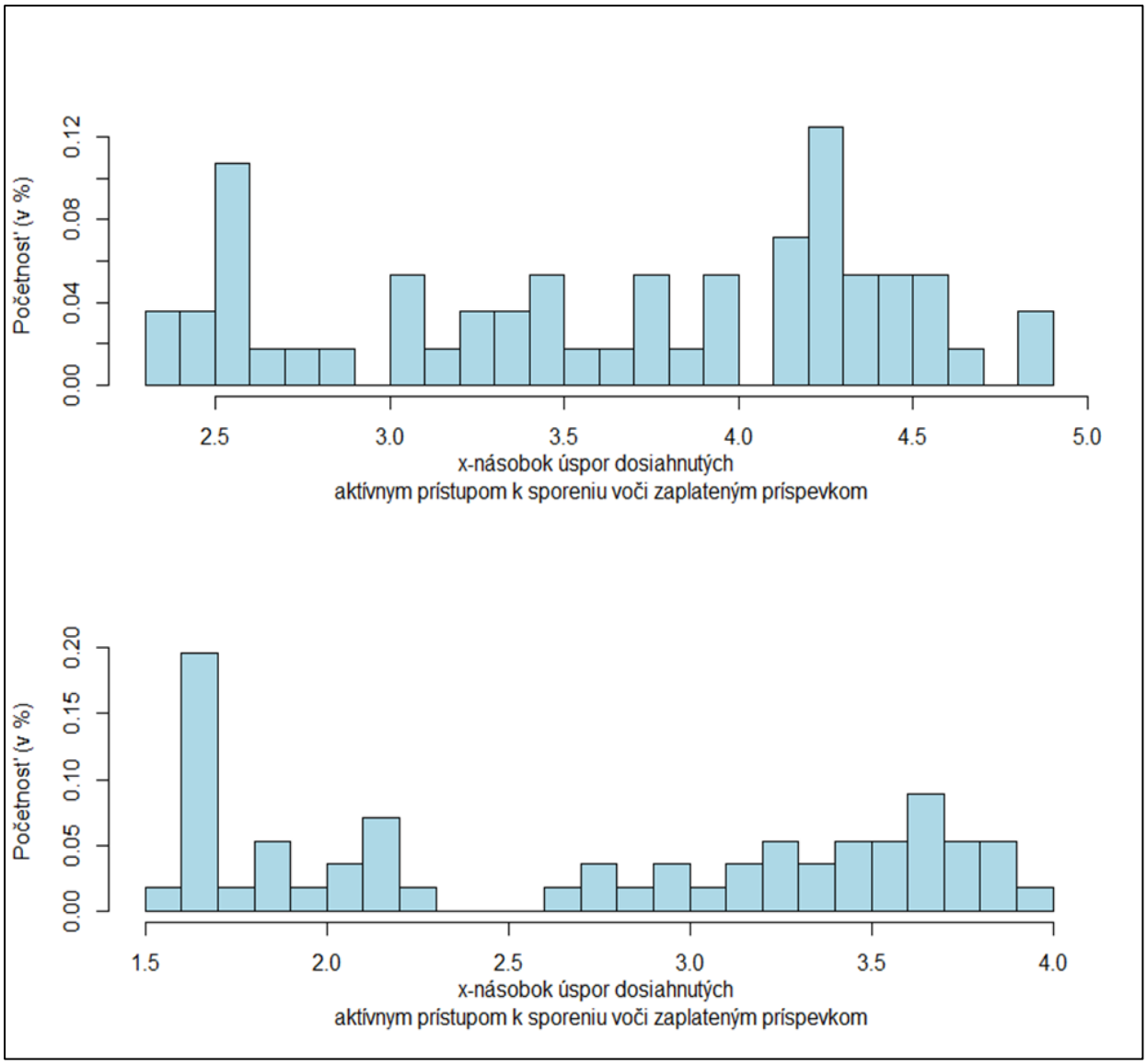
$\sum$ príspevky – suma zaplatených príspevkov na osobný dôchodkový účet vedený v dôchodkovej správcovskej spoločnosti počas celej doby sporenia,

Výsledky prezentujem prostredníctvom koeficientu, ktorý môžem interpretovať nasledovne :

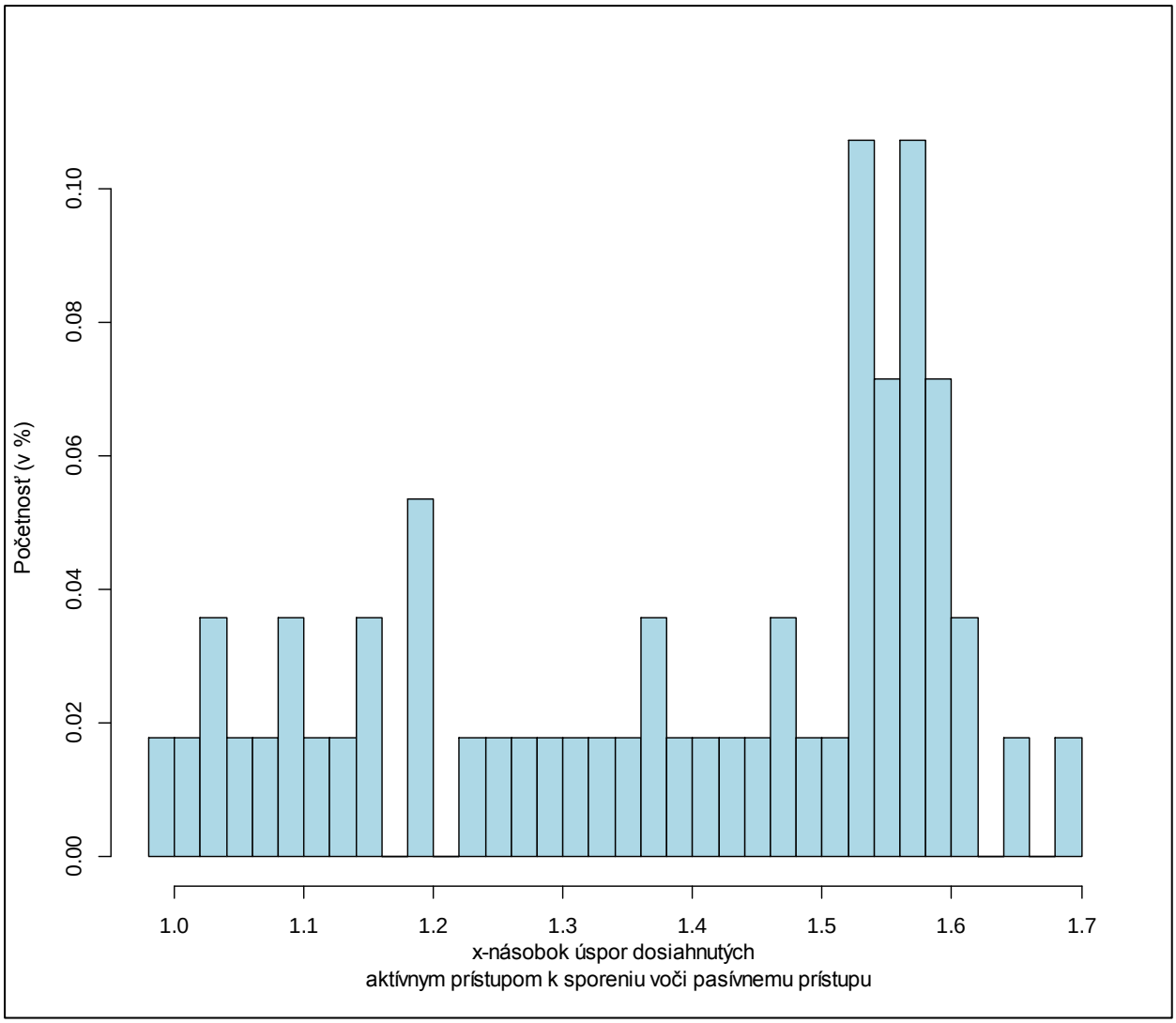
- Ak hodnota $K_t > 1$, aktívnym prístupom k úsporám sporiteľ dosiahol vyššiu hodnotu úspor ako pri pasívnom sporení iba do DDF,
- Ak hodnota $K_t < 1$, aktívny prístup k úsporám neprinesie sporiteľovi vyššiu nasporenú sumu ako pasívne sporenie do DDF,
- Ak hodnota $K_t = 1$, aktívnym prístupom k úsporám dosiahne sporiteľa rovnakú nasporenú sumu ako pri pasívnom sporení.

Na obrázku 3 zobrazujem dva histogramy, ktoré prezentujú o koľkonásobne viac zhodnotil sporiteľ svoje úspory pri uplatňovaní aktívneho prístupu k sporeniu, respektíve pri uplatňovaní pasívneho prístupu k sporeniu. Z výsledkov prezentovaných na obrázku 3 vyplýva, že v prípade aktívneho prístupu ku sporeniu sporiteľ dokázal zhodnotiť zaplatené príspevky do druhého piliera od 150 % (aktívny prístup mu priniesol 1,5 násobok zaplatených príspevkov) do 500 %. Pasívny prístup k sporeniu priniesol sporiteľovi zhodnotenie zaplatených príspevkov na úrovni 1,5 násobku respektíve 4 násobku zaplatených príspevkov, čo predstavovalo minimálne 50% zhodnotenie a maximálne 400 % zhodnotenie.

Obrázok 4 sumarizuje výsledky simulácií, prostredníctvom ktorých overujeme predpoklad, či aktívnym prístupom k sporeniu a zvolenou sporiacou stratégiou sporiteľ dosiahne vyššiu hodnotu úspor ako v prípade pasívneho prístupu. Môžeme konštatovať, že zvolená stratégia zabezpečila vo všetkých prípadoch vždy rovnakú alebo vyššiu hodnotou úspor ako pri pasívnom prístupe. Výsledky výskumu naznačujú, že v približne 3,5 % prípadov sporiteľ mohol dosiahnuť aktívnym prístupom rovnakú konečnú hodnotu úspor ako pri pasívnom prístupe, pričom vo zvyšných
96,5 % prípadov dosiahol relatívne vyššiu konečnú hodnotu úspor.



Obr. 24: Porovnanie nasporenia x-násobku úspor aktívnym prístupom a pasívnym prístupom k investovaniu voči zaplateným príspevkom.



Obr. 25: Nasporenie x-násobku úspor aktívnym prístupom k sporeniu voči pasívnemu sporeniu v dlhopisovom dôchodkovom fonde

5. Záver

Cieľom môjho výskumu bolo overenie predpokladu, či si sporiteľ aktívnym prístupom k svojim úsporám prostredníctvom zvolenej stratégie a signálov na prestup medzi indexovým dôchodkovým fondom a dlhopisovým dôchodkovým fondom dokáže zabezpečiť vyššie zhodnotenie svojich úspor ako v prípade pasívneho prístupu. Predpoklad z úvodu práce sa potvrdil. Z výsledkov môžem konštatovať, že v 96,5 % prípadoch dokázala zvolená stratégia dosiahnuť vyššiu hodnotu úspor ako v prípade pasívneho prístupu znamenajúceho sporenie výlučne v dlhopisovom dôchodkovom fonde, pričom vo všetkých pozorovaniach vytvorená sporiaca stratégia nedosiahla pre sporiteľa nižšiu hodnotu úspor ako sporenie v dlhopisovom dôchodkovom fonde. Aktivita v podobe prestupovania medzi fondmi na základe zvolenej sporiacej stratégie priniesla sporiteľovi vo väčšine prípadov vyššiu hodnotu úspor ako inertné správanie znamenajúce sporenie si v dlhopisovom fonde.

V ďalšej časti výskumu sa sústredím na dopracovanie súčastí modelu MikroSIM o rôzne determinanty vstupujúce do života jednotlivca, napr. nezamestnanosť, ochorenie, materská dovolenka a podobne. Pokračovať budem v modelovaní celoživotnej investičnej stratégie, prostredníctvom ktorej dokážem prepojiť akumulačnú fázu s výplatnou (deakumulačnou) fázou a v neposlednom rade budem venovať pozornosť testovaniu rôznych sporiacich stratégií, na ktorých výskume v súčasnosti pracujem.

Literatúra

BALÁŽ V. 2009. *Riziko a neistota – úvod do behaviorálnej ekonómie a financií.* Bratislava: Veda. ISBN 978-80-224-1082-3.

EVENSKY, H. – HORAN, M. S. – ROBINSON, R. T. 2011. *The New Wealth Management*. JohnWilley & Sons, Inc.ISBN 978-0-470-62400-5.

MELICHERČÍK, I. – ŠEVČOVIČ, D. 2012. Dynamic Model of Pension Savings Management with Stochastic Interest Rates and Stock Returns. In: *Mathematical and Statistical Methods for Actuarial Sciences and Finance. Springer*. ISBN 978-3-540-71296-1.

MEŠÁROVÁ, N. 2014. Strata úžitku sporiteľa v dôsledku poplatkovej politiky. In: *Sporenie & Investovanie* 3/2014. ISSN 1337-9704, s. 29-32.

SHARPE, W. 2011. Post-Retirement Financial Strategies: Forecasts and Valuation. In: *European Financial Management Association Annual Meeting*. Dostupné na internete: http://web.stanford.edu/~wfsharpe/retecon/PRS_20111001.pdf

ŠEBO, J. - VIRDZEK, T. 2013. Dismantling the myths about the pension funds' performance from the saver´s perspective. In: *Szczepanski, M. (ed.) "Pension Reforms – Comparison and Evaluation?"* . Poznan : Publishing House of Poznan University of Technolgy. pp. 109-123. ISBN 978-83-7775-280-7.

ŠEBO, J. - ŠEBOVÁ, Ľ. - VIRDZEK, T. 2014. Challenges in Slovak PAYG and DC schemes. In: *Szczepański-Brzeczek-Gajowiak (ed.). Social Security systems - against the challenges of demographics and market.* Poznan : Publishing House of Poznan University of Technology. 2014. ISBN 978-83-7775-344-6.

ŠEBO, J. - ŠEBOVÁ, Ľ. 2013. Pension Schemes as Complex Socio-Economic Dynamic Systems. In: *Journal of Economics and Social Research.* No. 2/2013. ISSN 1335-7069.

VIRDZEK, T. 2012. *Výkonnosť kapitalizačných systémov dôchodkového sporenia*. Banská Bystrica: Ekonomická fakulta UMB – Dizertačná práca: 0d50d1ac-3b63-4c5e-93d3-f8bb35f5a316

Adresa autora:

Michal Mešťan, Bc.
Ekonomická fakulta
Univerzita Mateja Bela
Tajovského 10, 975 90 Banská Bystrica
1. ročník, 2. stupeň
mestan.michal@gmail.com

Vzájomná súvislosť a predikcia vývoja miery inflácie a nezamestnanosti v podmienkach Slovenskej republiky

Mutual relation and prediction of the inflation rate and the unemployment rate in the conditions of Slovak Republic

Lukáš Ries

Abstract: In this paper I present the results of analysis causal relations between inflation rate and unemployment rate followed by verifying validation of Phillips curve in conditions of Slovakia in time 1997-2014. Also I predict the variables using by vector autoregression model for the next four quarterly time period, i use impulse response function, forecast error variance decomposition for variables. I visualize the time series using by chosen graphical components. The value added of this paper consist of description causal relations and prediction inflation rate and unemployment rate based on using VAR model.

Abstrakt: V tomto príspevku prezentujem výsledky analýzy vzájomnej súvislosti miery inflácie a nezamestnanosti s následným overením validity Phillipsovej krivky pre podmienky Slovenskej republiky za sledované obdobie 1997-2014. Taktiež predikujem dané premenné pomocou modelu vektorovej autoregresie na najbližšie štyri kvartálne obdobia, využívam funkciu reakcie impulzov, dekompozíciu rozptylu chyby prognózy pre premenné. Časové rady vizualizujem vybranými grafickými prvkami. Pridaná hodnota príspevku spočíva v deskripcii grangerovských kauzálnych väzieb a prognóze inflácie a nezamestnanosti pre podmienky Slovenskej republiky.

Key words: Inflation, unemployment, Phillips curve, vector autoregression, Granger simply causality, Granger instantaneous causality

Kľúčové slová: Inflácia, nezamestnanosť, Phillipsova krivka, vektorová autoregresia, Grangerova jednoduchá kauzalita, Grangerova okamihová kauzalita

Úvod

Náplňou príspevku je deskripcia súvislostí a grangerových kauzálnych väzieb v časovom horizonte rokov 1997 – 2014 inflácie a nezamestnanosti s cieľom následnej validácie platnosti Philipsovej krivky pre Slovenskú republiku.

V príspevku využívam pracovné prostredie štatistického programu R verzie 3.1.2. „Pumpkin Helmet“. (R Development Core Team, 2014). Pracujem s knižnicami *aplpack* (Wolf, 2014), *car* (Fox, Weisberg, 2011), *forecast* (Hyndman, Rasbash, Schmidt, 2012), *MASS* (Ripley, 2002), *MSBVAR* (Brandt, Davis, 2014), *MVN* (Korkmaz, 2014), *lmtest* (Hothorn, Zeichles, Farebrother, Cummins, Millo, 2014), *tcltk* (Grosjean, 2014), *tseries* (Trapletti, Hornik, LeBaron, 2013), *urca* (Pfaff, 2011), *vars* (Pfaff, 2008).

Modelujem sledované premenné pomocou vektorovej autoregresie (ďalej len VAR), ktorá je viacrovnicovým modelom, v ktorom sú všetky premenné považované za endogénne, a každá z nich je vysvetľovaná zvyšnými premennými oneskorenými v čase o určitú hodnotu - *p* (rád vektorovej autoregresie). Takýto model sa potom označuje VAR(p). Následne predikujem vývoj daných časových radov pomocou VAR s analýzou funkcií reakcií na impulz a dekompozíciou rozptylu chyby predikcie pre dané časové rady. Overujem taktiež kauzálne väzby v grangerovskom zmysle (okamihová a jednoduchá kauzalita).

Taktiež stručne popíšem podstatu Phillipsovej krivky ako súvislosť inflácie a nezamestnanosti, súčasné prístupy k modelovaniu Phillipsovej krivky. Pre odporúčanú a

maximálnu dĺžku príspevku neuvádzam samozrejme všetky grafické i štatistické výstupy ale len tie, ktoré nutné súvisia s vizualizovaním či doložením najvýznamnejších skutočností.

Skúmané časové rady pochádza z údajov z makroekonomickej databázy Národnej banky Slovenska. Jedná sa o sezónne očistené dáta, v kvartálnej periodicite a to od 1. kvartálu 1997 do 3. kvartálu 2014, sumárne 71 pozorovaní [cit. 16. 11. 2014]. Dostupné na http://www.nbs.sk/sk/menova-politika/makroekonomicka-databaza). Prvá premenná je miera inflácie, ktorú som vypočítal nasledovne:

$$\pi_t = 100 * \left(\frac{HICP_t}{HICP_{t-1}} - 1 \right) \qquad (1)$$

kde HICP je harmonizovaný index spotrebiteľských cien všetkých položiek sezónne očistených. Druhá premenná je miera nezamestnanosti vypočítaná Úradom práce, sociálnych vecí a rodiny ako nezamestnanosť z celkového počtu uchádzačov o pracovné miesto ktoré sú sezónne očistené.

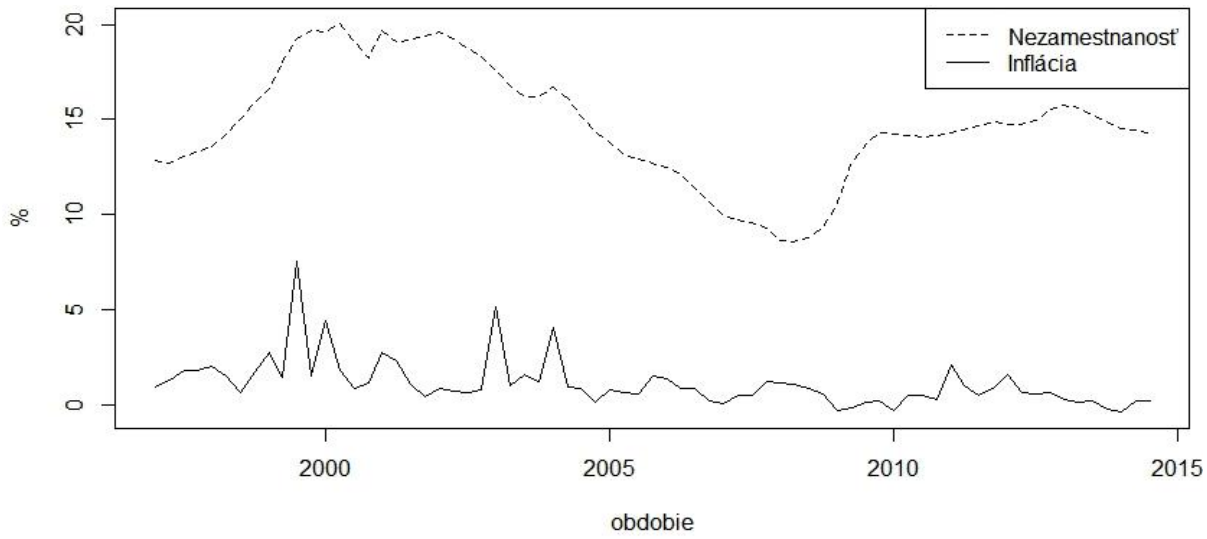
Náčrt do teórie Phillipsovej krivky

Phillipsova krivka v ekonomickej teórií vysvetľuje vzťah medzi infláciou a nezamestnanosťou asymetrickým substitučným klesajúcim inverzným exponenciálnym charakterom. Pôvodná Phillipsova krivka (1958, A. W. Phillips,) avšak zobrazuje vzťah medzi nezamestnanosťou a miernou zmeny nominálnych miezd pre pozorovania v spojenom kráľovstve (1861-1957). Neskôr však neokeynesovci modifikovali krivku o doplnenie vzťahom inflácie. Mala vyjadrovať substitučnú teóriu inflácie (určitá zmena inflácie sa „vykompenzuje-kúpi“ určitou zmenou nezamestnanosti). Za predpokladu miery zmeny zotrvačnej inflácie dôjde k posunu krivky v dôsledku adaptívnych očakávaní ekonomických subjektov (posun červená krivka obrázku 1). Platnosť tohto predpokladu sa potvrdila len čiastočne, následne nastala kritika zo strany monetaristov (M. Friedman), a následná vlastná teória tejto krivky. Friedman predpokladá vertikálny tvar krivky v dlhom období, nemôže existovať permanentná substitúcia inflácie a nezamestnanosti. V dlhom období len prirodzená miera nezamestnanosti je zlučiteľná so stabilnou infláciou.

Grafická analýza súvislostí a časového priebehu skúmaných premenných

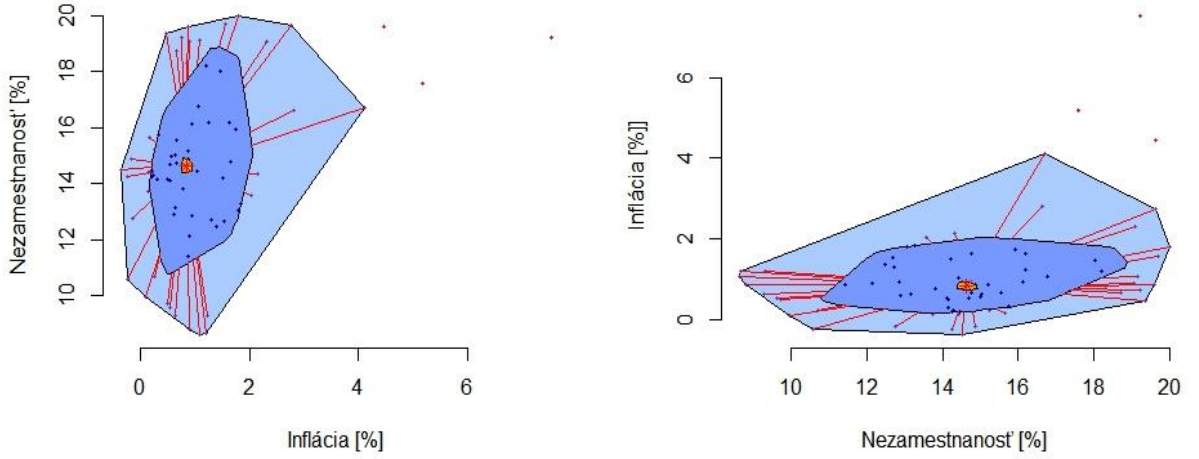
Ako prvý krok pri analýze dátovej vzorky som vykonal vizualizovanie dát pomocou grafu vývoja časového radu inflácie a nezamestnanosti zobrazený obrázkom 1. Na prvý pohľad je jasná nestacionarita časového radu nezamestnanosť, inflácia sa javí ako stacionárny heteroskedastický proces s dlhodobo klesajúcim trendom. Daný časový rad možno považovať za homoskedastický najmä za posledné obdobia cca od roku 2004, čo je spôsobené vplyvom vstupu do Európskej únie a následné dodržiavanie stanovenej hladiny inflácie.

Inflácia bola v celom sledovanom období v piatich prípadoch záporná (kvartálne zmeny boli analogicky záporné). Modelujem a analyzujem celé skúmané obdobie ako celok, na základe prvotného grafického nahliadnutia na priebeh časových radov by stálo za úvahu modelovať dané časové rady od roku 2009 po súčasnosť (majme ale na pamäti príliš málo pozorovaní).



Obr. 1: Časový priebeh inflácie a nezamestnanosti

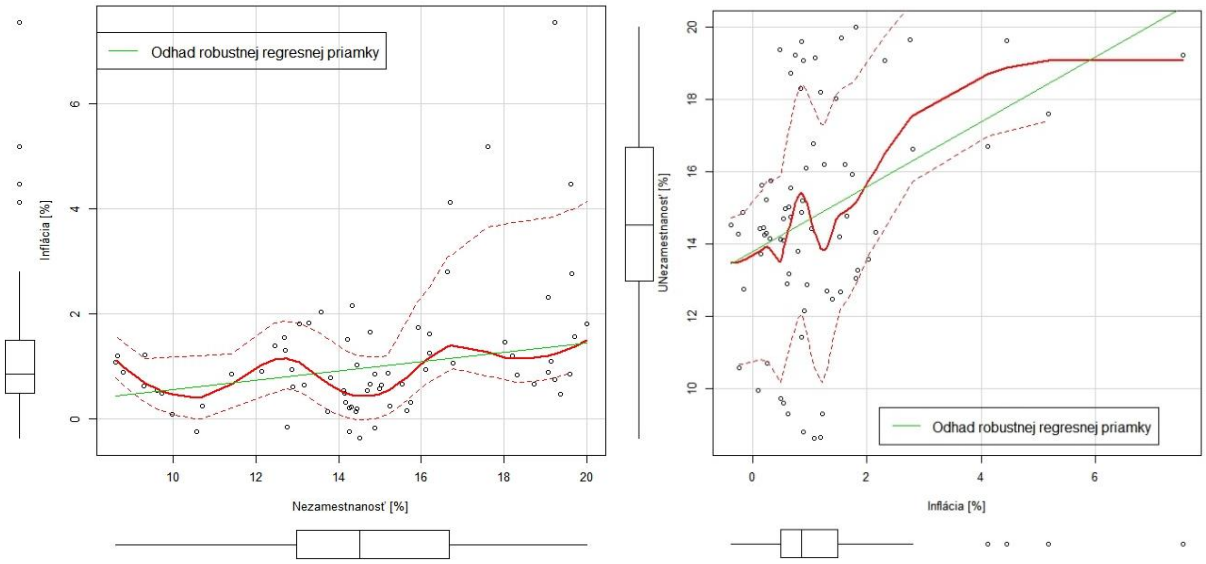
Na obrázku 2 je možné vidieť bagploty daných premenných avšak s rozdielnou orientáciou v grafe – ako závislú premennú som dátové vzorky navzájom vymenil a porovnal. Tmavším farebným odtieňom v grafe je cca 50 % hodnôt, zvyšných 50 % je označených svetlejším odtieňom. Pozorovania nezahrnuté do útvaru sú označené ako outliery. Takmer všetky pozorovania inflácie sú z rozmedzia 0 - 2 %, s pomerne viditeľným takmer „vertikálnym útvarom" dátovej vzorky, oproti priebehu nezamestnanosti ako nezávislej premennej, kde je priebeh pomerne rovnomerne rozptýlený v danom rozsahu inflácie (0 – 2 %). Inflácia je absolútne necitlivá na zmeny nezamestnanosti, čo sa potvrdí aj v ďalšom grafe. Dané prvotné úvahy z grafu sa môžu javiť mylne, nakoľko treba brať v úvahu pomer mierok na grafoch, avšak považujem grafickú analýzu a vizualizovanie vzoriek za primárnu analýzu a pochopenie súvislostí medzi danými premennými.



Obr. 2: Bagploty skúmaných premenných

Na obrázku 3 je možné vidieť scatterploty dátových vzoriek, kde som taktiež komparoval infláciu a nezamestnanosť separátne v 2 grafoch. Na základe prvotného náhľadu na dáta zhruba štyri pozorovania sa javia ako outliery (v prípade scatterplotu naľavo) ale v prípade výmeny premenných ako závislej a nezávislej možno považovať dané pozorovania za vplyvné (obrázok vpravo, influential observations). Rozdiely v sklone priamok odhadnutej regresnej robustnej a nerobustnej alternatívy boli veľmi minimálne. Táto skutočnosť ma pomerne prekvapila. Viditeľné zakrivenie je možné vidieť len na neparametrickom odhade trendovej krivky dátovej koncentrácie vyobrazený plnou hrubšou červenou čiarou s následnými

prerušenými čiarami zobrazujúci interval spoľahlivosti, t.j. pri akých hodnotách nezávislej premennej zoskupené dáta tvoria akýsi zhluk. Následne sa daná krivka „zakriví“ v smere tohto zhluku, resp. zhusteného výskytu pozorovaní.



Obr. 3: Scatter ploty skúmaných premenných s odhadom robustnej regresnej priamky

V prípade závislej premennej inflácia (obrázok vľavo) je sklon robustnej regresnej priamky len veľmi mierny a súvislosť medzi danými premennými je veľmi nízka. V tomto prípade je trend regresnej priamky pomerne očakávaný, t.j. mierny nárast, resp. očakával som že dané outlieri „vychýlia“ robustnú regresnú priamku v smere polohy outlierov. V prípade závislej premennej nezamestnanosť (obrázok vpravo) je viditeľný takmer vertikálny zhluk dát, ktorý „narušuje“ 4 outliery. Regresná priamka je o niečo „menej strmšia“ ako boli moje očakávania. Preložená odhadnutá robustná regresná priamka naznačuje priamu súvislosť. Tieto predbežné úvahy o sklone regresnej priamky a koncentrácií dát do určitých tvarov sa môžu javiť ako nepresné, nakoľko je rozdiel v osiach pomerne rôzny a prispôsobením osí by sa tieto úvahy možno zmenili.

V komparácií s ekonomickou teóriou pôvodnej Phillipsovej krivky (1958) sa vizualizovaním dátovej vzorky daný predpoklad nepreukázal (s rastúcou nezamestnanosťou by mala inflácia asymetricky klesať). V tomto prípade je inflácia takmer necitlivá na zmeny nezamestnanosti, t. j. validácia platnosti Philipsovej krivky sa nepotvrdila (uvažujem o pôvodnej krivke modifikovanú neokeynesovcami, kde platí asymetrický substitučný vzťah nezamestnanosť – inflácia).

Konštrukcia odhadu Phillipsovej krivky a predikcia skúmaných premenných

Následne zostrojím odhad VAR pre Phillipsovu krivku, overím grangerovské kauzálne väzby aj s testom na okamihový variant, predikciu pomocou rekurzívnej predikcie, ktorá využíva predikčnú chybu kovariančnej matice, grafické vizualizovanie reakcií na impulz (IRF), a tabuľkový a grafický výstup dekompozície rozptylu chyby prognózy (FEVD).

Prvým krokom k testu bolo overenie stacionarity dátovej vzorky (využil som kombináciu ADF a KPSS testu, taktiež vizualizácia pôvodných, 1x diferencovaných a 2x diferencovaných hodnôt spolu s výstupmi ACF a PACF). Časový rad inflácie nemá dôvod nebyť stacionárnym procesom, aj následné spomínané spôsoby validácie stacionarity potvrdili daný predpoklad. Časový rad nezamestnanosti možno považovať za I(1) proces (integrovaný rádu 1). Dané

časové rady možno považovať aj za kointegrované (na základe Phillips - Ouliaris testu ale aj Engel - Granger testovacej procedúre na overenie kointegrácie radov).

Pre overenie grangerovej kauzality v daných prípadoch testovania pracujem s pôvodnými hodnotami inflácie a prvými diferenciami nezamestnanosti. Využívam funkciu na testovanie jednoduchej a okamihovej grangerovej kauzality pre model VAR, ale testujem aj jednoduchú grangerovu kauzalitu len samotných časových radov inflácie a prvých diferencií nezamestnanosti (pomocou inej funkcie ako pre overenie kauzalít pre VAR). Overením kauzálnych väzieb odhadu VAR s konštantou sa nepodarilo jednoznačne potvrdiť daný typ kauzality pre rôzne rády oneskorenia (od 2 do 7, pri ráde 8 nezamietame predpoklad o žiadnom grangerovskom zmysle aj to veľmi tesne na 10 % hladine významnosti s využitím bootstrapovej simulácie s počtom 100 000). Pri následnom overovaní jednoduchej kauzality iba samotných časových radov je preukázateľný vplyv premennej inflácia na prvú diferenciu nezamestnanosti v grangerovskom zmysle až pri ráde 5 (na 10 % hladine významnosti) a pri ráde 6 (už na 5 % hladine významnosti), pri ráde 7 (aj na 1 % hladine významnosti). Treba mať na pamäti počet pozorovaní (70) a znižovanie stupňov voľnosti pre danú testovaciu štatistiku.

Pri voľbe rádu vektorovej autoregresie som sa opieral o 3 informačné kritériá (Akaikeho, Schwarzovo a Hannovo - Quinnovo) a o výslednú strednú kvadratickú chybu predikcie (FPE). Tieto kritériá v snahe o ich minimalizáciu za predpokladu združeného gaussovského rozdelenia inovácií viedli k voľbe oneskorenia 3 obdobia do minulosti. Odhad VAR (3) modelu má po prepise nasledovnou symbolikou nezamestnanosť = U , a infláciu = π tvar:

$$\hat{\pi} = -0{,}68138 + 0{,}08869 * \pi_{t-1} + 0{,}67203 * U_{t-1} + 0{,}34889 * \pi_{t-2} - 1{,}05344 * U_{t-2} - 0{,}00533 * \pi_{t-3} + 0{,}46955 * U_{t-3} \tag{4}$$

$$\hat{U} = 0{,}6025 - 0{,}05374 * \pi_{t-1} + 1{,}54599 * U_{t-1} - 0{,}01434 * \pi_{t-2} - 0{,}48037 * U_{t-2} + 0{,}11512 * \pi_{t-3} - 0{,}11019 * U_{t-3} \tag{5}$$

Uvedené odhady VAR(3) predstavujú odhad v aktuálnom čase, resp. v krátkodobom horizonte (short run). Daný odhad sa dá prepísať za určitých podmienok (predpoklad stacionarity modelu, predpoklad blížiacich sa veličín k svojim dlhodobým rovnovážnym ekvilibristickým úrovniam,...). Následne možno uvažovať odhad inflácie VAR (ako vzťah Phillipsovej krivky - inflácia ako závislá premenná) nasledovne:

$$\hat{\pi}_e = -0{,}68138 + 0{,}08869 * \pi_e + 0{,}67203 * U_e + 0{,}34889 * \pi_e - 1{,}05344 * U_e - 0{,}00533 * \pi_e + 0{,}46955 * U_e \tag{6}$$

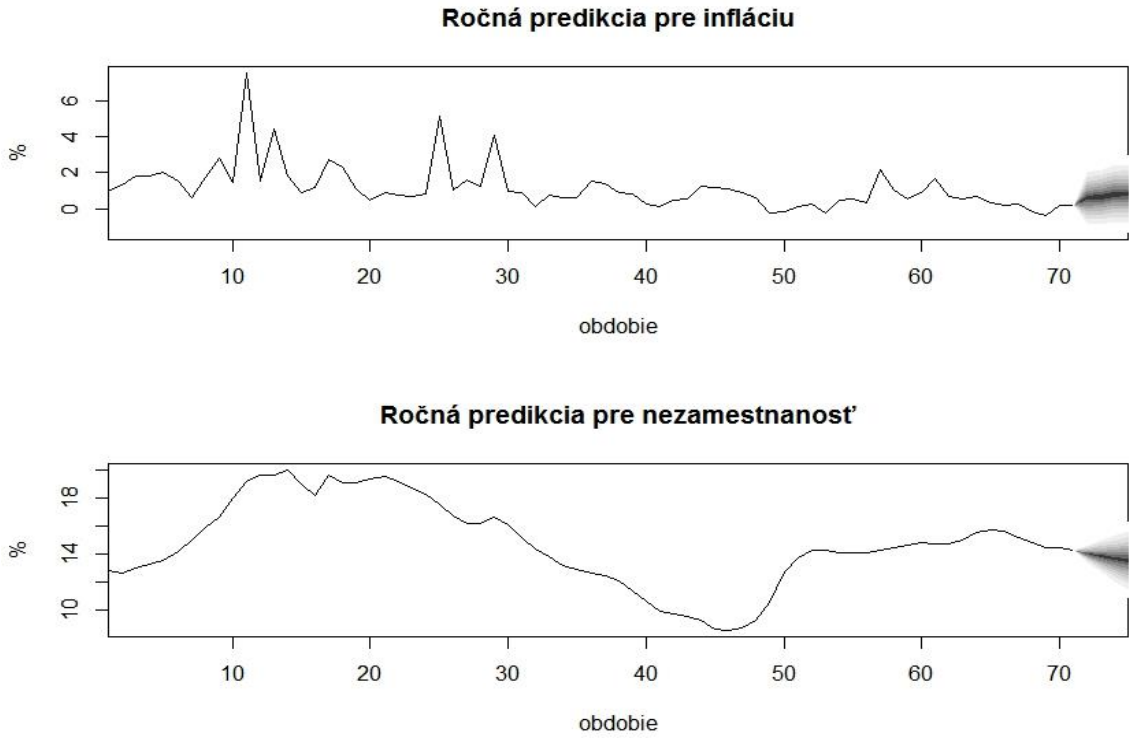
kde π_e a U_e sú hodnoty inflácie a nezamestnanosti v ekvilibriu (dlhodobá rovnovážna úroveň, long run). Následnými úpravami rovnice (6) dostávane dlhodobý tvar Phillipsovej krivky:

$$\hat{\pi}_e = -0{,}24306 + 0{,}155244 * U_e \tag{7}$$

Analogicky úpravou rovnice (5) VAR(3) modelu dostávame vzťah súvislosti nezamestnanosti od inflácie:

$$\hat{U}_e = 13{,}5181 + 1{,}05542 * \pi_e \tag{8}$$

Pri testovaní odhadnutého VAR(3) na stabilitu parametrov, a stabilitu modelu ako celku, som našiel korene charakteristického polynómu. Model je stabilný ako celok, nemá jednotkový. Následné CUSUM, MOSUM, RE a ME testy nepreukázali „defekty“ v stabilite parametrov VAR(3). Model možno považovať za stabilný vo svojich parametroch.



Obr. 4: Fancharty pre ročnú predikciu VAR s konfidenčnými pásmi spoľahlivosti 90 % (šedý odtieň)

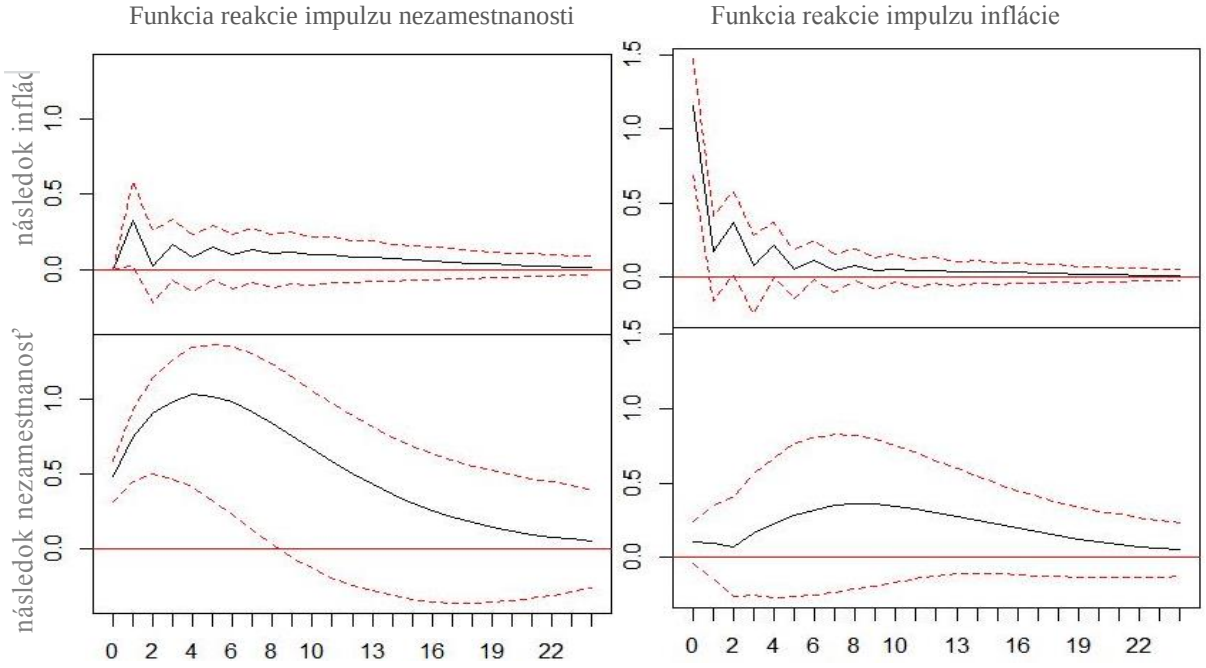
Rekurzívnu predikciu so zohľadnením predikčnej neistoty zobrazuje fanchart s 90 % intervalmi spoľahlivosti pre predikciu nasledujúcich štyroch kvartálov. Na obrázku 4 je možné vidieť pomerne vysokú chybu predikcie pri predikovaní inflácie. Možno usúdiť aj na základe „strmšieho a rýchlejšieho“ nárastu intervalu spoľahlivosti (zobrazený desiatimi stupňujúcimi sa odtieňmi šedej) predikovaných hodnôt už pri prvom kvartáli predikcie). Model predikuje veľmi mierne rastúce hodnoty inflácie pre nasledujúce obdobia. Možno usudzovať relatívne horšiu predikčnú schopnosť pre infláciu ako nezamestnanosť (čo vyplýva aj z grafického výstupu empirických a odhadovaných hodnôt pomocou VAR(3)) ale aj na základe fanchartu. Predikcie pre infláciu majú mierne rastúci trend (najtmavší odtieň predikcie) s vysokou variabilitou avšak cca od druhého kvartálu pomerne konštantným intervalom spoľahlivosti. Uvedené predikcie zobrazuje tabuľka 1.

Tab. 1: Predikcia hodnôt premenných pomocou VAR s 90 % intervalom spoľahlivosti

Predikcia pomocou VAR s intervalom spoľahlivosti				
kvartál	Inflácia [%]		Nezamestnanosť [%]	
	hodnota	90 % konfidenčný interval	hodnota	90 % konfidenčný interval
1	0,589	1,897	14,048	0,816
2	0,648	1,992	13,868	1,488
3	0,759	2,085	13,704	2,114
4	0,809	2,107	13,594	2,678

Náhodné inovácie (impulzy) jednotlivých premenných na iné premenné a samé seba som vyobrazil graficky (obrázok 5) pomocou funkcie reakcie premennej na impulz inej premennej vplyvom šokov v časovom horizonte 6 rokov s využitím bootstrapových konfidenčných pásov o spoľahlivosti 95 % o počte simulácií 10 000. Dané obrázky zachytávajú priebeh nekumulatívnej funkcie odoziev impulzov. Prerušované čiarkované= pásy predstavujú bootstrapový 95 % interval spoľahlivosti. Impulzné odozvy boli ortogonalizované. Je nutné zohľadniť nielen tvar funkcie reakcie impulzov premennej na danú premennú ale aj jej intenzitu vyjadrenú na vertikálnej osi.

Vrchné grafy predstavujú vplyv reakcií impulzov na infláciu, spodné grafy predstavujú vplyv reakcií impulzov na nezamestanosť. Grafy na ľavej strany sú reakcie impulzov nezamestnanosti, na pravej strane reakcie impulzov inflácie. V prípade reakcií impulzov nezamestnanosti na infláciu (ľavý vrchný graf) je vplyv najmenší v druhom oneskorení, následne sa vplyvy v párnych oneskoreniach vyskytujú v menšej miere ako v nepárnych, postupne sa však vplyv vytráca, zhruba v 4 roku je minimálny.

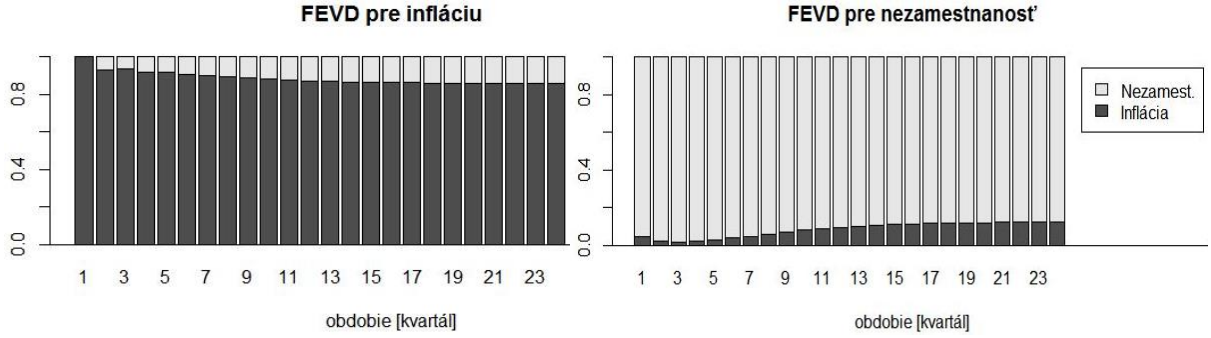


Obr. 5: Graf nekumulatívnych ortogonalizovaných funkcií reakcie impulzov nezamestnanosti a inflácie

V prípade vplyvu impulzov nezamestnanosti na samotnú premennú (ľavý spodný graf) je situácia pomerne „nežiadaná“, najsilnejší vplyv majú šoky z oneskorenia 5 kvartálu, následne polynomicky klesajú, efekt sa vytráca až v horizonte 5 rokov. V prípade efektu impulzov inflácie na samotnú premennú (pravý vrchný graf) je situácia vývoja „pomerne žiadaná“, t. j. neočakávané zmeny inflácie na samú seba sa relatívne „rýchlo stratia“, už po zhruba 5 kvartály je vplyv minimálny, najsilnejší je v prípade šoku oneskorenom v čase o jedno pozorovanie. Impulzy inflácie ma nezamestnanosť (pravý spodný graf) sú minimálne v prvom a druhom oneskorení, následne sa efekt stupňuje, je najsilnejší zhruba v 8 kvartály a následne mierne klesá. Efekt sa vytráca veľmi pomaly, stratí sa v horizonte zhruba 5 a viac rokov. Celkovo možno konštatovať, že šoky spôsobené samotnými premennými vplývajú na samé seba s rastúcim oneskorením klesajúco konvergentne k 0 (ak mierne abstrahujem v prípade šoku inflácie na nezamestnanosť opätovným sa „vrátením“ vplyvu šoku a následným poklesom). Šoky spôsobené premennými sa v čase vytrácajú, avšak šoky spôsobené

nezamestnanosťou na samotnú premennú sa vytrácajú pomalšie oproti vytrácaniu sa šokov spôsobených infláciou s rastúcimi oneskoreniami na samé seba (aj na základe kumulatívnych ortogonalizovaných funkcií reakcie impulzov). Priebeh kumulatívnych funkcií má konvergentný charakter vo všetkých prípadoch. Najsilnejší efekt šokov s následným pretrvávaním je v prípade nezamestnanosti na samotnú premennú a šokov inflácie na nezamestnanosť (avšak nie v takej silnej miere ako nezamestnanosť na samú seba)

Na obrázku 6 je vyobrazená dekompozícia percentuálneho podielu rozptylov chýb, ktoré nastanú pri predikcii jednotlivých premenných. Ako je možné vidieť na týchto dvoch grafoch, v horizonte 6 rokov je rozptyl chyby prognózy inflácie tvorený impulzmi inflácie v miere zhruba do 85%, miera vplyvu pri oneskorení 1 je nulová, následne veľmi mierne narastá a zotrváva na určitej relatívne konštantnej úrovni (obrázok vľavo). Veľmi podobný priebeh je viditeľný na situácií zobrazujúcej priebeh zloženia rozptylu chyby prognózy nezamestnanosti (obrázok vpravo). Najnižší vplyv impulzu (špecifického šoku) inflácie na prognózu nezamestnanosti je v 3 kvartáli a následne rastie na určité percento a zotrváva v časovom horizonte zhruba od 4 roku sa už vplyv nemení a je ustálený (obrázok vpravo).



Obr. 6: Graf dekompozície rozptylov chyby prognózy pre infláciu a nezamestnanosť

Priebeh je pomerne podobný, až na rozdiel v percentuálnom podiele rozptylu chyby na prognózy premenných do obdobia do 8 kvartálov. V obidvoch prípadoch je rozptyl chyby predikcie danej premennej spôsobený špecifickými šokmi inej premennej najnižší v 3 oneskorení (s výnimkou šokov nezamestnanosti na rozptyl predikcie inflácie, kde je nulový vplyv v oneskorení jedného kvartálu). Konkrétnejšie hodnoty rozptylov chyby predikcie vyobrazuje tabuľka 2. Rozptyl chyby, ktorý nastane pri prognóze premennej inflácie je do časového horizontu 1 roka spôsobený špecifickými šokmi nezamestnanosti v miere mierne prevyšujúcej 8% (ľavá časť tabuľky 1), zatiaľ čo v prípade rozptylu chyby predikcie nezamestnanosti vplyvom šokov inflácie (pravá časť tabuľky) je len v miere 2%, t. j. zhruba len štvrtinový vplyv šokov inflácie na rozptyl predikcie nezamestnanosti v komparácii vplyvu šokov nezamestnanosti na rozptyl predikcie inflácie. Ak berieme do úvahy časový horizont 2 roky, situácia je obdobná, avšak pomer šokov premenných nie je štvrtinový, ale menej ako polovičný. V horizonte 6 rokov je vplyv takmer rovnaký. Predikcia nezamestnanosti je menej „náchylná“ na špecifické šoky premennej inflácia do časového horizontu 2 rokov, následne sa vplyv postupne takmer vyrovná a jednotlivé vplyvy šokov na predikciu sú relatívne rovnaké v dlhom období, čo je aj v súlade s fanchartom predikcie (pozri obrázok 4),

Tab. 2: Dekompozícia rozptylu chyby prognózy

kvartál	Dekompozícia rozptylu chyby prognózy (FEVD) pre:			
	Inflácia [%]		Nezamestnanosť [%]	
	Inflácia[%]	Nezamestnanosť [%]	Inflácia[%]	Nezamestnanosť [%]
1	100,00	0,00	4,28	95,72
2	92,75	7,25	2,43	97,57
3	93,35	6,65	1,54	98,46
4	91,84	8,16	2,01	97,99
5	91,74	8,26	2,73	97,27
6	90,61	9,39	3,77	96,23
8	89,32	10,68	5,92	94,08
12	87,26	12,74	9,47	90,53
16	86,27	13,73	11,30	88,70
20	85,94	14,06	11,97	88,03
24	85,85	14,15	12,16	87,84

Záver

Na základe vykonaných analýz možno konštatovať pre podmienky Slovenskej republiky absolútnu neplatnosť Phillipsovej krivky (aj z krátkodobého neokeynesovského ale aj dlhodobého monetaristického poňatia). Skôr možno vzájomnú súvislosť chápať ako relatívne nepružnú s takmer horizontálnym priebehom, čo je spôsobené hlavným cieľom Európskej centrálnej banky stabilizácia miery inflácie do určitej hranice.

Hodnoty inflácie oscilujú v škále zhruba do 2 % čo je dôsledkom menovej politiky centrálnej banky. Dlhodobý trend vývoja inflácie za sledované obdobie je klesajúci, čo je následkom zlepšujúceho sa makroekonomického a politického prostredia, čo determinovala snaha o splnenie Maastrichtských kritérií pre vstup do Európskej menovej únie. Volatilita inflácie je nižšia a ustálenejšia (rozmedzie do cca 2 %) od vstupu Slovenskej republiky do Európskej únie, do tohto obdobia je viditeľná väčšia variabilita aj lokálne maximá inflácie.

Závery týkajúce sa vývoja a predikcie nezamestnanosti možno rozdeliť zhruba na tri obdobia („éra Mečiarizmu“, politika M. Dzurindu, dopad vplyvu hospodárskej krízy). Spomenuté prvé obdobie je charakteristické rastúcim trendom s globálnym maximom, čo bolo spôsobené najmä politickou nestabilitou, rizikovým bankovým sektorom, nepriaznivá situácia podnikateľského prostredia. Markantný obrat v vývoji trendu nastal zmenou politickej garnitúry a nastolením politicko-ekonomických opatrení vedúcich k stabilite ekonomiky, vnímania krajiny pre investorov, čo sa prejavilo klesajúcim trendom vývoja nezamestnanosti, samozrejme s tým spojená aj lepšia podnikateľská situácia. Za tretí zlom vo vývoji možno považovať dopad globálnej hospodárskej krízy spolu so zmenou politického vedenia v krajine spôsobené siginifikantným nárast nezamestnanosti, čo bolo následkom destabilizácie podnikateľského prostredia globálne. Následne sa rastúci trend stabilizoval, avšak zachoval si rastúci charakter. Posledné zhruba 2 roky je viditeľné čiastočné zlepšenie, čo sa prejavilo opätovným klesajúcim trendom.

Pri celkovej analýze súvislostí týchto dvoch premenných by stálo za úvahu zvážiť pridať iné premenné (hrubý domáci produkt, menový agregát M3, priemernú úrokovú mieru, ročná infláciu nominálnych miezd,...) a následne skúmať dané kauzálne vzťahy, ako je to aj obvyklé v medzinárodnom ponímaní ekonometrických analýz makroekonomických procesov.

Literatúra

ATKESON, A. – OHANIAN, L.E. 2001. *Are Phillips Curves Useful for Forecasting Inflation?* [online cit. 16.11. 2014]. Dostupné na http://www.andrew.cmu.edu/course/88-301/phillips/forecasting-with-phillips-frbm.pdf

EGGER, C. – FELBER, C. – WILDAUER, R. 2011. *Estimating The US Phillips Curve*. [online cit. 16.11. 2014]. Dostupné na http://homepage.univie.ac.at/robert.kunst/tser11_proj_egger_felber_wildauer.pdf)

HUŠEK, R. 2007. *Ekonometrická analýza.* Oeconomica. ISBN 978-80-245-1300-3. s. 368

KIRCHGÄSSNER, G. – WOLTERS, J. 2007. *Introduction to Modern Time Series Analysis. Springer*. ISBN 978-3-540-73290-7. s. 274

KOOP, G. – ONORANTE, L. 2012. *Estimating Phillips curves in turbulent times using The ESB surveys of proffesional forecasters*. (URL dostupné dňa 16.11. 2014 http://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1422.pdf

KULISH, M. – PAGAN, A, 2013. *Issues in Estimating New-Keynesian Phillips Curves in the Presence od Unknown Structural Change*. [online cit. 16.11. 2014]. Dostupné na http://www.rba.gov.au/publications/rdp/2013/pdf/rdp2013-11.pdf

LUKÁČIK, M. 2012. Vektorovo autoregresné modely v ekonometrii. In: *Ekonomika a informatika*, roč. 10, č. 1, s. 130 – 144

LUKÁČIK, M. – PEKÁR, J. 2006. Kointegračná analýza v ekonometrii. [online cit. 16.11. 2014]. Dostupné na http://www.fhi.sk/files/katedry/kove/veda-vyskum/prace/2006/Lukacik-Pekar2006.pdf

LÜTKEPOHL, H. 2005. *New Introduction to Multiple Time Series Analysis*. Springer. ISBN 3-540-40172-5. s. 764

MAKROEKONOMICKÁ DATABÁZA NBS. [online cit. 16.11. 2014]. Dostupné na http://www.nbs.sk/sk/menova-politika/makroekonomicka-databaza

PFAFF, B. 2008. *Analysis of Integrated and Cointegrated Time Series with R*. Springer. ISBN 978-0-387-75966-1. s. 188

R Development Core Team, 2014. R: *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, http://www.R-project.org/.

URAMOVÁ, M. – LACOVÁ, Ž. – HRONEC, M. 2010. *Makroekonómia 1.* Bratia Sabovci , s.r.o., Zvolen. 978-80-557-0043-4. s. 277

Adresa autora

Lukáš Ries
Ekonomická fakulta
Univerzita Mateja Bela Banská Bystrica
2. ročník, 2 stupeň
Lukas.Ries@yahoo.com

Rozšírenie pravdepodobnostného spánkového modelu
Innovation of probabilistic sleep model

Zuzana Rošťáková

Abstract: In this work we modify and validate the probabilistic sleep model (PSM) of Lewandowski, Rosipal and Dorffner (2012). Instead of a single channel electrophysiological (EEG) signal we consider information from three spatially different EEG channels represented with autoregressive model of order 10 or 5. For better detection of the REM stage we use electromyogram (EMG) signal. We analyze and validate these model extensions by comparing the obtained posterior curves with the curves of the original PSM. Finally, we study and analyze stability of the PSM training step. Addition of EMG to PSM and information from three different EEG channels led to improve PSM in comparison to original model. We identify the best model candidate for our further studies of relating the sleep process profiles with daytime behavior.

Abstrakt: V tejto práci sa zaoberáme zlepšením, resp. rozšírením pravdepodobnostného spánkového modelu (PSM) opísaného v článku Lewandowski, Rosipal a Dorffner (2012). Namiesto signálu z jednej elektródy sme uvažovali informácie z troch rôznych EEG kanálov reprezentované cez autoregresný model 10., resp. 5. rádu. Na zlepšenie detekcie REM stavov sme použili aj informácie z elektromyogramu (EMG). Rozšírenia tohto modelu sme analyzovali a overovali porovnaním kriviek posteriórnych pravdepodobností s krivkami pôvodného PSM. Nakoniec sme analyzovali stabilitu PSM v rámci jednotlivých trénovacích behov. Pridanie EMG ako aj informácie z 3 EEG kanálov viedli k zlepšeniu PSM oproti pôvodnému modelu. Z vytvorených modelov sme zvolili najlepšieho kandidáta pre ďalšie štúdie týkajúce sa spojenia spánkového profilu s každodenným správaním.

Key words: probabilistic sleep model, autoregressive model, spectral theory of time series, Rechtschaffen & Kales scores.

Kľúčové slová: pravdepodobnostný spánkový model, autoregresný model, spektrálna teória časových radov, Rechschaffen & Kales skóre.

1. Úvod

Spánok tvorí neoddeliteľnú súčasť života každého človeka. Pre normálne fungovanie človeka počas celého dňa, ako aj pre zdravie vo všeobecnosti, je dôležitá najmä jeho kvalita. Ako však charakterizovať zdravý spánok? Respektíve položme si možno „jednoduchšiu“ otázku: Ako charakterizovať spánok vo všeobecnosti?

V súčasnosti sa nielen v oblasti medicíny používa delenie spánku na niekoľko spánkových stavov v zmysle skórovania, ktoré zaviedli Allan Rechtschaffen a Anthony Kales v roku 1968 (tzv. RK skórovanie). Na základe spomínaného prístupu delíme spánok na dve základné fázy, tzv. REM a nonREM stavy.

REM alebo „rapid eye movement“ stav je charakteristický svalovým tonusom a, ako už napovedá samotný názov, rýchlymi, vedome nekontrolovateľnými pohybmi očí za zavretými viečkami. Táto fáza spánku sa spája so snami a opakuje sa približne 3- až 4-krát počas doby spánku. Prvá REM fáza sa zvyčajne vyskytuje približne 90 až 120 minút po zaspatí a trvá oveľa kratšie než REM stavy objavujúce sa pred koncom spánkového cyklu.

K nonREM stavom zaraďujeme stavy „wake“, S1, S2, S3 a S4. „Wake“ môžeme ešte čiastočne definovať ako stav bdelosti. Trvá pomerne krátko, osoba má pri ňom zatvorené oči, avšak je pri vedomí alebo je pomerne jednoduché sa z tohto stavu prebrať. Pri stave S1 môžeme už hovoriť o zaspávaní, zatiaľ čo stav S2 je akýmsi prechodom medzi ľahkým

a hlbokým spánkom. Pôvodné stavy S3 a S4 sa v súčasnosti súhrnne označujú ako SWS (slow wave sleep)[5] a reprezentujú hlboký spánok.

Napriek tomu, že tento systém sa aj dnes využíva v mnohých oblastiach zaoberajúcich sa spánkom, často sa stával terčom kritiky. Pomerne často sa ako jeden z jeho nedostatkov uvádza, že uvažuje delenie spánku na príliš malý počet fáz a nezohľadňuje možné odlišnosti v rámci jednotlivých stavov.

Z tohto dôvodu boli snahy o modelovanie, resp. charakterizáciu spánku rôznymi inými metódami, napríklad metódami matematickej štatistiky. Jednou z nich je práve pravdepodobnostný spánkový model (PSM) z článku Lewandowski, Rosipal a Dorffner (2012). Autori článku ukázali, že PSM môže priniesť viac doplňujúcich informácií o štruktúre spánku než samotné RK skórovanie.

V rámci tejto práce bolo naším cieľom okrem samotného popisu PSM aj jeho rozšírenie a možné zlepšenie. Všetky výsledky boli porovnané s klasickým skórovaním pomocou RK. K dispozícii sme mali EEG (elektroencephalogram) signál z 3 párov elektród Fp1 a Fp2, C3 a C4, O1 a O2. Okrem toho sme sa snažili zapojiť do modelu aj EMG (elektromyogram) signál z jednej elektródy umiestnenej na brade subjektu.

2. Pravdepodobnostný spánkový model

PSM sa snaží spánok modelovať pomocou posteriórnych pravdepodobností konečného počtu tzv. mikrostavov. Rovnako ako v článku Lewandowski, Rosipal a Dorffner (2012) sme uvažovali 20 mikrostavov.

V prvom kroku predspracovania EEG signálu dochádza k jeho deleniu na 3-sekundové okná. Už tu je možné vidieť rozdiel oproti klasickému prístupu cez RK skórovanie, kde sa zvyčajne uvažujú 20 až 30 sekundové intervaly. Na každom zo spomínaných 3-sekundových okien boli pomocou Burgovej metódy odhadnuté koeficienty autoregresného modelu rádu n (AR(n)). Vo vyššie spomínanom článku autori zvolili $n = 10$, my sme uvažovali 2 prípady: $n = 10$ a $n = 5$.

Okrem toho sme pomocou počítačového softwaru detekovali tzv. „vretienka“ (angl. spindles), ktoré sú typické pre stav S2. V prípade vretienok boli na základe diskriminačnej analýzy rozlišované 4 prípady:

- 0 = žiadny výskyt vretienok,
- 1 = možný výskyt vretienok,
- 2 = pravdepodobný výskyt vretienok,
- 3 = istý výskyt vretienok.

Pre každé 3-sekundové okno bola detekovaná prítomnosť artefaktov. Pri prechádzaní jednotlivých EEG sme sa snažili nahradiť autoregresné koeficienty a detekované typy vretienok v segmentoch s artefaktami pomocou AR koeficientov a typov vretienok z párovej elektródy[6]. Takto nahradené okno bolo potom považované za „čisté“. Ak to nebolo možné vykonať, t.j. artefakt sa nachádzal na tom istom mieste aj v rámci signálu z párovej elektródy, ponechali sme údaje získané z EEG primárnej elektródy, ale príslušné okno bolo označené za chybné a v nasledujúcich výpočtoch sme ho už neuvažovali.

V ďalšom kroku sme v priestore autoregresných koeficientov dimenzie n odhadli zmes normálnych rozdelení podľa vzorca:

---

5 podľa AASM (American Academy of Sleep Medicine)

6 V pôvodnom modeli bola primárnou elektródou elektróda C3, pod párovou elektródou bola chápaná elektróda C4.

$$p(a) = \sum_{i=1}^{20} \pi_i N(a|\mu_i, \Sigma_i), \quad (1)$$

kde a reprezentuje vektor autoregresných koeficientov a$N(a|\mu_i, \Sigma_i)$označuje distribučnú funkciu normálneho rozdelenia so strednou hodnotou μ_i a kovariančnou maticou Σ_i. $\pi_1, \pi_2, \ldots, \pi_{20}$ sú tzv. váhy, pre ktoré platí

$$\sum_{i=1}^{20} \pi_i = 1.$$

Okrem toho autori článku odvodili vzťahy medzi mikrostavmi PSM a pôvodnými spánkovými stavmi. Tento krok bol nutný nielen z dôvodu interpretácie mikrostavov, ktoré nie vždy musia odzrkadľovať klinicky jednoznačne definovaný fyziologický pôvod, ale aj kvôli snahe porovnať výsledky PSM s RK skórovaním.

Pred popísaním samotného trénovacieho algoritmu sme si pre zlepšenie prehľadnosti zaviedli nasledovné označenia. Latentnú premennú, ktorá charakterizuje mikrostavy, t.j. môže nadobúdať hodnoty 1,...,20, sme si označili *z*. Pod $p_R(r|z)$ z článku Lewandowski, Rosipal a Dorffner (2012) rozumieme pravdepodobnosť, že subjekt sa v danom 3-sekundovom segmente nachádza v stave r podľa RK skórovania ($r \in \{wake, S1, S2, SWS, REM\}$) za predpokladu, že sa nachádza v mikrostave *z*. Pravdepodobnosť, že vretienka v danom okne patria do triedy *s*, za predpokladu, že subjekt sa práve nachádza v mikrostave *z*, sme zapisovali v tvare $p_s(s|z)$. Podľa pravidiel elementárnej pravdepodobnosti a na základe nezávislosti medzi vektorom autoregresných koeficientov a, RK skórovaním r a triedou vretienok *s* platí:

$$p(z, a, r, s) = p(a, r, s|z)p(z) = p(z)p(a|z)p_R(r|z)p_s(s|z).$$

Potom

$$p(a, r, s) = \sum_{z=1}^{20} p\,(z, a, r, s) = \sum_{z=1}^{20} p\,(z)p(a|z)p_R(r|z)p_s(s|z), \quad (2)$$

kde $p(a|z) = N(a|\mu_z, \Sigma_z), z \in \{1, \ldots, 20\}$ z (1) a $p(z) = \pi_z$, čo možno zapísať aj v tvare:

$$p(z) = \prod_{i=1}^{20} \pi_k^{z_k},$$

kde $z_k = 1$, ak latentná premenná *z* nadobúda hodnotu *k*, inak $z_k = 0$.

Našou úlohou pri trénovaní modelu bolo na základe reálnych dát odhadnúť pravdepodobnosti uvedené v (2). Nech N je počet pozorovaní (pod pozorovaním tu rozumieme 3-sekundové okno). V rámci každého pozorovania i, $i \in \{1, \ldots, N\}$, sme mali k dispozícii vektor autoregresných koeficientov a_i, RK skóre r_ia zaradenie do triedy vretienok s_i. Nakoľko PSM je založený na nepozorovanej (latentnej) premennej *z*, vhodným nástrojom na získanie odhadov pravdepodobností uvedených v (2) bol „expectation-maximization" algoritmus (EM algoritmus). Ide o iteračný algoritmus opakujúci 2 kroky. Uvažujme logaritmus vierohodnostnej funkcie:

$$\begin{aligned} &\ln\left(p(A, R, S, Z|\Sigma, \mu, \pi, p_R, p_s)\right) = \\ &= \sum_{n=1}^{N} \sum_{i=1}^{20} z_{ni} \left[\ln \pi_i + \ln\left(N(a_n|\mu_i, \Sigma_i)\right) + \ln\left(p_{Rn}(r|i)\right) + \ln\left(p_{sn}(s|i)\right)\right], \end{aligned} \quad (3)$$

kde *A* je matica autoregresných koeficientov, *R* je vektor RK skóre a *S* je vektor tried vretienok pre *N* pozorovaní. Vektory $z_n = (z_{n1}, \ldots z_{n20}), n \in \{1, \ldots, N\}$, sme spojili do

matice Z. Z dôvodu väčšej prehľadnosti sme zvolili skrátené označenie $\mu, \Sigma, \pi, p_R, p_s$pre zodpovedajúce neznáme parametre.

V prvom kroku (E-kroku) sme odhadli (očakávanú) hodnotu logaritmu vierohodnostnej funkcie (3) s ohľadom na podmienenú pravdepodobnosť $(Z|X, \mu, \Sigma, \pi, p_R, p_s)$. V M-kroku sme túto hodnotu maximalizovali cez všetky možné hodnoty neznámych parametrov $\mu, \Sigma, \pi, P_R, P_S$. Získané odhady neznámych parametrov boli potom použité v ďalšom E-kroku. Podrobnejší popis EM algoritmu a priebehu trénovania modelu je možné nájsť v Dodatku A v článku Lewandowski, Rosipal a Dorffner (2012) .

3. PSM založený na EEG z viacerých elektród a EMG

V článku Lewandowski, Rosipal a Dorffner (2012) bol pri vytváraní modelu uvažovaný len EEG signál z centrálnej elektródy C3, resp. C4. Takto vytvorený model sa veľmi dobre zhodoval s RK skórovaním, najmä čo sa týka stavov S2 a SWS. Menší problém nastával pri stave REM, pretože PSM detekoval REM stavy aj na začiatku noci, ešte pred alebo počas zaspávania, čo je klinicky považované za patologické.

Tento nedostatok sme sa snažili odstrániť pridaním informácií z EMG signálu do modelu. Okrem toho sme v rámci tvorby a trénovania modelu použili aj EEG signály z frontálnych (Fp1, Fp2) a occipitálnych elektród (O1,O2). Viedla nás k tomu úvaha, že EEG signály z viacerých elektród môžu priniesť viac informácií o štruktúre spánku a mohli by viesť k zlepšeniu PSM. Napríklad už vyššie spomínané „spindles“ (vretienka) sú často pozorované nielen v centrálnej, ale aj vo frontálnej časti mozgu.

EEG signály sme spracovali analogicky ako v pôvodnom modeli, t.j. reprezentovali každý 3-sekundový segment pomocou AR(n) a vyhľadali sme „spindle“ a artefakty. Pri použití viacerých elektród bolo okno vyhlásené za chybné v prípade, že existoval aspoň jeden pár elektród (frontálnych, centrálnych alebo occipitálnych), v ktorom sa artefakt nachádzal v danom okne v EEG signáloch oboch elektród.

AR koeficienty z jednotlivých elektród sme spojili do jedného vektora dĺžky *3n*, resp. *2n*, ak sme uvažovali prípad dvoch elektród. Predpokladali sme totiž, že jednotlivé EEG signály nie sú nezávislé.

EMG signál bol meraný pomocou elektródy umiestnenej na brade subjektu. Táto elektróda zachytávala svalovú aktivitu subjektu počas noci. Pre REM stav je charakteristický svalový tonus, t.j. stuhnutie svalov a následné prudké uvoľnenie svalového tkaniva pred výskytom REM stavu. Na základe toho sme očakávali, že pri pohľade na EMG signál predchádzajúci REM stavu by malo dôjsť k nárastu amplitúdy EMG a následne k jej poklesu. Vďaka tomuto predpokladu by model mohol lepšie rozpoznávať skutočné REM stavy od miernych pohybov očí a svalovej aktivity na začiatku noci.

Podobne ako v prípade EEG signálu, aj EMG sme rozdelili na disjunktné okná dĺžky 3 sekundy. EMG bolo merané pri frekvencií 200 Hz, t.j. na každú sekundu merania sme mali k dispozícií 200 bodov. Otázkou teraz bolo, ako reprezentovať EMG na každom 3-sekundovom segmente.

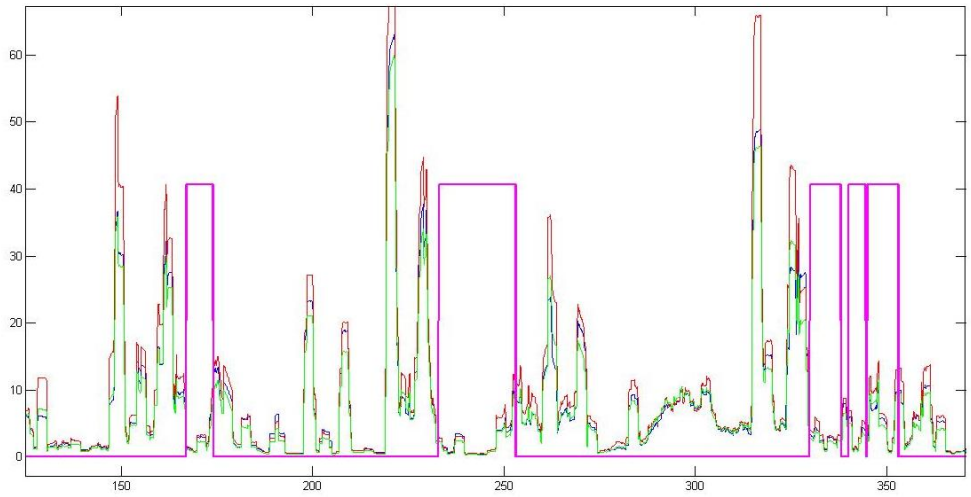
Prvým uvažovaných prístupom bol popis EMG pomocou výberovej variancie na každom okne. Viedla nás k tomu úvaha, že výberová variancia by mala dobre zachytávať nárast, resp. pokles amplitúdy EMG. Nakoľko sme však chceli pripojiť túto hodnotu pre každý 3-sekundový interval k vektoru autoregresných koeficientov z EEG, museli sme hodnoty variancie v rámci zachovania kontinuity s hodnotami autoregresných koeficientov preškálovať do rozpätia -10 až 10.

V tejto fáze prípravy modelu sa tiež vyskytli otázky, či je vhodné spájať autoregresné koeficienty EEG s varianciou EMG, keď ide v podstate o 2 rôzne charakteristiky signálov. Z toho dôvodu sme zvolili aj iný pohľad na reprezentáciu EMG a to práve pomocou

autoregresných koeficientov. V článku Shokrollahi et. al. (2009) autori využili na popis EMG autoregresný model rádu 25, avšak táto hodnota sa nám zdala príliš vysoká. Okrem toho sme museli brať ohľad aj na to, že autoregresné koeficienty EMG budeme pridávať k vektoru AR koefientov z viacerých elektród EEG. Ak by sme zvolili rád AR koeficientov EMG napr. m, pri trénovaní modelu by okrem iného bolo potrebné odhadovať kovariančnú maticu rozmeru $(30+m)\times(30+m)$ resp. $(15+m)\times(15+m)$, čo by pri príliš vysokých hodnotách m mohlo viesť k nestabilným odhadom.

Z týchto dôvodov sme za m zvolili najskôr hodnotu 2 a potom aj hodnotu 5. Koeficienty autoregresných modelov príslušných rádov boli rovnako ako v prípade EEG odhadované pomocou Burgovej metódy.

Pred použitím týchto reprezentácií EMG v modeli sme chceli overiť, ako dobre charakterizujú nárast, resp. pokles amplitúdy EMG. Pre každé 3-sekundové okno sme na základe AR koeficientov a spektrálnej teórie časových radov vypočítali spektrum v rozsahu 10 až 40 Hz. To znamená, že pre každý segment sme získali 31 hodnôt. Jednotlivé 3-sekundové okná sme potom reprezentovali ich mediánovou hodnotou. Na Obr. 1 sú vykreslené tieto hodnoty pre autoregresné modely 2. a 5. rádu. Okrem toho sme pre porovnanie pridali aj varianciu a stavy REM detekované na základe RK skórovania.



Obr. 1: Porovnanie mediánových hodnôt spektier AR(2) (modrá farba), AR(5) (červená farba) a variancie (zelená farba) EMG. REM stavy detekované pomocou RK sú označené fialovou farbou.

Z grafu na Obr. 1 je zrejmé, že krivky majú podobný priebeh, rozdiely sú len v rámci škály. Pri detailnejšom pohľade na segmenty predchádzajúce REM stavom sa nám ako najlepší prístup javila reprezentácia EMG cez autoregresný model rádu 2.

4. Trénovanie modelu a popis dát

Pri trénovaní sme uvažovali 4 typy modelu:

- model bez EMG,
- model s EMG reprezentovaným cez výberovú varianciu,
- model s EMG reprezentovaným cez AR(2),
- model s EMG reprezentovaným cez AR(5).

Čo sa týka EEG, vytvorili sme model z EEG signálu z 1 elektródy (C3, resp. C4), dvoch (C3, O1, resp. C4, O2) , resp. všetkých troch elektród.

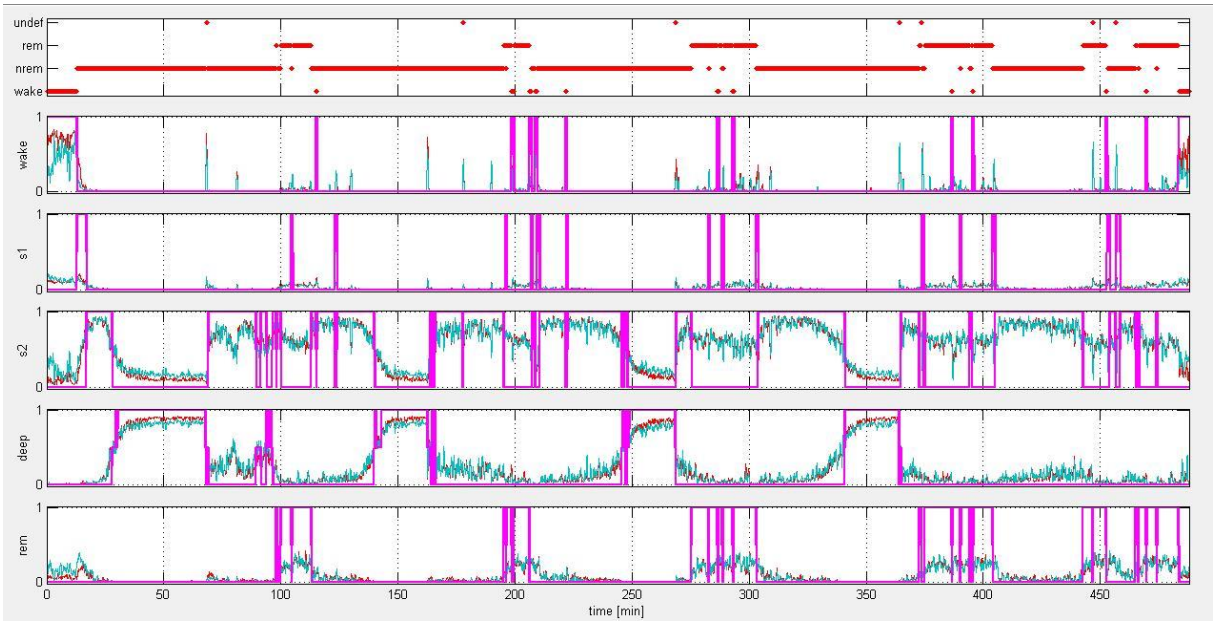
Na trénovanie modelu sme použili dáta z projektu Siesta, uvedené v Klösch, Kemp et. al. (2001). Pôvodne sme uvažovali 175 subjektov, pre každého z nich sme mali k dispozícii

merania z dvoch po sebe nasledujúcich nocí. Po spracovaní EEG a EMG signálov týchto subjektov, výpočte AR koeficientov a detekcii artefaktov, sme vypočítali percentuálny podiel artefaktov pre jednotlivé osoby a noci a na základe toho sme sa rozhodli pred trénovaním vylúčiť záznamy 2 subjektov z prvej noci. V prípade prvej noci prvého subjektu sme zistili 100 %-ný podiel artefaktov pri použití všetkých troch elektród, v druhom prípade bol podiel vyšší ako 50 %. Pri trénovaní modelu sme síce uvažovali len segmenty bez artefaktov, avšak pri vyššom podiele artefaktov v EEG zázname sme náchylní pochybovať o „vhodnosti“ použitia zvyšného záznamu.

Nakoniec sme teda pri trénovaní použili 348 záznamov, pričom trénovanie modelu sme vždy spustili 5-krát. Vytvorené PSM sme potom hodnotili nielen z hľadiska porovnania s RK skórovaním, ale aj v rámci stability spomínaných 5 behov. To znamená, že sme zisťovali, či daných 5 modelov obsahuje tých istých 20 mikrostavov, alebo zakaždým skonvergujú k iným mikrostavom.

5. Porovnanie modelov

Po natrénovaní sme mali k dispozícii niekoľko modelov. Každý z nich sme porovnali s pôvodným modelom a nakoniec definovali najvhodnejší model pre naše ďalšie štúdie. V jednotlivých prípadoch sme vždy natrénovali 5 modelov a pri porovnávaní sme použili ich priemer pre stavy „wake“, S1, S2, SWS a REM.



Obr. 2: Porovnanie pôvodného modelu (modrá farba) s modelom založeným na signáli z 3 elektród (červená farba)

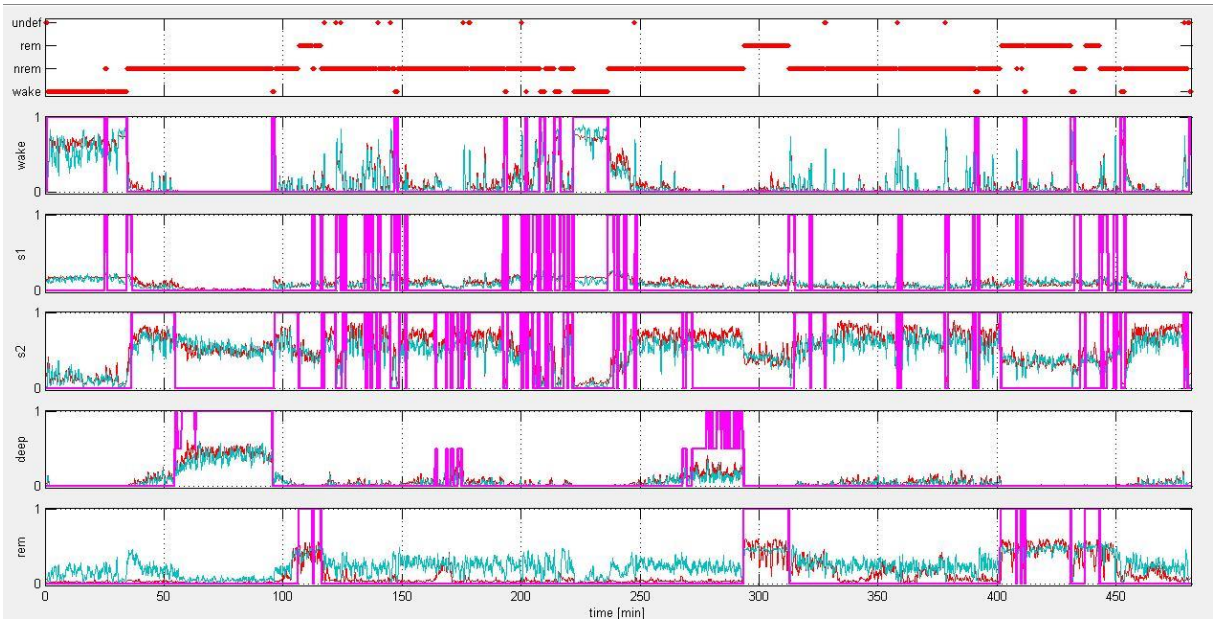
Na Obr. 2 je modrou farbou znázornený pôvodný model s jednou elektródou a červenou farbou model trénovaný na základe 3 elektród, v oboch prípadoch bez EMG.

Rozdiely sú zreteľné najmä na úplnom začiatku a konci noci v rámci stavu „wake“, ale aj S2, kedy model s 3 elektródami lepšie kopíruje RK skórovanie než pôvodný model. Ako sme neskôr zistili, prispievajú k tomu najmä informácie z frontálnych elektród. Vylúčením EEG signálu z Fp1, resp. Fp2 z modelu a porovnaním pôvodného modelu a modelu založeného len na zvyšných dvoch elektródach sme opäť zaznamenali zhoršenie na začiatku a konci merania. Celkovo tak môžeme konštatovať, že použitie EEG signálov z troch elektród, namiesto z jednej, nášmu modelu pomohlo.

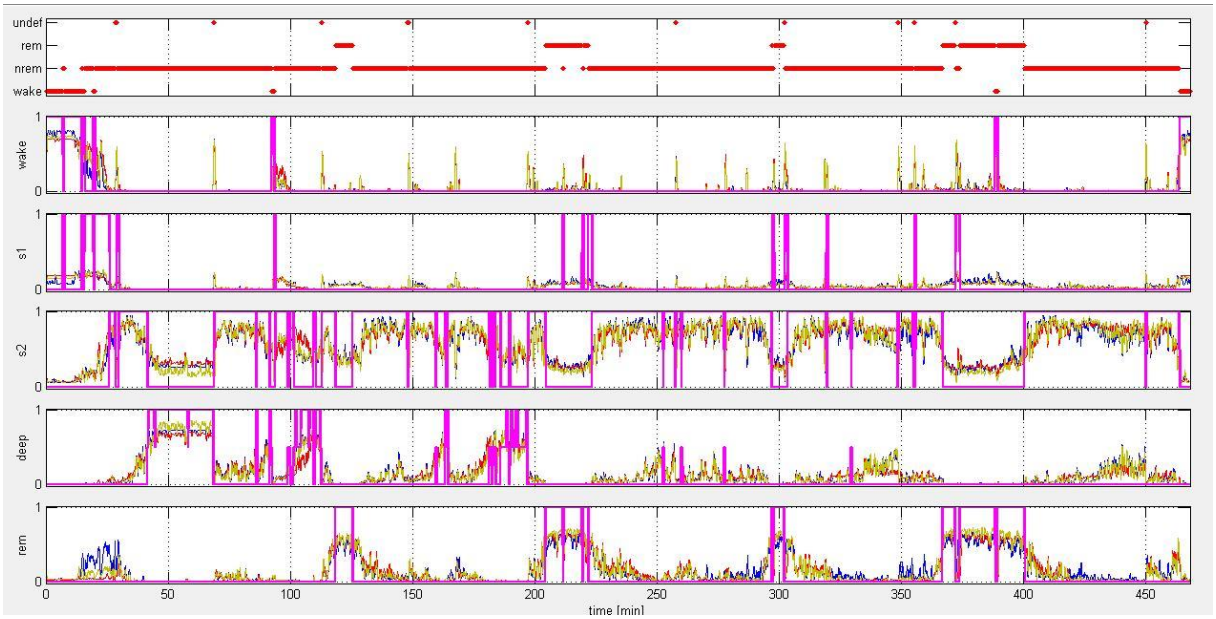
V druhom prípade sme porovnávali 2 modely založené len na centrálnej elektróde, pričom v jednom prípade sme uvažovali pôvodný model bez EMG a v druhom EMG reprezentované varianciou. Výsledok je uvedený na Obr. 3.

Už na prvý pohľad je zrejmé zlepšenie nielen v rámci stavu REM, ale aj v stave „wake“ na začiatku noci. Model založený na EEG z 2 elektród nadobúdal v úvode noci vyššie hodnoty než pôvodný model a viac sa tak priblížil k RK skórovaniu. Na základe toho môžeme skonštatovať pozitívny príspevok EMG signálu v modeli.

Ďalej sme uvažovali model s 3 elektródami a EMG. Reprezentáciu EEG sme vo všetkých 3 prípadoch ponechali na autoregresnom modeli rádu 10 a EMG postupne charakterizovali cez varianciu, AR(2) a AR(5). Všetky 3 modely sme porovnali navzájom a zhodnotili sme, či došlo k zlepšeniu alebo k zhoršeniu modelu. Výsledok je uvedený na Obr. 4. Porovnanie s pôvodným modelom už neuvádzame, nakoľko vyššie sme ukázali pozitívny prínos EMG ako aj EEG zo všetkých troch elektród.



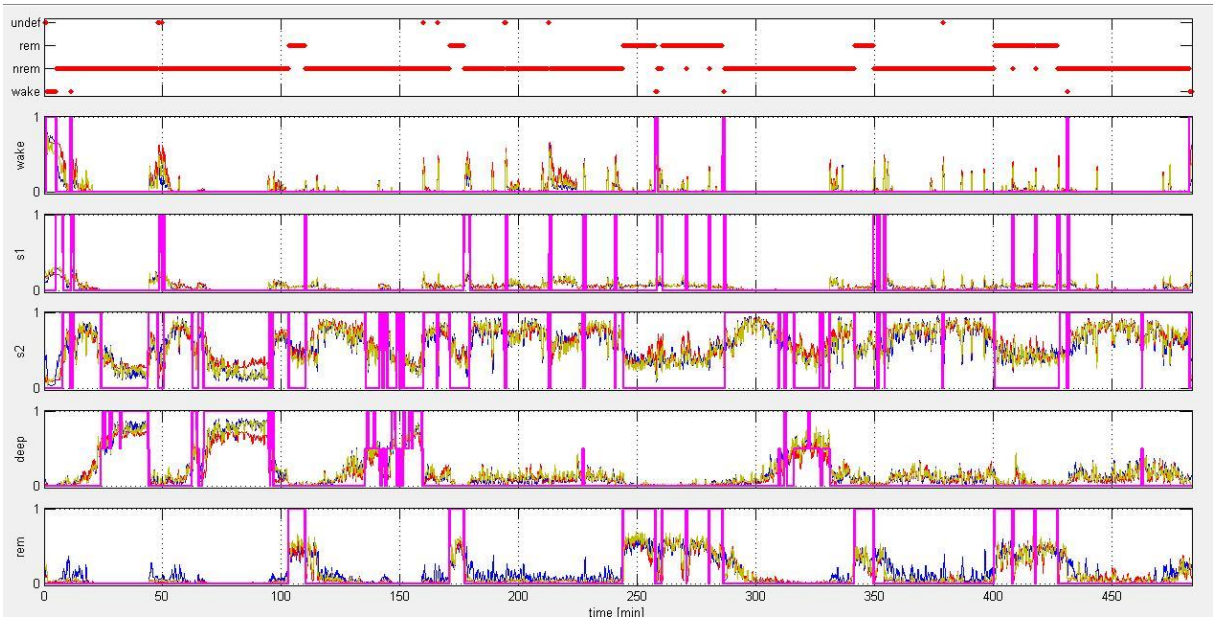
Obr. 3: Porovnanie pôvodného modelu (modrá farba) s modelom založeným na signáli z C3 a EMG reprezentovaným varianciou (červená farba)



Obr. 4: Modely s 3 elektródami (AR(10)) a EMG reprezentovanom varianciou (modrá farba), AR(2) (žltá farba) a AR(5) (červená farba)

Tri modely z Obr. 4 sa viac-menej zhodovali, rozdiely sme našli len v prípade REM stavu, kde bol prvý model (variancia EMG) o niečo horší než zvyšné dva. Medzi zvyšnými dvoma modelmi sme mohli rozhodnúť napr. na základe SWS alebo S2 stavu, kde sa model s EMG charakterizovaným cez AR(2) lepšie zhodoval s RK skórovaním než posledný spomínaný model.

V poslednom kroku sme uvažovali model s 3 elektródami, reprezentovanými však autoregresným modelom rádu 5. V prípade EMG sme vzali do úvahy všetky 3 vyššie spomínané prístupy.

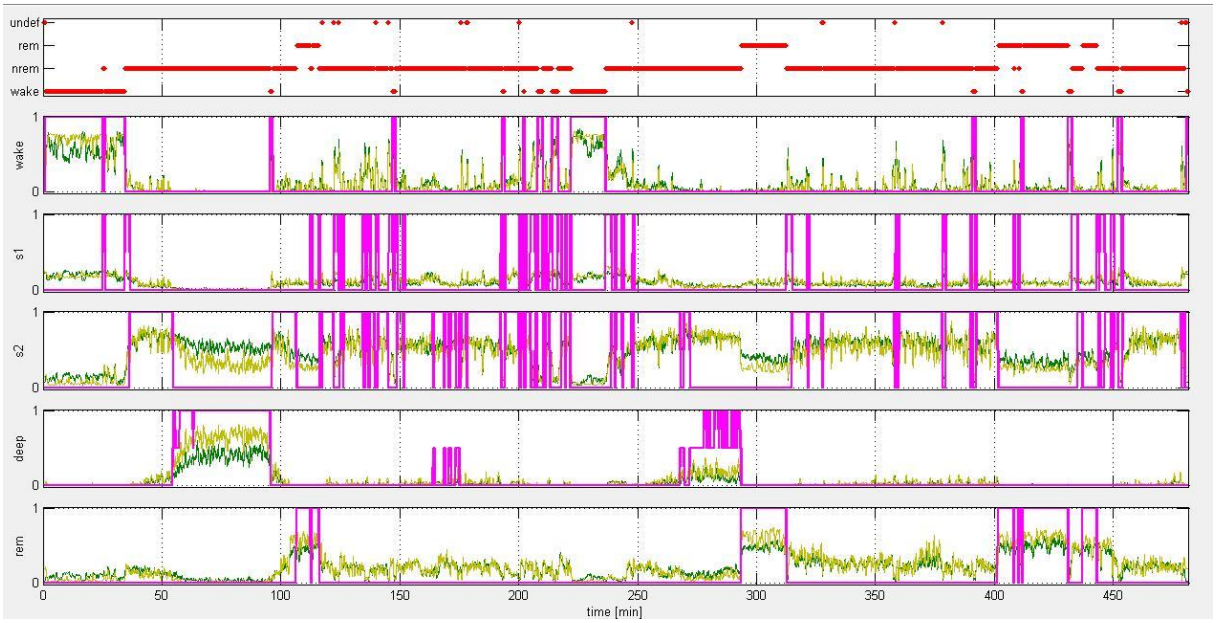


Obr. 5: Modely založené na 3 elektródach AR(5) a EMG reprezentovanom varianciou (modrá farba), AR(2) (žltá farba) a AR(5) (červená farba)

Opäť nám ako najlepší vyšiel model s EMG charakterizovaným cez autoregresný model 2. rádu. Na Obr. 5 si je možné všimnúť, že model s varianciou mal opäť problém s REM stavom na začiatku noci. Model s EMG AR(2) bol oproti modelu s EMG AR(5) viditeľne lepší v stavoch S2 a SWS.

Na poslednom grafe (Obr. 6) je znázornené porovnanie posledných dvoch modelov. Oba sú založené na EEG signály z 3 elektród a EMG AR(2), len v prvom prípade je EEG reprezentované pomocou AR(10) (zelená farba) a v druhom cez AR(5) (žltá farba).

Najlepšie sa nám teda javil model, kde EEG bolo popísané pomocou AR(5) v prípade všetkých troch elektród. Tento model bol aj pomerne stabilný. Pri porovnaní mikrostavov z jednotlivých piatich verzií tohto modelu sme našli až 17 mikrostavov z 20, ku ktorým trénovanie vždy dospelo.



Obr. 6: Porovnanie modelov s 3 elektródami reprezentovanými pomocou AR(10) (zelená farba) a AR(5) (žltá farba) a EMG AR(2)

6. Záver

V tomto príspevku sme sa zaoberali rozšírením a zlepšením pravdepodobnostného spánkového modelu opísaného v článku Lewandowski, Rosipal a Dorffner (2012). Postupne sme PSM vytvorili na základe EEG signálu z jednej (C3, C4), dvoch (C3, C4, O1, O2) alebo všetkých troch elektród, ktoré sme mali k dispozícii.

Porovnaním všetkých vytvorených modelov sme mohli skonštatovať, že model využívajúci všetky 3 elektródy sa lepšie zhoduje s RK skórovaním na začiatku a konci noci jednotlivých subjektov. Vynechaním frontálnej elektródy z modelu sme zaznamenali zhoršenie v rámci stavu „wake“.

Samotné použitie 3 elektród namiesto jednej však ešte nevyriešilo problém s detekciou REM stavov. Pri pohľade na priebehy nocí jednotlivých subjektov sme síce zaznamenali určité zlepšenie oproti pôvodnému modelu, ale ešte stále dochádzalo k detekcii REM stavu aj na začiatku noci. To nás viedlo k pridaniu EMG signálu do modelu. EMG signál sme reprezentovali pomocou variancie, AR(2) a AR(5).

Na základe porovnania priebehu variancie, spektier jednotlivých autoregresných modelov EMG a REM stavov určených cez RK skórovanie, sme usúdili ako najvhodnejšiu reprezentáciu EMG cez AR(2). Nakoľko však išlo len o subjektívne hodnotenie, pri tvorbe modelov sme použili všetky tri spomínané prístupy.

Konečným porovnaním všetkých dostupných modelov sa nám ako najlepší javil model s 3 elektródami charakterizovanými cez AR(5) a EMG reprezentovaným cez AR(2). Tento model vykazoval aj pomerne vysokú stabilitu čo sa týka trénovania.

Literatúra

KLÖSCH, G. – KEMP, B. – PENZEL, T. – SCHÖGL, A. – RAPPELSBERGER, P. – TRENKER, E. – GRUBER, G. – ZEITLHOFER, J. – SAETU, B. – HERRMANN, W.M. – HIMANEN, S.L. – KUNZ, D. – BARBANOJ, M.J. – RÖSCHKE, J. – VÄRRI, A. – DORFFNER, G. 2001. The SIESTA project polygraphic and clinical database. In: *Medicine and Biology Magazine,* roč. 20, č. 3, s. 51–57.

LEWANDOWSKI, A. – ROSIPAL, R. – DORFFNER, G. 2012 . Extracting more information from EEG recordings for a better description of sleep. In: *Computer methods and Programs in Biomedicine*, roč. 108, č. 3, s. 961 – 972.

RECHSCHAFFEN, A. – KALES, A. 1968. Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subject, U.S. Dept. of Health, Education, and Welfare, Bethesda, MD.

SHOKROLLAHI, M. – KRISHNAN, S. – JEWELL, D. – MURRAY, B. Autoregressive and Cepstral Analysis of Electromyogram in Rapid Movement Sleep. In: *IFMBE Proceedings,* roč. 25, č. 4, s. 1580 – 1583.

Adresa autorky:

Zuzana Rošťáková, Bc., 2. ročník, 2.stupeň
Fakulta matematiky, fyziky a informatiky
Univerzita Komenského v Bratislave
zuzana.rošťáková@gmail.com

Ústav merania SAV
Dúbravská cesta 9, 841 04 Bratislava

Finančná integrácia vs. vonkajšia nerovnováha? Aplikácia medzinárodnej investičnej pozície

Financial integration vs. external imbalances? Application of the international investment position

Ľubica Štiblárová

Abstract: The aim of this paper is to quantify international financial integration and external imbalances through created volume indicators. International investment position was selected as a data source. We use panel data analysis to examine indicators of financial integration in conditions of the European Union in time period 1999-2013. Results showed biggest influence of economic development and size of population. We also examined mutual relation between indicators of financial integration and external imbalances.

Abstrakt: Cieľom tohto príspevku je kvantifikácia medzinárodnej finančnej integrácie a vonkajšej nerovnováhy pomocou vyjadrenia objemových ukazovateľov. Za zdroj pre výpočet týchto ukazovateľov bol zvolený štatistický výkaz medzinárodná investičná pozícia. Na skúmanie týchto ukazovateľov sme využili analýzu panelových dát, ktorá bola zameraná na krajiny Európskej únie v časovom období 1999-2013. Analýzy preukázali, že najväčší vplyv má ekonomický rozvoj a veľkosť populácie. V závere sme skúmali ukazovatele finančnej integrácie a vonkajšej nerovnováhy a ich vzájomný vzťah.

Key words: Panel data, international investment position, financial integration, external imbalances, European Union.

Kľúčové slová: Panelové dáta, medzinárodná investičná pozícia, finančná integrácia, vonkajšie nerovnováhy, Európska únia.

1. Úvod

Otázke medzinárodnej finančnej integrácie je prikladaná v dnešnej dobe značná pozornosť, či už v celosvetovom meradle alebo v podmienkach Európskej únie. Väčšina štúdií sa zameriava na kvantifikáciu finančnej integrácie z pohľadu finančných trhov a sledovania príslušných výnosností (pozri napríklad Baele a kol., 2004; Kiehlborn a Mietzner, 2005), čím sa overuje zákon jednej ceny, a teda existencia integrovaného finančného trhu.

V rámci nášho príspevku však pre zhodnotenie tejto problematiky volíme v súčasnosti ešte málo v analýzach využívaný výkaz medzinárodnej investičnej pozície (Jánošík, 2010). Údaje čerpané z tohto výkazu vieme využiť nielen na výpočet ukazovateľov finančnej integrácie, rovnako tak čistá investičná pozícia môže slúžiť ako jeden z ukazovateľov vonkajšej nerovnováhy.

2. Metodika práce

a. Medzinárodná investičná pozícia a voľba ukazovateľov

Medzinárodná investičná pozícia poskytuje informácie o pozícií danej krajiny voči zahraničiu z hľadiska jej finančných investícií, prichádzajúcich a odchádzajúcich zo zahraničia. Čím viac cezhraničných finančných tokov krajina zaznamenáva, tým viac môžeme považovať integráciu jednotlivých krajín za intenzívnejšiu (Európska komisia, 2014).

Na druhej strane je dôležité, či dané investície vstupujú do krajiny (pasíva) alebo smerujú do zahraničia (aktíva). Ak prevažujú zahraničné finančné aktíva nad pasívami, krajina vystupuje voči zahraničiu ako čistý veriteľ, v opačnom prípade ako čistý dlžník (Tab.3). Toto rozlíšenie je kľúčové, pretože sa jedná sa o nerovnovážny stav krajiny vo vzťahu k zahraničiu, a teda môžeme čistú investičnú pozíciu využiť aj ako indikátor vonkajšej nerovnováhy.

Tab.3: Štruktúra medzinárodnej investičnej pozície

Čistá pozícia krajiny = Aktíva - Pasíva	
Aktíva (FA)	***Pasíva (FL)***
• Priame investície (DIA) • Portfóliové investície – majetková časť (PEA) – dlhová časť (PDA) • Ostatné investície • Rezervné aktíva	• Priame investície (DIL) • Portfóliové investície – majetková časť (PEL) – dlhová časť (PDL) • Ostatné investície

Zdroj: vlastné spracovanie podľa IMF(2009)

Pristúpme teda k zadefinovaniu jednotlivých ukazovateľov. Ukazovateľ vonkajšej nerovnováhy (tak ako ho používa napr. Európska komisia, 2014) predstavuje rozdiel medzi celkovými aktívami a pasíva, predelený nominálnym HDP:

$$Ukazovateľ\ vonkajšej nerovnováhy\left(VN\right)=\frac{FA-FL}{HDP_{nom.}} \tag{1}$$

Naopak, objemový ukazovateľ finančnej integrácie predstavuje súčet aktív a pasív predelený nominálnym HDP. Pre účely našej analýzy sme okrem základného tvaru vypočítali aj dve modifikácie. V prvej modifikácii uvažujeme len s priamymi a portfóliovými investíciami, pretože tvoria najvýznamnejšiu časť, v druhej modifikácii upúšťame aj od dlhovej zložky portfóliových investícii a zameriavame sa len na majetkovú časť. Cieľom takejto modifikácie je zistiť, ktorý variant najviac odráža stupeň finančnej integrácie, resp. prítomnosť ktorých podpoložiek je pre posúdenie finančnej integrácie kľúčová.

$$Ukazovateľ\ fin.in.\left(FI\ I.\right)=\frac{FA+FL}{HDP_{nom.}} \tag{2}$$

$$Ukazovateľ\ fin.in.\left(FI\ II.\right)=\frac{DIA+DIL+PEA+PEL+PDA+PDL}{HDP_{nom.}} \tag{3}$$

$$Ukazovateľ\ fin.in.\left(FI\ III.\right)=\frac{DIA+DIL+PEA+PEL}{HDP_{nom.}} \tag{4}$$

b. Analýza a použité údaje

V rámci tohto príspevku sa venujeme analýze krajín Európskej únie za časové obdobie 1999-2013, teda za celé obdobie fungovania spoločnej meny, euro. Ukazovatele finančnej integrácie a vonkajšej nerovnováhy boli vypočítané na základe údajov z databázy MMF.

Z dôvodu extrémnych hodnôt ukazovateľov finančnej integrácie sme sa rozhodli z analýzy vylúčiť Luxembursko a Írsko. V obidvoch prípadoch meranie pomocou objemového indikátora nie je vhodné, pretože sa domnievame, že neprimerane vysoká zásoba finančných aktív a pasív nie je spôsobená úrovňou ich finančnej integrácie, ale najmä výhodnou daňovou politikou (pozri napr. Riegert, 2013). Hodnoty ukazovateľa FI I. sa pre Luxembursko pohybujú v priebehu sledovaného obdobia od 100 až 160, pričom v ostatných krajinách nie je jeho hodnota vyššia ako 10, čo by mohlo výrazne ovplyvniť našu analýzu. Celkovo sme tak mali k dispozícii 390 pozorovaní za 26 krajín a 15 časových období.

Prehľad determinantov, ktoré boli zvolené na základe doteraz publikovaných analýz (pozri napr. Lane a Ferretti, 2008), vrátane ich vyjadrenia a zdrojov môžeme vidieť v Tab.4:

Tab.4: Prehľad skúmaných determinantov

	Vyjadrenie	Zdroj
Obchodná otvorenosť	$\frac{(Export + import)}{HDP}$	International Financial Statistics
Finančný rozvoj	$\frac{(Trhová\,kapitalizácia + bankové\,depozitá)}{HDP}$	World Development Indicators
Ekonomický rozvoj	HDP per capita	World Bank
Veľkosť populácie	Celková populácia	World Bank
Kapitálová otvorenosť	Chinn-Ito Index otvorenosti	Chinn-Ito Index dataset
Účasť v integračnom zoskupení	1- krajina je členom eurozóny 0- krajina nie je členom eurozóny	

Zdroj: vlastné spracovanie podľa Lane a Ferretti (2008)

c. Voľba modelu

Ako už vyplýva z opisu použitých údajov, budeme pracovať s prierezovými údajmi týkajúcimi sa krajín Európskej únie, rovnako tak budeme sledovať ako sa menia v priebehu časového obdobia, teda naša analýza bude zameraná na longitudinálne (panelové) dáta.

Pri použití takéhoto typu údajov môžeme v zásade uvažovať s dvoma typmi modelov, a to s modelom fixných efektov alebo modelom náhodných efektov. Predpokladom použitia modelu náhodných efektov je nekorelovanosť vysvetľujúcich premenných s náhodnou zložkou. Ak nevieme s istotou splniť tento predpoklad, je vhodné využiť tzv. „opatrnejší prístup" a použiť model fixných efektov (Želinský, 2013). Dopustíme sa tak menšej chyby pri použití modelu fixných efektov (ak je zároveň možné použiť model náhodných efektov), než by sme použili model náhodných efektov, keď to je neprípustné.

V oboch prípadoch rozhodujeme, či sú v modeli prítomné individuálne efekty, časové efekty alebo obidve naraz. Ak sú jednotlivé pozorovania špecifické, odlišné pre jednotlivé objekty, ide o individuálne efekty. Ak sa pozorovania líšia z dôvodu času, hovoríme o časových efektoch. V našom prípade predpokladáme, že jednotlivé krajiny majú špecifiká a pozorovania v závislosti od nich budú odlišné, teda v modeli by mali byť významné individuálne efekty, čo následne aj otestujeme.

Vo všeobecnosti takýto model môžeme podľa Wooldridga (2002) zapísať ako

$$y_{i,t} = x_{i,t} \cdot \beta + c_i + u_{i,t} \tag{5}$$

kde $y_{i,t}$ predstavuje vysvetľovanú premennú pre i-tu krajinu a časové obdobie t (v našom prípade príslušné ukazovatele),
$x_{i,t}$ vysvetľujúce premenné pre i-tu krajinu a časové obdobie t (viď Tab.4),
β regresné koeficienty,
$u_{i,t}$ náhodnú zložku a c_i špecifický efekt pre i-tu krajinu (heterogenitu)
pre i=1, 2…N a t=1, 2, …T. Uvažujeme s premennými v logaritmickom tvare.

Na odstránenie c_i z modelu je využívaná transformácia modelu Within, pri ktorej nestrácame žiadne pozorovanie (na rozdiel od využitia prvých diferencií), od modelu (1) odpočítame výraz:

$$\bar{y}_i = \bar{x}_i \cdot \beta + c_i + \bar{u}_i \tag{6}$$

pričom platí $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{i,t} \quad \bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{i,t} \quad \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T} u_{i,t}$ (7)

Následne dostávame výraz

$$y_{i,t} - \bar{y}_i = \left(x_{i,t} - \bar{x}_i\right) \cdot \beta + c_i - c_i + u_{i,t} - \bar{u}_i, \quad (8)$$

po finálnej úprave ako $\tilde{y}_{i,t} = \tilde{x}_{i,t}\beta + \tilde{u}_{i,t}$ (9)

pričom platí $\tilde{y}_{i,t} = y_{i,t} - \bar{y}_i \qquad \tilde{x}_{i,t} = x_{i,t} - \bar{x}_i \qquad \tilde{u}_{i,t} = u_{i,t} - \bar{u}_i.$ (10)

Jednotlivé modely, s ktorými budeme pracovať je potrebné a priori otestovať, či vyhovujú predpokladom kladeným na modely panelových dát. V prvom rade sa jedná o Chowov test tzv. „poolability“, na základe ktorého môžeme rozhodnúť, či je potrebné uvažovať s panelovou štruktúrou modelu. Na určenie štatistickej významnosti individuálnych, časových resp. dvojitých efektov bol použitý LM test (Honda), ďalej bolo potrebné vykonať testovanie poradovej korelácie pomocou Breusch-Godfrey/Wooldridgeovho testu poradovej korelácie, Pesaranov CD test prierezovej závislosti v paneloch, test jednotkového koreňa (Maddala-Wu test) a Breusch-Paganov test heteroskedasticity (Baltagi, 2005).

Ako už bolo spomenuté na začiatku, budeme uvažovať s modelom fixných efektov, teda volíme opatrnejší prístup. Na rozhodnutie aký model použiť je v praxi využívaný aj tzv. Hausmanov test, avšak ak sa v modeli potvrdí prítomnosť prierezovej závislosti alebo poradovej korelácie, estimátor náhodných efektov nie je efektívny, čo môže viesť k skreslenému výsledku tohto testu (Wooldridge, 2002). Keďže je to náš prípad, nepristúpime k tomuto testu v našej analýze.

Výpočty a odhady modelov boli uskutočnené v prostredí programu **R** (R Core Team, 2012). Na prácu s panelovými dátami bol použitý balík *plm* (Croissant a Millo, 2008), rovnako tak balíky *lmtest* (Zeileis a Hothorn, 2014) a *fBasics* (Rmetrics Core Team, 2014).

3. Výsledky a diskusia

a. Finančná integrácia- súhrnný model

Ako prvý bol odhadnutý model fixných efektov pre všetky skúmané krajiny a celé sledované obdobie 1999-2013. Vysvetľovanou premennou sú postupne ukazovatele *FI I.*, *FI II.* a *FI III.*, teda uvažujeme s jeho troma variantmi.

Tab.5: Finančná integrácia- testy súhrnných modelov (Modely fixných efektov)

	Model FI I.		**Model FI II.**		**Model FI III.**	
	test. štat.	p-hodnota	test. štat.	p-hodnota	test. štat.	p-hodnota
Test „poolability“	5,764	<0,001	3,279	<0,001	2,690	<0,001
Individuálne efekty	53,013	<0,001	38,465	<0,001	32,692	<0,001
Časové efekty	1,461	0,129	0,985	0,465	0,939	0,512
Dvojité efekty	74,918	<0,001	48,872	<0,001	34,397	<0,001
Poradová korelácia	3,407	0,001	2,625	0,008	2,941	0,003
Prierezová závislosť	19,660	<0,001	26,862	<0,001	22,586	<0,001
Stacionarita	161,42	<0,001	209,02	<0,001	322.77	<0,001
Heteroskedasticita	15,632	0,016	40,701	<0,001	20,922	0,002

Zdroj: vlastné spracovanie

Ako môžeme vidieť z Tab.5, modely sme podrobili jednotlivým testom pre modely panelových dát- uvedená je testovacia štatistika a príslušná p-hodnota. Podľa Chowovho testu „poolability“ platí, že pre všetky tri varianty modelov je potrebné uvažovať s ich panelovou štruktúrou. Rovnako tak sú všetkých troch prípadoch štatisticky významné individuálne efekty, časové efekty nevykazujú signifikantnosť. Ako štatisticky významné vychádzajú aj dvojité efekty, avšak je to predovšetkým spôsobené silnou významnosťou individuálnych efektov, teda ďalej budeme uvažovať len s individuálnymi efektmi.

Keďže je v modeloch prítomná prierezová závislosť, poradová korelácia, rovnako tak aj heteroskedasticita, odhad štandardných odchýlok odhadnutých koeficientov bude skreslený, teda testy významnosti odhadnutých koeficientov sú nesprávne. V takomto prípade na posúdenie významnosti koeficientov použijeme robustnú variačno-kovariačnú maticu (Croissant a Millo, 2008):

Tab.6: Finančná integrácia- súhrnné modely (Modely fixných efektov- individuálne efekty)

	Model FI I.			**Model FI II.**			**Model FI III.**		
	odhad β	p-hodnota		odhad β	p-hodnota		odhad β	p-hodnota	
obch. otvorenosť	0,094	0,643		0,198	0,307		0,134	0,539	
ekon. rozvoj	1,321	<0,001	***	1,241	<0,001	***	0,895	<0,001	***
populácia	3,726	<0,001	***	3,219	<0,001	***	1,161	0,073	.
finančný rozvoj	0,172	0,057	.	0,203	0,007	**	0,439	<0,001	***
kap. otvorenosť	0,066	0,229		0,191	0,003	**	0,314	<0,001	***
eurozóna- účasť	0,211	<0,001	***	0,237	<0,001	***	0,068	0,404	
R^2	0,6650			**0,7933**			0,6475		
R^2 korigované	0,6066			0,7497			0,5906		
Veľkosť vzorky: n=26, T=15, N=390									

Zdroj: vlastné spracovanie
*Označenie štatistickej významnosti koeficientov: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*
Veľkosť vzorky: n- počet krajín, T- počet časových radov, N- celkový počet pozorovaní.

Budeme teda vychádzať z výsledkov robustnej variačno-kovariačnej matice (Tab.6). Modifikácie ukazovateľa finančnej integrácie FI I., FI II. a FI III. ponúkajú odlišné výsledky. Zatiaľ čo v prvom modeli so základným ukazovateľom FI I. sú štatisticky významné 3 premenné, a to ekonomický rozvoj, populácia a účasť v eurozóne (uvažujeme s hladinou významnosti 0,05), v treťom modeli nemôžeme za štatisticky významnú označiť premennú účasť v eurozóne, ani populáciu. Rovnako nemôžeme za významnú označiť v prvom na rozdiel od ďalších dvoch modelov kapitálovú otvorenosť a finančný rozvoj.

Ako najvhodnejší nám teda pre posúdenie finančnej integrácie vychádza druhý model s uvažovaným ukazovateľom FI II., pomocou ktorého vieme vysvetliť približne 79 percent variability daného ukazovateľa. Ukazovateľ FI I. predstavuje príliš hrubé vyjadrenie, obsahuje v sebe zložky, ktoré neodzrkadľujú úroveň finančnej integrácie (napríklad rezervné aktíva). Tretí ukazovateľ zase vynecháva dlhovú zložku portfóliových investícii, ktorá má podstatný vplyv a zastúpenie, a teda nie je vhodné ju zanedbať. Z tohto dôvodu, rovnako tak z dôvodu obmedzeného rozsahu príspevku budeme v ďalších analýzach uvažovať už len s FI II. ukazovateľom.

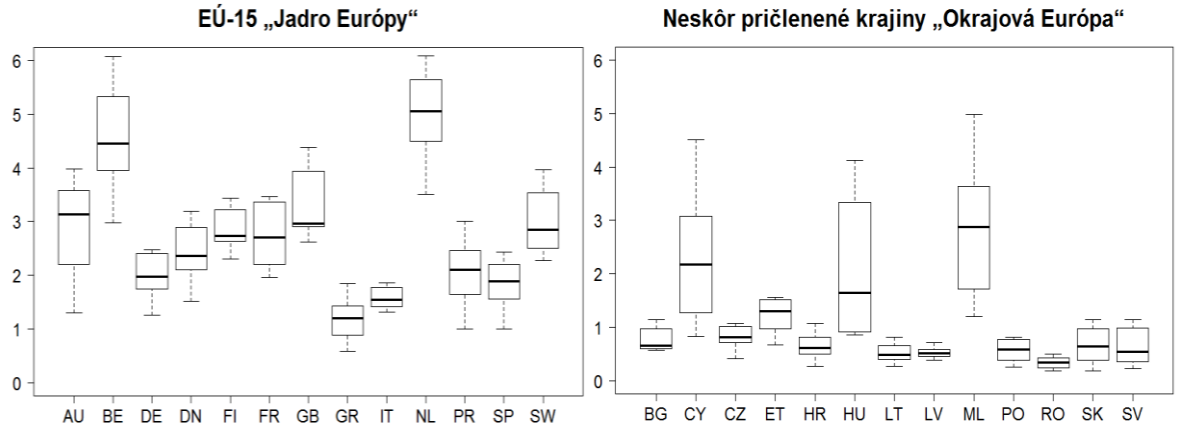
Ak sa pozrieme na hodnotu jednotlivých odhadnutých regresných koeficientov, dostávame opäť zaujímavé výsledky. Predpokladali sme pozitívny vplyv ekonomického, finančného rozvoja, kapitálovej otvorenosti či účasti krajiny v eurozóne na príslušný ukazovateľ finančnej integrácie, čo sa nám aj potvrdilo. Avšak neočakávaný výsledok dostávame pri premennej veľkosť populácie, ktorá má v našich modeloch najväčší pozitívny vplyv na hodnotu ukazovateľa finančnej integrácie. Predpokladali sme, že sa preukáže nepriama závislosť medzi veľkosťou populácie krajiny a úrovňou finančnej integrácie, pretože populačne menšie krajiny majú sklon byť obchodne otvorenejšie a závislejšie na zahraničí, teda mali by mať rozsiahlejšie kapitálové toky vo vzťahu k zahraničiu. Avšak na našej vzorke sa tento predpoklad nepotvrdil, na základe nami definovaných ukazovateľov populačne väčšie krajiny prejavujú vyššiu úroveň finančnej integrácie než menšie.

b. Čiastkové modely

Z dôvodu rešpektovania nerovnomernej úrovne finančnej integrácie v rámci krajín EÚ sme sa rozhodli pre hlbšiu analýzu vykonať dezagregáciu súhrnného modelu dvoma spôsobmi.

Prvý spôsob spočíva v rozdelení analyzovaných krajín na 2 skupiny- na skupinu EÚ-15[7], ktorú v rámci našej analýzy označíme ako „jadro Európy“ (Core Europe) a druhú skupinu, ktorú budú tvoriť neskôr pričlenené krajiny, teda tzv. „okrajová Európa“[8]. Pre takéto delenie sme sa rozhodli z dôvodu, že pri krajinách EU-15 očakávame vyšší stupeň finančnej integrácie oproti neskôr pričleneným krajinám, nakoľko by malo takéto dlhoročné zoskupenie zohrávať významnú úlohu. Druhým rozdelením je rozdelenie časového obdobia. V súhrnom modeli analyzujeme celé obdobie 1999-2013, čím abstrahujeme od vplyvu krízy na analyzované dáta. Avšak po roku 2008 môžeme sledovať výrazné zmeny, či už v prípade vysvetľovanej premennej alebo jednotlivých vysvetľujúcich faktorov. Preto preskúmame nakoľko kríza ovplyvňuje daný model.

Už z Obr.1 a 2 môžeme vidieť, že krajiny EÚ-15 vykazujú oproti neskôr pričleneným krajinám vyššie hodnoty ukazovateľa finančnej integrácie. Zatiaľ čo u krajín EÚ-15 ukazovateľ FI II. dosahuje hodnotu v rozmedzí od 2 po 6, v prípade neskôr pričlenených krajín je úroveň daného ukazovateľa omnoho nižšia a nepresahuje hodnotu 2 (s výnimkou Maďarska, Cypru a Malty).



Obr.7 a Obr.8: Vývoj ukazovateľa FI II. v krajinách EÚ-15 a neskôr pričlenených krajinách, obdobie 1999-2013 (vlastné spracovanie)

Odhad čiastkových modelov pri členení krajín môžeme vidieť v Tab.7. V oboch prípadoch sa ako významné potvrdili individuálne efekty, rovnako tak sme zaznamenali prítomnosť prierezovej závislosti, poradovej korelácie a heteroskedasticity, teda opäť pri posudzovaní významnosti odhadovaných koeficientov vychádzame z robustnej variačno-kovariačnej matice.

V prípade krajín EÚ- 15 sú v našom modeli významné 4 parametre, a to ekonomický a finančný rozvoj, kapitálová otvorenosť a účasť v integračnom zoskupení. V modeli s neskôr pričlenenými krajinami sa premenná finančný rozvoj nepotvrdila ako významná, oproti tomu však vykazuje štatistickú významnosť premenná populácia.

---

[7] EÚ-15 predstavuje 12 zakladajúcich krajín EÚ- Belgicko (BE), Dánsko (DN), Francúzsko (FR), Taliansko (IT), Írsko (IR), Luxembursko (LU), Nemecko (DE), Holandsko (NL), Portugalsko (PR), Veľká Británia (GB), Španielsko (SP), spolu s Rakúskom (AU), Švédskom (SW) a Fínskom (FI), ktoré pristúpili k EÚ v roku 1995.

[8] „Okrajová Európa“: BG- Bulharsko, CY- Cyprus, CZ- Česká republika, ET- Estónsko, HR- Chorvátsko, HU- Maďarsko, LT- Litva, LV- Lotyšsko, ML- Malta, PO- Poľsko, RO- Rumunsko, SK- Slovenská republika, SV- Slovinsko

Tab.7: Finančná integrácia- čiastkové modely podľa rozdelenia krajín (Modely fixných efektov- individuálne efekty)

	EÚ-15 „Jadro Európy“			***Neskôr pričlenené krajiny***		
	odhad β	p-hodnota		odhad β	p-hodnota	
obchodná otvorenosť	0,088	0,834		0,137	0,211	
ekonomický rozvoj	1,649	<0,001	***	1,071	0,004	**
populácia	2,900	0,327		-0,237	0,005	**
finančný rozvoj	0,236	<0,001	***	0,912	0,212	
kap. otvorenosť	0,578	0,006	**	0,200	0,008	**
eurozóna- účasť	0,217	<0,001	***	0,251	0,009	**
R^2/ R^2 korigované	0,633/0,567			0,764/0,681		
Veľkosť vzorky	**n=13, T=14, N=182**			**n=13, T=14, N=182**		
Testy:	test. štatistika	p-hodnota		test. štatistika	p-hodnota	
Individuálne efekty	22,58	<0,001		27,55	<0,001	
Časové efekty	1,328	0,201		1,47	0,129	
Dvojité efekty	36,31	<0,001		36,79	<0,001	
Poradová korelácia	110,5	<0,001		126,64	<0,001	
Prierezová závislosť	13,56	<0,001		6,219	<0,001	
Heteroskedasticita	15,78	0,014		14,264	0,026	

Zdroj: vlastné spracovanie

Na vzorke „okrajovej Európy“ sa nám potvrdili očakávania ohľadom negatívneho vzťahu populácie a finančnej integrácie. Rovnako tak vidíme, že práve premenná ekonomický rozvoj zohráva pri tejto skupine významný vplyv na finančnú integráciu- jedná sa o krajiny, ktoré v hlavnej miere prijímajú kapitál zo zahraničia, sú závislé na ekonomicky vyspelejších krajinách a čím viac sú ekonomicky výkonné, tým viac sú zaujímavejšie pre investorov kapitálu, čo sa v konečnom dôsledku odrazí na ich úrovni finančnej integrácie.

Na druhej strane krajiny EÚ- 15 v zásade poskytujú tejto skupine krajín kapitál a sú ekonomicky vyspelejšie, rovnako tak sa nachádzajú na vyššej úrovni finančnej integrácie. Jedná sa tiež o populačne rozsiahlejšie krajiny, a teda predpoklad o nepriamej závislosti medzi populáciou a úrovňou finančnej integrácie nemohol byť dokázaný. V zásade nám vychádzajú logické výsledky pri štatistickej významnosti premenných, lebo v tejto skupine krajín má štatistický význam ekonomický rozvoj, kapitálová otvorenosť, finančný rozvoj, pretože vo väčšine ide o krajiny s rozvinutým finančným systémom čo napomáha ich úrovni integrácie, rovnako tak účasť v menovej únii. Tá v obidvoch skupinách krajín podporuje ich cezhraničnú kooperáciu a zvyšuje tak veľkosť zahraničných finančných aktív a pasív.

Rozdelením časového obdobia sme dostali odlišné výsledky pre obdobie pred krízou a po nej. Od roku 2008 môžeme sledovať odlišný vývoj vysvetľovanej premennej, ako aj skúmaných determinantov. Prepad nastal už v roku 2008, preto sme tento rok zaradili až do druhého časového obdobia. V oboch sledovaných časových obdobiach boli potvrdené štatisticky významné individuálne efekty. Najlepší výsledok v rámci celej analýzy a potvrdenie významnosti všetkých skúmaných premenných reprezentuje model všetkých krajín uvažujúc časové obdobie pred krízou- model vysvetľuje približne 81 percent celkovej variability ukazovateľa finančnej integrácie (viď Tab.8).

Tab.8: Finančná integrácia- čiastkové modely podľa rozdelenia času (Modely fixných efektov- individuálne efekty)

	Obdobie 1999-2007			***Obdobie 2008-2013***		
	odhad β	p-hodnota		odhad β	p-hodnota	
obchodná otvorenosť	0,512	0,007	**	0,585	0,014	*
ekonomický rozvoj	1,316	<0,001	***	0,98	0,002	**
populácia	6,359	<0,001	***	1,472	0,243	
finančný rozvoj	0,156	0,018	*	0,050	0,815	
kapitálová otvorenosť	0,196	<0,001	***	0,166	0,628	
eurozóna- účasť	0,327	<0,001	***	-0,095	0,448	
R^2/ R^2 korigované	**0,818**/0,706			0,228/0,164		
Veľkosť vzorky	**n=26, T=9, N=234**			**n=26, T=6, N=156**		
Testy:	test. štatistika	p-hodnota		test. štatistika	p-hodnota	
Individuálne efekty	61,70	<0,001		20,57	<0,001	
Časové efekty	1,192	0,304		1,09	0,364	
Dvojité efekty	57,14	<0,001		102,2	<0,001	
Poradová korelácia	149,1	<0,001		46,90	<0,001	
Prierezová závislosť	3,55	<0,001		19,51	<0,001	
Heteroskedasticita	31,86	<0,001		18,91	0,004	

Zdroj: vlastné spracovanie

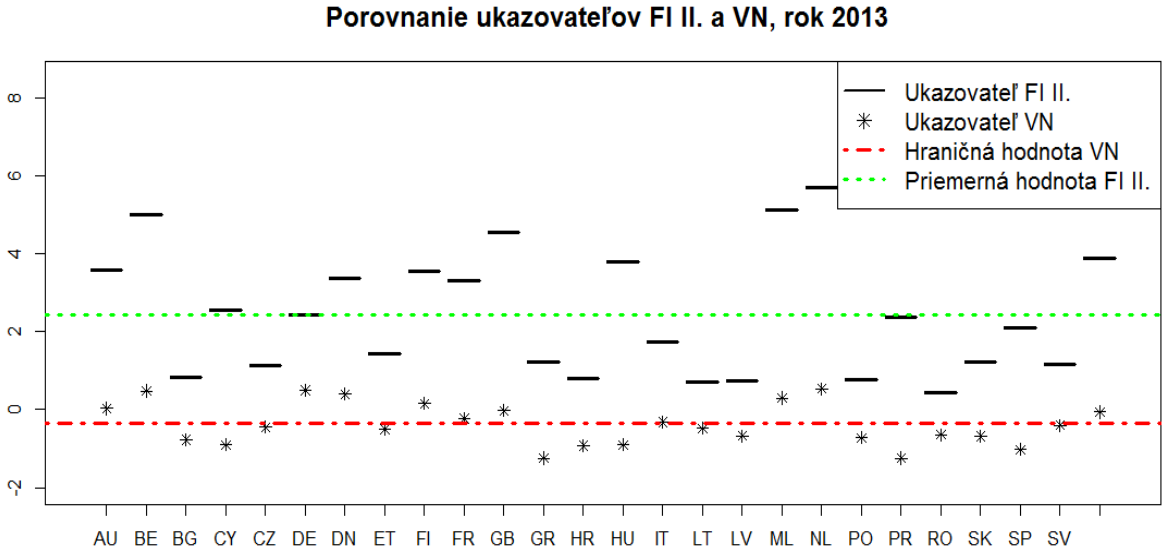
Pôsobenie krízy má zásadný vplyv na skúmané premenné a celý náš model. Ako môžeme vidieť, za štatisticky významné v období 2008-2013 môžeme považovať len premenné obchodná otvorenosť a ekonomický rozvoj. V predchádzajúcich prípadoch sme štatistickú významnosť obchodnej otvorenosti okrem tohto časového delenia vôbec nezaznamenali, model po kríze celkovo ani z pohľadu koeficientu determinácie nie je vyhovujúci.

4. Porovnanie finančnej integrácie a vonkajšej nerovnováhy

Na posúdenie vzájomného vzťahu finančnej integrácie a vonkajšej nerovnováhy sme na začiatok zvolili grafické znázornenie. Na Obr.9 môžeme vidieť vývoj ukazovateľa FI II. a VN v najaktuálnejšom pozorovaní z roku 2013. Pre lepšiu ilustráciu a pochopenie ich vzájomného vzťahu je znázornená priemerná hodnota ukazovateľa FI II. pre Európsku úniu v danom roku, rovnako tak hraničná hodnota pre ukazovateľ VN. Hraničnú hodnotu sme stanovili na úroveň -0,35 na základe odporúčania Európskej komisie. V prípade, že krajina dosahuje nižšiu hodnotu, vonkajšia nerovnováha je v značne kritickom pásme (Európska komisia, 2014).

V sledovanom období vidíme, že sa prejavuje istá závislosť medzi ukazovateľom finančnej integrácie a vonkajšej nerovnováhy. Nami vytvorené hranice vytvárajú vo väčšine prípadov pravidlo, že ak je krajina v značnej nerovnováhe, jej ukazovateľ finančnej integrácie je nižší ako priemer, resp. krajina s vysokou úrovňou finančnej integrácie nevykazuje výraznejšiu vonkajšiu nerovnováhu. Čiže nedostávame sa do situácie trade-off, keď by mala krajina snahou zvýšiť finančnú integráciu vykazovať vyššiu vonkajšiu nerovnováhu.

Pre posúdenie vzájomného vzťahu týchto dvoch ukazovateľov sme sa ďalej rozhodli ísť cestou skúmania determinantov, ktoré ovplyvňujú dané ukazovatele a cez ne sa pokúsiť odvodiť vzájomnú spojitosť. Ak sa preukázal v prechádzajúcich analýzach štatistický význam daných premenných v prípade finančnej integrácie, zaujíma nás, či budú vykazovať štatistickú významnosť aj v prípade vonkajšej nerovnováhy.



Obr.9: Porovnanie ukazovateľov FI II. a VN, rok 2013

Použijeme obdobný model, avšak v tomto prípade vysvetľovanou premennou bude ukazovateľ vonkajšej nerovnováhy. Robíme tak z toho dôvodu, aby sme identifikovali, či existuje spojitosť medzi vplyvom faktorov na finančnú integráciu a vonkajšou integráciou, resp. či pôsobia v rovnakom alebo odlišnom smere.

Výsledky sa vzťahujú k modelu všetkých krajín EÚ pre časové obdobie do krízy (1999-2007), keďže pri danom modeli sme dosiahli najlepšie výsledky (neuvádzame už vykonané testy, opäť vychádzame z variačno-kovariačnej matice kvôli prítomnosti poradovej korelácie, prierezovej závislosti a heteroskedasticity).

Tab.9: Porovnanie finančnej integrácie a vonkajšej nerovnováhy (Modely fixných efektov-individuálne efekty)

	Finančná integrácia			***Vonkajšia nerovnováha***		
	odhad β	p-hodnota		odhad β	p-hodnota	
obchodná otvorenosť	0,512	0,007	**	0,573	0,004	**
ekonomický rozvoj	1,316	<0,001	***	4,8e-05	0,022	*
populácia	6,359	<0,001	***	-4,2e-08	0,080	.
finančný rozvoj	0,156	0,018	*	-3,5e-03	0,049	*
kapitálová otvorenosť	0,196	<0,001	***	-0,199	0,049	*
eurozóna- účasť	0,327	<0,001	***	-0,373	0,031	*
R^2/ R^2 korigované	0,818/0,706			0,562/0,503		
Veľkosť vzorky	**n=26, T=9, N=234**			**n=26, T=9, N=234**		

Zdroj: vlastné spracovanie

Výsledky po predchádzajúcej grafickej analýze sú o trochu zaujímavejšie, potvrdila sa významnosť všetkých determinantov okrem populácie. Čím je ukazovateľ vonkajšej nerovnováhy nižší, tým je situácia horšia, teda ak by boli koeficienty pri determinantoch kladné, nárast premenných by neprehlboval nerovnováhu. To však nemôžeme tvrdiť pri premenných finančný rozvoj, kapitálová otvorenosť, či účasť v eurozóne, výsledky na základe skúmaných determinantov je ťažko zovšeobecňovať na vzájomný vzťah finančnej integrácie a vonkajšej nerovnováhy. Vidíme, že medzi nimi existujú prepojenia, avšak ich vzájomné hodnotenie si vyžaduje hlbšiu analýzu.

5. Záver

Aj keď sme dosiahli odlišné výsledky pri rôznych čiastkových variantoch modelu, analýzy preukázali, že najväčší vplyv na úroveň finančnej integrácie z daných determinantov má veľkosť populácie a ekonomický rozvoj meraný ako HDP per capita. Pre obmedzenú dĺžku príspevku sme v závere len jednoduchou obmenou modelu poukázali na možný vzájomný vzťah ukazovateľov finančnej integrácie a vonkajšej nerovnováhy meraných na základe MIP.

Predmetom ďalšieho skúmania bude preto bližšia analýza štruktúry jednotlivých zahraničných finančných aktív a pasív krajiny, pridanie nových premenných do modelu pri posudzovaní vonkajšej nerovnováhy a finančnej integrácie, rovnako tak úprava modelu na jeho dynamickú podobu.

Literatúra

BAELE, L. a kol. 2004. *Measuring financial integration in the euro area*. Frankfurt: ECB.

BALTAGI, B. H. 2005. *Econometric Analysis of Panel Data- Third edition*. Chichester: John Wiley & Sons. ISBN 978-0-470-01456-1.

CROISSANT Y. – MILLO, G. 2008. Panel Data Econometrics in R: The plm Package. In: *Journal of Statistical Software*, roč. 27, č. 2, s. 1 – 51. ISSN 1548-7660.

EUROPEAN COMMISSION 2014. *Alert Mechanism Report 2014* [online]. EC: Brussels. Dostupné na:<http://ec.europa.eu/europe2020/pdf/2014/amr 2014_en.pdf>.

IMF 2009. *Balance of Payments and International Investment Position Manual*. Washington: International Monetary Fund. ISBN 978-1-58906-812-4.

JÁNOŠÍK, P. 2010. Medzinárodná investičná pozícia. In: *Biatec*, roč. 18, č. 5, s. 25 – 28.

KIEHLBORN, T. – MIETZNER, M 2005. *EU financial integration: Is there a 'core Europe'?* [online]. Frankfurt. [cit. 2014-11-16]. Dostupné na:<http://papers.ssrn.com/sol3/ papers.cfm? abstract_id=675822>.

LANE, P. R. – MILESI- FERRETTI, G. M. 2008. *The Drivers of Financial Globalization* [online]. Washington: IMF. [cit. 2014-10-18]. Dostupné na:<https://www.aeaweb.org/ annual_mtg_ papers/2008/2008_576.pdf>.

R CORE TEAM 2012. *R: A language and environment for statistical computing.* Viedeň: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

RIEGERT, B. 2013. *Luxembourg: the EU's top tax haven*. Deutsche Welle.

RMETRICS CORE TEAM 2014. *Rmetrics - Markets and Basic Statistics* [online]. [cit. 2014-11-07]. Dostupné na:<http://cran.r-project.org/web/packages/fBasics/fBasics.pdf>.

WOOLDRIDGE, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press. ISBN 0-262-23219-7.

ZEILEIS, A. – HOTHORN, T. 2014. *Testing Linear Regression Models* [online]. [cit. 2014-11-16]. Dostupné na:<http://cran.r-project.org/web/packages/lmtest/lmtest.pdf>.

ŽELINSKÝ, T. 2013. Environmentálna Kuznetsova krivka a pozícia Slovenska a Českej republiky v Európskej únii. In: *E+M Ekonomie a Management*, roč. 16, č. 3, s. 4 – 19.

Informácie o autorovi:

Ľubica Štiblárová, Bc.
Ekonomická fakulta
Technická univerzita v Košiciach
2. ročník, 2. stupeň
lubica.stiblarova@gmail.com

Modelovanie daňovej konkurencie/kooperácie v systéme dvoch krajín
Tax competition/cooperation modelling in two countries system

Martin Zoričák

Abstract: Poor condition, as well as the development of public finances currently forcing the countries to consolidation of it. Process of consolidation in the public finance is accompanied with changes in tax policy, where country decides whether to join to other countries to create coherent system – it is the case of tax coordination, resp. harmonisation or on the another hand country might choose to compete with other countries with its own tax policy. In the paper, we examine a model, which consists of set of rules of behaviour for countries on the basis of economic indicators. Results imply that it is better for country to focus on tax cooperation, where countries establish equilibrium with higher tax rate.

Abstrakt: Nepriaznivý stav, ako aj vývoj verejných financií v súčasnej dobe núti krajiny k ich konsolidácii. Konsolidácia verejných financií je sprevádzaná zmenou daňovej politiky, kde sa krajina môže rozhodnúť, či sa zapojí do vytvárania jednotnejšieho systému viacerých krajín – teda do daňovej koordinácie, resp. harmonizácie alebo bude ostatným krajinám konkurovať svojou daňovou politikou. V článku sme skúmali model, v ktorom boli vytvorené pravidlá správania sa krajín na základe ekonomických ukazovateľov. Výsledky použitého modelu naznačujú, že pre krajiny je výhodnejšie zamerať sa na daňovú kooperáciu, pri ktorej sa krajiny vytvoria rovnováhu pri vyššej daňovej sadzbe.

Key words: tax competition, tax cooperation, tax harmonisation, economic model, mobile tax base.

Kľúčové slová: daňová konkurencia, daňová kooperácia, daňová harmonizácia, ekonomický model, mobilný daňový základ.

1. Úvod

Krajiny na celom svete citlivo nastavujú svoje daňové systémy, ako súčasť hospodárskych politík. Daňovou politikou môže krajina výrazne zmeniť ekonomické prostredie v rámci vnútra krajiny, ale či už úmyselne, alebo neúmyselne, pôsobí aj na ostatné krajiny. Zmeny daňových systémov sú o to podstatnejšie v krajinách, ktoré sú nejakou formou zoskupené, či už ide o OSN, OECD alebo EÚ (Tanzi, V. – Davoodi, H 1995). Čím užšie zoskupenie, tým väčší dopad má každá zmena na ostatných členov zoskupenia. V zoskupeniach ako je Európska menová únia, kde menová politika je pod centralizovanou kontrolou ECB, je fiškálna, resp. daňová politika jediným nástrojom, ktorým krajina vie ovplyvňovať ekonomiku svojej krajiny. Zoskupenia sa preto môžu rozhodnúť uplatniť jednu z dvoch stratégií a to buď daňovú konkurenciu, alebo daňovú kooperáciu (Troeger, V. E. 2013).

2. Teoretické aspekty

Daňová konkurencia

K daňovej konkurencii môže dochádzať tam, kde subjekty môžu alokovať svoje zdroje, resp. môže dôjsť k presunu samotných subjektov do jurisdikcií s nižším daňovým zaťažením oproti domácej, pôvodnej krajine. V tomto bode treba podotknúť, že nezáleží od nominálnej daňovej sadzby, ale od efektívnej daňovej sadzby. Efektívna daňová sadzba zohľadňuje nominálnu daňovú sadzbu, ako aj daňovú základňu, teda predmet zdanenia. Efektívna daňová sadzba je vhodná aj na medzinárodné porovnávania a má lepšiu vypovedaciu schopnosť ako nominálna daňová sadzba. Dobrým príkladom je daňová politika v USA (Desai, M. A. 2012), kde došlo k zníženiu nominálnej daňovej sadzby, ale zároveň došlo k rozšíreniu daňovej

základne, aby pokryla čo najväčšiu časť hospodárstva. Daňová konkurencia sa týka hlavne mobilných základov dane, ide hlavne o príjmy z finančného kapitálu, zisky spoločností a pod (Láchová, L. 2007).

Daňovú konkurenciu môžeme členiť na (Kovács, L. 2005):

- **poctivú** - vyplývajúcu z rozdielnej úrovne celkového daňového zaťaženia obyvateľstva, ktorá je výsledkom preferencií voličov, ktoré prikladajú stupni zabezpečovaných verejných statkov,
- **škodlivú** - ak je zameraná na preťahovanie zahraničného kapitálu (daňových základov) a vytvára tak podmienky pre nezamestnanosť,
- **spornú** – taká, ktorá sa hlavne týka vlažnosti odovzdávania daňových informácií a prípadne prelomenia bankového tajomstva.

Daňová kooperácia a harmonizácia

Spolupráca krajín v oblasti daňovej politiky môže byť na rôznych úrovniach, v závislosti od viacerých faktorov, ako napríklad už spomínané členstvo v zoskupeniach, ekonomická situácia krajiny, ale aj hospodársky cyklus.

Prvým stupňom je daňová kooperácia. Daňová kooperácia je typická pre zóny spoločného, resp. voľného obchodu, vo vnútri OECD, MMF ako aj pre EÚ. Jej podstatou sú bilaterálne a multilaterálne modely zdanenia s cieľov obmedziť arbitrážne obchody, stanovujú sa odporúčania napríklad proti škodlivej daňovej konkurencii. Základné prvky sú transparentnosť a informovanosť štátov o príjmoch rezidentov (Nerudová, D. 2005).

Druhým, hlbším stupňom je daňová harmonizácia. Tá predstavuje tému viacerých diskusií členských štátov EÚ, v snahe o harmonizáciu daňových systémov vo vnútri EÚ. Avšak kroky v poslednej dobe naznačujú odklon od tejto stratégie z dlhodobejšieho hľadiska. Dôvodom je, že cieľ harmonizácie je zbližovanie a prispôsobovanie jednotlivých daní v rámci zoskupenia, ako aj daňových systémov na základe spoločných pravidiel, na základe čoho by jednotlivé štáty prišli o ďalšiu časť autonómie (Nerudová, D. 2005).

3. Ekonomický model

Použitý model predstavuje teoretické východisko pre skúmanie dopadov daňovej konkurencie. Model uvažuje o dvoch krajinách, medzi ktorými je možné presúvať kapitál bez obmedzení. Subjektmi modelu sú dve krajiny a firma, o ktorej kapitál súperia dané krajiny.

Krajiny

Každá z krajín formuluje vlastnú daňovú politiku a sleduje produkt krajiny (1), daňové príjmy (2), investície do infraštruktúry (7), úrokovú mieru [9](8), dlh (9), bohatstvo (10), deficit a daňovú sadzbu krajiny (obr. 1). Základným východiskom pre krajinu je maximalizácia bohatstva krajiny, pri čo najvyšších daňových príjmov a najnižšom dlhu, resp. zachovaní pomeru medzi daňovými príjmami a dlhom. Model uvažuje krajinu ako 100 % racionálny subjekt, čo sa odráža na tom, že celé daňové príjmy investuje do infraštruktúry a jej hlavným cieľom je maximalizácia bohatstva krajiny (nemá teda postranné, politické zámery). Súčasťou

[9] Pre porovnanie je vhodnejšie použiť reálnu úrokovú sadzbu, aby sme porovnanie mohli uskutočniť v stálych cenách. Z toho dôvodu sme fixovali úrokovú sadzbu krajiny B a úroková sadzba krajiny A je tak naviazaná na cenovú hladinu v krajine B.

modelu sú parametre a^i – potenciálny produkt *t*-tej krajiny, d^i – odpisová sadzba v *i*-tej krajine, r^2 – úroková miera krajiny B.

$$Y_t^i = a^i.(1 - e^{-K^i}), \tag{1}$$

$$R_t^i = \tau_t^i.Y_t^i, \tag{2}$$

$$\Pi_t^i = \left[2 - \max\left(\frac{Yinf_t^j}{Yinf_t^i}, 2\right)\right].Y_t^i - R_t^i, \tag{3}$$

$$\Pi_t = \Pi_t^i + \Pi_t^i, \tag{4}$$

$$I_{t+1}^i = \Pi_t^i.\frac{\tau_t^j}{\tau_t^i + \tau_t^j}, \tag{5}$$

$$K_{t+1}^i = max\{[(1 - d^i).K_t^i + I_t^i], 0\}, \tag{6}$$

$$Yinf_{t+1}^i = (1 - d^i).Yinf_t^i + def_t^i.R_t^i, \tag{7}$$

$$r_{t+1} = -\left(\frac{Dlh_t^j + (def_t^j - 1).R_t^j}{Dlh_t^i + (def_t^i - 1).R_t^i}\right), \tag{8}$$

$$Dlh_{t+1}^i = [Dlh_t^i + (def_t^i - 1).R_t^i].r_{t+1}, \tag{9}$$

$$W_{t+1}^i = K_{t+1}^i + Yinf_{t+1}^i - Dlh_{t+1}^i. \tag{10}$$

Krajiny na dosiahnutie maximálneho blahobytu môžu upravovať dva parametre, ktoré sú v kompetencií ich vlád. Krajiny môžu upravovať daňovú sadzbu a koeficient deficitu, čím aktívne môžu meniť hospodárenie krajiny, ako aj výsledný blahobyt. Rozhodovací proces krajín je zobrazený v obr. 2.

Základným prvkom rozhodovacieho procesu pre daňovú sadzbu , ako aj pre deficitný koeficient je vzťah podielu daňový príjmov na dlhu krajiny v období *t* a *t +1*. Pri sledovaní tohto ukazovateľa je krajina schopná v relatívne krátkej dobe reagovať na zmeny v ekonomike. Ako už bolo uvedené, základným cieľom krajiny je maximalizácia blahobytu. Použitím vybraných ukazovateľov vie krajina reagovať rýchlejšie na meniace sa prostredie oproti prípadu, ak by sa rozhodovala na základe zmeny blahobytu.

Rozhodovacie kritéria pre výpočet zmeny deficitného koeficientu:

$$def_{t+1}^i = \begin{cases} 0{,}9\,.\,def_t^i\,, & \frac{R_t^i}{Dlh_t^i} \leqq \frac{R_{t-1}^i}{Dlh_{t-1}^i} \\ 1{,}1\,.\,def_t^i, & \frac{R_t^i}{Dlh_t^i} > \frac{R_{t-1}^i}{Dlh_{t-1}^i} \end{cases}$$

Rozhodovacie kritéria pre výpočet zmeny daňovej sadzby:

$$\tau_{t+1}^i = \begin{cases} \begin{cases} 0{,}98\,.\,\tau_t^i, & \tau_t^i > \tau_{t-1}^i \\ \tau_t^i, & \tau_t^i \leqq \tau_{t-1}^i \end{cases}, & \frac{R_t^i}{Dlh_t^i} \leqq \frac{R_{t-1}^i}{Dlh_{t-1}^i} \\ \frac{1}{0{,}98}\,.\,\tau_t^i, & \frac{R_t^i}{Dlh_t^i} > \frac{R_{t-1}^i}{Dlh_{t-1}^i} \end{cases}$$

Obr. 26: Rozhodovacie procesy pre deficitný koeficient a daňovú sadzbu

V modeli vystupujú premenné:

Y_t^i - produkt i-tej krajny v čase t,
R_t^i - daňové príjmy i-tej krajny v čase t,
Π_t, Π_t^i - celkový zisk,zisk v i-tej krajine v čase t,
I_{t+1}^i - investície i-tej krajny v čase t,
K_t^i - kapitál i-tej krajny v čase t,
$Yinf_t^i$ - investície do infraštruktúry i-tej krajny v čase t,
r_t - úroková sadzba v čase t,
Dlh_t^i - dlh i-tej krajny v čase t,
W_t^i - bohatstvo i-tej krajny v čase t,
τ_t^i - daňová sadzba i-tej krajny v čase t,
def_t^i - deficitný koeficient i-tej krajny v čase t.

Firma

Firma vystupuje v modeli ako reprezentatívny subjekt, ktorý disponuje mobilným základom dane vo forme kapitálu. Východiskom je, že firma začína z určitým objemom kapitálu, ktorý môže zvyšovať cez investície. Firma vytvára zisk, na ktorého výšku pôsobí infraštruktúra produkt ako aj daňové výnosy(daňové zaťaženie z pohľadu firmy). Tento zisk vie následne firma presúvať medzi jednotlivými krajinami a rozhoduje sa iba na základe daňovej sadzby, teda presúva ho tam, kde ho vie lepšie zhodnotiť.

4. Aplikácia modelu a výsledky

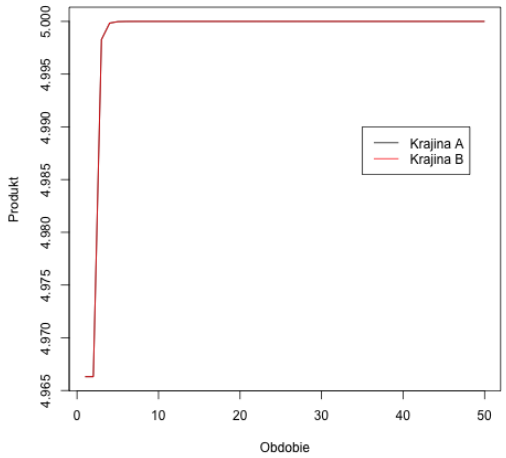
Po aplikovaní vyššie uvedených vzťahov v časovom horizonte 50 období sme získali vývoj jednotlivých ukazovateľov a vzťahy medzi ukazovateľmi, ako aj krajinami navzájom. V modeli boli nasledovné parametre prednastavené, ich počiatočné hodnoty pre obdobie $t = 1$ sú uvedené v tab. 1.

Tab. 5:Počiatočné hodnoty parametrov a premenných

Parameter/Premenná	Hodnota
a^1, a^2	5,0
d^1, d^2	0,2
r^1, r^2	1,0
K^1, K^2	5,0
I^1, I^2	1,0
$Yinf^1, Yinf^2$	1,0
Dlh^1, Dlh^2	1,0
def^1, def^2	1,1
τ^1, τ^2	0,2
W^1, W^2	1,0
R^1, R^2	0,0

Produkt

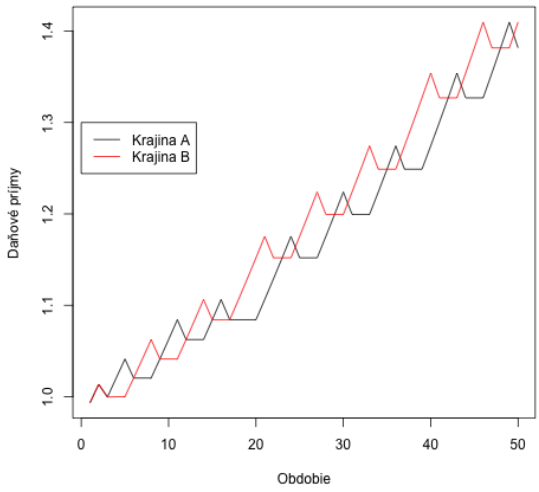
Produkt oboch krajín (obr. 2) sa v priebehu prvých piatich období dostal na úroveň potenciálneho produktu. Vývoj kapitálu, ktorý priamo vplýva na výšku produktu prestáva od určitého objemu vplývať na vývoj produktu v danej krajiny. Tento fakt je zreteľný, nakoľko ako je ďalej ukázané (obr. 6), kapitál po 20 období začne klesať, avšak tento pokles sa neprejaví na produkte krajín. Túto skutočnosť môžeme vysvetliť tak, že ak krajina, ktorá sa orientuje na kumulovanie kapitálu, prekročí určitý objem, ktorý vie daná ekonomika absorbovať, prestane kapitál vplývať na produkt danej krajiny.



Obr. 27: Produkt krajiny A a B

Daňové výnosy

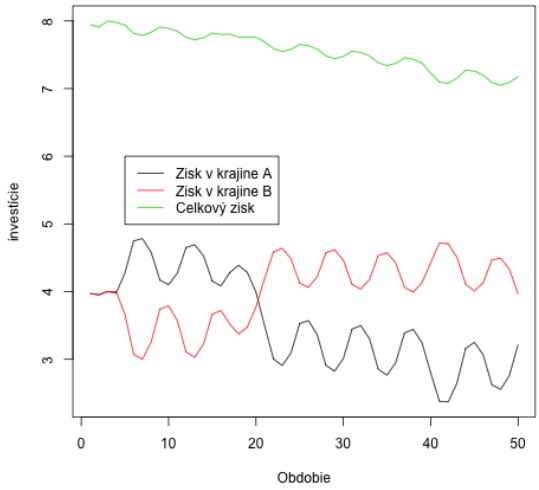
Z vyššie uvedeného vzťahu pre daňové výnosy (2) môžeme konštatovať, že tie tvoria určitý podiel z produktu v závislosti od daňovej sadzby. Vývoj daňových príjmov(obr. 3), ako môžeme vidieť, závisí od 5 obdobia (keď obe krajiny dosiahli potenciálny produkt) sú v plnej miere závislé od daňovej sadzby.



Obr. 28: Daňové príjmy krajiny A a B

Zisk firmy

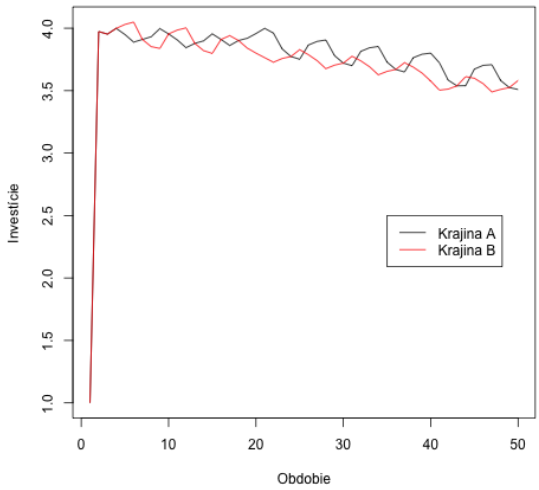
Firma generuje v oboch krajinách zisk, ktorý následne prerozdeľuje formou investícií. Zisk závisí od produktu, daňových príjmov ale aj pomerového ukazovateľa infraštruktúry oboch krajín, ktorý odzrkadľuje podmienky, za akých v danej krajine dosahuje spoločnosť zisk v porovnaní s inou krajinou. Priebeh tvorbi zisku je zobrazený v obr. 4



Obr. 29: Zisk firmy v krajine A, B a celkový zisk

Investície firmy

Investície sú ovplyvňované dvoma atribútmi a to výškou zisku, ktorý firma dosiahla a daňové sadzby v jednotlivých krajinách. Nárast daňových sadzieb, ktorý sa odzrkadľuje v zvýšení daňových príjmov (obr. 5) a následne v poklese zisku firmy, má negatívny dopad na vývoj objemu investícií v oboch krajinách. Firma umiestňuje celý svoj zisk vo forme investícií do krajiny A alebo B. Z toho dôvodu tu nedochádza k špekulatívnym dôvodom poklesu investícií, tento pokles je výlučne spôsobený poklesom disponibilného zisku.

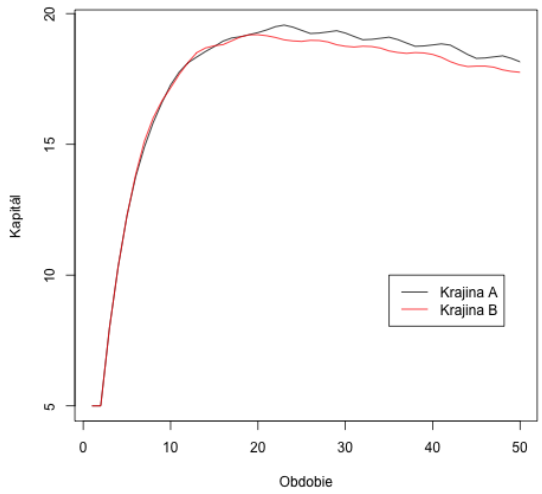


Obr. 30: Investície v krajine A a B

Kapitál firmy

Na výšku kapitálu firmy v danej krajine vplýva výška odpisovej sadzby, ktorá je v kompetencii krajiny a výška investícií, o ktorých rozhoduje firma, avšak nepriamo na ne vplýva aj krajina prostredníctvom daňovej politiky(konkrétne pomerom daňovej sadzby voči

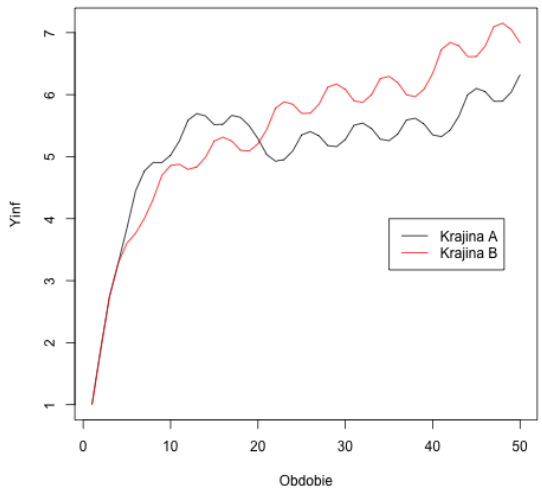
druhej krajine). Pre zachovanie konštantnej výšky kapitálu je potrebné, aby výška investícii bola rovná výške odpisom. Keďže ,ako už oblo vyššie uvedené, objem investícií začne klesať, tento pokles sa prejaví aj vo výške kapitálu v oboch krajinách (obr. 6). Pokles kapitálu je však oneskorený, nakoľko maximálna výška investícií je nad spomínaným rovnovážnym bodom.



Obr. 31: kapitál v krajine A a B

Infraštruktúra krajiny

V našom modeli uvažujeme, že krajiny sa správajú 100 % racionálne. Úplná racionalita sa prejavuje v rozhodovaní použitia daňových príjmov a to buď na tvorbu novej infraštruktúry, alebo na splatenie deficitu[10]. Na základe obr. 4 a obr. 7 môžeme konštatovať, že existuje závislosť medzi ziskom firmy v danej krajine a investíciami do infraštruktúry. Príčinou tejto závislosti je zmena daňovej sadzby.

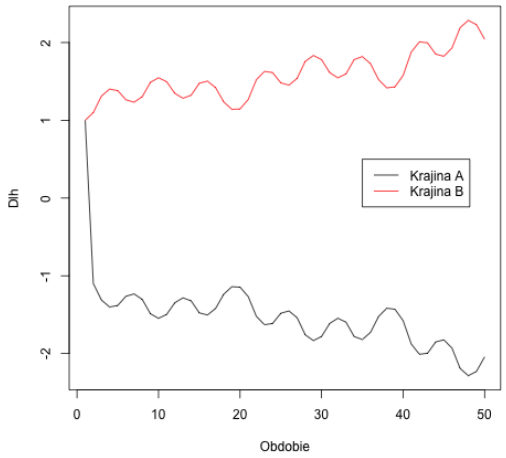


Obr. 32: Investície do infraštruktúry v krajine A a B

[10] K splácaniu deficitu dochádza v prípade, ak koeficient deficitnej sadzby je menej ako 1.

Dlh krajiny

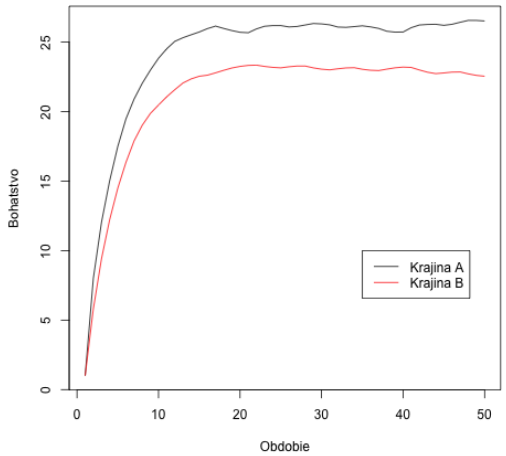
Predpokladom modelu je, že existujú dve krajiny v uzavretom systéme. Z uvedeného vyplýva, že jedna krajina vystupuje v pozícii veriteľa (v našom prípade ide o krajinu A) a druhá krajina v pozícii dlžníka (krajina B). Túto skutočnosť môžeme pozorovať v inverznom vývoji dlhu jednej krajiny vzhľadom na dlh druhej krajiny (obr. 8).



Obr. 33: Dlh krajiny A a B

Bohatstvo krajiny

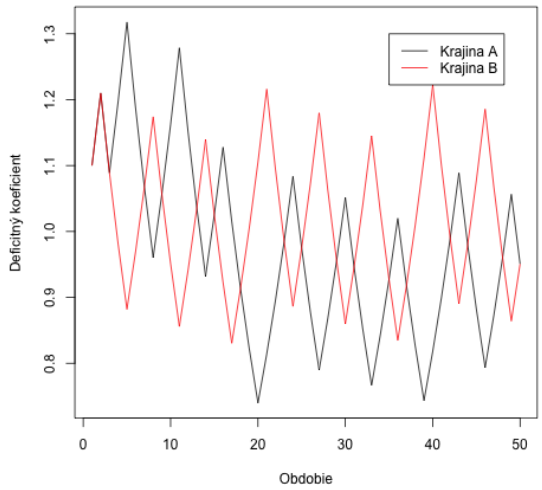
Model uvažuje o troch veličinách vplývajúcich na bohatstvo krajiny, tieto veličiny sú kapitál, investície do infraštruktúry a dlh krajiny. Snahou krajiny je maximalizácia bohatstva krajiny. Maximalizáciu môže dosiahnuť priamym ovplyvňovaním dlhu prostredníctvom deficitnej sadzby alebo investíciami do infraštruktúry cez zmenu daňovej sadzby. Do určitej miery vie krajina nepriamo ovplyvniť aj výšku kapitálu nastavením daňovej sadzby s ohľadom na daňovú sadzbu druhej krajiny. V súčasnej implementácii sa však krajina rozhoduje výhradne na základe vnútornej situácie (na základe podielu daňových príjmov na dlhu krajiny) pri určení daňovej sadzby. Ako môžeme vidieť na obr. č. krajina A dosahuje vyššiu úroveň bohatstva 90 % sledovaného obdobia. Vyššia úroveň bohatstva je spôsobená priaznivejšími výsledkami kapitálu, ako aj dlhu (presnejšie záporného dlhu nakoľko ide v danom systéme o veriteľa). Počiatočný pozitívny vývoj v oblasti investícií do infraštruktúry ako aj veriteľská pozícia má silnejší vplyv ako nižšia úroveň investícií do infraštruktúry v ostatnom období (obr. 9).



Obr. 34: Bohatstvo krajiny A a B

Deficitný koeficient

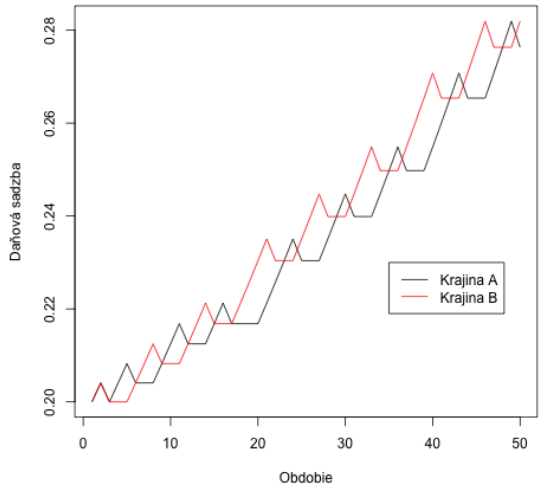
Deficit oboch krajín osciluje približne okolo hodnoty 1 (obr. 10). Priemerný deficitný koeficient za sledované obdobie je v krajine A 0,955, krajine B 0,996, z čoho vyplýva že z dlhodobého hľadiska krajiny hospodária s vyrovnaným rozpočtom.



Obr. 35: Deficitný koeficient krajiny A a B

Daňová sadzba

Daňová sadzba oboch krajín za sledované obdobie prevažne rastie, existujú však určité hladiny, na ktorých sa dočasne ustáli (obr. 11).



Obr. 36: Daňová sadzba krajiny A a B

5. Záver

V modelovanom systéme dvoch krajín sa jedna z krajín vyprofiluje do pozície veriteľa a druhá krajina do pozície dlžníka. Veriteľská pozícia prináša krajine viacero výhod a to v prvom rade vyššie bohatstvo krajiny. Ďalšou výhodou je väčší priestor v možnosti zadlženia sa cez vyšší deficitný koeficient v prvých obdobiach, ako aj vyššie daňové príjmy. Vývoj oboch krajín je však podobný a je viditeľná závislosť krajín na sebe, napr. v oblasti presúvania kapitálu, ako aj financovania dlhu.

Literatúra

LÁCHOVÁ, L. 2007. *Daňové systémy v globálním světě*. 1. vyd. Praha: ASPI

NERUDOVÁ, D. 2005. *Harmonizace daňových systémů zemí Evropské unie*. Praha: ASPI

KOVÁCS, L. 2005. *Tax harmonisation versus tax competition in Europe*. SPEECH/05/624. Dostupné na: http://europa.eu/rapid/press-release_SPEECH-05-624_en.pdf

TANZI, V. – DAVOODI, H. 1995. *Corruption, Public Investment, and Growth*. WP/97/139. Dostupné na: http://www.imf.org/external/pubs/ft/wp/wp97139.pdf

TROEGER, V. E. 2013. *Tax Competition and the Myth of the ´Race to the Bottom´*. Dostupné na: http://www.ifcreview.com/restricted.aspx?articleId=6824

VARTIA, L. 2008. *How do Taxes Affect Investment and Productivity? An Industry-Level Analysis of OECD Countries. OECD working papers: Eoconomic Department working papers*. Vydanie 656. Dostupné na: http://www.oecd-ilibrary.org/economics/how-do-taxes-affect-investment-and-productivity_230022721067

DESAI, M. A. 2012. *A Better Way to Tax U. S. Businesses.* Dostupné na: https://hbr.org/2012/07/a-better-way-to-tax-us-businesses

KOMJATIOVÁ, G. 2014. *Daňová konkurencia a daňové prispôsobovanie v Európskej únii*. Diplomová práca. Košice: TUKE, EkF

Adresa autora:

Martin Zoričák, Bc.
Ekonomická fakulta
Technická univerzita v Košiciach
2. ročník, 2. stupeň
martin.zoricak@student.tuke.sk



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

Kontakt:
petra.judicakova@t-systems.sk

Pridaj sa k tým najlepším!

T · · Systems ·



SAS ONLINE

SAS® onDemand for Academics

Pre akademické účely sprístupňujeme SAS aj online prostredníctvom cloudu. Spracovanie analýz prebieha na vzdialenom hostiteľskom 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 vysokorýchlostnému 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™

www.facebook.com/SASslovakia

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 vzídu. Rozhodovať iba na základe intuície sa dnes nedá. Analýzy dát a získavanie užitočných informácií zohrávajú v biznisových rozhodnutiach čoraz väčšiu úlohu."

doc. Ing. Iveta Stankovičová, 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 Prokeinová PhD.
vysokoškolský pedagóg
Fakulta ekonomiky a manažmentu
Slovenská poľnohospodárska univerzita v Nitre

sas® | 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 rýchlejším a presnejším strategickým podnikateľským rozhodnutiam, získavejším vzťahom 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.**

Mená výrobkov alebo služieb SAS a SAS Institute Inc. sú registrovanými obchodnými značkami alebo obchodnými značkami SAS Institute Inc. v USA a iných krajinách.
® označuje USA registráciu. Ostatné mená značiek a produktov sú registrovanými obchodnými značkami alebo značkami k nim patriacich spoločností. Copyright © 2014, SAS Institute Inc. Všetky práva vyhradené.



ISBN 978-80-88946-66-3

