Posouzení vlivu vybraných makroekonomických veličin na vývoj systému sociálního zabezpečení

Jana Zahálková [1]

Abstrakt

Příspěvek se zabývá posouzením závislosti finančního vývoje salda systému sociálního zabezpečení v ČR na vývoji některých makroekonomických veličin (míra nezaměstnanosti, tempo růstu HDP, průměrné tempo růstu reálné mzdy). Nejdříve je věnována pozornost matematickému vymezení vícenásobné regresní a korelační analýzy, která je používána k popisu statistických závislostí. Dále jsou aplikována skutečná data. Zjištěné výsledky jsou následně znázorněny a interpretovány.

Klíčová slova

Systém sociálního zabezpečení, míra nezaměstnanosti, tempo růstu HDP, průměrné tempo růstu reálných mezd, regresní a korelační analýza, vysvětlující proměnná, vysvětlovaná proměnná.

1. Úvod

Systém sociálního zabezpečení, a zvláště pak sociálního důchodového pojištění, se v posledních letech potýká s nedostatkem finančních prostředků pro krytí svých potřeb. Jinými slovy výdaje z tohoto systému začínají převyšovat příjmy do systému.

Sociální důchodové pojištění je v České republice financováno průběžně, čili PAYG metodou (pay-as-you-go systém), která je založena na mezigenerační solidaritě a redistribuci. To znamená, že generace ekonomicky aktivních lidí platí příspěvky na sociální zabezpečení, ze kterých se vyplácejí důchody a dávky pro generaci současných penzistů. Až se dnešní ekonomicky aktivní lidé dostanou do důchodového věku, budou jejich penze hrazeny z příspěvků generace dnešních dětí atd.

Výhodou těchto systémů je již zmíněná solidarita mezi generacemi, dále pak účinnější obrana proti inflaci a oproti fondovým systémům se nemusí čekat na akumulaci kapitálu. Velkou nevýhodou těchto systémů však představuje silná závislost na příznivém demografickém a ekonomickém vývoji.

Hlavní příčina narůstajících schodků systému důchodového pojištění je spatřována především v populačním stárnutí. ČR dnes patří k populačně nejstarším zemím Evropy. Na každou ženu v reprodukčním věku připadá méně než 1,2 živě narozených dětí, což je hodnota hluboko pod úrovní potřebnou k prosté reprodukci obyvatelstva. Zlepšující se zdravotní péče, snižující se úmrtnost, klesající porodnost prodlužuje očekávanou délku života a v konečném důsledku vede ke stárnutí populace.

Průběžně financovaný systém je ovlivněn nejen populačním stárnutím, ale i vývojem makroekonomických veličin.

Cílem příspěvku je posouzení závislosti finančního vývoje salda systému sociálního zabezpečení na vývoji tempa růstu HDP, míry nezaměstnanosti a tempa růstu reálné mzdy.

[1] Ing. Jana Zahálková, katedra Financí, Ekonomická fakulta VŠB-TU Ostrava, Sokolská 33, Ostrava 701 21, jana.zahalkova.ekf@vsb.cz.

2. Metodologická část

K poznání a matematickému popisu statistických závislostí slouží metody regresní a korelační analýzy. **Regresní analýza** se zabývá jednostrannými závislostmi, to znamená, že se zkoumají obecné tendence ve změnách vysvětlovaných proměnných (závisle proměnné y) vzhledem ke změnám vysvětlujících proměnných (nezávisle proměnné x). **Korelační analýza** se zabývá vzájemnými závislostmi. Větší důraz je kladen na intenzitu vzájemného vztahu mezi veličinami než na posouzení příčinné závislosti.

Pro vystižení vývoje závisle proměnné je v tomto příspěvku použita mnohonásobná lineární regrese.[2]

Závislost lze charakterizovat rovnicí

$$y_i = \eta_i + \varepsilon_i, \tag{1}$$

kde y_i je i-tá hodnota vysvětlované proměnné y, η_i je i-tá hodnota teoretické regresní funkce a ε_i jsou náhodné odchylky (odchylka y_i od η_i) , které lze interpretovat jako důsledek působení náhodných vlivů včetně eventuální nedokonalosti zvolené regresní funkce. *Teoretická regresní funkci* η lze vyjádřit ve tvaru[3]

$$\eta_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i}, \tag{2}$$

kde β_0, β_1, β_2, β_3 jsou neznámé parametry, x_1, x_2, x_3 jsou vysvětlující proměnné.

Odhadnutou regresní funkci Y lze zapsat ve tvaru

$$Y_i = b_0 + b_{yx_1 \cdot x_2 x_3} x_{1i} + b_{yx_2 \cdot x_1 x_3} x_{2i} + b_{yx_3 \cdot x_1 x_2} x_{3i}\,. \tag{3}$$

Parametry $b_{yx_1 \cdot x_2 x_3}$, $b_{yx_2 \cdot x_1 x_3}$, $b_{yx_3 \cdot x_1 x_2}$ jsou *dílčí regresní koeficienty*. Udávají, jak se změní závisle proměnná y při jednotkové změně vysvětlující proměnné x před tečkou za předpokladu, že proměnné x uvedené za tečkou zůstávají neměnné.

Aby odhadnutá regresní funkce co nejlépe vystihovala danou závislost, je zaveden požadavek na minimalizaci součtu čtverců chyb[4] ε_i, platí tedy

$$Q = \sum_{i=1}^{n} \varepsilon_i^{\,2} = \sum_{i=1}^{n} (y_i - \eta_i)^2 \text{.....min}\,. \tag{4}$$

Dosadíme-li do podmínky (4), dostaneme

$$Q = \sum_{i=1}^{n} \left(y_i - b_0 - b_{yx_1 \cdot x_2 x_3} x_{1i} - b_{yx_2 \cdot x_1 x_3} x_{2i} - b_{yx_3 \cdot x_1 x_2} x_{3i}\right)^2 \text{.....min.}, \tag{5}$$

kde n představuje počet pozorování.

Hodnota Q (5) je minimální, jestliže jsou všechny parciální derivace podle jednotlivých parametrů *b* rovny nule.

Po převodu normálních rovnic do maticového tvaru

$$\begin{pmatrix} \sum y_i \\ \sum y_i \cdot y_{1i} \\ \sum y_i \cdot y_{2i} \\ \sum y_i \cdot y_{3i} \end{pmatrix} = \begin{pmatrix} n & \sum x_{1i} & \sum x_{2i} & \sum x_{3i} \\ \sum x_{1i} & \sum x_{1i}^{\,2} & \sum x_{1i} \cdot x_{2i} & \sum x_{1i} \cdot x_{3i} \\ \sum x_{2i} & \sum x_{1i} \cdot x_{2i} & \sum x_{2i}^{\,2} & \sum x_{2i} \cdot x_{3i} \\ \sum x_{3i} & \sum x_{1i} \cdot x_{3i} & \sum x_{2i} \cdot x_{3i} & \sum x_{3i}^{\,2} \end{pmatrix} \cdot \begin{pmatrix} b_0 \\ b_{yx_1 \cdot x_2 x_3} \\ b_{yx_2 \cdot x_1 x_3} \\ b_{yx_3 \cdot x_1 x_2} \end{pmatrix} \tag{6}$$

[2] Při použití mnohonásobné lineární regresní funkce platí, že závisle proměnná y je lineárně závislá na každém z vysvětlujících proměnných x a všechny tyto proměnné x jsou vzájemně nezávislé nebo alespoň ovlivňují změny závisle proměnné všechny jedním směrem.

[3] Teoretická vymezení jsou vztažena na 3 vysvětlující proměnné vyskytující se v příspěvku.

[4] Tzv. metoda nejmenších čtverců.

lze získat odhady parametrů *b* řešením soustavy rovnic (např. použitím Cramerova pravidla[5]).

Pro účely srovnání a posouzení individuálního vlivu jednotlivých vysvětlujících proměnných na závisle proměnnou lze použít také *normalizované regresní koeficienty*, tzv. B-koeficienty. Platí vztah

$$B_{yx_1 \cdot x_2 x_3} = \frac{s_{x_1}}{s_y} b_{yx_1 \cdot x_2 x_3}$$

$$B_{yx_2 \cdot x_1 x_3} = \frac{s_{x_2}}{s_y} b_{yx_2 \cdot x_1 x_3}$$

$$B_{yx_3 \cdot x_1 x_2} = \frac{s_{x_3}}{s_y} b_{yx_3 \cdot x_1 x_2} , \tag{7}$$

kde s_{x_1}, s_{x_2}, s_{x_3}, s_y jsou směrodatné odchylky jednotlivých proměnných.

Velikost dílčích regresních koeficientů je ovlivněna volbou měrné jednotky. Naproti tomu B-koeficienty představují bezrozměrné číslo, tím je možné jejich vzájemné srovnání, které slouží především k určení intenzity vlivů jednotlivých vysvětlujících proměnných na závisle proměnnou.

Dále je nezbytné propočty doplnit o *index determinace* I^2_{yx} (8) a o *párové korelační koeficienty* r_{yx} (9), které měří těsnost závislosti popsané lineární regresní funkcí. Z toho vyplývá, že zjištěné párové korelační koeficienty, podobně jako B-koeficienty, pomohou rozhodnout o vhodnosti zavedení jednotlivých veličin do regresního modelu.

$$I^2_{yx} = \frac{1 - \frac{\sum (y_i - Y_i)^2}{n}}{\frac{\sum (y_i)^2}{n} - \left(\frac{\sum y_i}{n} \right)^2} \tag{8}$$

$$r_{yx} = \frac{s_{xy}}{\sqrt{s_x^2 s_y^2}} \tag{9}$$

V případě, že B-koeficienty a párové korelační koeficienty poskytují odlišné výsledky týkající se vhodnosti zavedení veličin do modelu, je nezbytné prozkoumat *multikolinearitu*. Multikolinearita představuje závislost mezi vysvětlujícími proměnnými a informace o ní lze čerpat z matice korelačních koeficientů R. Při posuzování je důležitá hodnota determinantu. Jsou-li všechny dvojice vysvětlujících proměnných párově nekorelované, je determinant korelační matice roven jedné a s narůstající multikolinearitou se přibližuje nule. V zásadě platí, že multikolinaerita je škodlivá, pokud některý z koeficientů korelační matice překročí hodnotu zhruba 0,75.

Exaktněji to lze zjistit prostřednictvím *Farrarova-Glauberova testu*. Testovým kritériem je výraz

$$B = -\left[(n-1) - \frac{1}{6}(2p+5) \right] \cdot \ln|R| , \tag{10}$$

[5] Dílčí regresní koeficienty jsou rovny podílu D_i/D, kde D je determinant matice na pravé straně rovnice (6) (matice A) a D_i je determinant matice, která vznikne z matice A výměnou i-tého sloupce za sloupec levé strany rovnice.

kde n je rozsah výběru, p je počet vysvětlujících proměnných zařazených do modelu a $|R|$ je determinant korelační matice. Testové kritérium má rozdělení χ^2 s [p(p-1)/2] stupni volnosti. Hypotéza H_0 znamená nezávislost vysvětlujících proměnných, H_1 závislost vysvětlujících proměnných. Kritickým oborem jsou ty hodnoty testového kritéria B, které překročí příslušný kvantil rozdělení χ^2, tj.

$$B \geq \chi^2_{1-\alpha}[p(p-1)/2], \tag{11}$$

v tomto případě by byla již multikolinearita považována za statisticky významnou a zařazení veličiny do modelu by nebylo vhodné.

3. Využití regresní a korelační analýzy ke zjištění vlivu vybraných makroekonomických ukazatelů na vývoj systému sociálního zabezpečení

Vstupní údaje byly získány z webových stránek Českého statistického úřadu a ze statistických ročenek za jednotlivé roky (Tab. č.1). Data jsou za 11 po sobě jdoucích let ($i = 11$), počínaje rokem 1993[6].

Symbol	Položka	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
	Příjmy na sociální zabezpečení (mil. Kč)	108968	132731	154318	174316	191004	203910	210888	222176	242320	258513	272366
	Dávky sociálního zabezpečení (mil. Kč)	117280	136261	154974	153442	204802	221088	237109	255073	273301	292242	305030
y	SALDO (mil. Kč)	-8312	-3530	-656	20874	-13798	-17178	-26221	-32897	-30981	-33729	-32664
x_1	Míra nezaměstnanosti (%)	4,30%	4,30%	4,00%	3,90%	4,80%	6,50%	8,70%	8,80%	8,10%	7,30%	7,80%
x_2	Tempo růstu HDP (%), reálně	0,10%	2,20%	5,90%	4,20%	-0,70%	-1,10%	1,20%	3,90%	2,60%	1,50%	3,70%
x_3	Průměrné tempo růstu reálné mzdy (%)	7,60%	7,80%	8,70%	8,70%	1,30%	-1,30%	6,20%	2,40%	3,80%	5,40%	6,70%

Tab.č.1: Vstupní údaje

[6] V předcházejících letech byla metodika výpočtu příjmové a výdajové strany sociálního zabezpečení odlišná.

i	y_i	x_{1i}	x_{2i}	x_{3i}	$y_i \cdot x_{1i}$	$y_i \cdot x_{2i}$	$y_i \cdot x_{3i}$	Y_i	$y_i - Y_i$	$(y_i - Y_i)^2$
(0)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	-8312	0,043	0,001	0,076	-357,4	-8,31	-631,7	-4221,633	-4090,37	16731103
2	-3530	0,043	0,022	0,078	-151,8	-77,7	-275,3	-1485,228	-2044,77	4181091,8
3	-656	0,04	0,059	0,087	-26,24	-38,7	-57,07	5632,352	-6288,35	39543371
4	20874	0,039	0,042	0,087	814,09	876,7	1816	4181,736	16692,3	278631664
5	-13798	0,048	-0,01	0,013	-662,3	96,59	-179,4	-9313,305	-4484,7	20112490
6	-17178	0,065	-0,01	-0,01	-1117	189	223,31	-22824,24	5646,24	31879975
7	-26221	0,087	0,012	0,062	-2281	-315	-1626	-36194,37	9973,37	99468126
8	-32897	0,088	0,039	0,024	-2895	-1283	-789,5	-33603,58	706,576	499250,11
9	-30981	0,081	0,026	0,038	-2509	-806	-1177	-29929,23	-1051,77	1106219,3
10	-33729	0,073	0,015	0,054	-2462	-506	-1821	-25228,91	-8500,09	72251548
11	-32664	0,078	0,037	0,067	-2548	-1209	-2188	-26105,6	-6558,4	43012604
$\sum$	-179092	0,685	0,235	0,573	- 14196	-3080	-6707	-179092	-3,8E-11	607417444

Tab.č.2: Výpočtová tabulka

Hodnoty salda systému sociálního zabezpečení odhadnuté tímto modelem lze nalézt v 8. sloupci tabulky č.2. Důležitý je i sloupec 9, kde jsou zachyceny odchylky od skutečnosti.

Odhadnutou regresní funkci Y (3) lze tedy zapsat ve tvaru

$$Y = 27916{,}06 - 757799 \cdot x_1 + 129906{,}8 \cdot x_2 + 4181{,}283 \cdot x_3 .$$

b_0	27916,06	s_{x_2}	0,020921	r_{yx_1}	$-0{,}87779$
$b_{yx_1 \cdot x_2 x_3}$	-757799	s_{x_3}	0,031271	r_{yx_2}	0,195588
$b_{yx_2 \cdot x_1 x_3}$	129906,8	$B_{yx_1 \cdot x_2 x_3}$	$-0{,}86953$	r_{yx_3}	0,460011
$b_{yx_3 \cdot x_1 x_2}$	4181,283	$B_{yx_2 \cdot x_1 x_3}$	0,163982	$\lvert R \rvert$	0,4844986
s_y	16573,58	$B_{yx_3 \cdot x_1 x_2}$	0,007889	B	5,9179
s_{x_1}	0,019017	I^2_{yx}	79,9%	$\chi^2_{0.95}[3]$	7,8147

Tab.č.3: Vypočtené hodnoty

Řešením normálních rovnic byly vypočteny tyto dílčí regresní koeficienty:

$b_0 = 27916{,}06$

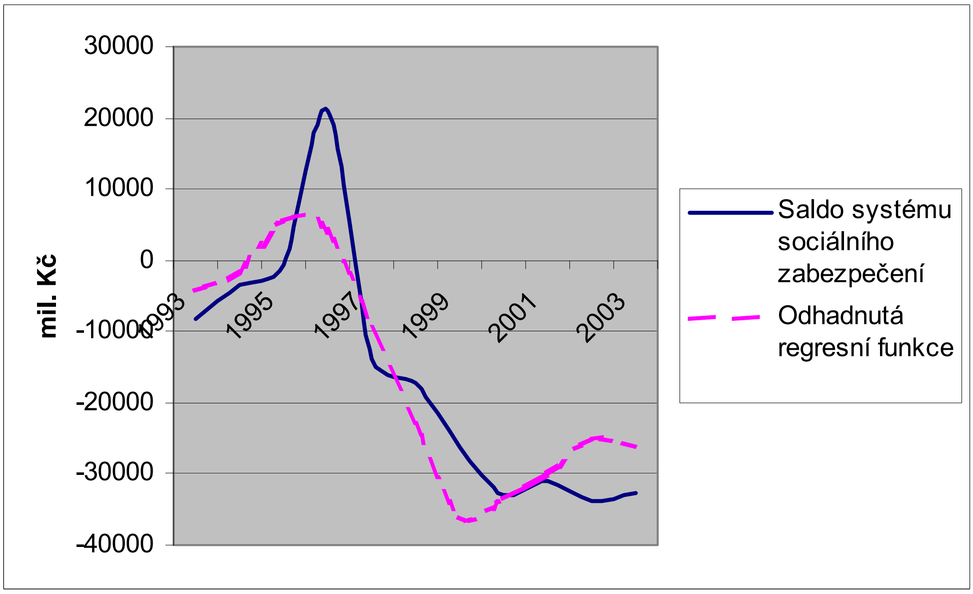
$b_{yx_1 \cdot x_2 x_3} = -757799 \quad \left(\dfrac{mil.Kč}{\%} \right)$, což znamená, že zvýší-li se míra nezaměstnanosti o 1 %, sníží se saldo systému sociálního zabezpečení o 775 779 mil. Kč, jestliže vyloučíme vliv tempa růstu HDP a tempa růstu reálné mzdy. Saldo je rozdílem mezi příjmy a výdaji, snižující se velikost tohoto ukazatele tedy nesvědčí nic pozitivního, je to následek snižujících se příjmů a/nebo zvyšujících se výdajů.

$b_{yx_2 \cdot x_1 x_3} = 129906{,}8 \quad \left(\dfrac{mil.Kč}{\%} \right)$ udává, že zvýší-li se tempo růstu HDP o 1 %, zvýší se saldo systému sociálního zabezpečení o 129 906,8 mil. Kč za předpokladu, že se nezmění ani tempo růstu mezd, ani míra nezaměstnanost.

$b_{yx_3 \cdot x_1 x_2} = 4181{,}283 \left(\frac{mil.Kč}{\%}\right)$ udává, že jestliže se zvýší průměrné tempo růstu reálných mezd o 1 %, vzroste saldo systému sociálního zabezpečení o 4 181,283 mil. Kč, pokud se míra nezaměstnanosti i tempo růstu HDP nezmění.

Index determinace I_{yx}^2 dle (8) vyšel 79,9%, což znamená, že daný model vystihuje závislost z téměř 80 % (80 % rozptylu znaku y je způsobeno vlivem x_1, x_2, x_3), zbylých 20% disperze znaku y je způsobeno náhodnými vlivy. Vývoj salda systému sociálního zabezpečení skutečného i zjištěného pomocí odhadnuté regresní funkce ukazuje obr. č. 1.

Obr.č. 1: Saldo systému sociálního zabezpečení – skutečnost a odhadnutá regresní funkce



Dle (7) lze vypočíst B-koeficienty

$$B_{yx_1 \cdot x_2 x_3} = -757799 \cdot \frac{0{,}019017}{16573{,}58} = -0{,}86953,$$

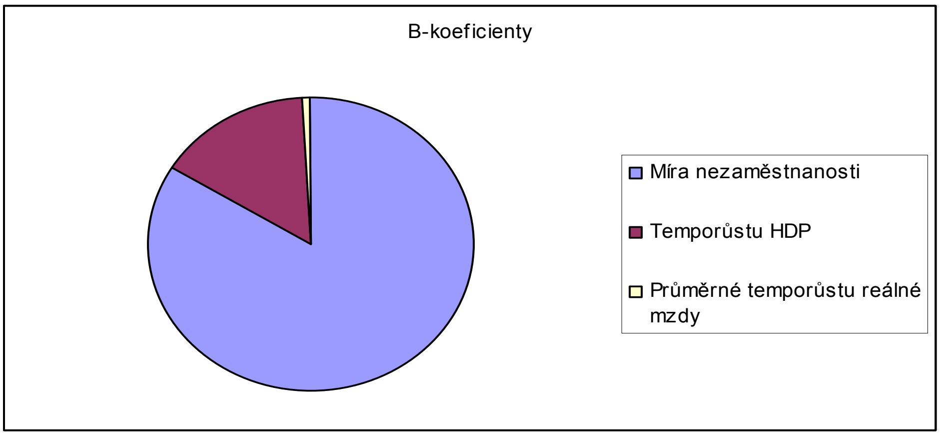
$$B_{yx_2 \cdot x_1 x_3} = 129906{,}8 \cdot \frac{0{,}020921}{16573{,}58} = 0{,}163982,$$

$$B_{yx_3 \cdot x_1 x_2} = 4181{,}283 \cdot \frac{0{,}031271}{16573{,}58} = 0{,}007889.$$

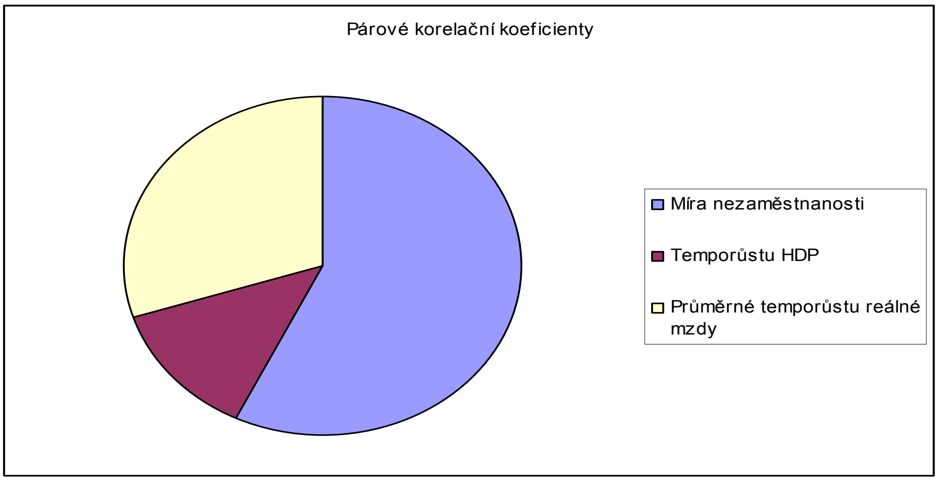
Zatímco dílčí regresní koeficienty jsme vzájemně nemohli srovnávat, vypočtené B-koeficenty již srovnatelné jsou. Výsledky ukazují, že na saldo systému sociálního zabezpečení nejvíce působí míra nezaměstnanosti, tempo růstu HDP má výrazně menší vliv. Z výsledku dále vyplývá, že tempo růstu reálné mzdy na saldo systému sociálního zabezpečení téměř nepůsobí, nemá v modelu význam, viz obr. č. 2.

Hodnoty párových korelačních koeficientů zjištěných dle (9) potvrdily, tak jako B-koeficienty, vhodnost zavedení míry nezaměstnanosti do modelu a poměrně malý význam tempa růstu HDP v modelu (viz obr. č. 3).

Obr.č. 2: Podlí vlivu vybraných makroekonomických veličin na vývoj systému sociálního zabezpečení (dle B-koeficientů)



Obr.č. 3: Podlí vlivu vybraných makroekonomických veličin na vývoj systému sociálního zabezpečení (dle párových korelačních koeficientů)



B-koeficienty ukázaly velmi malý význam průměrného tempa růstu reálných mezd, což je však poněkud v rozporu se zjištěním, které poskytly párové korelační koeficienty. Z dosažených dat není možné rozhodnou o vhodnosti zavedení dané vysvětlující proměnné do modelu.

Řešení tohoto problému lze nalézt po prozkoumání *multikolinearity*. Matice korelačních koeficientů je

$$R = \begin{pmatrix} 1 & -0{,}03087 & -0{,}40606 \\ -0{,}03087 & 1 & 0{,}603993 \\ -0{,}40606 & 0{,}603993 & 1 \end{pmatrix}, \tag{12}$$

determinant korelační matice je $|R| = 0{,}4844986$. Žádný z koeficientů korelační matice vysvětlujících proměnných nepřekročil hodnotu 0,75, je možné se tedy domnívat, že multikolinearita není v modelu škodlivá. Exaktnější zjištění přináší Farrarův-Glauberův test. Dle (10) je

$$B = -\left[(11-1) - \frac{1}{6}(2 \cdot 3 + 5)\right] \cdot \ln|0{,}4844986| = 5{,}9179 .$$

Kritická hodnota $\chi^2_{0.95}[3] = 7{,}8147$.

Protože $B = 5{,}9197 < 7{,}8147$, to znamená, že testové kritérium se nenalézá v kritickém oboru, přijímáme hypotézu H_0 o nezávislosti vysvětlujících proměnných a na 5% hladině významnosti nepovažujeme multikolinearitu za statisticky významnou. Zařazení proměnné x_3 (průměrné tempo růstu reálných mezd) do modelu je tedy vhodné.

4. Závěr

Příspěvek se zabýval posouzením vlivu vývoje míry nezaměstnanosti, tempa růstu HDP a tempa růstu reálných mezd na vývoj salda systému sociálního zabezpečení. Závislost byla sledována prostřednictvím regresní a korelační analýzy. Byla zvolena lineární regresní funkce, která podle indexu determinace vystihuje sledované závislosti z 80%, takže se dá hovořit o poměrně těsné závislosti.

Dosažené výsledky ukázaly, že systém sociálního zabezpečení je nejvíce ovlivněn *mírou nezaměstnanosti* (při vzrůstu míry nezaměstnanosti o 1% by saldo systému pokleslo o téměř 780 000 mil. Kč). *Tempo růstu HDP* ovlivňuje saldo systému sociálního zabezpečení menší měrou. Saldo systému je nejméně ovlivněno změnami *tempa růstu reálných mezd*. Původní domněnku o nevhodnosti zařazení veličiny do modelu vyvrátil Farrarův-Glauberův test, který vhodnost zařazení této veličiny do modelu prokázal.

Literatura

[1] BEZDĚK, V. Penzijní systémy obecně i v kontextu české ekonomiky (současný stav a potřeba reforem). ČNB, Sekce měnová, Praha, 2000.

[2] GRIFFITH – JONES, S. et al. Reforms of the pension system and national savings. Czech National Bank, Prague, 1998.

[3] HINDLS, R., HRONOVÁ, S., SEGER, J.: Statistika pro ekonomy. Professional publishing, Praha, 2003.

[4] KREBS, V a kol. Sociální politika. ASPI Publishing, s.r.o., Praha, 2002.

[5] Statistická ročenka České republiky. Český statistický úřad, Praha, 1993 – 2003.

[6] www.cssz.cz

[7] www.czso.cz

Summary

Examination of the influence of selected macroeconomic quantities upon social security system development

The article deals with a review of the dependance of the financial development of the Social security system in the Czech Republic on some macroeconomic variables (the unemployment rate, the GDP growth rate, the average real wages growth rate). First the attention is paid to the mathematical definition of the multiple regression and correlation analysis which is used for a description of the statistical dependences. Further there are real dates applied. Finding results are demonstrated and interpreted.