



Ekonomická
fakulta
Faculty
of Economics

Jihočeská univerzita
v Českých Budějovicích
University of South Bohemia
in České Budějovice

Jihočeská univerzita v Českých Budějovicích
Ekonomická fakulta
Katedra aplikované matematiky a informatiky

Diplomová práce

Ekonometrická analýza demografických aspektů vývoje ekonomiky

Vypracoval: Bc. Václav Quang Le
Vedoucí práce: Mgr. Michal Houda, Ph.D.

České Budějovice 2022

JIHOČESKÁ UNIVERZITA V ČESKÝCH BUDĚJOVICÍCH

Ekonomická fakulta

Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: Bc. Václav Quang LE
Osobní číslo: E20451
Studijní program: N0488A050006 Finance a účetnictví
Studijní obor:
Téma práce: Ekonometrická analýza demografických aspektů vývoje ekonomiky
Zadávající katedra: Katedra aplikované matematiky a informatiky

Zásady pro vypracování

Cílem práce je uplatnit ekonometrické metody za účelem vyhodnocení demografických aspektů (např. vývoje porodnosti či procesu stárnutí populace) na vývoj ekonomiky, financí a souvisejících oblastí.

Metodický postup:

1. Vyhledání literárních pramenů související se studovanou problematikou, studium odpovídajících ekonometrických a demografických metod, sběr dat.
2. Konstrukce vhodných ekonometrických modelů a odhad jejich parametrů.
3. Vyhodnocení výsledků a jejich ekonomická interpretace.
4. Závěr a diskuse.

Rozsah pracovní zprávy: 50 – 60 stran
Rozsah grafických prací: dle potřeby
Forma zpracování diplomové práce: tištěná

Seznam doporučené literatury:

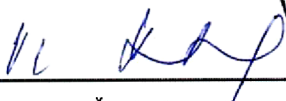
1. ALHO, J., & SPENCER, B. D. (2005). Statistical demography and forecasting. Springer series in statistics. New York: Springer.
2. CIPRA, T. (2013). Finanční ekonometrie. 2., upravené vydání. Praha: Ekopress.
3. WOOLDRIDGE, J. M. (2016). Introductory econometrics: A modern approach. 6th edition. Boston: Cengage Learning.

Další odborná časopisecká a knižní literatura dle konkrétního zaměření práce.

Vedoucí diplomové práce: **Mgr. Michal Houda, Ph.D.**
Katedra aplikované matematiky a informatiky

Datum zadání diplomové práce: 25. ledna 2021
Termín odevzdání diplomové práce: 15. dubna 2022

JIHOČESKÁ UNIVERZITA
V ČESKÝCH BUDĚJOVICÍCH
EKONOMICKÁ FAKULTA
Studentská 13 (1)
370 05 České Budějovice



doc. Dr. Ing. Dagmar Škodová Parmová
děkanka



doc. RNDr. Tomáš Mrkvička, Ph.D.
vedoucí katedry

Prohlašuji, že svou diplomovou práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

16. září 2022

Bc. Václav Quang Le

Poděkování

Touto cestou chci poděkovat svému vedoucímu práce Mgr. Michalovi Houdovi, Ph.D. za vedení práce, užitečné rady a flexibilní přístup. Poděkování náleží rovněž mé rodině, bez které by tato diplomová práce nebyla dokončena.

Obsah

1	Úvod	11
2	Celosvětový vznik a vývoj penzijních systémů	13
3	Penzijní systém ČR	17
3.1	Krátká historie penzijního systému	17
3.2	Struktura penzijního systému	19
3.3	Důchodový systém	19
3.4	Penzijní fondy	20
4	Demografický vývoj obyvatelstva ČR	21
4.1	Věková struktura obyvatelstva	21
4.2	Zdravá délka života	22
5	Metodika	25
5.1	Cíle práce	25
5.2	Data	25
5.3	Statistický software	26
5.4	Statistické metody	26
5.4.1	Vícenásobná lineární regrese	26
5.4.2	Problém zdánlivé regrese	27
5.4.3	Statistické testy	27
5.5	Hodnocení předpovědí modelů	29
6	Analýza dat	31
6.1	Závislé proměnné	31

6.1.1	Příjmy na důchodové pojištění	31
6.1.2	Výdaje systému důchodového pojištění	33
6.1.3	Úprava dat	34
6.2	Nezávislé proměnné	39
6.2.1	Počet ekonomicky aktivních, zaměstnaných a nezaměstnaných .	39
6.2.2	Hrubý domácí produkt	40
6.2.3	Index spotřebitelských cen	40
6.2.4	Náhrady zaměstnancům a čistý smíšený důchod	41
6.2.5	Počet důchodců	42
6.2.6	Průměrná mzda	43
6.3	Jednotkový kořen	43
7	Příjmové modely	45
7.1	Příjmový model 1	47
7.2	Příjmový model 2	49
7.3	Příjmový model 3	51
7.4	Shrnutí	53
8	Výdajové modely	55
8.1	Výdajový model 1	56
8.2	Výdajový model 2	58
8.3	Výdajový model 3	60
8.4	Shrnutí	62
9	Hodnocení modelů a předpovědi	65
9.1	Vyhodnocování představených modelů	65
9.1.1	Příjmové modely	65
9.1.2	Výdajové modely	67
9.2	Výběr a předpověď modelů	68
9.2.1	Výběr modelů pro predikci	68
9.2.2	Predikce příjmů a výdajů	68
10	Závěr	73

Kapitola 1

Úvod

Důchodový systém ČR je postaven na principu solidarity mezi starší a mladší generací. Mladší přispívají tedy na důchod starším. Jenže tento systém stojí na předpokladu, že bude existovat více mladých než starších v důchodu. V následujících desetiletích hrozí přehození počtu mladší a starší generace, které je způsobeno vlivem nízké mortality a porodnosti, které s sebou přináší nová doba. Výsledkem je pyramidová věková struktura stojící na hlavě. Za několik desítek let odejde jedna z nejpočetnějších generací do důchodu, což určitě vytvoří velký tlak na veřejné finance. Hrozí nedostatek prostředků pro vyplácení starobního důchodu, které by se projevilo ve snížení vyplácených dávek, což může přivést nepřipravenou generaci do chudoby. Tlak na veřejné finance by měl zmírnit třetí pilíř doplňkové penzijní spoření. Jeho nevýhodou je, že výnosy jsou úměrné vloženým prostředkům. To znamená, že sociálně slabší si nemusí tolik našetřit. Navíc je dle reportu OECD (2020) reálná vnitřní míra návratnosti záporná.

Cílem diplomové práce je ekonometrická analýza demografických aspektů vývoje ekonomiky. Předmětem analýzy se stala udržitelnost prvního pilíře penzijního systému. Analýza je provedena z pohledu příjmů z pojistného a příslušenství na důchodové pojištění a výdajů na dávky důchodového pojištění prostřednictvím techniky vícenásobné lineární regrese. Dílčím cílem je na základě nalezených regresních modelů předpovědět tři následující čtvrtletí do budoucnosti.

Teoretická část práce se nejprve zabývá všeobecným vývojem systému sociální ochrany a penzijního systému, poté následuje penzijní systém ČR a demografický vývoj ČR. Praktická část se věnuje regresním modelům, testování modelů v nekrizovém a krizovém období a nakonec je provedena predikce příjmů a výdajů.

Kapitola 2

Celosvětový vznik a vývoj penzijních systémů

Rozvoj systémů sociální ochrany a národních penzijních systémů probíhal nezávisle na jednotlivých územích. Jejich organizace a fungování se odvíjelo od, na území specifických, politických, ekonomických a sociálně-kulturních podmínek. Penzijní systémy spolu se sociálními službami jsou součástí systému sociální ochrany a sociálního zabezpečení. Mezinárodní organizace práce definuje systém sociální ochrany jako formu pojištění jedince, které zajišťuje základní práva jako právo na zdravotní péči, právo na finanční zabezpečení dětí, starých a lidí neschopných pracovat kvůli nemoci či postižení (Jimon, Dumiter & Baltés, 2021).

První formy sociální ochrany vznikaly již ve starověku. První zmínky o systému sociální ochrany vedou do Egypta. Zde existoval fond financovaný kameníky, ze kterého byla jejím členům poskytována finanční a materiální pomoc v případě sociálních rizik. V Římské říši byla sociální ochrana nejchudší vrstvě společnosti zajištěna bezplatným rozdělováním potravin. Gaius Julius Caesar jako konzul uděloval chudým občanům s alespoň třemi dětmi a válečným veteránům vlastnická práva k určitým pozemkům (Jimon et al., 2021).

Vývoj moderních systémů sociální ochrany začal v době průmyslové revoluce a reagoval na kritiku kapitalistického ekonomického systému, který podporoval sociální nerovnost a chudobu. V této době přišel Adam Smith s konceptem ekonomiky, která by se dokázala samaregulovat. Na jeho myšlenky navázali další autoři, kteří došli k závěru, že zamezení sociální nerovnosti včetně chudoby a udržení ekonomického růstu se neobejde bez státní intervence. To nasměrovalo vývoj systémů sociální ochrany k dodržování sociálních práv občanů za účelem udržení ekonomického vývoje a růstu zemí (Jimon et al., 2021). Ginsburg (1979) na základě toho uvádí hlavní směry, jakými by měl stát vydat. Jedná se

o vzdělání, zdravotní služby, snižování chudoby, sociální pojištění a sociální služby.

Na konci 19. století nastal zásadní vývoj v oblasti sociální ochrany. Německo totiž jako první stát přijalo sociální politiku ohledně sociální ochrany. Kancléř Otto von Bismarck v této době zavedl penzijní systém a regulaci zdravotního a pracovního pojištění. Zavedený penzijní systém se ale týkal drobné menšiny, protože byl určen pro osoby nad 70 let. Součástí starobní penze byla i invalidní penze (Döring, 1997). Tento systém byl určen pro průmyslové pracovníky, kteří čelili různým ohrožením. Systém byl postaven na příspěvcích od pracovníků a na solidaritě mezi nimi. Předpokládá tedy výběr sociálních příspěvků od pracovníků a jejich redistribuci, dojde-li ke splnění kritéria věku a doby přispívání. Postupem času tento systém podchytil ostatní kategorie pracovníků a rozšiřoval se po Evropě (Jimon et al., 2021).

V polovině 20. století přinesl William Henry Beveridge nový pohled na systém sociální ochrany. Navrhoval finanční asistenci všem občanům, kteří se nacházeli v delikátních situacích. Ve výsledku se jedná o systém sociálního pojištění s celostátní působností. Rozpočet penzijního systému nepocházel z příspěvků, ale ze státního rozpočtu. Penze byla vyplácena tehdy, kdy byly splněny podmínky týkající se věku, příjmu nebo majetku. Tento systém navrhovaný Beveridgem byl zaveden Velkou Británií, Dánskem, Švédskem a dalšími státy. Navíc vedle toho tyto státy zavedli příspěvkový penzijní systém (Jimon et al., 2021).

Na počátku 20. století představuje americký ekonom Milton Friedman (1912) alternativu k financování penzijního systému. Autor byl odpůrcem myšlenky veřejné redistribuce prostředků a zastával myšlenku, že obstarávání finančních prostředků potřebných ve stáří bylo mělo být osobní odpovědností. Autor navrhoval, aby penzijní prostředky byly investovány na finančním trhu za účelem jejich zhodnocení. José Piñera zavedl tento systém v Chile, a tím došlo k celosvětovému rozšiřování tohoto systému (Jimon et al., 2021).

Vzhledem k ekonomické a demografické situaci navrhuje Světová banka více pilířový penzijní model. Jedná se o 5 pilířů, kterými jsou (Holzmann, Paul, Dorfman & Dorfman, 2008):

1. mandatorní nultý pilíř, který má za cíl snížit chudobu starších lidí, tedy je finančně zabezpečit. Tento pilíř zajišťuje, aby lidé s celoživotně nízkými příjmy včetně osob,

- kteřé se málo podílely na ekonomice, měli minimální úroveň ochrany;
2. mandatorní první pilíř založený na sociálních příspěvcích od poplatníků během pracovních činností. Tento pilíř reaguje na krátkozrakost jedinců, nízké příjmy, na chybějící finanční trhy či jejich rizika apod. Tato komponenta je obvykle financována na principu pay-as-you-go (PAYG), a tudíž podléhá demografickým a politickým rizikům;
 3. mandatorní druhý pilíř, který zahrnuje soukromě spravované penzijní fondy. Nevýhodou je, že účastníci mohou být vystaveni riziku fondu a trhu, různým vysokým poplatkům a riziku dlouhověkosti v případě jednorázové výplaty;
 4. dobrovolný třetí pilíř, který se podobá druhému pilíři. Má kompenzovat nepružnost ostatních pilířů a má stejné nevýhody jako druhý pilíř;
 5. nefinanční čtvrtý pilíř obsahující neformální podporu (např. rodinná podpora), jiné formální sociální programy (jako zdravotní péče) a jiná individuální finanční a nefinanční aktiva (např. vlastnictví domů).

Podle metody financování systému rozlišujeme průběžné financování (pay-as-you-go, PAYG), kam patří systém navrhovaný Bismarckem, a fondové financování, které je charakteristické pro systém navrhovaný Friedmanem. Systémy založené na příspěvcích můžeme dále dělit na dávkově (DB) a příspěvkově (DC) definované systémy. Rozlišují se pouze v použité bázi pro výpočet výše dávky. U DB systémů závisí výše dávek na výši výdělku, ze kterého jsou odváděny příspěvky, zatímco u DC systémů je výše dávek ovlivněna nakumulovaným kapitálem v penzijním fondu. Kromě těchto tradičních mechanismů existují další varianty, které kombinují rysy DB a DC. Mezi nejznámějšími patří systém založený na penzijních bodech a notional defined contributions (NDC). V případě prvního jmenovaného systému je výše dávek stanovena na základě nakumulovaných penzijních bodů, které korelují s výší výdělků, a hodnotě těchto bodů při odchodu do důchodu (Jimon et al., 2021). NDC kombinuje DC s PAYG. Systém funguje tak, že občané přispívají na fiktivní spořicí účet, ze kterého jsou vypláceny dávky. Odvedené částky se zhodnocují mírou stanovenou vládou. Fiktivní účty jsou užívány z toho důvodu, že odvedené příspěvky jsou využity na výplatu současných dávek (Bezděk, 2000).

Kapitola 3

Penzijní systém ČR

3.1 Krátká historie penzijního systému

Co se týče systému sociální ochrany, tak první regulace, která se toho týkala, je datována k roku 1887 a zaměřovala se na nehody pracovníků při výkonu práce. V následujícím roce, tedy v 1888, byly dále zavedeno zdravotní pojištění pro nemoc a mateřství. Až o 57 let později, v roce 1945, byly představeny rodinné příspěvky. A o dalších 46 let byly implementovány dávky v nezaměstnanosti (Jimon et al., 2021).

Penzijní systém byl založen na principech kancléře Bismarcka, tedy na sociální solidaritě a na přispívání. Podle toho se orientuje organizace a fungování systému. Systém byl financován na základě příspěvků, které financovali zaměstnanci a zaměstnavatelé. Dávky byly vypláceny po splnění příslušného kritéria, v případě starobní penze se jednalo o důchodový věk (Jimon et al., 2021). Penzijní systém byl převzat z Rakousko-Uherska. Významným rokem byl rok 1924, kdy bylo zavedeno starobní a invalidní pojištění dělníků. Poté přibýly další formy důchodů (Klepárník, Štípek, Bittner & Pícl, 2017). V tom samém roce byl důchodový věk stanoven na 65 let (Jimon et al., 2021).

Významným milníkem byl také rok 1948, kdy byl přijat zákon č. 99/1948 Sb, o národním pojištění, který stanovil odvod ve výši 10 % vyměřovacího základu zaměstnance do systému. Toto pojistné měl platit zaměstnavatel. Než zákon nabyl účinnost bylo vydáno rozhodnutí, podle kterého se zaměstnavatel a zaměstnanec podíleli každý 50 % na platbě pojistného (Klepárník et al., 2017).

V roce 1956 byl snížen důchodový věk na 60 let u mužů a na 55 let u žen (Jimon et al., 2021). O 8 let později byla představena diferenciací důchodového věku u žen na 53 až 57 let v závislosti na počtu vychovaných dětí. V tomto roce byla také zavedena progresivní daň z důchodu, která byla následně o 10 let později zrušena (Klepárník et al., 2017).

Na počátku 90. let byly provedeny reformy, které si kladly za cíl sjednotit celý systém po organizační stránce a přizpůsobit existující systém sociálního zabezpečení tehdejšími podmínkami tržního hospodářství. V této době byl přijat zákon č. 210/1990 Sb., o o změnách v působnosti orgánů České republiky v sociálním zabezpečení a o změně zákona č. 20/1966 Sb., o péči o zdraví lidu, na jehož základě vznikla Česká správa sociálního zabezpečení (ČSSZ), která má působnost v oblasti sociálního zabezpečení. V roce 1991 došlo k implementaci pravidelné valorizace důchodů, indexace činila 10 % inflace a růstu reálných mezd (Klepárník et al., 2017).

V tabulce 3.1 jsou uvedeny některé změny penzijního systému ČR po 1993 s důrazem na důchodový systém.

Tabulka 3.1: Některé změny penzijního systému

Rok	Změna
1994	Zavedení penzijního připojištění
1995	Přijetí zákona č. 155/1995 Sb., o důchodovém pojištění, které zavedlo základní a procentní výměru důchodu.
2010	Doba pojištění pro vznik nároku na starobní důchod se zvyšuje z 25 let na 35 let. Důchodcům, kteří vedle starobního důchodu vykonávají výdělečnou činnost, se zvyšuje procentní výměra o 0,4 % výpočtového základu za každých 360 dní při pobírání plné výše důchodu nebo o 1,5 % za každých 180 dní při pobírání poloviční výše.
2011	Jednoznačně byly stanoveny redukční hranice ve vazbě na průměrnou mzdu. Zároveň se zvyšuje důchodový věk o 2 měsíce pro každou mužskou kohortu a až o 6 měsíců pro ženskou kohortu od roku 2019 bez horního limitu. Dále se mění redukční hranice procentní výměry předčasného důchodu.
2013	Zavedení důchodového spoření (2. pilíř) a nahrazení penzijního připojištění doplňkovým penzijním spořením.
2015	Ukončení důchodového spoření.
2018	Zastropování důchodového věku na 65 let.
2019	Základní výměra důchodu se zvyšuje z 9 % na 10 % a zároveň se zvyšuje procentní výměra o 1 000 Kč měsíčně pro důchodce nad 85 let.

2022	Zavádění tzv. výchovného, které zvyšuje procentní výměru starobního důchodu o 500 Kč, pokud důchodce pečoval o dítě do jeho zletilosti.
------	---

Zdroj: Klepárník et al. (2017), MPSV (2022)

3.2 Struktura penzijního systému

V současné době funguje 1. a 3. pilíř penzijního systému. První pilíř představuje dávky starobního důchodu, které jsou součástí důchodového systému. Vedle těchto dávek je důchodový systém tvořen dále invalidní a pozůstalostní penzí. Tyto uvedené dávky spadají pod působnost ČSSZ, která podřízena Ministerstvu práce a sociálních věcí ČR (MPSV). Na druhé straně máme soukromě spravované penzijní fondy, kam patří třetí pilíř nazývaný doplňkové penzijní spoření, které nahradilo penzijní připojištění prostřednictvím transformovaných fondů. Tyto fondy patří do působnosti Ministerstva financí ČR.

3.3 Důchodový systém

Důchodový systém je primárně financován placeným sociálním pojištěním vypočítaném z vyměřovacího základu. V případě závislé činnosti se jedná zejména o mzdu či plat a v případě OSVČ rozdíl mezi příjmy a výdaji.

Dávky důchodového systému upravuje zákon č. 155/1995 Sb., o důchodovém pojištění. Podle něho nárok na starobní důchod vzniká splněním dvou podmínek, a to dosažením starobního věku a získáním potřebné doby důchodového pojištění.

Invalidní důchod se dle stupně invalidity rozděluje na tři stupně. Pozůstalostní důchod, ke kterému patří vdovský, vdovecký a sirotčí důchod, se poskytuje osobám, jejichž partner/ka nebo rodič zemřeli a současně pobírali nebo měli nárok na ostatní důchody.

Výše důchodu se skládá ze základní a procentní výměry. Základní výměra je stejná pro všechny důchody, zatímco procentní výměra se odvíjí od výše dosažených výdělků a rovněž od délky doby důchodového pojištění. Důchody se zvyšují vždy k 1. lednu, ledaže růst cen překročí hranici 5 %. Základní výměra se zvyšuje dle reálné průměrné mzdy. Naproti tomu se procentní výměra zvyšuje dle vyššího z indexu spotřebitelských cen za domác-

nosti celkem nebo za domácnosti důchodců a z poloviny dle reálné průměrné mzdy.

3.4 Penzijní fondy

V roce 1994 bylo zavedeno penzijní připojištění zákonem č. 42/1994 Sb., o penzijním připojištění se státním příspěvkem. Takto se ČR zařadila mezi prvními transformujícími se ekonomikami, které zaváděli třetí pilíř. Zpočátku se systém jevil jako úspěšný, ale s postupem času stagnoval, protože průměrné příspěvky byly nízké a zároveň neexistovaly známky, že by je účastníci chtěli zvyšovat. Charakteristické pro tyto fondy bylo, že byla garantováno nezáporné zhodnocení (Klepárník et al., 2017).

Od roku 2013 byl tento systém nahrazen doplňkovým penzijním spořením, které se odlišuje především tím, že účastníci nemají garantováno nezáporné zhodnocení, ale jako kompenzaci mají účastníci každoročně bezplatnou možnost vybírat investiční strategii. Tyto fondy jsou upraveny zákonem č. 427/2011 Sb., o doplňkovém penzijním spoření.

Dle reportu OECD (2020) činí průměrná nominální míra návratnosti penzijních fondů za roky 2014–2019 0,9 %, zatímco reálná míra -1 %. Reálná vnitřní míra návratnosti za rok 2019 je $-1,4$ %. Důvodem nízké míry návratnosti penzijních fondů je investiční portfolio, které se z velké části skládá z vládních dluhopisů a bankovních vkladů. Celkem tyto položky tvoří 95,4 % všech investic. Míra účasti na dobrovolných penzijních fondech je v roce 2019 64,11 %.

Kapitola 4

Demografický vývoj obyvatelstva ČR

V této kapitole se práce bude věnovat demografickému vývoji obyvatelstva v ČR. První část se orientuje na věkovou strukturu a na střední dobu života a druhá část na zdravou délku života.

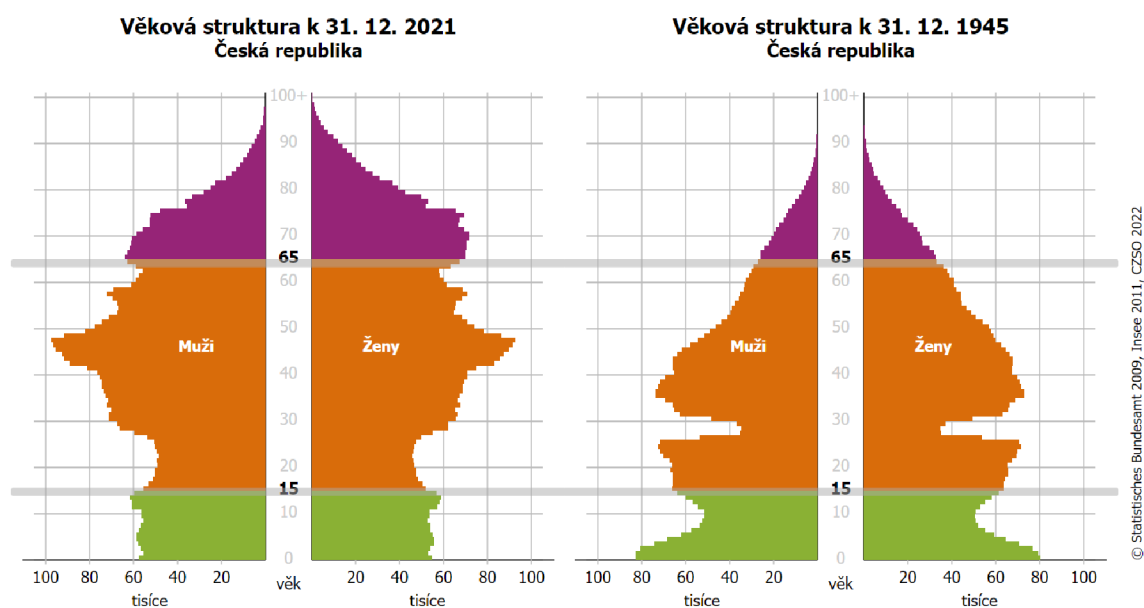
4.1 Věková struktura obyvatelstva

Sundbärg rozděluje populaci podle věku do tří základních skupin, kterými jsou dětská, ekonomicky aktivní složka a složka lidí v důchodovém věku. V závislosti na velikosti dětské složky a složky starobních důchodců se rozlišují různé typy věkových struktur obyvatelstva (Klufová & Poláková, 2010).

Na obrázku 4.1 lze spatřit věkové struktury po jednotlivých letech na konci roku 2021 a roku 1945. Porovnáme-li obě věkové struktury, zjistíme, že dětská složka (zelená) je menší než ostatní složky. To je typické pro regresivní věkovou strukturu, kdy porodnost klesá, čímž dochází k stále nižšímu podílu mladého obyvatelstva. Věková struktura z roku 1945 má proti tomu pyramidový tvar, protože dětská složka je širší než ostatní složky. Tato podoba je obvyklá pro progresivní věkovou strukturu, kdy natalita je vyšší než mortalita. Lidé se dnes také dožívají vyššího věku, jak lze spatřit z obrázku. V porovnání s minulostí se dnes o hodně víc lidí dožívá vyššího věku.

Malý podíl mladých lidí v předproduktivním věku vysvětlují Klufová a Poláková tím, že v dnešní době vznikly poprvé podmínky umožňující zcela svobodné životní rozhodování, které není tak ovlivněno společenským postavením, příjmem či majetkem. Dnes mají lidé široko škálu možností se seberealizovat, diferencované možnosti získávání vzdělávání a kvalifikace. Taková svoboda se také odráží v rodinném chování. Dříve byla životní kariéra taková, že lidé měli omezenější přístup ke vzdělání a zároveň v mladém věku uzavírali sňatek, brzy následovaným narozením převážně dvou dětí po sobě (Klufová & Poláková,

Obrázek 4.1: Věková struktura ČR v letech 2021 a 1945



Zdroj: ČSÚ (2022)

2010). Šprocha (2022) vysvětluje odkládání porodů do vyššího věku jednak na základě negativních jevů souvisejících s transformací v 90. letech (např. inflace, nezaměstnanost, ztráta jistot apod.), ale také restrukturalizací a rostoucí soutěživostí na trhu práce, potřebou zvyšování kvalifikace a flexibility, růstem nejistoty v zaměstnání apod. Dále uvádí, že odkládání může být způsobeno změnami norem, hodnot k reprodukci a rodině. Autor rovněž zjistil, že v nejmladších generacích dochází ke stagnaci míry odkládání a k utlumení mezigeneračního zvyšování objemu odkládání porodů.

Velký počet starobních důchodců je dán vyšší střední dobou života neboli nadějí na dožití. V posledních 40 letech se střední doba života zvýšila v průměru o zhruba 15 let. Nejvíce se na růstu podílela kvalitnější zdravotní péče (Montero & Simões, 2021). Hübelová, Manea a Machučová (2021) dodávají, že k růstu hodnoty také přispívá pozitivní vývoj kvality života a životních podmínek.

4.2 Zdravá délka života

Zdravá délka života je ukazatel vydávaný Eurostatem, který vyjadřuje zbývající počet let, který může osoba prožít v dobrém zdraví. Obecně vyšší hodnota ukazatele vyjadřuje zdravější pracovní sílu, méně odchodů z povolání kvůli zdravotním problémům a sníženou či

odloženou potřebu pro dlouhodobou péči. Nevýhodou tohoto ukazatele je, že vychází z vlastního subjektivního vnímání zdraví osob. To znamená, že data mohou být ovlivněná subjektivitou osob, ale také sociálními a kulturními faktory. Například lidé s nižším vzděláním mají větší tendenci nahlásit zdravotní omezení než lidé s vyšším vzděláním (OECD & Unie, 2020).

Kapitola 5

Metodika

5.1 Cíle práce

Diplomová práce si klade za cíl ekonometricky analyzovat demografické aspekty vývoje ekonomiky. Analyzována bude udržitelnost systému starobního důchodu. Udržitelnost je zkoumána z pohledu příjmů a výdajů. K dosažení stanoveného cíle bude využita technika vícenásobné lineární regrese, pomocí níž budou stanoveny tři modely pro každou zkoumanou proměnnou. Dílčím cílem práce je pomocí nalezeného regresního modelu predikovat 2. až 4. čtvrtletí roku 2022.

5.2 Data

Data mají charakter časové řady a začínají prvním čtvrtletím roku 2000 a končí prvním čtvrtletím roku 2022. Druhé čtvrtletí roku 2022 nebylo zahrnuto do analýzy, protože u některých časových řad chyběla data. Periodicita je jedno čtvrtletí. Data jsou čerpána z veřejně dostupných statistik České správy sociálního zabezpečení, z veřejné databáze a z databáze národních účtů Českého statistického úřadu. Bližší informace ohledně zdrojů dat jsou uvedeny na konci každé podkapitoly analýzy dat. Data pro predikci na základě nalezeného modelu byla získána z makroekonomické predikce vydávanou Ministerstvem financí ČR (dále jen MF).

5.3 Statistický software

Pro statistické zpracování dat je využít statistický program R spolu s open-source vývojovým prostředím RStudio. Pro statistickou analýzu jsou využity standardní a některé rozšiřující balíčky, konkrétně `car`, `lmtest`, `sandwich` a `fpp3`. Součástí posledně vyjmenovaného jsou následující balíčky:

- `tibble`
- `dplyr`
- `tidyr`
- `lubridate`
- `ggplot2`
- `tsibble`
- `tsibbledata`
- `feasts`
- `fable`

5.4 Statistické metody

5.4.1 Vícenásobná lineární regrese

K analýze demografických a ostatních veličin, které ovlivňují systém starobního důchodu, se v práci se bude aplikovat technika vícenásobné lineární regrese. S tím nástrojem souvisí několik předpokladů, které umožňují interpretovat určité aspekty modelu. Nejsou-li však splněny, jsou odhady parametrů systematicky vychýlené vzhledem k populačním parametrům. Níže jsou uvedeny předpoklady podle Wooldridge (2009), pokud se lineární regrese aplikuje na časové řady.

Předpoklady při konečném počtu pozorování

1. Model je lineární v parametrech
2. Žádná perfektní kolinearita nezávislých proměnných
3. Nulová podmíněná střední hodnota reziduí
4. Homoskedasticita reziduí
5. Žádná autokorelace reziduí

Pokud platí předpoklady 1 až 3, nejsou odhadované parametry systematicky vychýlené. Za platnosti předpokladů 4 a 5 nevykazuje statistická inference modelu systematické vychýlení. Homoskedasticita reziduí bude testována pomocí Breuschova–Paganova testu a autokorelace reziduí prostřednictvím Ljungova–Boxova testu.

Heteroskedasticitu lze vyřešit Whiteovými standardními chybami robustními vůči hete-

roskedasticitě. Autokorelace reziduí lze korigovat v případě autokorelace prvního řádu metodou Cochrane–Orcutt nebo Prais–Winsten. Pro korigování inference se lze využít heteroskedasticitě a autokorelačně konstantní estimátory kovarianční matice (HAC) navrhované autory Newey & West (Wooldridge, 2015). Tyto vyjmenované metody jsou v práci aplikovány okrajově.

5.4.2 Problém zdánlivé regrese

Zdánlivá regrese (angl. spurious regression) nastává mimojiné v případě, kdy proměnné v regresním modelu nejsou stacionární. Zjednodušeně řečeno jsou data stacionární, pokud se časová řada pohybuje okolo nějaké střední hodnoty s konstantním rozptylem, neměnnou kovarianční strukturou a zároveň nevykazuje žádnou sezónnost. Tento problém lze odstranit např. zařazením časového trendu mezi nezávislé proměnné, první či sezónní diferencí dat apod. (Wooldridge, 2009). Pokud lze diferencí docílit stacionaritu časové řady, má tato řada s vysokou pravděpodobností jednotkový kořen (angl. unit root) (Cipra, 2013).

V práci se nestacionární data převedou na stacionární prostřednictvím první či sezónní difference, přičemž první diferencí se rozumí rozdíl mezi dvěma po sobě chronologicky následujícími datovými body. Hyndman a Athanasopoulos (2021) uvádějí, že pořadí diferencování není důležité. Avšak doporučují provést sezónní diferencii před první diferencí, když data vykazují silnou sezónnost.

Cipra (2013) doporučuje provést rozšířený Dickeyův–Fullerův test spolu s KPSS testem pro ověření existence jednotkového kořene. Těmto a ostatním testům je věnována následující podkapitola.

5.4.3 Statistické testy

K otestování normality reziduí regresního modelu se využije **Shapirův–Wilkův test**. Nulovou hypotézou je, že rezidua mají normální rozdělení. Naproti tomu alternativní hypotéza neguje tvrzení nulové hypotézy, tedy rezidua nemají normální rozdělení.

Ljungův–Boxův test patří mezi portmanteau testy a je přesnější variantou Boxova–Pierceova testu. Testová statistika Q^* Ljungova–Boxova testu se počítá dle níže uvedeného vzorce.

Nulovou hypotézou je, že rezidua nelze rozeznat od bílého šumu. Alternativní hypotéza na druhou stranu tvrdí opak. Za platnosti nulové hypotézy sleduje testová statistika $Q^* \sim \chi^2$ rozdělení o $(\ell - K)$ stupních volnosti, kde ℓ představuje rozsah řádu autokorelace, který má být otestován, a K počet parametrů v modelu. Hyndman a Athanasopoulos (2021) doporučují, aby ℓ nabývalo hodnoty $2m$, obsahují-li data sezónní složku. V tomto případě představuje m periodicitu dat. Práce se bude řídit doporučením autorů, tzn. že $\ell = 8$. Pokud však je $K \geq \ell$, použije se $\ell = 10$.

$$Q^* = T(T + 2) \sum_{k=1}^{\ell} (T - k)^{-1} r_k^2 \quad (5.1)$$

Breuschův–Paganův test patří mezi testy Langrangeových multiplikátorů (dále jen LM). Obecně se preferuje testová statistika Koenkera díky jeho větší použitelnosti (Wooldridge, 2009). Níže je uvedena testová statistika LM, kterou navrhl Koenker.

$$LM = n \cdot R_{\hat{u}^2}^2 \quad (5.2)$$

Rozšířený Dickeyův–Fullerův test (dále jen ADF) ověřuje nulovou hypotézu, že v autoregresní časové řadě je přítomen jednotkový kořen. Tento test má tři varianty alternativní hypotézy, že v autoregresní časové řadě není jednotkový kořen. První varianta předpokládá autoregresní proces s nulovou úrovní a bez lineárního časového trendu, druhá varianta proces s nenulovou úrovní a bez časového trendu a třetí varianta proces s nenulovou úrovní a časový trend. Tento test oproti klasickému testu řádně zohledňuje autokorelovanost proměnné (Cipra, 2013). V práci je stanoven řád autokorelace na 4 vzhledem k čtvrtletní periodicitě dat a podle 4. řádu se bude vyhodnocovat test.

KPSS test navrhli Kwiatkowski, Phillips, Schmidt a Shin (1992). Na rozdíl od ostatních testů ověřujících přítomnost jednotkového kořene předpokládá nulová hypotéza, že časová řada je stacionární kolem deterministického trendu. Naproti tomu předpokládá alternativní hypotéza přítomnost jednotkového kořene. Tento test byl navržen jako doplněk k (A)DF a k podobným testům, protože při klasickém testování se vychází z platnosti nulové hypotézy.

Hladina významnosti α u výše uvedených všech testů je stanovena na úrovni 0,05. Bude-li vypočítaná p-hodnota nižší než hladina významnosti α , bude nulová hypotéza zamítnuta ve prospěch alternativní hypotézy.

5.5 Hodnocení předpovědí modelů

U představených modelů bude hodnocena předpověď v krizovém a v nekrizovém období. Za nekrizové období se považuje období 2018–2019 a za krizové období 2020–2022 Q1. Pro otestování předpovědi v krizovém období sahají testová data do roku 2019 včetně. Pro hodnocení předpovědí v rámci nekrizového období zahrnují tréninková data období do roku 2018 včetně. Testovými daty jsou pak následující dva roky, resp. zbytek dat.

Bodová předpověď modelů bude hodnocena na základě směrodatné odchylky chyby (RMSE) a střední absolutní škálované chyby (MASE) a směrodatné odchylky škálované chyby (RMSSE). Intervalová předpověď modelů bude hodnocena prostřednictvím Winklerova skóre a také bude porovnávána s modelem, který využívá sezónní naivní metodu předpovědi.

Pro předpověď následujících tří čtvrtletí bude využit model, který vykázal nejlepší výsledek dle následujících kritérií: křížová validace („vynech jeden“, CV), korigované Akaikeho informační kritérium (AICc) a Schwarzovo bayesovské informační kritérium (BIC). Nejlepší předpovědní schopnost má ten model, který vykazuje nejnižší hodnoty kritérií. Hodnota CV je vyjádřena střední čtvercovou chybou (MSE). AICc se počítá na základě rovnice (5.3) a BIC rovnicí (5.4), kde T představuje počet pozorování použitých k odhadu modelu, k je počet nezávislých proměnných bez absolutního členu a SSE součet čtvercových chyb.

$$AICc = T \ln \left(\frac{SSE}{T} \right) + 2(k + 2) + \frac{2(k + 2)(k + 3)}{T - k - 3} \quad (5.3)$$

$$BIC = T \ln \left(\frac{SSE}{T} \right) + (k + 2) \ln(T) \quad (5.4)$$

Kapitola 6

Analýza dat

V této kapitole budou analyzována data, která budou využita v regresní analýze. Tato kapitola se skládá ze dvou částí. V první části se práce bude zabývat daty závislých proměnných, v druhé části daty nezávislých proměnných.

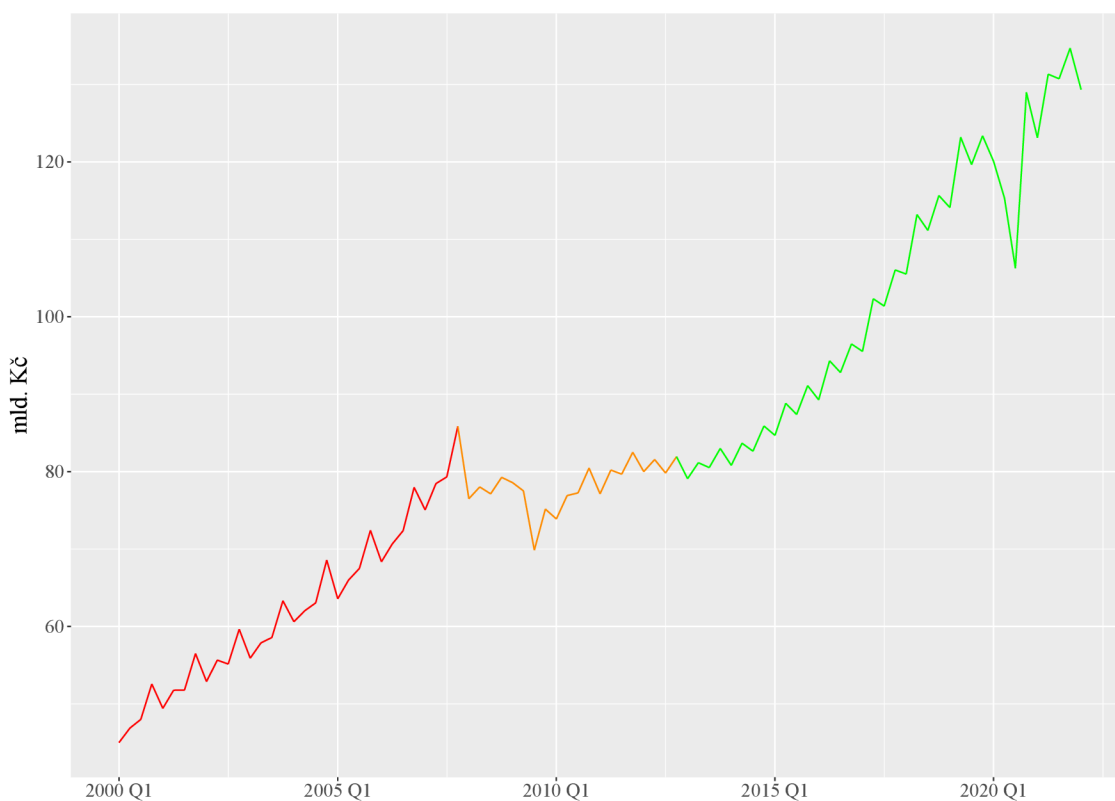
6.1 Závislé proměnné

V následujících odstavcích se práce bude věnovat analýze a úpravám závislých proměnných, kterými jsou příjmy na důchodové pojištění a výdaje na dávky důchodového pojištění.

6.1.1 Příjmy na důchodové pojištění

Vývoj příjmů z pojistného a příslušenství na důchodové pojištění zachycuje obrázek 6.1. Na obrázku lze spatřit rychlý počáteční růst příjmů až do roku 2008, kdy nastala celosvětová finanční krize. Příjmy se propadly a začaly mírně růst. Od roku 2013 začala česká ekonomika prudce růst, což se také projevilo na příjmech. Prudký růst zachycuje zeleně obarvená křivka. Velký propad příjmů lze spatřit téměř na konci křivky, kdy v roce 2020 nastala pandemie Covid-19. V tomto období česká vláda se snažila podpořit firmy a OSVČ tak, že jim umožnila odložit platby sociálního pojištění. Právě toto způsobilo obrovský propad příjmů důchodového pojištění.

Obrázek 6.1: Vývoj příjmů z pojistného a příslušenství na důchodové pojištění



Zdroj: ČSSZ, vlastní zpracování

Ačkoli zabarvení křivky perfektně, až na výjimku, koreluje s růstem a poklesem příjmů a vypadá, jakoby by byla křivka rozdělena na období před, v rámci a po celosvětové finanční krizi, plní obarvení křivky jinou funkci. Zabarvení se týká datového souboru. Zelená barva představuje data, u kterých byly k dispozici čtvrtletní data. U oranžové křivky byly pouze roční data k dispozici. Za období červené křivky nebyla nalezena data ohledně příjmů důchodového pojištění. Červená část křivky koresponduje s obdobím 2000–2007, oranžová část s obdobím 2008–2013 a zelená část s obdobím 2014–2022 Q1.

Data pro červenou a oranžovou křivku byla čerpána ze statistiky výběru pojistného na sociální zabezpečení. Tyto příjmy obsahovaly jednak příjmy na důchodové pojištění, tak i příjmy na nemocenské pojištění a příjmy z příspěvku na státní politiku zaměstnanosti. Z tohoto důvodu bylo rozhodnuto o úpravě dat.

Úprava dat na čtvrtletní data byla provedena následovně. Data zachycující vývoj příjmů a úspěšnosti v jednotlivých letech měla měsíční periodicitu. Příjmy za jednotlivé měsíce byly nejdříve nasčítány do příslušného čtvrtletí. Součet jednotlivých čtvrtletí se pak

porovnal s ročními příjmy. Pomocí podílů ročních příjmů na důchodové pojištění a ročních příjmů na sociální zabezpečení bylo získáno 5 hodnot. Tyto vypočítané hodnoty byly poté použity na přepočítání čtvrtletních příjmů sociálního zabezpečení na čtvrtletní příjmy důchodového pojištění tak, že se roční příjmy SZ vydělily vypočítaným podílem. Tento postup byl použit pro data představující oranžovou křivku. Pro transformaci červených dat byl využit obdobný proces. Odlišoval se tím, že vypočtené hodnoty se nejdříve zprůměrovaly. Uvedený proces a vypočítané podíly zachycuje tabulka A.1. Odchylka příjmů na důchodové pojištění od příjmů na sociální zabezpečení se pohybuje z podstatné části pod 0,30 %. Tabulka A.1 zahrnuje pouze přepočítání dat, které přísluší k oranžové křivce.

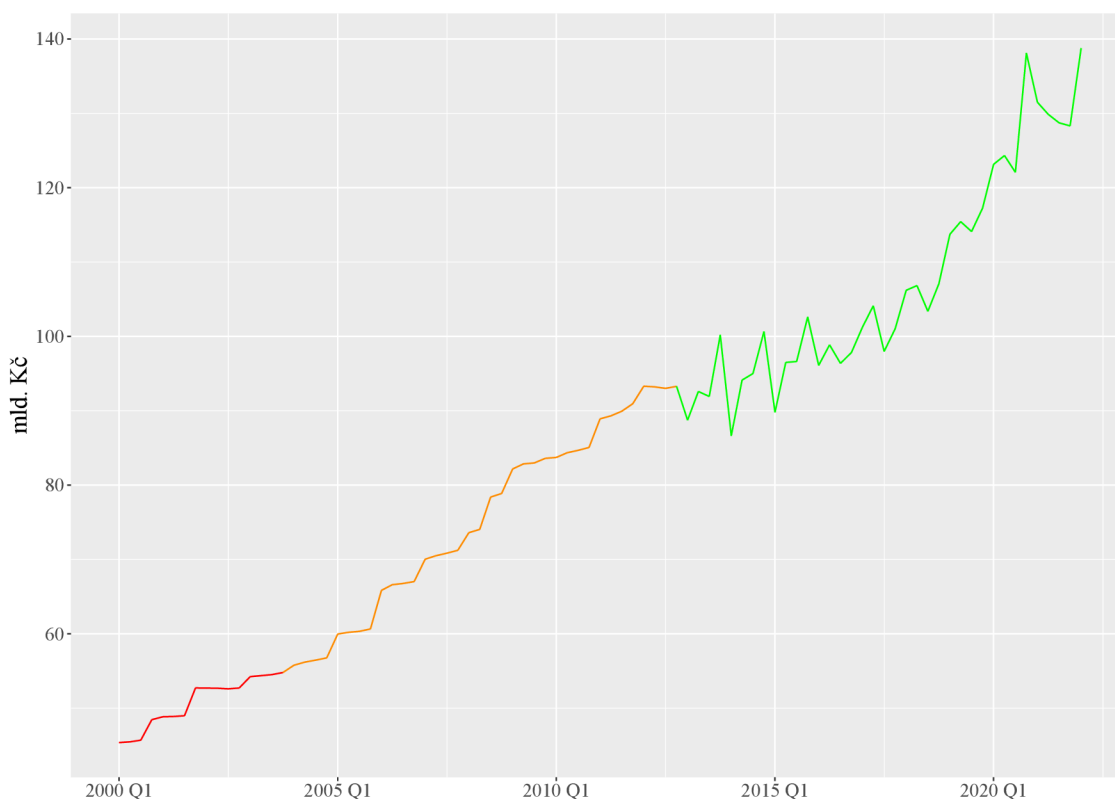
Co se týče zelených dat, tak byly získány z přehledu vybraných statistických ukazatelů z agend ČSSZ pod záložkou „Souhrnný přehled ukazatelů“.

6.1.2 Výdaje systému důchodového pojištění

Vývoj výdajů na dávky důchodového pojištění lze spatřit na obrázku 6.2. Z obrázku vyplývá, že výdaje od roku 2016 začaly růst prudkým tempem. Největší nárůst výdajů nastal ve 4. čtvrtletí roku 2022, kdy výdaje činily 138,1 mld. Kč.

Zabarvení křivky plní stejnou funkci jako u příjmové křivky. Červené části křivky připadá období 2000–2003, oranžové části 2004–2012 a zelené části 2013–2022 Q1. Chybějící data výdajů (oranžová a červená) byla získána následujícím způsobem. Počet důchodců se vynásobil trojnásobkem příslušné průměrné výše sólo důchodů (tj. jediné vyplácené důchody). Tento postup však nezahrnuje situace, kdy důchodce pobírá více důchodů současně. Pro srovnání vypočtených dat se využily výdaje ze statistických ročenek vydávaných ČSSZ. Tyto výdaje byly nejdříve upraveny, protože byly sestaveny na účetním principu věcné a časové souvislosti „nákladů“. Výdaje totiž zahrnovaly zálohu na výplatu důchodu z prosince minulého roku a naopak nezahrnovaly zálohu na výplatu důchodů na následující rok. Stejně jako příjmy zelené křivky byly výdaje získány z přehledu vybraných statistických ukazatelů z agend ČSSZ, které vice versa nezahrnují zálohu z minulého roku a zahrnují zálohu na následující rok. Následná úprava dat byla provedena stejným způsobem jako u příjmů s jediným rozdílem, že pro průměr vypočtených podílů bylo využito prvních pět. Tabulka A.2 zachycuje přepočítání dat za období prvních pěti let, které korespondují k oranžové křivce. Dle níže umístěné tabulky činí odchylka vypočítaných výdajů

Obrázek 6.2: Vývoj výdajů na dávky důchodového pojištění



Zdroj: ČSSZ, vlastní zpracování

od skutečných okolo 7 %. Největší odchylka nastala u roku 2006 a činí 9,21 %.

6.1.3 Úprava dat

Před regresní analýzou závislých proměnných jsou nejdříve tyto proměnné upraveny. Cílem úprav je snížit rozptyl dat nebo vytvořit jednodušší anebo konzistentnější vzory vývoje dat pro snadnější modelování a zpřesnění predikcí.

Hyndman a Athanasopoulos (2021) uvádí čtyři metody úpravy historických dat. Jedná se o úpravu kalendářních variací, počtu obyvatel, inflace nebo úpravu prostřednictvím matematických transformací. Bylo uvažováno o užití úpravy vzhledem ke kalendáři a matematické transformace.

Mezi matematické transformace patří logaritmická a mocninná transformace. Výhodou logaritmické transformace je, že změna logaritmické hodnoty se přibližně rovná relativní neboli procentní změně původní hodnoty. Pod mocninné transformace spadají odmocninová anebo kubická transformace. Ačkoli mohou být užitečné, dají se transformovaná

data stěží interpretovat.

Existují však modely, které zahrnují jak logaritmickou, tak i mocninnou transformaci. Jedním z nich je Boxova–Coxova transformace, která byla publikována v roce 1964 a která je pojmenovaná po jejích autorech. Tato metoda umožňuje převést rozdělení dat na normální rozdělení při správné hodnotě transformačního parametru λ . Transformace dat spočívá v tom, že se podle hodnoty lambda λ aplikují různé modely transformace. Původně Box a Cox (1964) navrhovali následující modely:

$$w_t = \begin{cases} \frac{y_t^\lambda}{\lambda} & \text{pokud } \lambda \neq 0; \\ \ln(y_t) & \text{pokud } \lambda = 0. \end{cases} \quad (6.1)$$

Z rovnice (6.1) vyplývá, že při $\lambda \neq 0$ se na data aplikuje mocninná transformace a při $\lambda = 0$ přirozený logaritmus. Ten je i nevýhodou tohoto modelu, protože se nedá aplikovat na záporné hodnoty y_t . Je-li však $\lambda = 1$, nedochází k žádné transformaci a data si zachovávají svůj původní tvar. Další model dokáže pracovat se zápornými hodnotami y_t , ale to platí pro hodnoty y_t vyšší než $-\lambda_2$.

$$w_t = \begin{cases} \frac{(y + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & \text{pokud } \lambda_1 \neq 0; \\ \ln(y + \lambda_2) & \text{pokud } \lambda_1 = 0. \end{cases} \quad (6.2)$$

V práci se bude využívat model navrhovaný autory Bickel a Doksum (1981), který je uveden níže. Výhodou tohoto modelu nad výše uvedenými je, že ho lze uplatnit na záporné hodnoty y_t za předpokladu, že $\lambda \neq 0$.

$$w_t = \begin{cases} \ln(y_t) & \text{pokud } \lambda = 0; \\ (\text{sign}(y_t)|y_t|^\lambda - 1)/\lambda & \text{pokud } \lambda \neq 0. \end{cases} \quad (6.3)$$

Na rozdíl od transformace (6.1) dochází u transformace (6.3) k posunutí dat směrem dolů o jednotku a zároveň nedochází ke změně tvaru časové řady, pokud $\lambda = 0$.

Cílem u Boxovy–Coxovy transformace je tedy nalézt takovou hodnotu λ , při které je sezónní rozptyl přibližně shodný napříč celou časovou řadou. K dobré hodnotě λ lze dospět mnoha způsoby. Jedním ze způsobů je užít metodu navrhovanou autorem Guerrero (1993). Tato metoda není závislá na zvoleném modelu a snaží se stabilizovat rozptyl.

Výběr optimální hodnoty λ probíhá tak, že se nejdříve časová řada rozdělí do h skupin se zachováním chronologického uspořádání. Poté se počítá průměr \bar{Z}_h a směrodatná odchylka S_h pro každou skupinu. Následně se volí taková hodnota λ , při které

$$S_h / \bar{Z}_h^{1-\lambda} = a, \quad h = 1, \dots, H. \quad (6.4)$$

Podmínkou je, že hodnota nějaké konstanty $a > 0$. Nakonec se hodnota λ optimalizuje hledáním minima variačního koeficientu rovnice (6.4) napříč všemi skupinami nebo prostřednictvím modelu lineární regrese rovnice (6.4), která je uvedena níže. Balíček `feasts` hledá hodnotu λ na základě minima variačního koeficientu.

$$\ln(S_h) = \ln(a) + (1 - \lambda) \cdot \ln(\bar{Z}_h) + \epsilon_h, \quad h = 1, \dots, H \quad (6.5)$$

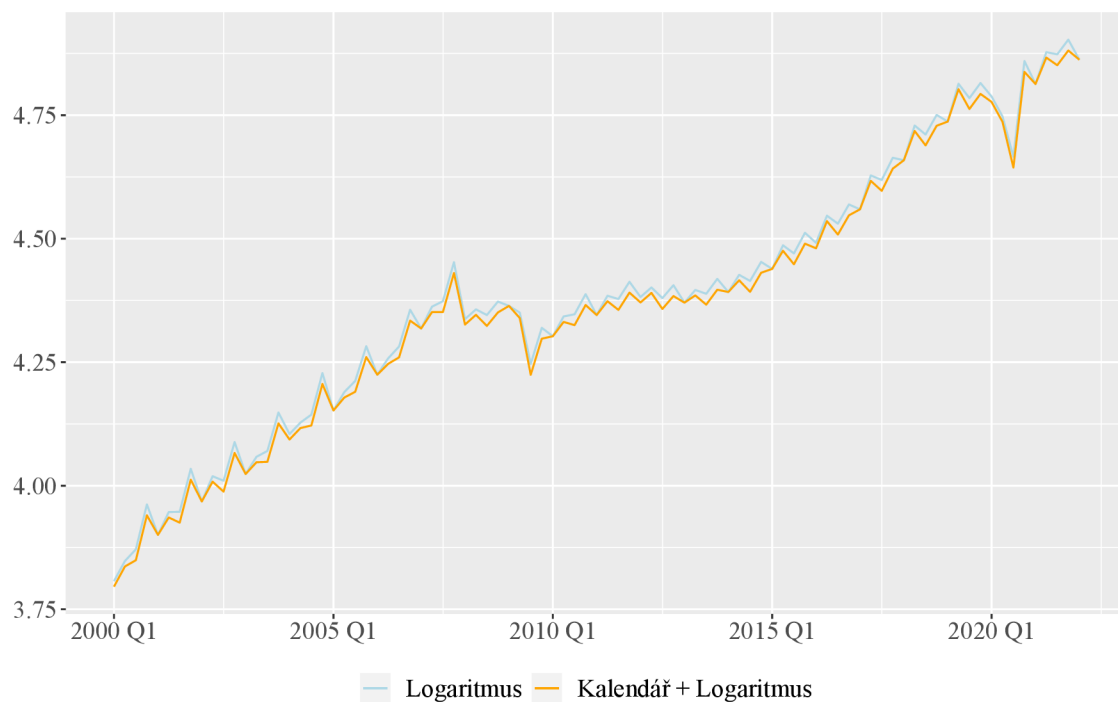
Pokud byla časová řada transformována Boxovou–Coxovou transformací a na základě této řady byla provedena předpověď, tak zpětná transformace bodové předpovědi není průměrem rozdělení předpovědi, ale obvykle jeho mediánem za předpokladu symetrického rozdělení po transformaci. Autoři Hyndman a Athanasopoulos (2021) navrhnou rovnici (6.6) pro korigování tohoto vychýlení. Proměnná $\hat{w}_{T+h|T}$ představuje h -krokový průměr předpovědi a σ_h^2 h -krokový rozptyl předpovědi po zpětné transformaci.

$$\hat{y}_{T+h|T} = \begin{cases} \exp(\hat{w}_{T+h|T}) \left[1 + \frac{\sigma_h^2}{2} \right] & \text{pokud } \lambda = 0; \\ (\lambda \hat{w}_{T+h|T} + 1)^{1/\lambda} \left[1 + \frac{\sigma_h^2(1-\lambda)}{2(\lambda \hat{w}_{T+h|T} + 1)^2} \right] & \text{pokud } \lambda \neq 0. \end{cases} \quad (6.6)$$

Kalendářní úprava byla provedena tak, že se čtvrtletní údaj vydělil počtem dní, který přísluší danému čtvrtletí. Tento podíl se poté vynásobil čtvrtletím o 90 dnech.

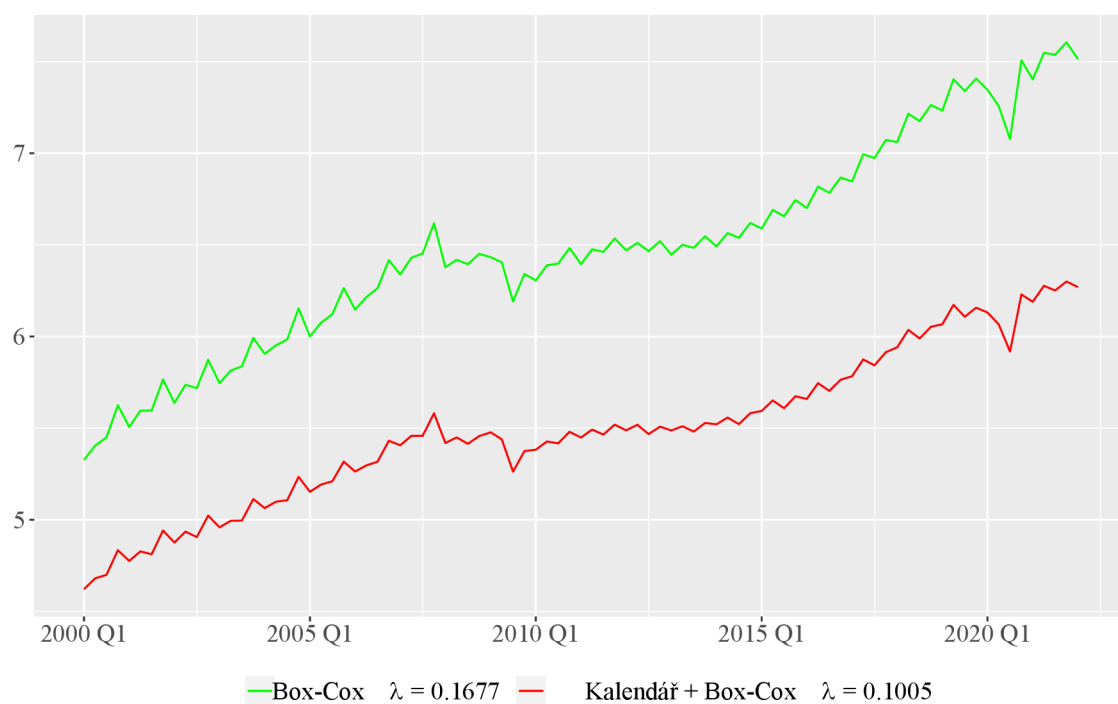
Logaritmická transformace časové řady příjmů včetně její kombinace s kalendářní úpravou zachycuje obrázek 6.3. Modrá křivka představuje logaritmickou transformaci a oranžová křivka logaritmickou transformaci po kalendářní úpravě časové řady. Na obrázku vidíme, že kalendářní úprava zmírnila sezónní variaci. Transformace dat pomocí přirozeného logaritmu přinesla určité snížení rozptylu v porovnání s původní časovou řadou.

Obrázek 6.3: Logaritmická transformace příjmů na důchodové pojištění



Zdroj: ČSSZ, vlastní zpracování

Obrázek 6.4: Boxova–Coxova transformace příjmů na důchodové pojištění



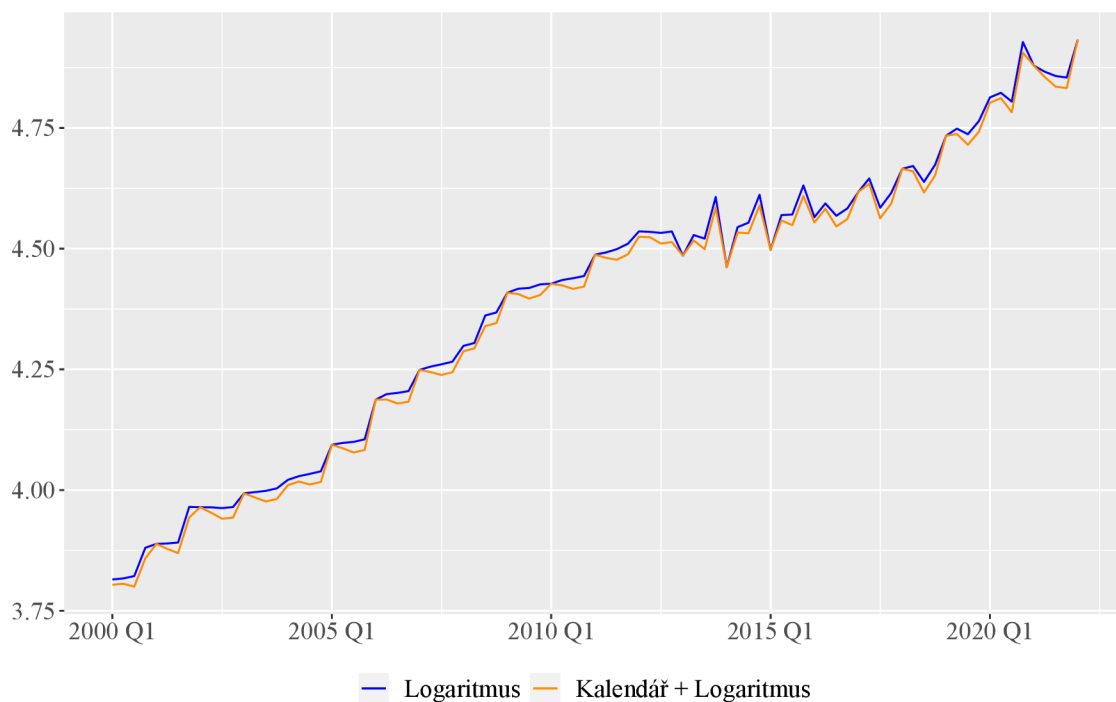
Zdroj: ČSSZ, vlastní zpracování

Transformace časové řady prostřednictvím upravené Boxovy–Coxovy transformace zob-

razuje obrázek 6.4. Zelená křivka znázorňuje Boxovu–Coxovu transformaci a červená křivka stejnou transformaci kalendářně upravených dat. Hodnota λ pro zelenou křivku činí 0,1677 a pro červenou křivku 0,1005. Tyto hodnoty jsou blízké 0, a tudíž je transformace blízká logaritmováním hodnot. Červená křivka se projevuje nižší variací a stabilnějšími tvary v porovnání se zelenou křivkou. Srovnáme-li obě transformace je Boxova–Coxova metoda účinnější co se týče snížení rozptylu napříč časovou řadou.

Transformace výdajů na dávky důchodového pojištění lze nalézt na obrázcích 6.5 a 6.6. Zabarvení křivek má stejný význam jako v předešlých dvou obrázcích. Kalendářní úprava u obou obrázků přinesla určitou variaci v rámci období 2000–2012. Hodnota λ Boxovy–Coxovy transformace je $-0,4953$, resp. $-0,5712$ v případě uplatnění kalendářní úpravy. Uvedené hodnoty jsou vzdálené od 0, tudíž nebude transformace dat podobná logaritmické transformaci.

Obrázek 6.5: Logaritmická transformace výdajů na dávky důchodového pojištění

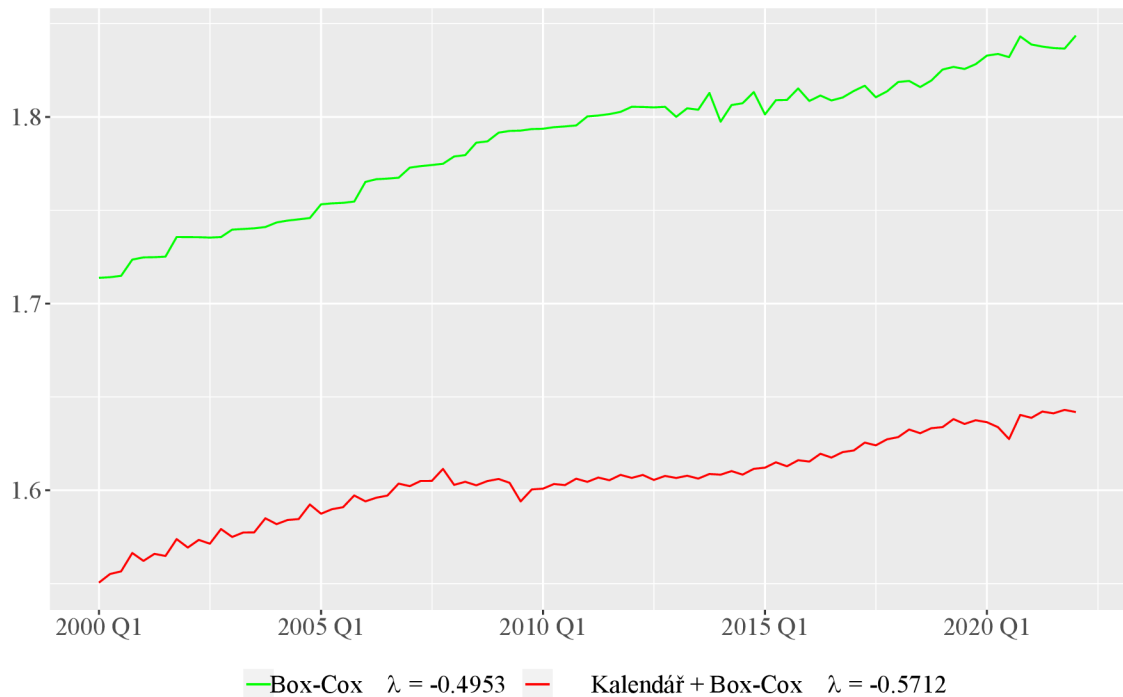


Zdroj: ČSSZ, vlastní zpracování

Porovnáním obou obrázků lze zjistit, že Boxova–Coxova transformace silněji potlačuje rozptyl. Podíváme-li se na obrázek 6.6, vidíme, že aplikací kalendářní úpravy a Boxovy–Coxovy transformace je jednak rozptyl menší a jednak tvary křivky konstantnější než v případě aplikace pouhé Boxovy–Coxovy transformace.

Na základě uvedených poznatků bylo rozhodnuto, že v regresní analýze bude užitá časová řada příjmů a výdajů, která byla modifikovaná Boxovou–Coxovou transformací s výše uvedenými hodnotami λ .

Obrázek 6.6: Boxova–Coxova transformace výdajů na dávky důchodového pojištění



Zdroj: ČSSZ, vlastní zpracování

6.2 Nezávislé proměnné

V následující části se práce bude věnovat analýzou dat jednotlivých nezávislých proměnných, které byly zařazeny do uvažovaných regresních modelů.

6.2.1 Počet ekonomicky aktivních, zaměstnaných a nezaměstnaných

Počet ekonomicky aktivní pracovní síly, zaměstnaných a nezaměstnaných má v příjmovém modelu představovat potenciální počet poplatníků sociálního pojištění, resp. má za část HDP vysvětlovat závislou proměnnou. Ekonomicky pracovní síla se skládá ze zaměstnaných a nezaměstnaných ve věku starší 15 let. Mezi zaměstnané patří jednak zaměstnanci, tak i podnikatelé a pomáhající rodinní příslušníci ad. Podrobnější informace lze nalézt v metodice Výběrového šetření pracovních sil.

Vývoj ekonomicky aktivní pracovní síly, zaměstnaných a nezaměstnaných udává obrázek A.1. Proměnné jsou rozděleny dle pohlaví. Trhu práce se účastní více mužů než žen. Počet zaměstnaných u obou pohlaví pocítil nárůst v posledních letech. Největší nárůst vzhledem k počátku zaznamenaly ženy. Co se týče nezaměstnaných, neexistuje žádný významný rozdíl.

Data byla získána z veřejné databáze ČSÚ pod kategorií „Zaměstnanost, nezaměstnanost“ pod názvem tabulky „Základní charakteristiky ekonomického postavení obyvatelstva ve věku 15 a více let“.

6.2.2 Hrubý domácí produkt

Hrubý domácí produkt byl zařazen mezi nezávislémi proměnnými, protože má představovat zbylou část ekonomiky, poté co byly počet ekonomicky aktivní pracovní síly a příjmy zaměstnanců, OSVČ apod. zařazeny mezi nezávislémi proměnnými. V příjmovém modelu bude využit sezónně očištěný HDP ve stálých cenách roku 2015. Sezónně neočištěná varianta této veličiny bude užitá ve výdajovém modelu.

Čtvrtletní vývoj nominálního HDP, HDP ve stálých cenách roku 2015 a sezónně očištěného HDP ve stálých cenách roku 2015 udává obrázek A.2. Nominálnímu HDP je přiřazena černá barva, reálnému červená a sezónně očištěnému zelená barva. Hodnoty veličin jsou uvedeny v mil. Kč. Zpočátku rostlo HDP prudce, ale po globální finanční krizi rostlo reálné HDP pomalejším tempem než na začátku. Nominální HDP naproti tomu rostlo rychlejším tempem než na začátku. Z obrázku lze vyčíst, že růst HDP v posledních letech byl převážně způsoben inflací.

Časová řada HDP včetně jejích variant byla získána z databáze národních účtů ČSÚ pod záložkou „Čtvrtletní odhad HDP“.

6.2.3 Index spotřebitelských cen

Index spotřebitelských cen měří růst cenové hladiny. Mezi nezávislémi proměnnými byl zařazen, protože jsou závislé proměnné vyjádřené v nominálních hodnotách. Další důvod se týká valorizace důchodů. Jak již bylo uvedeno je procentní výměra zvyšována na základě růstu indexu spotřebitelských cen za domácnosti celkem nebo za domácnosti

důchodců. Pro zvyšování se použije ten, který je pro důchodce výhodnější.

Ve výdajovém modelu bude zahrnut pouze index za domácnosti celkem. Důvodem je velmi vysoký korelační koeficient mezi dvěma indexy, který nabývá hodnot okolo 0,99 a který by zanesl do modelu multikolinearitu či (v extrémním případě) porušil předpoklad neexistence dokonalé kolinearity mezi nezávislými proměnnými. Korelace se počítala za období 2016–2022 a vycházela z indexů se základem průměru roku 2015. Dalším důvodem je zamezení přeparametrizace modelu. Navíc časová řada začíná rokem 2010.

Vývoj indexu spotřebitelských cen za domácnosti celkem zobrazuje obrázek A.3. Základem indexů je průměr roku 2015. Na obrázku lze spatřit, že v obdobích krizí docházelo k většímu růstu cenové hladiny. Největší nárůst v roce 2022, kdy inflace je tažena nabídkou. Jedním z důvodů bylo omezení činností v době pandemie a s ním související obnovení činností, které se nevyrovnává poptávce. Od roku 2018 se index spotřebitelských cen počítá na základě klasifikace ECOICOP, která zavedla podrobnější členění do spotřebního koše.

Data pochází z veřejné databáze ČSÚ, kategorie „Ceny, inflace“, ze záložky „Inflace, spotřebitelské ceny“, z tabulek „Indexy spotřebitelských cen podle klasifikace COICOP - časové řady“ a „Indexy spotřebitelských cen podle klasifikace ECOICOP - časové řady“.

6.2.4 Náhrady zaměstnancům a čistý smíšený důchod

Náhrady zaměstnancům se skládají z mezd, platů a sociálních příspěvků, které poskytují zaměstnavatelé. Mzdy a platy jsou uvedeny v brutto hodnotách, tedy před odpočtem daně z příjmů a zákonných příspěvků na sociální a zdravotní pojištění apod. Zastřešují jak peněžité, tak i naturální odměny. Sociální příspěvky zahrnují příspěvky zaměstnavatele na všeobecné sociální a zdravotní pojištění za jeho zaměstnance, penzijní a zdravotní pojištění, povinné úrazové pojištění a nemocenské a další příspěvky.

Čistý smíšený důchod představuje souhrn důchodu z podnikání a důchodu z činností osob samostatně výdělečně činných. Podnikatelský důchod je běžný zisk před zdaněním a rozdělením. Tyto důchody se týkají sektoru domácností a nedají se od sebe statisticky odlišit. Čistý smíšený důchod vzniká jako rozdíl hrubého smíšeného důchodu a spotřeby fixního kapitálu, který představují snížení hodnoty dlouhodobého majetku z důvodu fyzického

nebo morálního zastarání. Zjednodušeně se tedy jedná o odpisy dlouhodobého majetku.

Vývoj výše uvedených veličin lze spatřit na obrázku A.4. Nominální hodnota veličin v mil. Kč byla přepočtena výše uvedeným indexem spotřebitelských cen. Z obrázku vyplývá, že mzdy a platy jsou výrazně vyšší než čistý smíšený důchod a zároveň rostou rychlejším tempem. Reálná hodnota čistého smíšeného důchodu postupem času klesá. Co se týče sezónního očištění náhrad zaměstnancům a sociálních příspěvků, není tato úprava velice významná.

Data byla čerpána z databáze národních účtů ČSÚ pod kategorií „Čtvrtletní odhad HDP“ z tabulek „Mzdy a platy“ a „Sociální příspěvky zaměstnavatelů“ a pod kategorií „Čtvrtletní sektorové účty“ z tabulek „Časová řada (sezónně očištěných) sektorových účtů za S14“.

6.2.5 Počet důchodců

Počet důchodců zahrnuje tři nezávislé proměnné. Jedná se o počet starobních důchodců podle pohlaví a počet ostatních důchodců, do kterého spadají invalidní a pozůstalostní důchodci.

Časový vývoj počtu důchodců je vidět na obrázku A.5, kde oranžová křivka představuje vývoj počtu starobních důchodkyň, modrá starobních důchodů a černá ostatních důchodů. Z obrázku lze spatřit, že je více starobních důchodkyň než starobních důchodců, což není překvapivé, protože se ženy v průměru dožívají vyššího věku. Zajímavé je, že počet ostatních důchodců mezi roky 2000 a 2010 mírně rostl a od roku 2010 konstantně klesá. Podíváme-li se na vývoj starobních důchodců, vidíme, že v posledních letech rostl počet starobních důchodců rychleji než u starobních důchodkyň, kde je křivka konstantní. V posledních letech ubývá počet starobních důchodkyň. Výrazný pokles ostatních důchodců, resp. růst starobních důchodců v roce 2010 byl způsoben převedením invalidních důchodců nad 65 let na starobní důchodce.

Časová řada vývoje počtu důchodců byla získána ze statistik ČSSZ pod kategorií „Důchodová statistika“ pod záložkou „Základní přehledy“, která nachází na jejich internetových stránkách.

6.2.6 Průměrná mzda

Průměrná mzda má ve výdajovém modelu představovat jeden z faktorů používaných pro zvýšení důchodů vedle indexu spotřebitelských cen. Platí to jak pro základní výměru, tak i pro procentní výměru. V práci nebude zohledněno sledované časové období pro stanovení růstu reálné mzdy.

V regresním modelu bude použita reálná průměrná měsíční mzda na přepočtené počty zaměstnanců za čtvrtletí. Reálná průměrná mzda byla vypočítána jako podíl nominální průměrné měsíční mzdy na přepočtené počty zaměstnanců za čtvrtletí a výše uvedeného indexu spotřebitelských cen.

Vývoj nominální a reálné průměrné měsíční mzdy na přepočtené počty zaměstnanců ukazuje obrázek A.6, na kterém je nominální průměrná mzda vyznačena červenou barvou a reálná průměrná mzda zelenou barvou. Na obrázku lze vidět, že ve 3. čtvrtletí je průměrná měsíční mzda nižší než ve 2. čtvrtletí daného roku. Podíváme-li se na období krizí, tak po globální finanční krizi v roce 2008 došlo pouze ke stagnaci reálných průměrných mezd a po pandemii Covid-19 nelze s jistotou určit, zda bude krátkodobě reálná průměrná mzda stagnovat, nebo klesat. Vypadá to, že ve 4. čtvrtletí roku 2021 průměrná reálná mzda dosáhla vrchol a začíná klesat.

Časová řada nominální průměrné měsíční mzdy byla získána z veřejné databáze ČSÚ pod kategorií „Mzdy a náklady práce“ z tabulky „Počet zaměstnanců a průměrné hrubé měsíční mzdy“.

6.3 Jednotkový kořen

Společně ADF a KPSS testy potvrdily přítomnost jednotkového kořene ve výše uvedených časových řadách. U některých proměnných zamítl ADF nulovou hypotézu při 1. anebo 2. řádu.

Zajímavé je, že po provedení první diference testy stále prokazovaly přítomnost jednotkového kořene u některých proměnných. Pokud však byla provedena sezónní diference jako první, tak ji testy již neprokázaly. Například u indexu spotřebitelských cen musela být provedena jak sezónní diference, tak i první diference. Důvodem bylo, že ADF neprokázal stacionaritu, zatímco KPSS stacionaritu prokázal. Pokud se však testuje v rámci období

od roku 2013 do 2022 Q1, tak u mnohých proměnných po sezónní diferenci prokazují testy přítomnost jednotkového kořene.

Kapitola 7

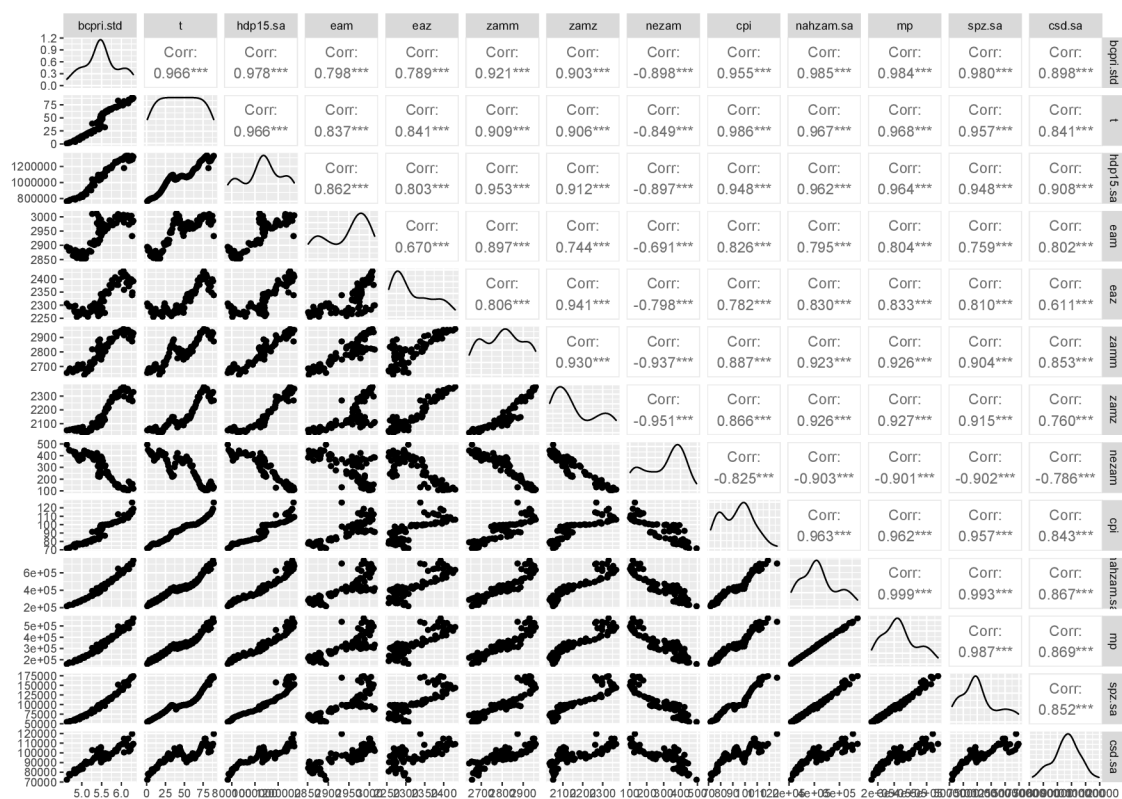
Příjmové modely

V této kapitole se diplomová práce bude věnovat regresním modelům, které se týkají příjmů na důchodové pojištění. Dle minulé kapitoly jsou závislou proměnnou příjmy po kalendářní úpravě a Boxově–Coxově transformaci s uvedenou hodnotou λ . Před regresní analýzou je vhodné se podívat na vztahy mezi proměnnými.

Obrázek 7.1 je maticí bodových grafů a umožňuje tyto vztahy zkoumat. Pod obrázkem se nachází legenda vysvětlující zkratky jednotlivých proměnných. První sloupec obrázku zachycuje vztahy mezi závislou proměnnou a nezávislými proměnnými. Na druhou stranu vyjadřuje první řádek sílu těchto vztahů prostřednictvím korelace mezi proměnnými a také obsahuje statistickou významnost hodnot korelace. Z prvního sloupce vyplývá, že mezi závislou proměnnou a téměř všemi nezávislými proměnnými existuje pozitivní vztah. Pouze u počtu nezaměstnaných ukazuje obrázek negativní vztah s příjmy. Co se týče síly těchto vztahů, jsou všechny hodnoty korelace významné. S téměř všemi nezávislými proměnnými činí korelace s příjmy více než 0,97.

V ostatních sloupcích jsou zobrazeny vztahy mezi jednotlivými nezávislými proměnnými. Obdobně to platí pro ostatní řádky vyjadřující korelaci mezi nezávislými proměnnými. Nejvyšší hodnota korelace 0,999 je mezi náhradami zaměstnancům a proměnnými navzájem, ze kterých se skládá, což není překvapivé. Toto platí také pro reálné hodnoty těchto proměnných. Pokud by byly tyto proměnné zahrnuty do modelu, hrozí porušení předpokladu neexistence dokonalé kolinearity mezi nezávislými proměnnými. Dále lze vysoké hodnoty korelace spatřit mezi proměnnou vyjadřující trend (t) a reálným HDP sezónně očištěným, náhradami zaměstnancům a proměnnými, ze kterých se skládá. V těchto případech je korelace vyšší než 0,966. U některých vztahů je korelace vyšší, resp. nižší než 0,90, resp. $-0,90$.

Obrázek 7.1: Matice korelačních polí nezávislých proměnných příjmových modelů



Zdroj: Vlastní zpracování

Legenda:

<i>bcpri.std</i>	Příjmy na DP po kalendářní úpravě a Boxově–Coxově transformaci
<i>t</i>	Lineární trend
<i>hdp15.sa</i>	Reálné HDP sezónně očištěné
<i>eam</i>	Ekonomicky aktivní muži
<i>eaz</i>	Ekonomicky aktivní ženy
<i>zamm</i>	Zaměstnaní muži
<i>zamz</i>	Zaměstnané ženy
<i>nezam</i>	Nezaměstnaní celkem
<i>nahzam.sa</i>	Náhrady zaměstnancům sezónně očištěné
<i>mp</i>	Mzdy a platy
<i>spz.sa</i>	Sociální příspěvky zaměstnavatelů sezónně očištěné
<i>cpi</i>	Index spotřebitelských cen se základem průměr roku 2015
<i>csd.sa</i>	Čisté smíšené důchody sezónně očištěné

7.1 Příjmový model 1

Při modelování příjmů na důchodové pojištění bylo nalezeno mnoho modelů. V následujícím textu bude představeno několik nalezených modelů. Jedním z modelů je uveden níže:

$$\widehat{bcprī.stđ}_t = \beta_0 + hdp15.sa_t + eam_t + eaz_t + cpi_t + mzda15_t \quad (7.1)$$

Rovnice (7.1) vysvětluje příjmy na důchodové pojištění po kalendářní úpravě a Boxově–Coxově transformaci na základě reálného HDP sezónně očištěného (dále jen SO), ekonomicky aktivní pracovní síly pro obě pohlaví, indexu spotřebitelských cen a reálné mzdy. Odhadnuté hodnoty proměnných, p-hodnoty a další statistiky obsahuje tabulka 7.1. Z tabulky vyplývá, že všechny nezávislé proměnné jsou významné. Podíváme-li se na koeficienty, zjistíme, že mnohé proměnné mají očekávané znaménko. U ekonomicky pracovní síly jsou koeficienty záporné. To znamená, že při zvýšení ekonomicky aktivních mužů o 10 000 se za jinak stejných podmínek sníží transformovaný příjem o 0,001654 jednotek. Absolutní člen β_0 je velmi vysoký vzhledem k tomu, že se transformované příjmy pohybují přibližně od 4,62 do 6,30. Odhadované koeficienty jsou uvedeny v desetinných číslech.

Tabulka 7.1: Výstup z R pro příjmový model 1

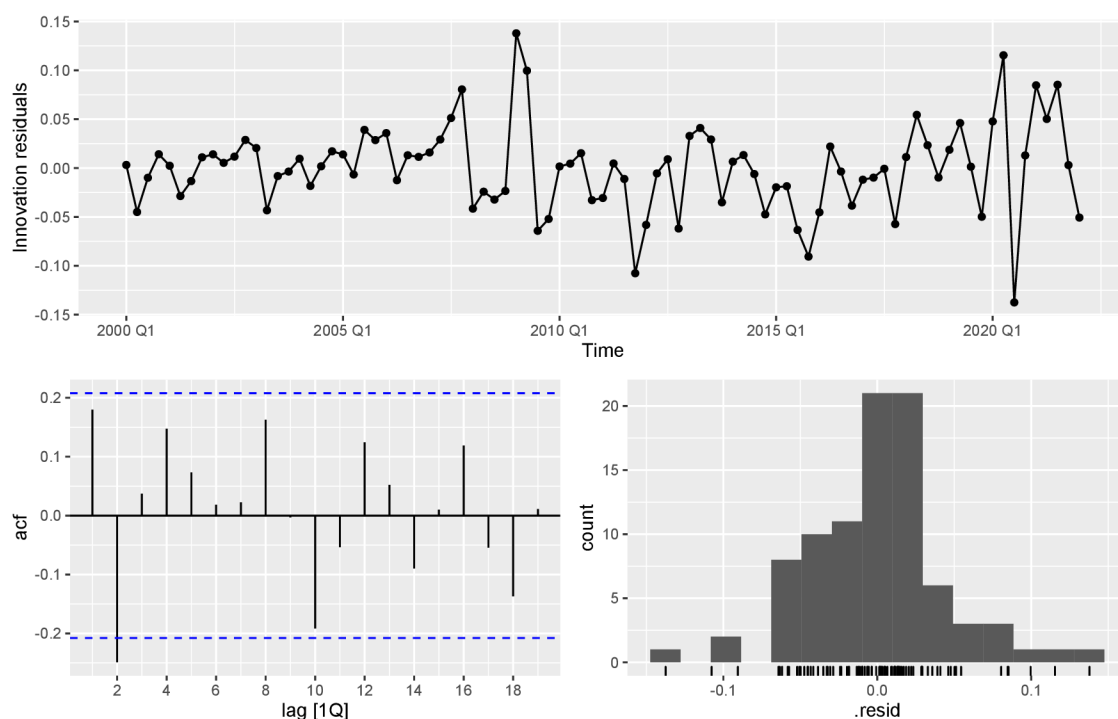
Rezidua:					
Min	1Q	Medián	3Q	Max	
-0,137507	-0,024194	0,002365	0,018717	0,137884	
Koeficienty:					
	Odhad	Std. chyba	t-statistika	p-hodnota	
β_0	7,866	6,35E-01	12,384	< 2E-16	***
$hdp15.sa_t$	1,520E-06	1,25E-07	12,206	< 2E-16	***
eam_t	-1,654E-03	1,92E-04	-8,598	4,17E-13	***
eaz_t	-3,416E-04	1,63E-04	-2,100	0,0388	*
cpi_t	6,955E-03	1,24E-03	5,624	2,45E-07	***
$mzda15_t$	4,002E-05	3,65E-06	10,958	< 2E-16	***

Kódy význ.:	0 ‘****’	0,001 ‘***’	0,01 ‘**’	0,05 ‘.’	0,1 ‘ ’	1
Rezidální standardní chyba: 0,04579 při 83 stupních volnosti						
Koef. determinace R^2 :	0,9887			Adj. koef. det. \bar{R}^2 :	0,9880	
F-statistika:	1453 při 5 a 83 st. vol.			p-hodnota:	< 2,22E-16	

Zdroj: Vlastní zpracování

Obrázek 7.2 zahrnuje tři grafy. Jedná se o graf vývoje reziduí, autokorelační funkci a histogram reziduí. Co se týče rozptylu reziduí, lze rozeznat heteroskedasticitu. Podíváme-li se na dolní levý graf, zjistíme, že pouze autokorelace 2. řádu je významná. Ověřením autokorelace Ljungovým–Boxovým testem dostáváme testovou statistiku (dále jen TS) ve výši 14,2 a p-hodnota činí 0,0008. To znamená, že test prokázal významnost autokorelace do 8. řádu na hladině významnosti 0,05.

Obrázek 7.2: Grafy reziduí příjmového modelu 1



Zdroj: Vlastní zpracování

Jelikož byla prokázána významnost autokorelace, musí brát test heteroskedasticity reziduí s rezervou. Dle Breuschova–Paganova testu činí TS 17,301 a p-hodnota 0,0040. Výsledkem tedy je, že byla prokázána heteroskedasticita. Co se týče normality reziduí, tak podle Shapirova–Wilkova testu byla prokázána nenormalita, protože p-hodnota se rovná 0,0369

(TS = 0,97). Rozhodl jsem se ignorovat nenormalitu reziduí vzhledem k velikému počtu pozorování (91) z důvodu centrální limitní věty. Máme-li dostatek dat, tak součty a průměry se chovají dle normálního rozdělení bez ohledu na původní rozdělení dat. Čím více máme dat, tím lépe to funguje. Aplikujeme-li na model původní HAC, nedojde ke změně významnosti proměnných, zvyšuje se pouze standardní chyba.

7.2 Příjmový model 2

Další model představuje rovnice (7.2), která vysvětluje transformované příjmy lineárním časovým trendem, zaměstnanými ženami a muži, nezaměstnanými, mzdami a platy, sociálními příspěvky, indexem spotřebitelských cen a reálným HDP SO.

$$\widehat{bcpri.std}_t = \beta_0 + t + zamm_t + zamz_t + nezam_t + mp15_t + spz15.sa_t + cpi_t + hdp15.sa_t \quad (7.2)$$

Výstup z R pro výše uvedený model obsahuje tabulka 7.2. Z tabulky lze vyčíst, že odhadované koeficienty mají opět očekávané znaménko kromě počtu zaměstnaných u obou pohlaví. Absolutní člen β_0 je opět vysoký a odhadované koeficienty jsou desetinnými čísly. Všechny nezávislé proměnné jsou významné. Součástí tohoto modelu byl čistý smíšený důchod SO, u kterého ve výsledku nebyla prokázána významnost, a tudíž byl z modelu vynechán. Ve srovnání s předchozím modelem je adjustovaný koeficient determinace \bar{R}^2 vyšší.

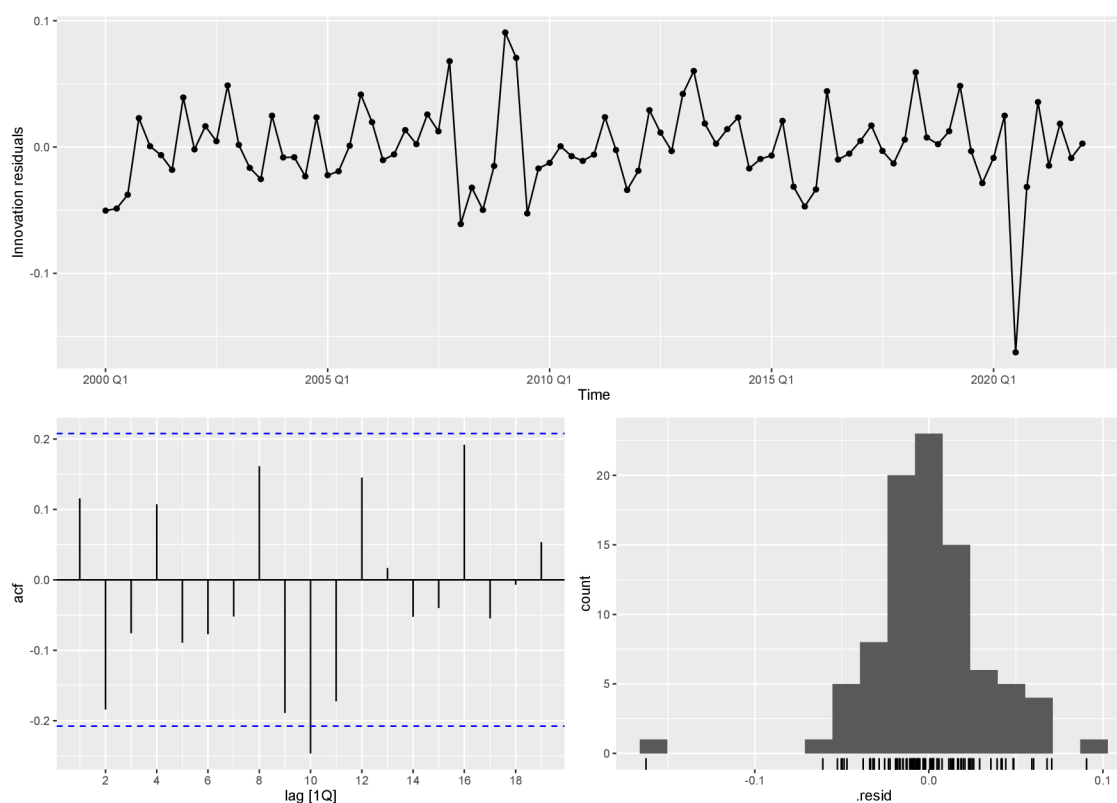
Tabulka 7.2: Výstup z R pro příjmový model 2

Rezidua:					
Min	1Q	Medián	3Q	Max	
-0,162472	-0,016474	-0,002346	0,018605	0,090638	
Koeficienty:					
	Odhad	Std. chyba	t-statistika	p-hodnota	
β_0	6,69E+00	7,08E-01	9,445	1,17E-14	***
t	3,63E-03	1,47E-03	2,461	0,016006	*
$zamm_t$	-7,20E-04	1,86E-04	-3,866	0,000224	***
$zamz_t$	-7,68E-04	1,79E-04	-4,284	5,07E-05	***

$nezam_t$	-6,01E-04	1,71E-04	-3,522	0,000711	***	
$mp15_t$	1,69E-06	2,89E-07	5,837	1,09E-07	***	
$spz15.sa_t$	6,42E-06	9,50E-07	6,762	2,02E-09	***	
cpi_t	4,39E-03	2,12E-03	2,074	0,041264	*	
$hdp15.sa_t$	7,60E-07	1,42E-07	5,369	7,57E-07	***	
Kódy význ.:	0 '***'	0,001 '**'	0,01 '*'	0,05 '.'	0,1 ''	1
Rezidální standardní chyba: 0,03557 při 80 stupních volnosti						
Koef. determinace R^2 :	0,9934		Adj. koef. det. \bar{R}^2 :		0,9928	
F-statistika:	1513 při 8 a 80 st. vol.		p-hodnota:		< 2,22E-16	

Zdroj: Vlastní zpracování

Obrázek 7.3: Grafy reziduí příjmového modelu 2



Zdroj: Vlastní zpracování

Podíváme-li na autokorelační funkci reziduí nacházející se na obrázku 7.3, vidíme, že pouze autokorelace 10. řádu je významná. Ljungův–Boxův test do 10. řádu vyšel signifikantní s p-hodnotou 0 (TS = 20,1). U vývoje reziduí by se dalo říct, že platí homoskedasticita.

Dle Breuschova–Paganova testu nebyla prokázána heteroskedasticita, jelikož je p-hodnota 0,3683 s TS = 8,6992 při 8 stupních volnosti (dále jen SV). Histogram reziduí napovídá o možné normalitě reziduí. Shapirův–Wilkův test vyšel signifikantní s p-hodnotou 0 (TS = 0,9268). To znamená, že data nemají normální rozdělení.

7.3 Příjmový model 3

Tento model je modelem diferencí. Zde je meziroční změna (sezónní diference) příjmu funkcí meziroční změny ekonomicky aktivních mužů, reálné hodnoty náhrady zaměstnancům SO a změny meziroční změny indexu spotřebitelských cen. Model je znázorněn rovnicí (7.3). Tento model vychází z proměnných předchozího modelu. Všechny tyto proměnné jsou na rozdíl od předchozích modelů stacionární.

$$\widehat{sdpri}_t = \beta_0 + sdeam_t + sdnahzam15.sa_t + sddcpi_t \quad (7.3)$$

Pokud se podíváme na sloupec odhadnutých koeficientů z tabulky 7.3, zjistíme, že meziroční změna ekonomicky aktivních mužů má záporný koeficient. Znaménka ostatních proměnných mají logický význam. Všechny ostatní proměnné v modelu 2, jejichž stacionární varianty byly použity v tomto modelu, se prokázaly být nevýznamné od nuly. Adjustovaný koeficient determinace \bar{R}^2 je u tohoto modelu nižší než u ostatních modelů a činí 0,6199. Vypadá to, jako by zahrnutím časového trendu nebyla vyřešena problematika zdánlivé regrese.

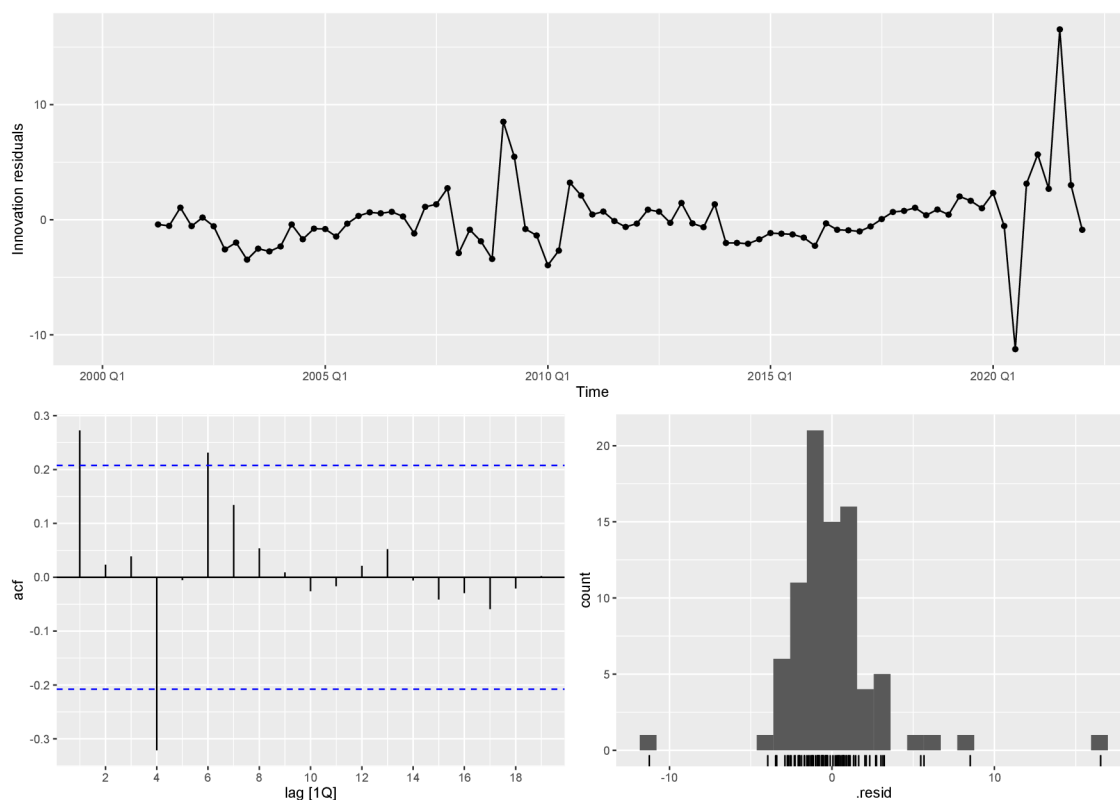
Tabulka 7.3: Výstup z R pro příjmový model 3

Rezidua:					
Min	1Q	Medián	3Q	Max	
-11,2433	-1,2966	-0,3730	0,8777	16,5382	
Koeficienty:					
	Odhad	Std. chyba	t-statistika	p-hodnota	
β_0	1,07E+00	4,29E-01	2,485	0,01505	*
<i>sdeam</i>	-4,27E-02	1,90E-02	-2,244	0,0276	*
<i>sdnahzam15.sa</i>	2,27E-04	2,07E-05	10,938	< 2E-16	***

<i>sddcpi</i>	9,84E-01	3,58E-01	2,749	0,00739	**	
Kódy význ.:	0 ‘***’	0,001 ‘**’	0,01 ‘*’	0,05 ‘.’	0,1 ‘ ’	1
Rezidální standardní chyba: 3,049 při 80 stupních volnosti						
Koef. determinace R^2 :	0,6337		Adj. koef. det. \bar{R}^2 : 0,6199			
F-statistika:	46,13 při 3 a 80 st. vol.		p-hodnota: < 2,22E-16			

Zdroj: Vlastní zpracování

Obrázek 7.4: Grafy reziduí příjmového modelu 3



Zdroj: Vlastní zpracování

Autokorelační funkce na obrázku 7.4 ukazuje 3 významné řády autokorelace reziduí, a to 1., 4. a 6. řád. Ljungův–Boxův test prokázal významnost autokorelace reziduí při p-hodnotě 0,0001 ($TS = 22,9$), $\ell = 8$ a $K = 4$. Na první pohled to vypadá, že rezidua mají normální rozdělení. Shapirův–Wilkův test prokázal, že rezidua nemají normální rozdělení při p-hodnotě 0 a $TS 0,7860$. Co se týče homoskedasticity reziduí, tak z vyššího grafu to z toho vyplývá. Pro úplnost je uveden Breuschův–Paganův test, podle kterého nelze

zamítnout nulovou hypotézu, neboť p-hodnota činí 0,6572 (TS = 1,6099, SV = 3).

7.4 Shrnutí

Dle výše představených modelů jsou nominální příjmy z pojistného a příslušenství na důchodové pojištění pozitivně korelovány s reálným příjmem poplatníků, reálnými sociálními příspěvky zaměstnavatelů, s indexem spotřebitelských cen a s reálným HDP. Naopak jsou příjmy negativně korelovány se základnou poplatníků sociálního pojištění a s nezaměstnanými, tedy s ekonomicky aktivní pracovní silou. Negativní korelace s počtem zaměstnaných by se dala interpretovat jako délka odpracovaných let, kdy s jejím prodloužením se zhoršují dovednosti pracujících, pokud to přeženeme.

Každopádně dle těchto modelů lze příjmy zvýšit růstem příjmů zaměstnanců a růstem ekonomiky anebo zvýšením sazeb sociálního pojištění. Navíc se příjmy OSVČ prokázaly jako nevýznamné kvůli jejich celkové velikosti, a tudíž představují další potenciál růstu příjmů na důchodové pojištění. Příjmy poplatníků souvisejí s produktivitou a ta, z určité části, se zdravou délkou života. To dává smysl, jelikož zdraví poplatníci mohou snadněji a hlouběji pronásledovat další činnosti, které jsou předmětem pojistného. To však znamená zvyšování kvality zdravotní péče, provádění preventivních prohlídek, podporování zdravého životního stylu apod. ze strany státu. Určitou roli ve výši příjmu bude hrát i vzdělání a kvalifikace.

Kapitola 8

Výdajové modely

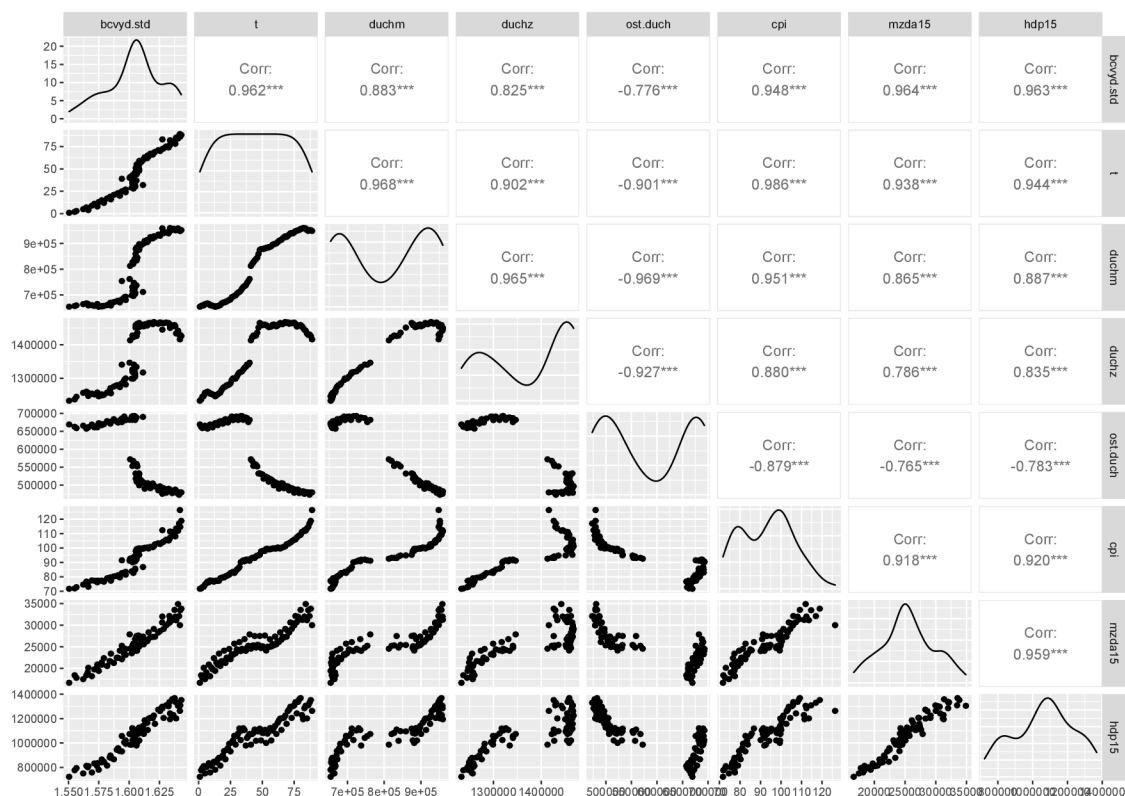
V této kapitole se práce bude zabývat regresním modelům, jejichž závislou proměnnou jsou výdaje na dávky důchodového pojištění. Jak již bylo výše uvedeno probíhají výdaje nejdříve kalendářní úpravou a poté podléhají Boxově–Coxově transformaci. Nyní se podívejme na vztahy mezi proměnnými.

Obrázek 8.1 představuje matici bodových grafů, která podává informaci o vztazích mezi závislou a nezávislými proměnnými a o síle těchto vztahů. Legenda vysvětlující zkratky proměnných je uvedena níže. Z prvního sloupce vyplývá, že můžeme od téměř všech proměnných očekávat kladný vztah a od ostatních důchodů záporný vztah. Největší korelace je mezi lineárním časovým trendem a indexem spotřebitelských cen. Korelace tohoto páru proměnných činí 0,986, což může indikovat možnou perfektní kolinearitu. Silné vztahy vykazují starobní důchodci se starobními důchodkyněmi, s ostatními důchodci a s lineárním časovým trendem. Zde činí korelace kolem 0,967, resp. $-0,969$. Všechny korelace jsou významné na hladině významnosti 0,05.

Legenda:

<i>bcvyd.std</i>	Výdaje na dávky DP po kalendářní úpravě a Boxově–Coxově transformaci
<i>t</i>	Lineární časový trend
<i>duchm</i>	Starobní důchodci
<i>duchz</i>	Starobní důchodkyně
<i>ost.duch</i>	Ostatní důchodci – zahrnuje invalidní a pozůstalostní důchodce
<i>cpi</i>	Index spotřebitelských cen se základem průměr roku 2015
<i>mzda15</i>	Reálná mzda
<i>hdp15</i>	Reálné HDP

Obrázek 8.1: Matice korelačních polí nezávislých proměnných výdajových modelů



Zdroj: Vlastní zpracování

Opět bylo nalezeno mnoho regresních modelů. V následujících kapitolách bude několik těchto modelů představeno.

8.1 Výdajový model 1

Model podle rovnice (8.1) vyjadřuje, že výdaje na dávky důchodového pojištění po kalendářní úpravě a Boxově–Coxově transformaci jsou funkcí lineárního časového trendu, starobních důchodců, starobních důchodkyň, současného indexu spotřebitelských cen, indexu spotřebitelských cen před dvěma čtvrtletími, současné reálné mzdy a minulé reálné mzdy.

$$\widehat{bcvyd.std}_t = \beta_0 + t + duchm_t + duchz_t + cpi_t + cpi_{t-2} + mzda15_t + mzda15_{t-1} \quad (8.1)$$

Výstup z R zachycuje tabulka 8.1. Odhadované koeficienty jsou z větší části intuitivní. Dle tohoto modelu se výdaje snižují, pokud se zvýší počet starobních důchodců nebo pokud dojde ke zvýšení předminulého indexu spotřebitelských cen, což v prvně jmenovaném případě nedává smysl a v druhém jmenovaném případě by se dalo o tom diskutovat. Dalo by se to tím, že se předminulý index je při výpočtu inflace ve jmenovateli, a tudíž jeho zvýšení by představovalo deflaci, kdyby čítec se neměnil. Pokud indexy posoudíme společně, tak v konečném výsledku dochází ke snížení výdajů při zvýšení obou indexů. Z tabulky vyplývá, že všechny proměnné jsou významné. Součástí nezávislých proměnných byli ostatní důchodci, ale na základě t-testu nebyla prokázána jejich významnost. Z tohoto důvodu byla tato proměnná z modelu vyloučena. Adj. \bar{R}^2 činí 0,9878.

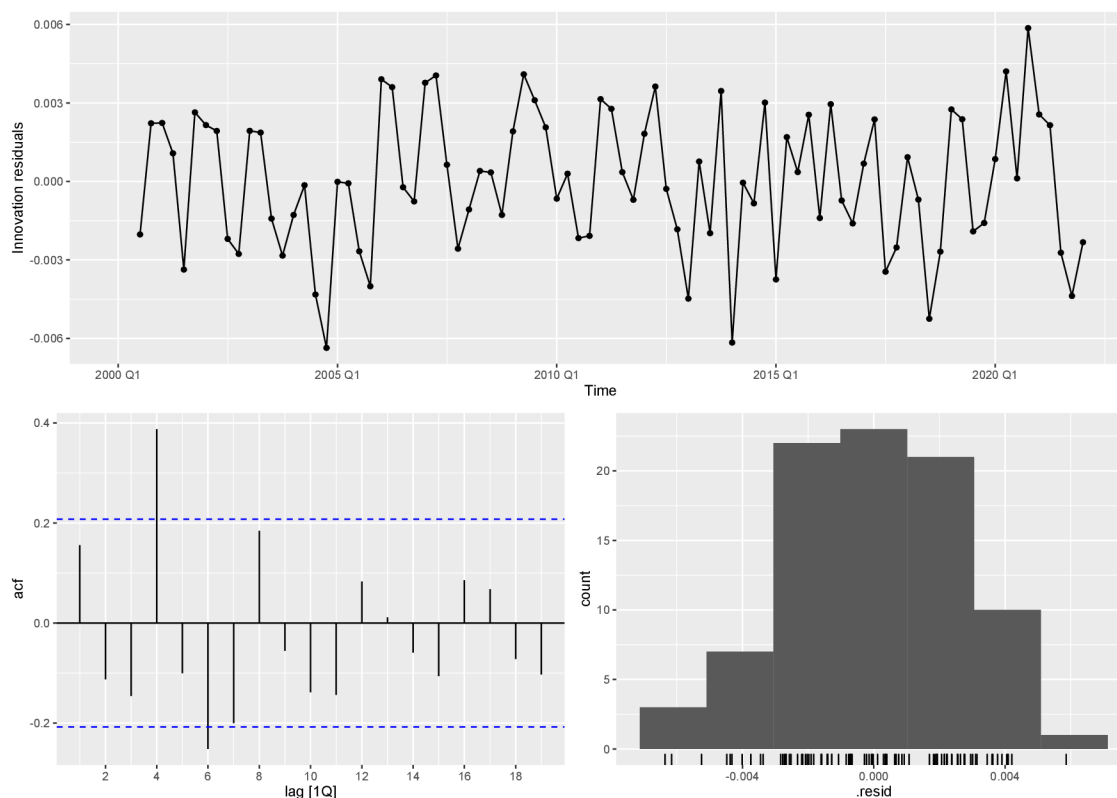
Tabulka 8.1: Výstup z R pro výdajový model 1

Rezidua:						
Min	1Q	Medián	3Q	Max		
-0,006361	-0,002002	-0,000014	0,002188	0,005861		
Koeficienty:						
	Odhad	Std. chyba	t-statistika	p-hodnota		
β_0	1,275	1,96E-02	64,924	< 2E-16	***	
t	3,64E-04	1,37E-04	2,664	0,00936	**	
$duchm_t$	-2,40E-07	2,34E-08	-10,252	3,6E-16	***	
$duchz_t$	2,42E-07	1,66E-08	14,572	< 2E-16	***	
cpi_t	5,62E-04	2,55E-04	2,203	0,03050	*	
cpi_{t-2}	9,66E-04	3,29E-04	2,934	0,00437	**	
$mzda15_t$	5,82E-07	2,48E-07	2,345	0,02154	*	
$mzda15_{t-1}$	7,21E-07	2,49E-07	2,898	0,00486	**	
Kódy význ.:	0 '***'	0,001 '**'	0,01 '*'	0,05 '.'	0,1 ' '	1
Rezidální standardní chyba: 0,002766 při 79 stupních volnosti						
Koef. determinace R^2 :		0,9888		Adj. koef. det. \bar{R}^2 : 0,9878		
F-statistika:		996,8 při 7 a 79 st. vol.		p-hodnota: < 2,22E-16		

Zdroj: Vlastní zpracování

Autokorelace reziduí 4. a 6. řádu jsou významné dle grafu na obrázku 8.2. Ljungův–Boxův test prokázal významnost autokorelace a p-hodnota činí 0 (s TS = 33,5) při $\ell = 8$ a $K = 8$. Pro úplnost Breuschův–Paganův test neprokázal heteroskedasticitu, protože p-hodnota je 0,7876 (TS = 3,9321; SV = 7). Shapirův–Wilkův test neprokázal nenormalitu reziduí, protože p-hodnota = 0,2729 (TS = 0,9821).

Obrázek 8.2: Grafy reziduí výdajového modelu 1



Zdroj: Vlastní zpracování

8.2 Výdajový model 2

Tento model se snaží vysvětlit výdaje po kalendářní úpravě a Boxově–Coxově transformaci lineárním časovým trendem, transformovanými výdaji minulého roku, starobními důchodci, ostatními důchodci, předminulým indexem spotřebitelských cen, minulou reálnou průměrnou mzdou a jeho předešlou hodnotou, viz rovnice (8.2).

$$bcvyd.\widehat{std}_t = \beta_0 + t_t + bcvyd_{t-4} + duch_t + ost.duch_t + cpi_{t-2} + mzda15_{t-1} + mzda15_{t-2} \quad (8.2)$$

Ze sloupce odhadovaných koeficientů v tabulce 8.2 vyplývá, že všechny proměnné mají logické znaménko, tedy hodnoty koeficientů jsou kladné (vyjma lineárního časového tren-

du) a dávají smysl. Mimo lineární časový trend (p-hodnota = 0,09) jsou všechny proměnné významné. Ve srovnání s předchozím modelem je adjustovaný koeficient determinace \bar{R}^2 vyšší a činí 0,9931.

Tabulka 8.2: Výstup z R pro výdajový model 2

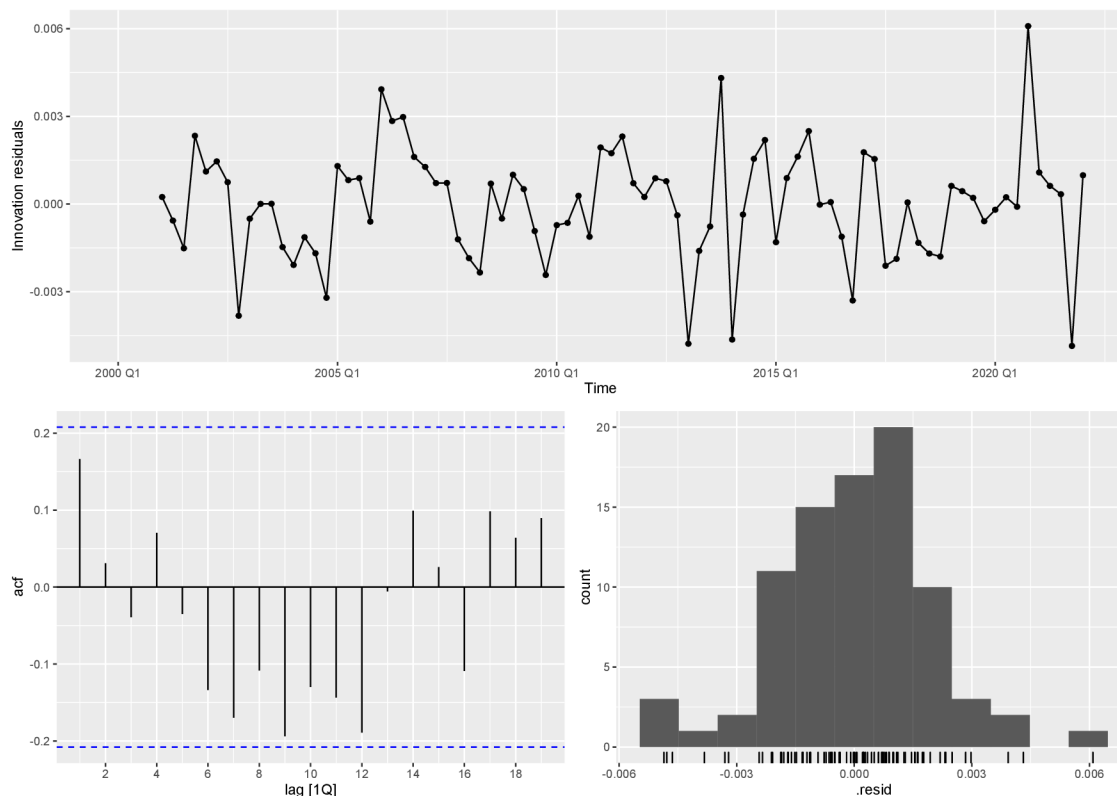
Rezidua:						
Min	1Q	Medián	3Q	Max		
-0,0048587	-0,0011366	0,0002123	0,0010020	0,0060907		
Koeficienty:						
	Odhad	Std. chyba	t-statistika	p-hodnota		
β_0	6,338E-01	7,79E-02	8,136	5,56E-12	***	
t_t	-1,521E-04	8,86E-05	-1,717	0,09002	.	
$bcvyd.std_{t-4}$	4,436E-01	6,60E-02	6,718	2,83E-09	***	
$duch_t$	5,423E-08	8,77E-09	6,187	2,75E-08	***	
$ost.duch_t$	8,831E-08	1,27E-08	6,930	1,13E-09	***	
cpi_{t-2}	7,531E-04	1,63E-04	4,610	1,57E-05	***	
$mzda15_{t-1}$	5,648E-07	1,82E-07	3,111	0,00261	**	
$mzda15_{t-2}$	5,688E-07	1,86E-07	3,054	3,10E-03	**	
Kódy význ.:	0 '***'	0,001 '**'	0,01 '*'	0,05 '.'	0,1 ''	1
Rezidální standardní chyba: 0,001998 při 77 stupních volnosti						
Kof. determinace R^2 :		0,9937		Adj. koef. det. \bar{R}^2 : 0,9931		
F-statistika:		1724 při 7 a 77 st. vol.		p-hodnota: < 2,22E-16		

Zdroj: Vlastní zpracování

Podíváme-li se na autokorelační funkci na obrázku 8.3, zjistíme, že žádný řád autokorelace není významný. Avšak dle Ljungova–Boxova testu byla prokázána významnost autokorelace při p-hodnotě 0,0009 (TS = 9,58; $\ell = 10$, $K = 8$). Rezidua vykazují určitou podobnost s normálním rozdělením. Shapirův–Wilkův test neprokázal nenormalitu reziduí. P-hodnota testu činí 0,1567 (TS = 0,9781). Dle Breuschova–Paganova testu vykazují

rezidua homoskedasticitu, jelikož vyšla p-hodnota 0,3715 (TS = 7,5751; SV = 7).

Obrázek 8.3: Grafy reziduí výdajového modelu 2



Zdroj: Vlastní zpracování

8.3 Výdajový model 3

Tento model využívá data z období 2013–2022 Q1. V rámci tohoto modelu je výdaj po dvou diferencích vysvětlován počtem důchodců, indexem spotřebitelských cen, současnou a minulou reálnou mzdou po dvou diferencích. Všechny tyto proměnné jsou tedy stacionární. Model je znázorňován rovnicí (8.3).

$$s\widehat{d}v_{ydt} = \beta_0 + sdduch_t + sddcpi_t + sddmzda15_t + sddmzda15_{t-1} \quad (8.3)$$

Z odhadnutých koeficientů z tabulky 8.3 vyplývá, že absolutní člen β_0 je záporný a ostatní proměnné jsou kladné. Avšak β_0 a počet důchodců celkem vyšly jako nevýznamné. To znamená, že model vysvětluje vývoj výdajů na základě proměnných, které jsou brány v úvahu při zvyšování důchodu, a tudíž nebere v potaz počet důchodců. Hodnotově má největší vliv rychlost meziroční změny inflace. Adjustovaný koeficient determinace \bar{R}^2 je v porovnání s ostatními modely nízký, jelikož činí 0,3114. Může to také znamenat, že

zahrnutí časového trendu zdánlivou regresi nevyřešilo.

Tabulka 8.3: Výstup z R pro výdajový model 3

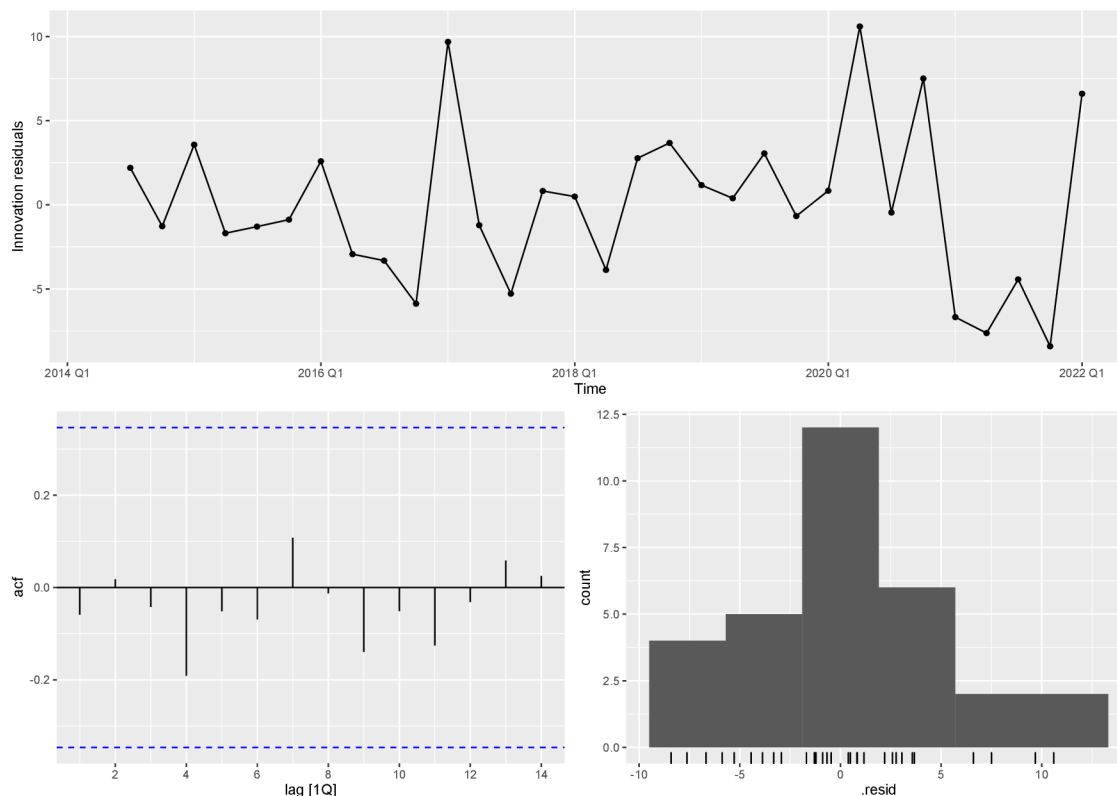
Rezidua:						
Min	1Q	Medián	3Q	Max		
-8,4100	-3,1257	-0,4592	2,6765	10,5940		
Koeficienty:						
	Odhad	Std. chyba	t-statistika	p-hodnota		
β_0	-0,266498	0,992156	-0,269	0,79035		
$sddduch_t$	0,000110	0,000171	0,643	0,52562		
$sddcpi_t$	2,683193	0,883278	3,038	0,00537	**	
$sddmzda15_t$	0,004209	0,001310	3,212	0,00350	**	
$sddmzda15_{t-1}$	0,004230	0,001414	2,992	0,00600	**	
Kódy význam.: 0 '****'		0,001 '***'	0,01 '*'	0,05 '.'	0,1 ''	1
Rezidální standardní chyba: 5,049 při 26 stupních volnosti						
Kof. determinace R^2 :		0,4032	Adj. koef. det. \bar{R}^2 :		0,3114	
F-statistika:		4,391 při 4 a 26 st. vol.		p-hodnota:		0,0076089

Zdroj: Vlastní zpracování

Tento model nemá žádnou významnou autokorelaci reziduí, jak lze spatřit z autokorelační funkce na obrázku 8.4. Pro úplnost je uveden Ljungův–Boxův test. Test nepřekvapivě neprokázal autokorelaci reziduí s velkou p-hodnotou 0,4967 (TS = 2,38). Podíváme-li se na histogram reziduí, vidíme, že tvar má určitou podobnost k normálnímu rozdělení. Provedeme-li Shapirův–Wilkův test, zjistíme, že test neprokázal nenormalitu reziduí při p-hodnotě 0,6708 (TS = 0,9752). Co se týče heteroskedasticity reziduí, byla prokázána Breuschovým–Paganovým testem s p-hodnotou 0,0361 (TS = 10,272; SV = 4). Podíváme-li se na horní graf, vykazují rezidua na určitých místech homoskedasticitu a na stejných místech jako v předchozím modelu se nachází výkyvy. Na základě toho jsem se rozhodl heteroskedasticitu ignorovat. Pokud aplikujeme původní Whiteovy chyby, zjistíme, že

proměnná $sddmzda15_{t-1}$ se stala nevýznamnou (p-hodnota = 0,0525).

Obrázek 8.4: Grafy reziduí výdajového modelu 3



Zdroj: Vlastní zpracování

8.4 Shrnutí

Dle výše uvedených modelů jsou nominální výdaje na dávky důchodového pojištění pozitivně korelovány s úrovní výdajů minulého roku, s celkovým počtem starobních důchodců včetně ostatních důchodců, s indexem spotřebitelských cen a s průměrnou reálnou měsíční mzdou.

Pokud se výdaje řídí těmito modely, lze výdaje na dávky důchodového pojištění snížit snížením počtu důchodců. Ve výsledku se má tedy zkrátit období pro vyplácení důchodů. Toho lze dosáhnout např. manipulací hranice pro přiznání starobního důchodu anebo hypoteticky ovlivněním naděje dožití apod. Na růst výdajů má samozřejmě vliv také inflace, a proto by neměla být vysoká. Úroveň minulého výdaje může znamenat, že problém vysokých výdajů se může do kumulovat.

Jimon et al. (2021) zdůrazňují, že snížení úrovně dávek za účelem snížení výdajů na DP

přináší riziko chudoby a materiální deprivace mezi starobními důchodci. Autoři na základě ekonometrické analýzy zjistili, že výdaje lze mimo jiné snížit poklesem státního deficitu, růstem veřejného dluhu a počtu všech starobních důchodců. Vliv růstu veřejného dluhu na pokles výdajů vysvětlují tím, že veřejné prostředky mohou být alokovány na jejich splácení, čímž může dojít k poklesu úrovně dávek, což by vedlo k poklesu výdajů na DP.

Kapitola 9

Hodnocení modelů a předpovědi

V této kapitole se práce bude v první části věnovat vyhodnocování představených modelů, jak přesně dokážou předpovídat v nekrizovém a krizovém období. V druhé části se budou prostřednictvím křížové validace vybírat vhodné modely k předpovídání a pomocí těchto vybraných modelů se bude předpovídat druhé až čtvrté čtvrtletí roku 2022. Obrázky jsou uvedeny v příloze.

9.1 Vyhodnocování představených modelů

9.1.1 Příjmové modely

Obrázky A.7 a A.8 zobrazují predikci příjmových modelů na období 2018–2020. Na obou obrázcích jsou zobrazeny 80% a 90% predikční intervaly, přičemž 90% interval je označen světlejší barvou.

Obrázek A.7 ukazuje skutečný vývoj příjmů po kalendářní úpravě za období 2015 Q1 – 2022 Q1 a bodových včetně intervalových předpovědí modelů 1 a 2. Model 1 je označen červenou barvou a model 2 modrou barvou. Vidíme, že bodová předpověď modelu 1 je podhodnocená, ale je bližší ke skutečnému vývoji. Na druhé straně bodová předpověď modelu 2 je nadhodnocená a lze spatřit, že na konci roku 2019 je velký rozdíl mezi bodovou předpovědí a skutečným vývojem ve srovnání s modelem 1. Z velké části je skutečný vývoj příjmů pokryt predikčními intervaly.

Obrázek A.8 ukazuje předpovědi modelu 3 a vývoj meziročního růstu, resp. poklesu příjmů. Lze spatřit, že bodová předpověď je ve srovnání se skutečným vývojem podhodnocená. Na rozdíl od předchozích modelů zachycuje predikční interval všechna skutečná pozorování.

Podíváme-li se na chyby předpovědí zachycené různými ukazateli v tabulce 9.1, vidíme,

že nejnižší chyby má příjmový model 3. Winklerovo skóre je nejnižší u třetího modelu. Sloupce srovnání představuje srovnání Winklerova skóre daného modelu se sezónní naivní metodou předpovědi. Hodnoty v tomto sloupci vyjadřují, jak jsou dané modely lepší (+) nebo horší (–) než naivní metoda. Z tohoto sloupce vyplývá, že model 1 je ze všech uvedených modelů lepší než naivní metoda.

Tabulka 9.1: Hodnocení bodové a intervalové předpovědi před krizovým obdobím

	RMSE	MASE	RMSSE	Winkler	Srovnání
Příjmový model 1	3,93	0,889	0,883	17,2	0,897
Příjmový model 2	5,18	1,14	1,16	30,2	0,82
Příjmový model 3	2,33	0,845	0,575	7,56	0,606

Zdroj: Vlastní zpracování

Bodové a intervalové předpovědi modelů v krizovém období (2020 Q1 – 2022 Q1) zachycují obrázky A.9 a A.10. Na obrázcích lze vidět, že žádný z modelů nedokáže zachytit skutečný vývoj. Z modelů zachytil vývoj nejlépe model 3.

Z tabulky 9.2 vyplývá, že nejlepší bodovou a intervalovou predikci dává model 1 a za ním model 3. Co se týče intervalové předpovědi, tak ve srovnání s modelem využívajícím sezónní naivní metodu předpovědi jsou modely 1 a 2 horší, ale model 3 je lepší.

Tabulka 9.2: Hodnocení bodové a intervalové předpovědi za krizové období

	RMSE	MASE	RMSSE	Winkler	Srovnání
Příjmový model 1	7,51	1,55	1,45	80,84	–1,090
Příjmový model 2	9,78	2,04	1,89	115,52	–1,986
Příjmový model 3	7,71	2,25	1,99	112,48	0,256

Zdroj: Vlastní zpracování

9.1.2 Výdajové modely

Obrázky A.11, A.12 a A.13 v příloze zachycují předpovědi pro nekrizové období. V rámci tohoto období jsou bodové předpovědi celkově lepší než u příjmových modelů. Kromě výdajového modelu 2 nejsou bodové předpovědi nadhodnocené ani podhodnocené. Skutečný vývoj výdajů se nachází v rámci 95% predikčního intervalu, u modelů 2 a 3 dokonce v 80% intervalu. Predikční interval modelu 2 je užší než u ostatních modelů. Nejširší je interval u třetího modelu kvůli malému datovému souboru. Podíváme-li se na tabulku 9.3, zjistíme, že nejlepší bodovou předpověď dává model 3 a nejlepší intervalovou předpověď model 2. Ve srovnání se sezónní naivní metodou jsou všechny modely lepší.

Tabulka 9.3: Hodnocení předpovědi výdajových modelů za nekrizové období

	RMSE	MASE	RMSSE	Winkler	Srovnání
Výdajový model 1	3,95	0,935	0,968	21,07	0,700
Výdajový model 2	3,28	0,712	0,804	16,04	0,771
Výdajový model 3	2,27	0,828	0,630	19,54	0,133

Zdroj: Vlastní zpracování

Předpověď výdajových modelů v rámci krizového období jsou zde přesnější než u příjmových modelů. Předpovědi jsou zobrazeny na obrázcích A.14, A.15 a A.16 v příloze. Nejlépe zachytil skutečný vývoj model 1, protože téměř všechna pozorování leží v rámci 95% predikčního intervalu. Bodová předpověď tohoto modelu má tendenci ve srovnání se skutečností nadhodnocovat. Nejhůře zachytil vývoj model 3. Predikční interval 1. modelu je velmi široký, zejména na konci časové řady. Tudiž rovněž zachytil téměř všechna skutečná pozorování. Z tabulky 9.4 vyplývá, že nejlépe predikoval bodově model 2 a intervalově model 1. Intervalová předpověď 3. modelu je stejně dobrá jako sezónní naivní intervalová předpověď.

Tabulka 9.4: Hodnocení předpovědi výdajových modelů za krizové období

	RMSE	MASE	RMSSE	Winkler	Srovnání
Výdajový model 1	10,76	2,14	2,37	35,68	0,769
Výdajový model 2	6,45	1,32	1,42	46,86	0,697
Výdajový model 3	12,25	4,07	3,36	100,30	0,029

Zdroj: Vlastní zpracování

9.2 Výběr a předpověď modelů

9.2.1 Výběr modelů pro predikci

Pro předpovídání byly vybrány jak u příjmových, tak i u výdajových modelů model 2. Důvodem je, že dosahují minimálních hodnot v jednotlivých kritériích. Jednotlivé výsledky kritérií jsou uvedeny níže v tabulce 9.5.

Tabulka 9.5: Výsledky vybraných kritérií

Kritérium	CV	AICc	BIC
Příjmový model 1	0,00230	-539,712	-523,674
Příjmový model 2	0,00140	-580,534	-558,468
Příjmový model 3	9,98271	193,962	205,347
Výdajový model 1	8,51E-06	-1012,96	-993,100
Výdajový model 2	4,37E-06	-1044,62	-1025,04
Výdajový model 3	21,4151	233,953	247,361

Zdroj: Vlastní zpracování

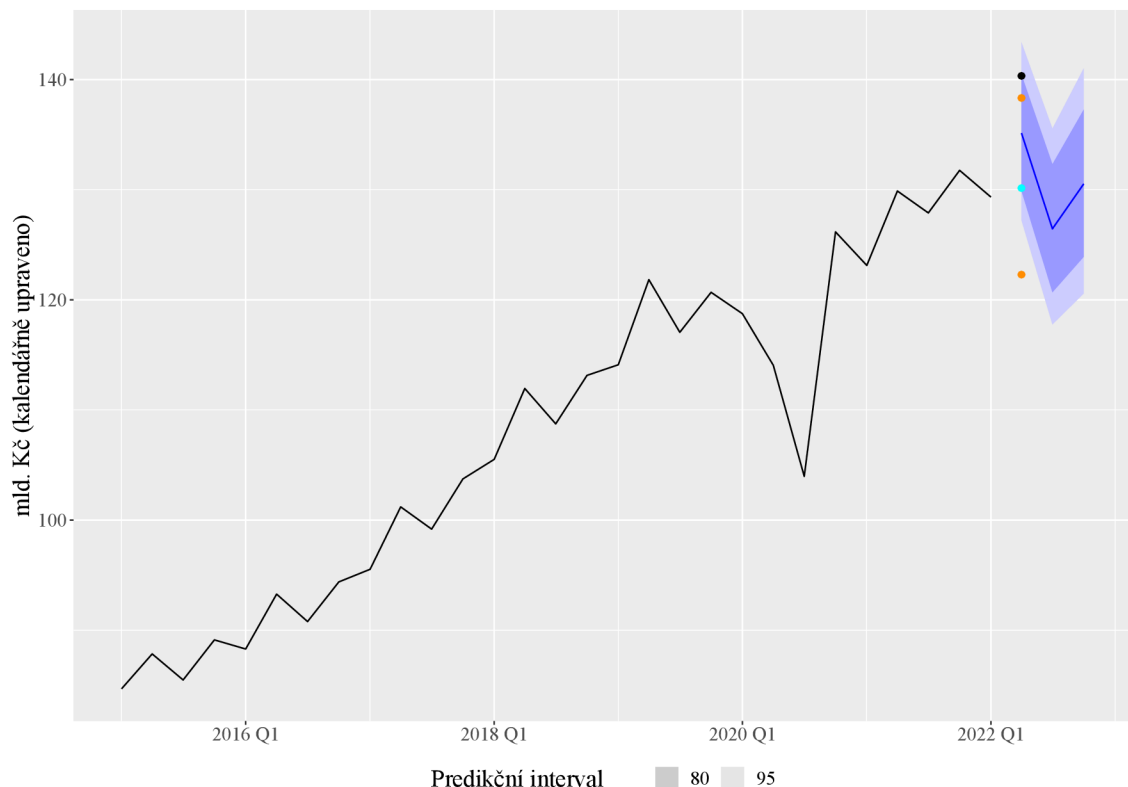
9.2.2 Predikce příjmů a výdajů

Data pro predikci byla čerpána z makroekonomické predikce MF. Avšak ne všechny proměnné obsahovala tato predikce. Jedná se o sezónně očištěná data, ostatní důchodce a proměnné dělené dle pohlaví jako např. důchodci a důchodkyně. Sezónně očištěná data byla získána tak, že se predikovaná sezónně neočištěná data vynásobila průměrným podílem historických sezónně očištěných a sezónně neočištěných dat příslušného čtvrtletí za poslední dva roky. Predikovaný celkový počet zaměstnaných a starobních důchodců byl rozdělen na pohlaví podle průměrného historického poměru mužů a žen v příslušném čtvrtletí za poslední dva roky. Počet ostatních důchodců byl stanoven na poslední známou hodnotu, jelikož autor nepředpokládá velké změny v počtu.

Obrázek 9.1 zobrazuje bodovou a intervalovou predikci čtvrtletních příjmů (kalendářně upravených) na následující tři čtvrtletí. Na obrázku je zahrnut jednak 80%, tak 95% predikční interval. Dle modelu příjmy ve 2. čtvrtletí porostou a v následujícím klesnou a poté opět porostou. Černý bod představuje skutečný příjem ve 2. čtvrtletí po kalendářní úpravě

a činí 140,33 mld. Kč (po zpětném přepočtení 141,33 mld. Kč). Lze vidět, že černý bod se nachází v 80% intervalu. Azurový bod a oranžové body představují bodovou předpověď, resp. 95% predikční interval modelu při využití skutečných dat za 2. čtvrtletí s výjimkou počtu zaměstnaných. Skutečné příjmy se v tomto případě nachází mimo 95% predikční interval. Bodová předpověď je také nižší. V tabulce 9.6 lze nalézt obě bodové (BP) a intervalové (PI) predikce včetně zpětně přepočtených hodnot. Na konci každé části tabulky se nachází hodnoty získané na základě skutečných dat za 2. čtvrtletí 2022.

Obrázek 9.1: Předpověď upravených příjmů na následující tři čtvrtletí



Zdroj: Vlastní zpracování v R

Tabulka 9.6: Bodová a intervalová předpověď příjmů na důchodové pojištění

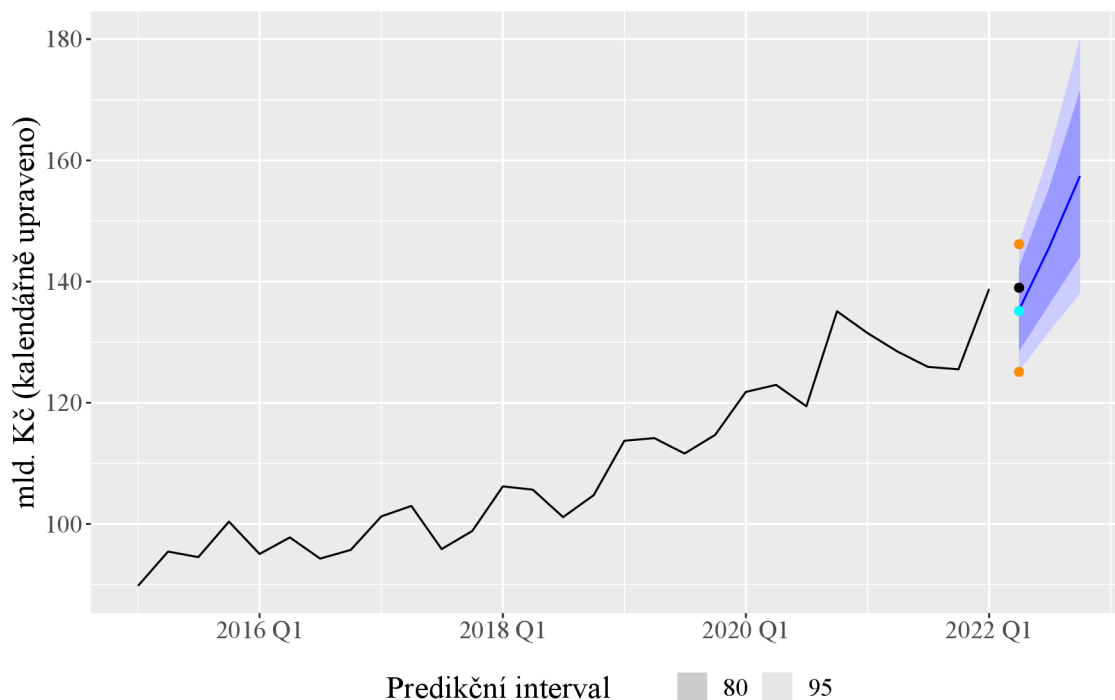
Kalendářně upravené			
	BP	95% PI	80% PI
Q2	135,14	<127,22; 143,37>	<129,90; 140,45>
Q3	126,44	<117,76; 135,54>	<120,68; 132,30>
Q4	130,52	<120,57; 141,01>	<123,90; 137,26>
Q2 – SK	130,14	<122,29; 138,33>	<124,94; 135,42>

Zpětně přepočítané			
	BP	95% PI	80% PI
Q2	136,64	<128,63; 144,96>	<131,34; 142,01>
Q3	127,84	<119,07; 137,05>	<122,02; 133,77>
Q4	131,97	<121,91; 142,58>	<125,28; 138,79>
Q2 – SK	131,59	<123,65; 139,87>	<126,33; 136,92>

Zdroj: Vlastní zpracování

Predikce upravených čtvrtletních výdajů na dávky důchodového pojištění na následující tři čtvrtletí zachycuje obrázek 9.2. Dle tohoto obrázku dojde v následujících čtvrtletích k nárůstu výdajů na dávky DP. Model pro 2. čtvrtletí predikoval pokles výdajů. Barevné body na obrázku mají stejný význam jako v předešlém případě. Černý bod představuje upravené výdaje, které činí 138,99 mld. Kč (ve skutečnosti 140,53 mld. Kč), a nachází se ještě v 80% predikčním intervalu. Ve skutečnosti však došlo k minimálnímu růstu výdajů. Pokud použijeme skutečná data, zjistíme, že azurový bod je blízko bodové předpovědi. V tabulce 9.7 lze nalézt předpovědi výdajového modelu včetně přepočtených hodnot. Poslední řádek v tabulce obsahuje predikované hodnoty při využití skutečných dat.

Obrázek 9.2: Předpověď upravených výdajů na dávky DP na následující tři čtvrtletí



Zdroj: Vlastní zpracování v R

Tabulka 9.7: Bodová a intervalová předpověď výdajů na dávky DP

Kalendářně upravené			
	BP	95% PI	80% PI
Q2	135,32	<125,25; 146,35>	<128,55; 142,32>
Q3	145,71	<131,83; 161,32>	<136,3; 155,52>
Q4	157,41	<138,01; 180,06>	<144,11; 171,45>
Q2 – SK	135,16	<125,1; 146,17>	<128,39; 142,14>
Zpětně přepočítané			
	BP	95% PI	80% PI
Q2	136,82	<126,64; 147,98>	<129,98; 143,9>
Q3	148,95	<134,76; 164,9>	<139,33; 158,98>
Q4	160,91	<141,08; 184,06>	<147,31; 175,26>
Q2 – SK	136,66	<126,49; 147,79>	<129,82; 143,72>

Zdroj: Vlastní zpracování

Pro zajímavost jsou v příloze přidány obrázky A.17 a A.18, které znázorňují předpověď regresních modelů s autokorelovanými rezidui. Tyto modely mají téměř stejné proměnné jako modely využitě k predikci. Pouze byl vypuštěn lineární časový trend. Porovnáním obrázků zjistíme, že predikční interval je užší a že bodová předpověď je bližší skutečným hodnotám. Tyto modely nebyly analyzovány, protože program z nějakého důvodu nedokázal vypočítat standardní chyby pro inferenci, což bylo považováno za podezřelé. Z tohoto důvodu nebyl tato technika aplikována v práci.

Kapitola 10

Závěr

Diplomová práce si kladla za cíl ekonometricky analyzovat udržitelnost systému starobního důchodu a predikovat jeho budoucí vývoj. Udržitelnost byla vnímána z dvojího pohledu, jednak z pohledu příjmů na důchodové pojištění, tak z pohledu výdajů na dávky důchodového pojištění. Analýza byla provedena prostřednictvím mnohonásobné lineární regrese na datech za období 2000 – 2022 Q1 a na jejím základě byla provedena předpověď budoucího vývoje příjmů a výdajů.

V práci byly představeny tři příjmové a tři výdajové modely. Bylo zjištěno, že příjmy jsou pozitivně korelovány s HDP, indexem spotřebitelských cen a příjmy zaměstnaných osob. Na druhou stranu jsou však příjmy negativně korelovány s počtem zaměstnaných. Co se týče výdajů, jsou pozitivně korelovány s minulými výdaji, s počtem celkových důchodců, indexem spotřebitelských cen a reálnou průměrnou měsíční mzdou.

Následně byly analyzovány předpovědi představených modelů. Zkoumala se předpověď modelů za dvouleté období před pandemií Covid-19 (2018 – 2019) a za období krize až do konce datového souboru (2020 – 2022 Q1). Výsledkem bylo, že v nekrizovém období, dokázaly modely dobře zachytit dynamiku vývoje. Avšak predikce v krizovém období se ukázala být ne tolik přesná.

Poté byla provedena křížová validace modelů pro výběr modelu na predikování. Výsledkem bylo, že nejlepší výsledky dosáhl model 2 jak pro příjmový, tak pro výdajový model. Dle vybraných modelů jdou čtvrtletní příjmy v následujících 3 čtvrtletích do strany a dojde k výraznému růstu čtvrtletních výdajů. Pro predikci byla využita data z makroekonomické predikce MF. Skutečné příjmy a výdaje za 2. čtvrtletí leží uvnitř 80% predikčního intervalu, avšak na hraně.

Summary

The first pillar of the pension system of the Czech Republic is based on principles of solidarity and cotribuity, which are taken over from the social protection system of chancellor Otto von Bismarck. The system may not be as sustainable as in the past given the demographic aging of the Czech population.

The objective of this work is to realize an econometric analysis of demographic aspects of economic growth. The analysed object is the aforementioned sustainability of the first pillar. The analysis have been performed from two views, revenue and expenditure of the system. The technique of multiple linear regression have been used for the analysis. The partial objective is the prediction of these variables for the next three quarters. Three models have been introduced for each variable. The data are time series with quarterly seasonality. These reach from year 2000 to 2022 Q1.

The models mentioned above have been then tested how they performed in the period two years before the pandemic Covid-19 in 2020 (non-crisis period) and the rest period (crisis period). The results are that the models performed better in the non-crisis period which is not surprising. In the crisis period expenditure models performed better with the prediction.

For each dependent variable, prediction model II have been chosen because they have better results in the following criteria: leave-one-out cross validation, Corrected Akaike's Information Criterion, Schwarz's Bayesian Information Criterion. The revenue model predicts that the revenue will go sideways and expenditure one that the expenditures will grow. For Q4 the expenditure of the system should be higher than the revenue.

The results of linear regression model are that revenues are positively correlated with CPI, real GDP and total income of employees and negatively correlated with the level of

employment and unemployment. The first mentioned doesn't make any sense. Expenditures are positively correlated with past expenditures, the level of beneficiaries, CPI, real monthly average gross wage.

Keywords: Econometric analysis, Linear regression, Sustainability, Pension system, Forecasting

Bibliografie

- Bezděk, V. (2000). *Penzijní systémy obecně i v kontextu české ekonomiky: (současný stav a potřeba reform)* (I. díl). Česká národní banka. Dostupné 15. září 2022 z https://www.cnb.cz/export/sites/cnb/en/economic-research/.galleries/research_publications/mp_wp/download/vp2500.pdf
- Bickel, P. J. & Doksum, K. A. (1981). An Analysis of Transformations Revisited. *Journal of the American Statistical Association*, 76(374), 296–311. doi:10.1080/01621459.1981.10477649
- Box, G. E. P. & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–243. doi:10.1111/j.2517-6161.1964.tb00553.x
- Cipra, T. (2013). *Finanční ekonometrie* (2., upr. vyd). Praha: Ekopress.
- Döring, D. (1997). Is the German Welfare State Sustainable? *Restructuring the Welfare State*, 38–61. doi:10.1007/978-3-642-60652-6_4
- Ginsburg, N. (1979). Class, Capital and Social Policy. doi:10.1007/978-1-349-16169-0
- Guerrero, V. M. (1993). Time-series analysis supported by power transformations. *Journal of Forecasting*, 12(1), 37–48. doi:10.1002/for.3980120104
- Holzmann, R., Paul, R., Dorfman, H. & Dorfman, M. (2008). Pension Systems and Reform Conceptual Framework, World Bank. Dostupné 13. září 2022 z <https://documents1.worldbank.org/curated/en/716871468156888545/pdf/461750NWP0Box334081B01PUBLIC10S.pdf>
- Hübelová, D., Manea, B. C. & Machučová, B. (2021). Vývoj naděje dožití a odvrátitelná úmrtnost v krajích České republiky v období let 2006–2016 a jejich determinanty. *Demografie*, 63(1), 48–65. Dostupné 14. září 2022 z https://www.czso.cz/documents/10180/143550791/13005321q1_48-65.pdf/1dfbcc9c-5d8b-492b-9abf-a9ff7398b7fc?version=1.1

- Hyndman, R. J. & Athanasopoulos, G. (2021). *Forecasting: principles and practice*. Dostupné 26. srpna 2022 z <https://otexts.com/fpp3>
- Jimon, S. A., Dumiter, F. C. & Baltes, N. (2021). *Financial Sustainability of Pension Systems*. doi:10.1007/978-3-030-74454-0
- Klepárník, V., Štípek, V., Bittner, J. & Pícl, M. (2017). *Penzijní reformy: od Chile po středovýchodní Evropu*. Praha: Úřad vlády ČR. Dostupné z <https://www.vlada.cz/assets/media-centrum/dulezite-dokumenty/Penzijni-reformy.pdf>
- Klufová, R. & Poláková, Z. (2010). *Demografické metody a analýzy: demografie české a slovenské populace*. Praha: Wolters Kluwer Česká republika.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3), 159–178. doi:10.1016/0304-4076(92)90104-Y
- Montero, A. U. & Simões, O. A. (2021). Life Expectancy Improvements by Age, Class, and Mortality Chapter in France, Czechia, and the United States. *Demografie*, 63(4), 228–245. doi:10.54694/dem.0286
- MPSV. (2022). Hlavní změny v důchodovém pojištění. Dostupné 16. září 2022 z <https://www.mpsv.cz/web/cz/hlavni-zmeny-v-duchodovem-pojisteni>
- OECD. (2020). *Pension Markets in Focus 2020*. Dostupné 13. září 2022 z <http://www.oecd.org/finance/pensionmarketsinfofocus.htm>
- OECD & Unie, E. (2020). *Health at a Glance: Europe 2020: State of Health in the EU Cycle*. doi:10.1787/82129230-en
- Šprocha, B. (2022). Bezdetnosť a odkladanie rodenia prvých detí v krajinách Vyšehradskej skupiny. *Demografie*, 64(2), 139–157. doi:10.54694/dem.0302
- Wooldridge, J. M. (2009). *Introductory Econometrics: A Modern Approach* (Fourth Edition). Mason: South-Western Cengage Learning.
- Wooldridge, J. M. (2015). *Introductory Econometrics: A Modern Approach* (Sixth Edition). Boston: Cengage Learning.
- Zákon č. 155/1995 Sb., o důchodovém pojištění.
- Zákon č. 427/2011 Sb., o doplňkovém penzijním spoření.

Seznam obrázků

4.1	Věková struktura ČR v letech 2021 a 1945	22
6.1	Vývoj příjmů z pojistného a příslušenství na důchodové pojištění	32
6.2	Vývoj výdajů na dávky důchodového pojištění	34
6.3	Logaritmická transformace příjmů na důchodové pojištění	37
6.4	Boxova–Coxova transformace příjmů na důchodové pojištění	37
6.5	Logaritmická transformace výdajů na dávky důchodového pojištění	38
6.6	Boxova–Coxova transformace výdajů na dávky důchodového pojištění	39
7.1	Matice korelačních polí nezávislých proměnných příjmových modelů	46
7.2	Grafy reziduí příjmového modelu 1	48
7.3	Grafy reziduí příjmového modelu 2	50
7.4	Grafy reziduí příjmového modelu 3	52
8.1	Matice korelačních polí nezávislých proměnných výdajových modelů	56
8.2	Grafy reziduí výdajového modelu 1	58
8.3	Grafy reziduí výdajového modelu 2	60
8.4	Grafy reziduí výdajového modelu 3	62
9.1	Předpověď upravených příjmů na následující tři čtvrtletí	69
9.2	Předpověď upravených výdajů na dávky DP na následující tři čtvrtletí	70
A.1	Vývoj ekonomicky aktivní pracovní síly, zaměstnaných a nezaměstnaných	85
A.2	Vývoj HDP	85
A.3	Vývoj indexu spotřebitelských cen	86
A.4	Vývoj náhrad zaměstnancům a čistého smíšeného důchodu	86
A.5	Vývoj počtu důchodů	87

A.6	Vývoj nominální a reálné průměrné měsíční mzdy	87
A.7	Předpovědi příjmových modelů 1 a 2 před krizovým obdobím	88
A.8	Předpovědi příjmového modelu 3 před krizovým obdobím	88
A.9	Předpovědi příjmových modelů 1 a 2 za krizové období	89
A.10	Předpovědi příjmového modelu 3 za krizové období	89
A.11	Předpovědi výdajového modelu 1 před krizovým obdobím	90
A.12	Předpovědi výdajového modelu 2 před krizovým obdobím	90
A.13	Předpovědi výdajového modelu 3 před krizovým obdobím	91
A.14	Předpovědi výdajového modelu 1 za krizové období	91
A.15	Předpovědi výdajového modelu 2 za krizové období	92
A.16	Předpovědi výdajového modelu 3 za krizové období	92
A.17	Předpověď příjmů na DP regr. m. s autokorelovanými rezidui	93
A.18	Předpověď výdajů na dávky DP regr. m. s autokorelovanými rezidui	93

Seznam tabulek

3.1	Některé změny penzijního systému	18
7.1	Výstup z R pro příjmový model 1	47
7.2	Výstup z R pro příjmový model 2	49
7.3	Výstup z R pro příjmový model 3	51
8.1	Výstup z R pro výdajový model 1	57
8.2	Výstup z R pro výdajový model 2	59
8.3	Výstup z R pro výdajový model 3	61
9.1	Hodnocení bodové a intervalové předpovědi před krizovým obdobím	66
9.2	Hodnocení bodové a intervalové předpovědi za krizové období	66
9.3	Hodnocení předpovědi výdajových modelů za nekrizové období	67
9.4	Hodnocení předpovědi výdajových modelů za krizové období	67
9.5	Výsledky vybraných kritérií	68
9.6	Bodová a intervalová předpověď příjmů na důchodové pojištění	69
9.7	Bodová a intervalová předpověď výdajů na dávky DP	71
A.1	Přepočtení příjmů SZ na příjmy DP	83
A.2	Přepočtení výdajů na dávky důchodového pojištění	84

Příloha

Tabulka A.1: Přepočtené příjmy SZ na příjmy DP

Čas	Příjmy SZ	Roční příjmy SZ	Roční příjmy DP	Podíl příjmů DP a SZ	Upravené příjmy
2008 Q1	82,56				76,50
2008 Q2	84,21				78,02
2008 Q3	83,24				77,13
2008 Q4	85,54	335,55	310,9	0,9265	79,26
2009 Q1	78,78				78,57
2009 Q2	77,72				77,51
2009 Q3	70,05				69,86
2009 Q4	75,36	301,91	301,1	0,9973	75,16
2010 Q1	74,1				73,89
2010 Q2	77,13				76,91
2010 Q3	77,47				77,25
2010 Q4	80,69	309,39	308,5	0,9971	80,46
2011 Q1	77,23				77,13
2011 Q2	80,31				80,20
2011 Q3	79,78				79,67
2011 Q4	82,61	319,93	319,5	0,9987	82,50
2012 Q1	80,09				80,00
2012 Q2	81,65				81,56
2012 Q3	79,91				79,82
2012 Q4	82,02	323,67	323,3	0,9989	81,93
			Průměr	0,9980	

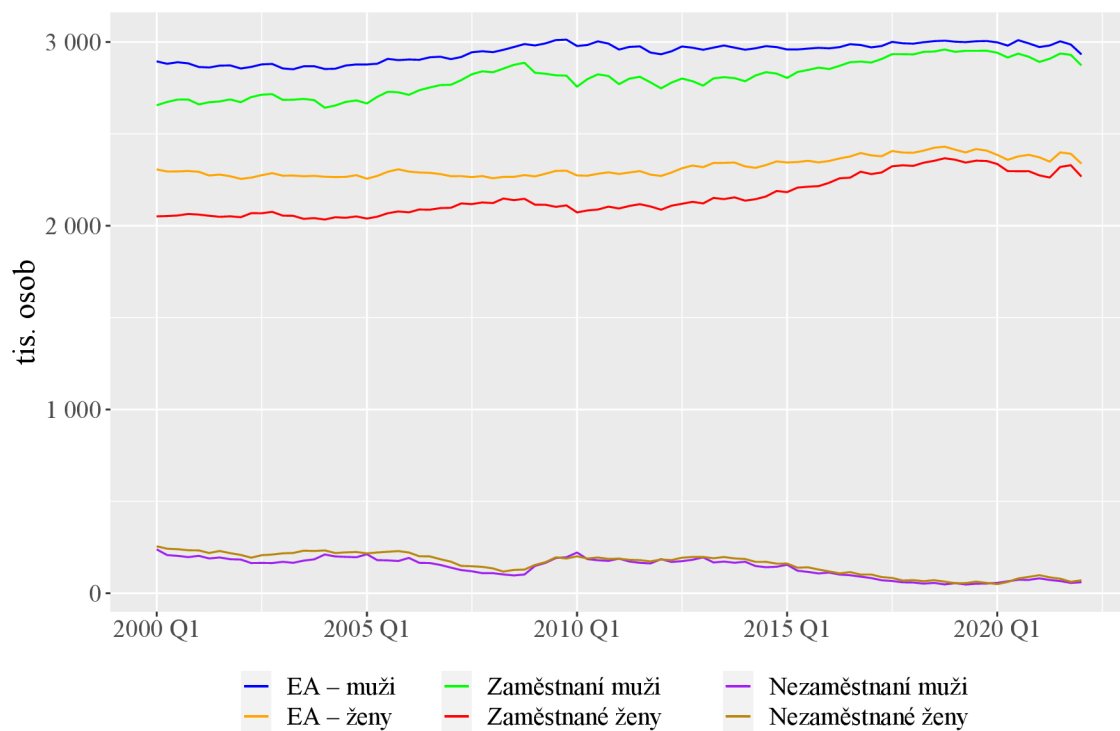
Zdroj: Vlastní zpracování

Tabulka A.2: Přepočtené výdaje na dávky důchodového pojištění

Čas	Vypočítané výdaje	Roční vyp. výdaje	Skutečné výdaje	Podíl vyp. a sk. výdajů	Přepočtené výdaje
2004 Q1	52,57	215,88	225,19	0,9427	55,77
2004 Q2	52,97				56,19
2004 Q3	53,22				56,46
2004 Q4	53,51				56,76
2005 Q1	56,77	232,06	241,16	0,9465	59,97
2005 Q2	56,98				60,20
2005 Q3	57,11				60,33
2005 Q4	57,41				60,65
2006 Q1	60,50	248,63	266,23	0,9189	65,84
2006 Q2	61,20				66,60
2006 Q3	61,36				66,77
2006 Q4	61,59				67,02
2007 Q1	65,54	268,72	282,60	0,9358	70,03
2007 Q2	65,97				70,50
2007 Q3	66,29				70,84
2007 Q4	66,66				71,23
2008 Q1	69,64	292,97	304,90	0,9463	73,60
2008 Q2	70,07				74,04
2008 Q3	74,18				78,39
2008 Q4	74,64				78,87
			Průměr	0,9381	

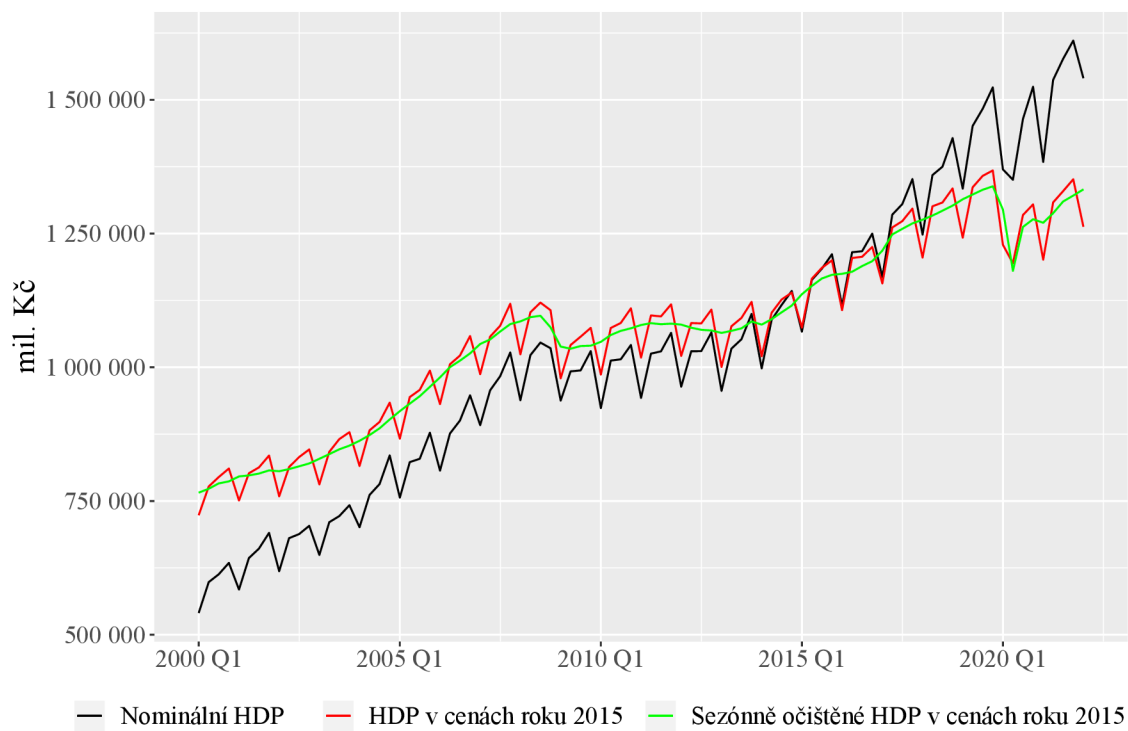
Zdroj: Vlastní zpracování

Obrázek A.1: Vývoj ekonomicky aktivní pracovní síly, zaměstnaných a nezaměstnaných



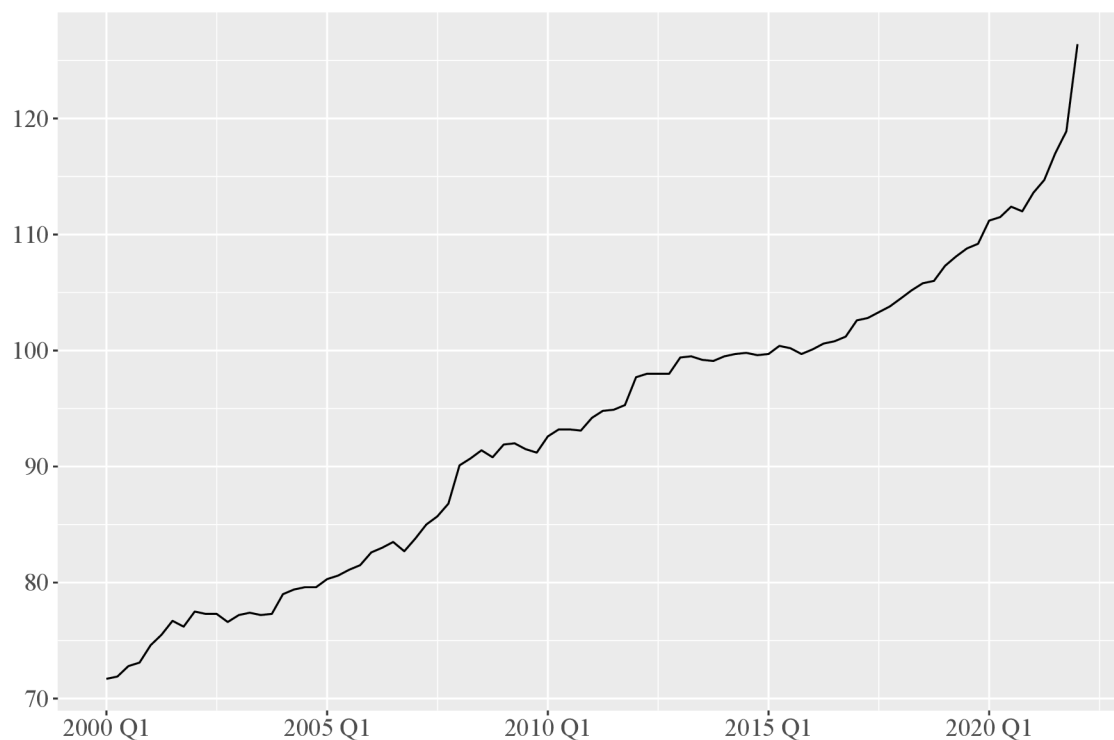
Zdroj: ČSÚ, vlastní zpracování

Obrázek A.2: Vývoj HDP



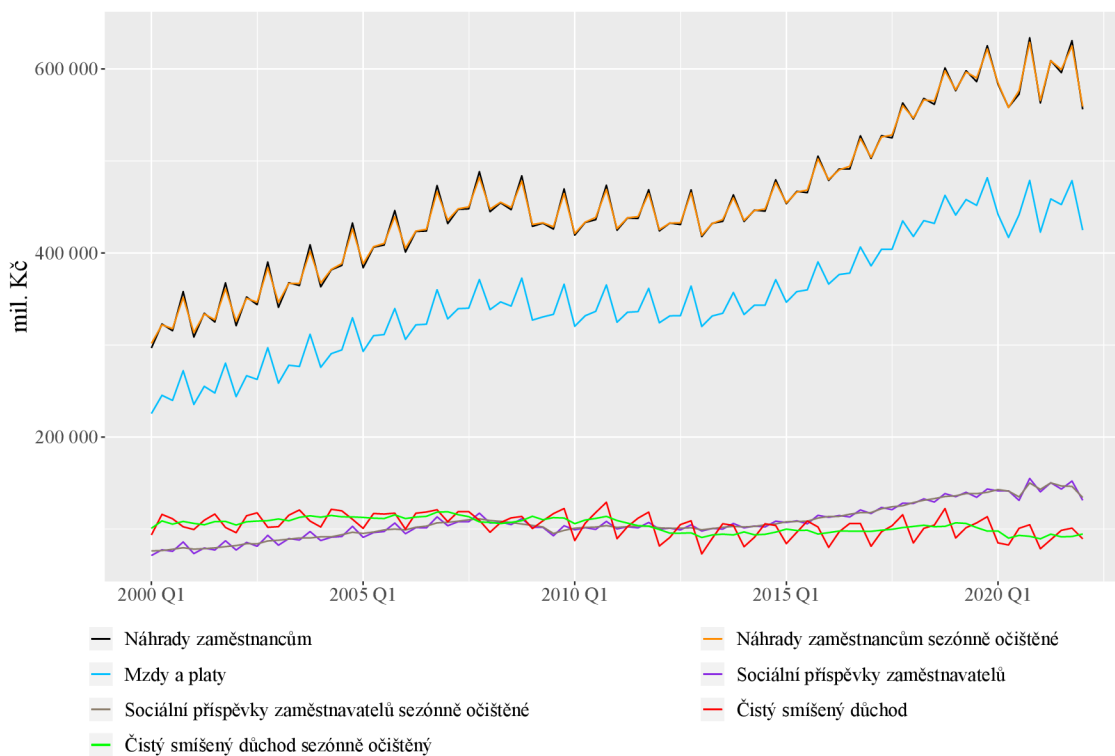
Zdroj: ČSÚ, vlastní zpracování

Obrázek A.3: Vývoj indexu spotřebitelských cen



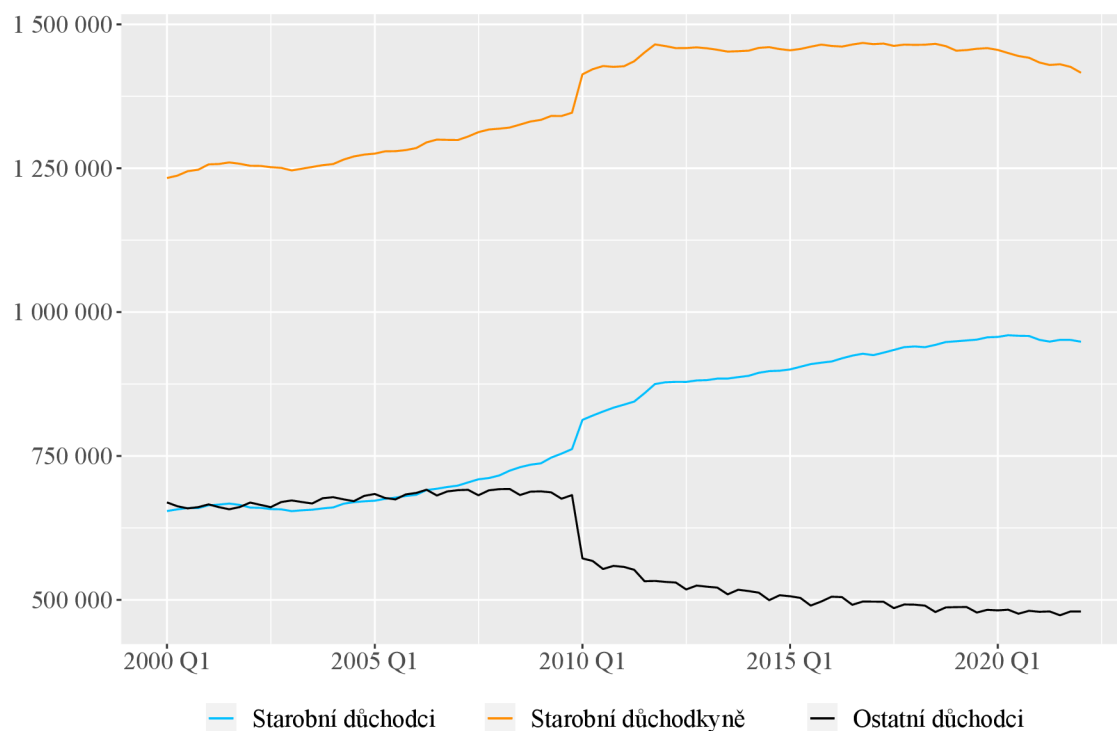
Zdroj: ČSÚ, vlastní zpracování

Obrázek A.4: Vývoj náhrad zaměstnancům a čistého smíšeného důchodu



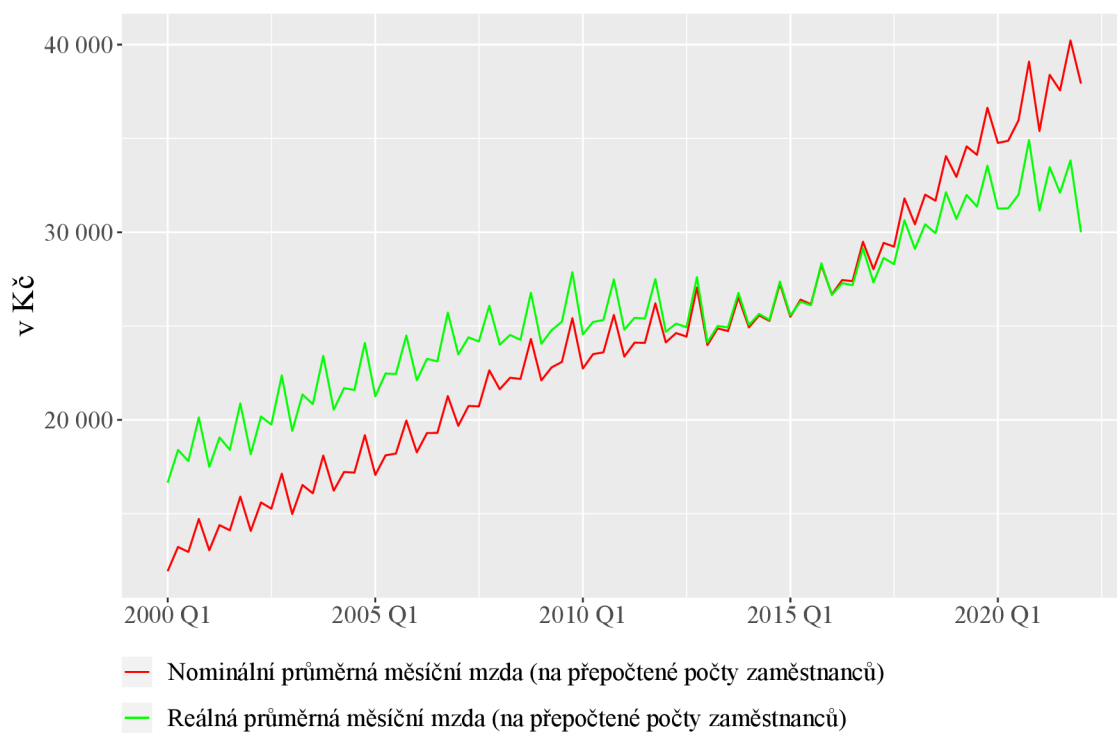
Zdroj: ČSÚ, vlastní zpracování

Obrázek A.5: Vývoj počtu důchodů



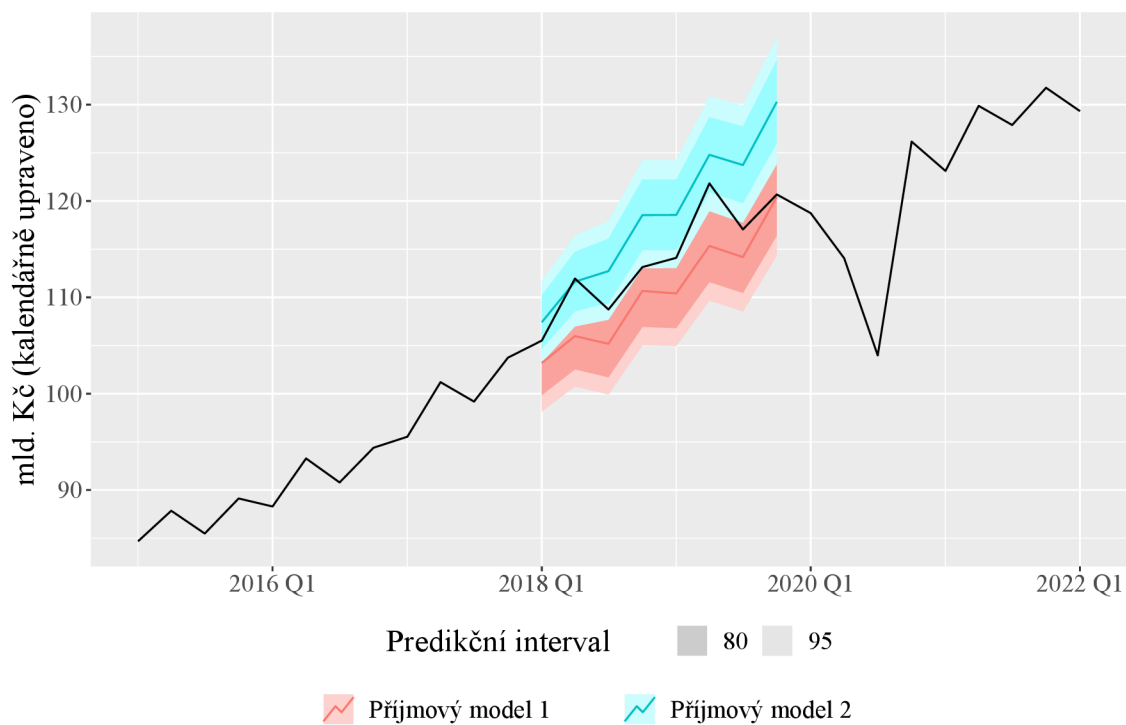
Zdroj: ČSSZ, vlastní zpracování

Obrázek A.6: Vývoj nominální a reálné průměrné měsíční mzdy



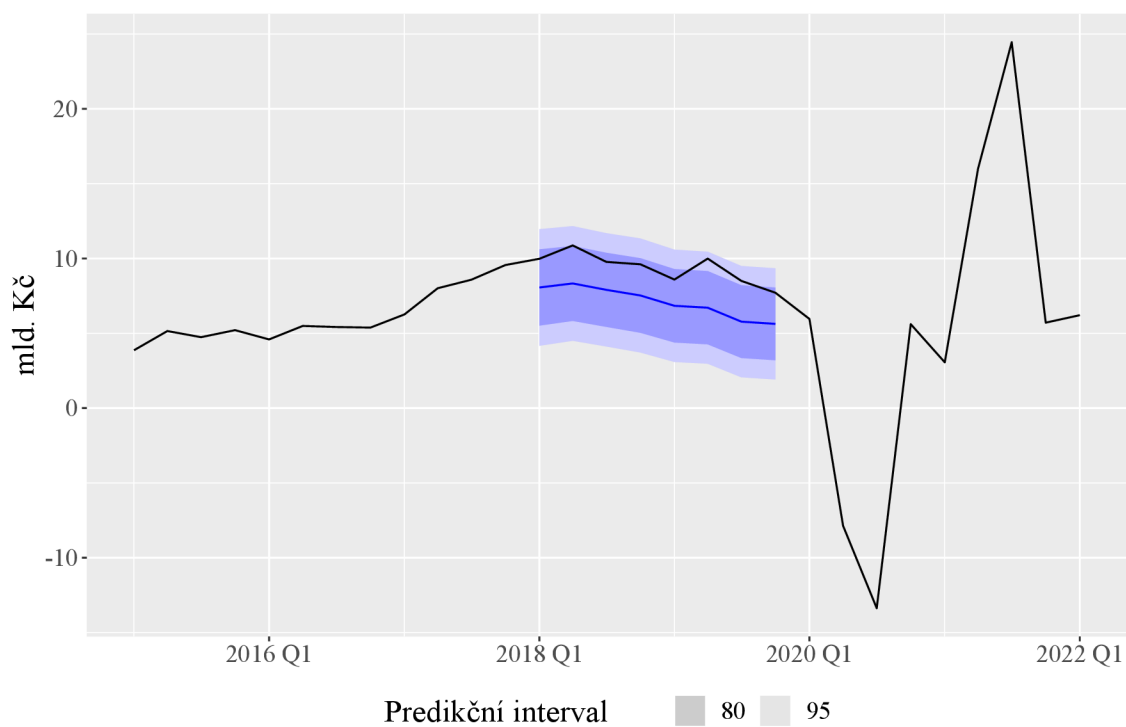
Zdroj: ČSÚ, vlastní zpracování

Obrázek A.7: Předpovědi příjmových modelů 1 a 2 před krizovým obdobím



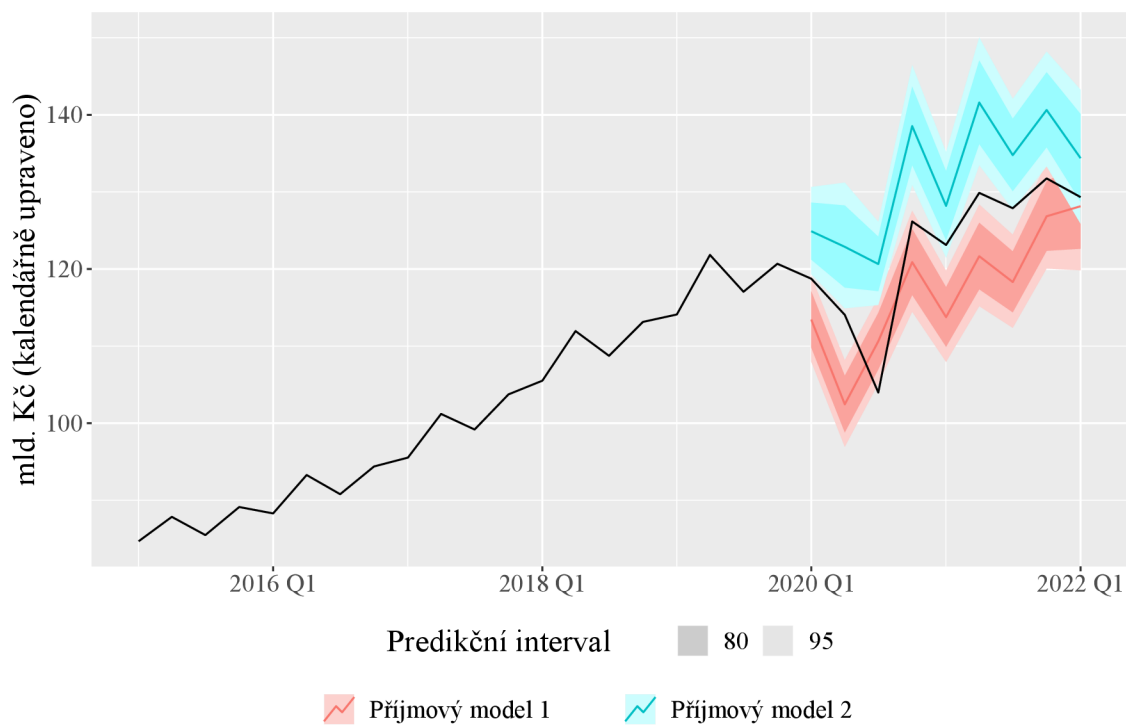
Zdroj: Vlastní zpracování

Obrázek A.8: Předpovědi příjmového modelu 3 před krizovým obdobím



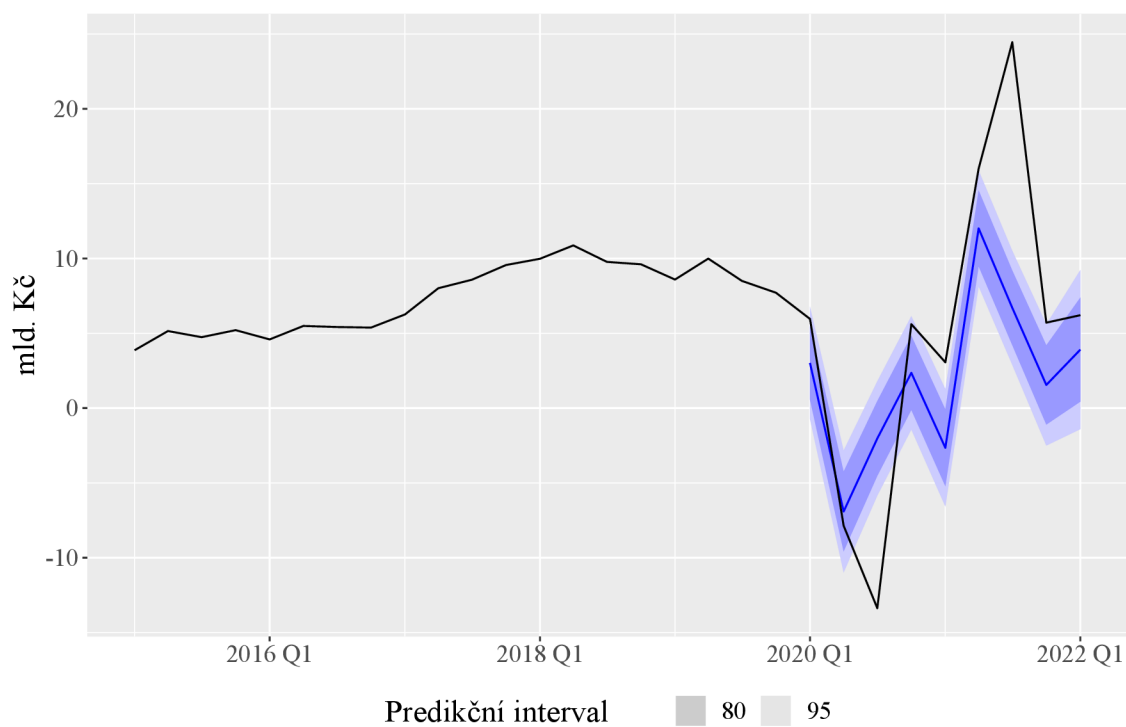
Zdroj: Vlastní zpracování

Obrázek A.9: Předpovědi příjmových modelů 1 a 2 za krizové období



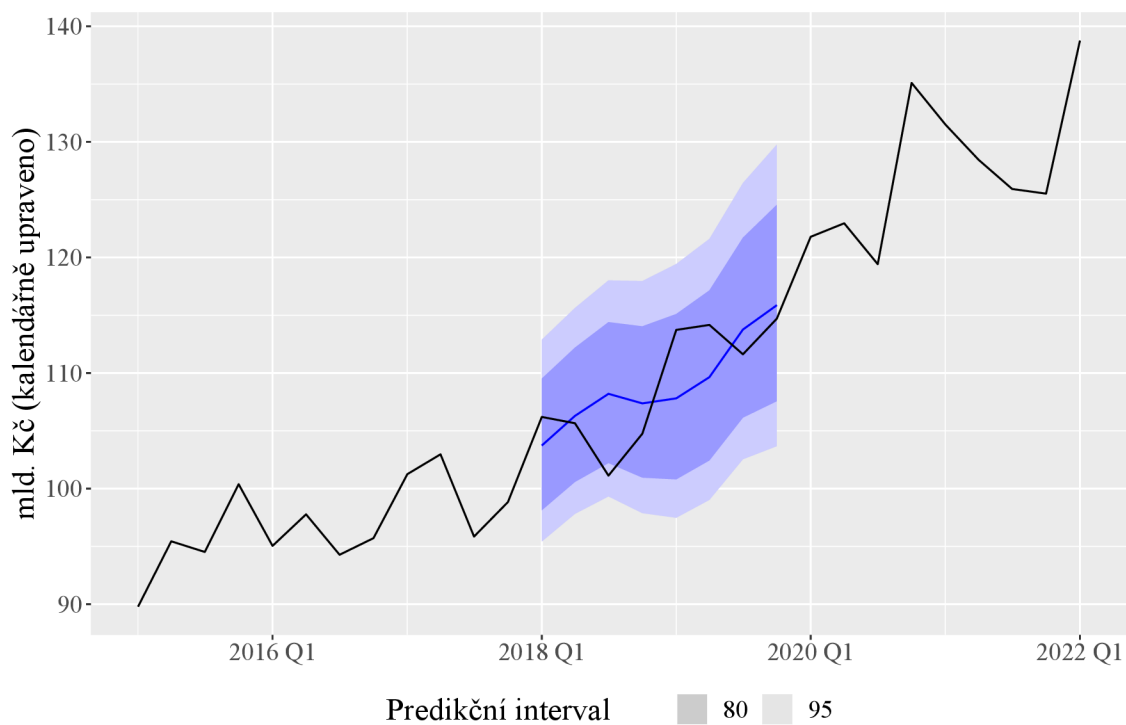
Zdroj: Vlastní zpracování

Obrázek A.10: Předpovědi příjmového modelu 3 za krizové období



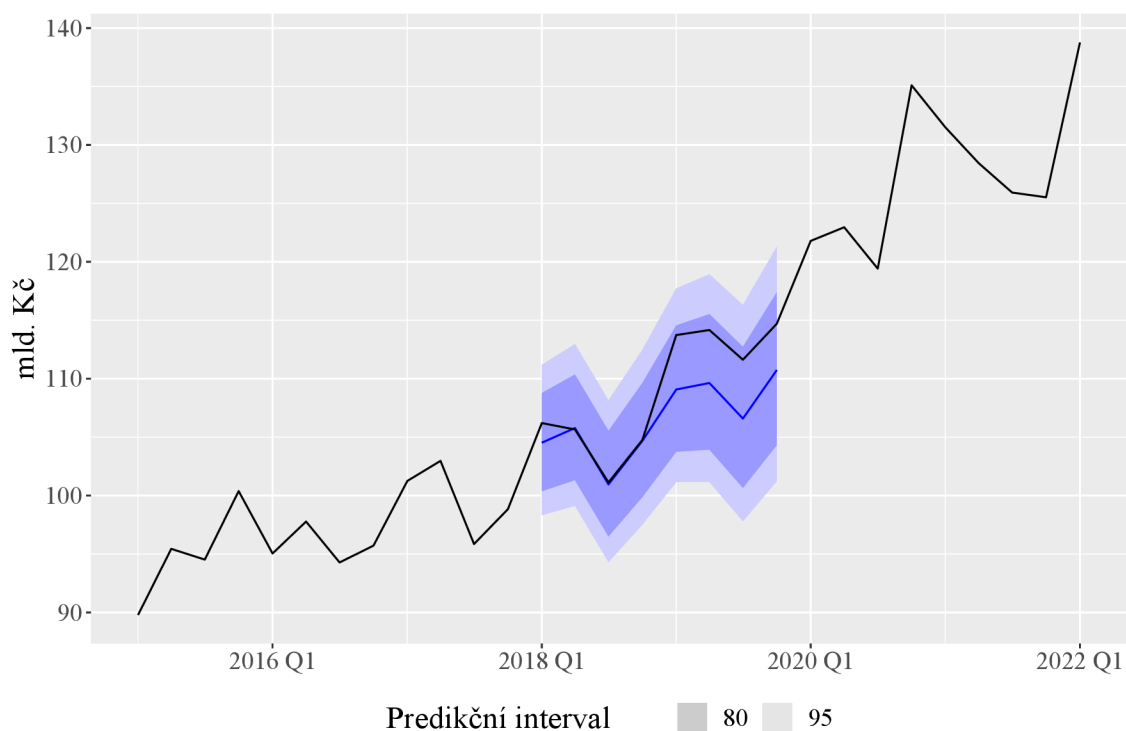
Zdroj: Vlastní zpracování

Obrázek A.11: Předpovědi výdajového modelu 1 před krizovým období



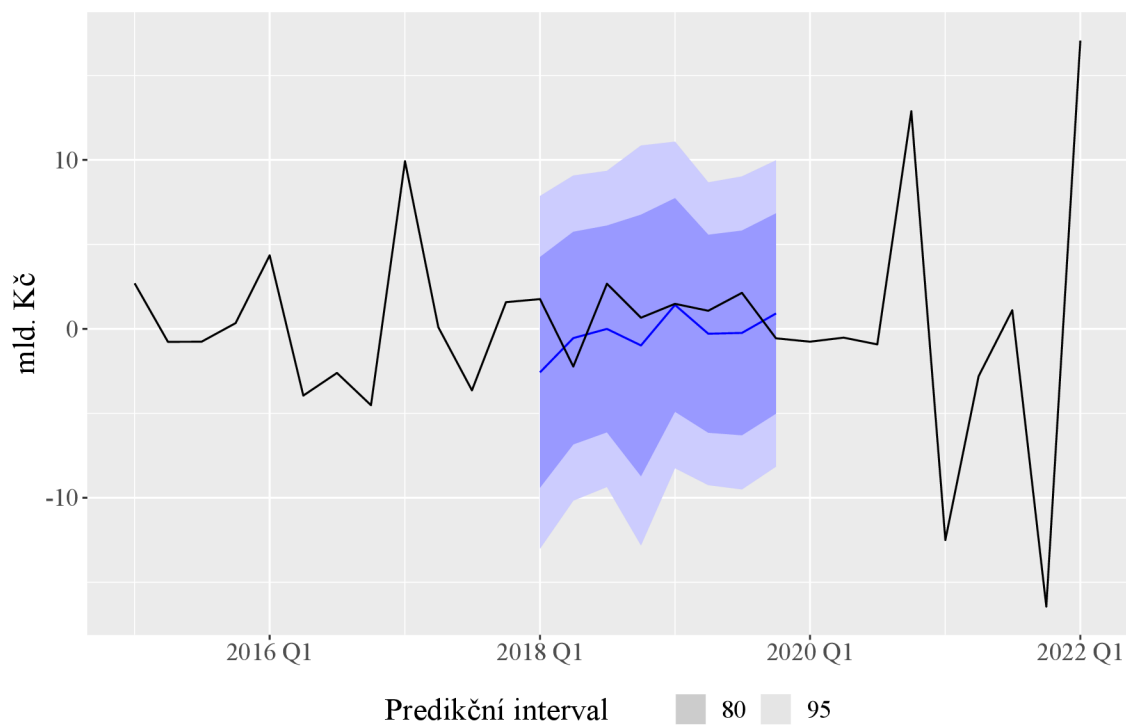
Zdroj: Vlastní zpracování

Obrázek A.12: Předpovědi výdajového modelu 2 před krizovým období



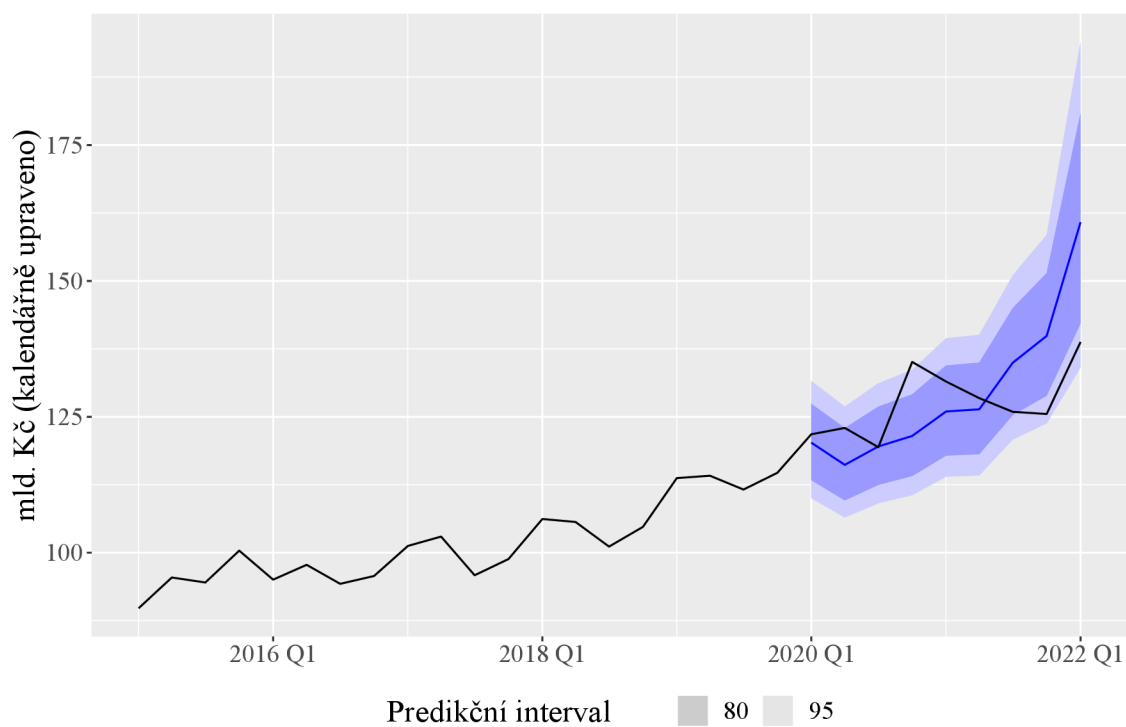
Zdroj: Vlastní zpracování

Obrázek A.13: Předpovědi výdajového modelu 3 před krizovým obdobím



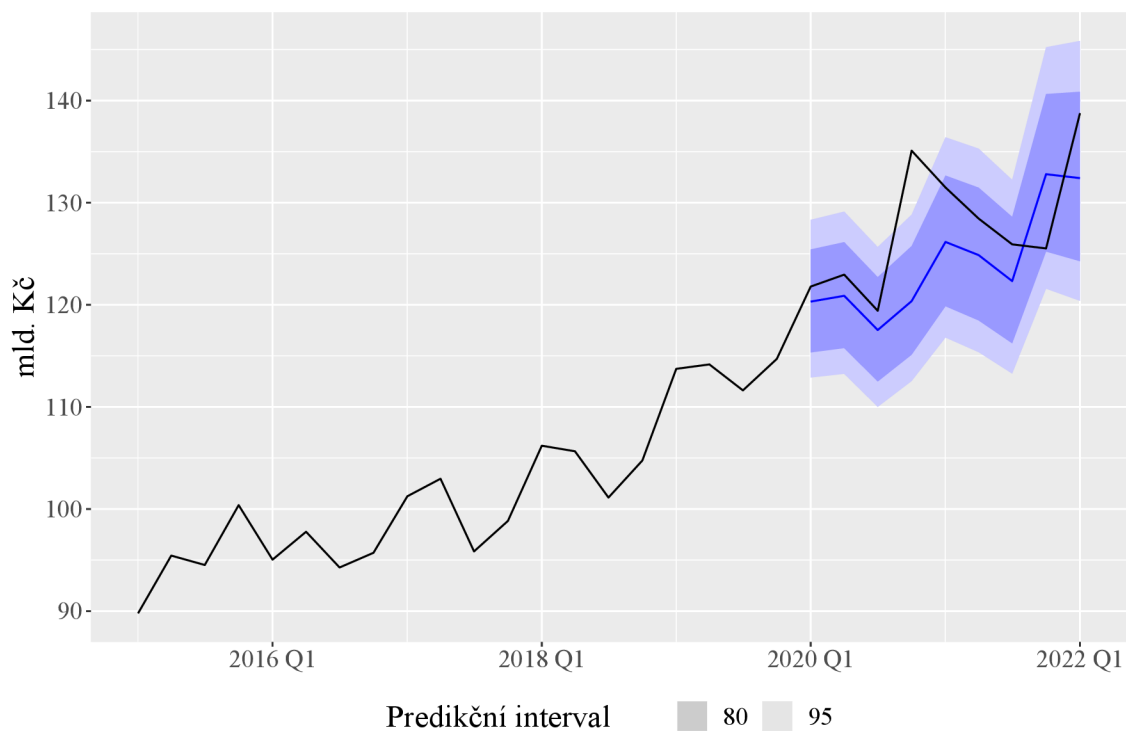
Zdroj: Vlastní zpracování

Obrázek A.14: Předpovědi výdajového modelu 1 za krizové období



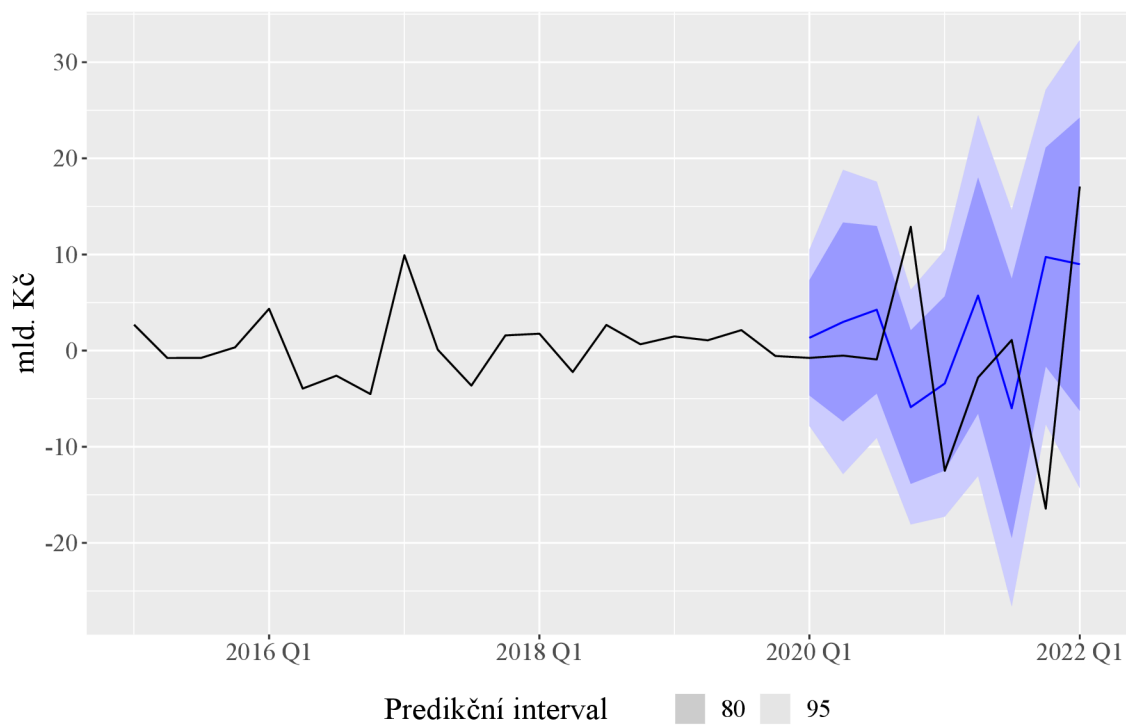
Zdroj: Vlastní zpracování

Obrázek A.15: Předpovědi výdajového modelu 2 za krizové období



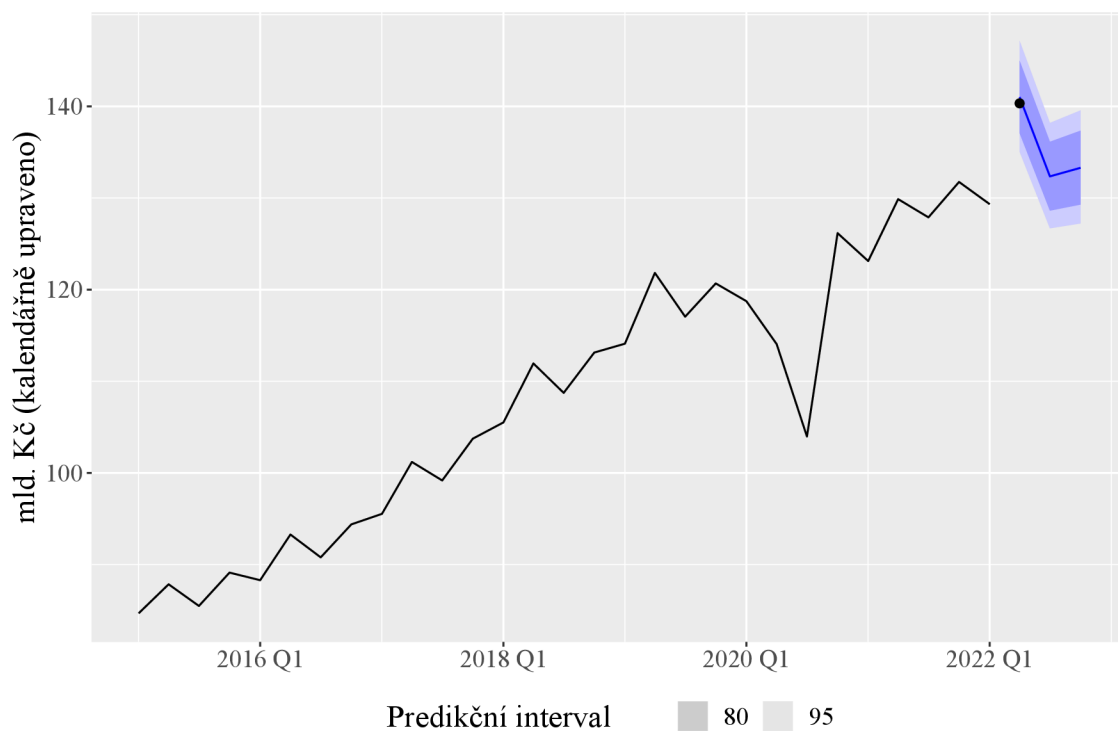
Zdroj: Vlastní zpracování

Obrázek A.16: Předpovědi výdajového modelu 3 za krizové období



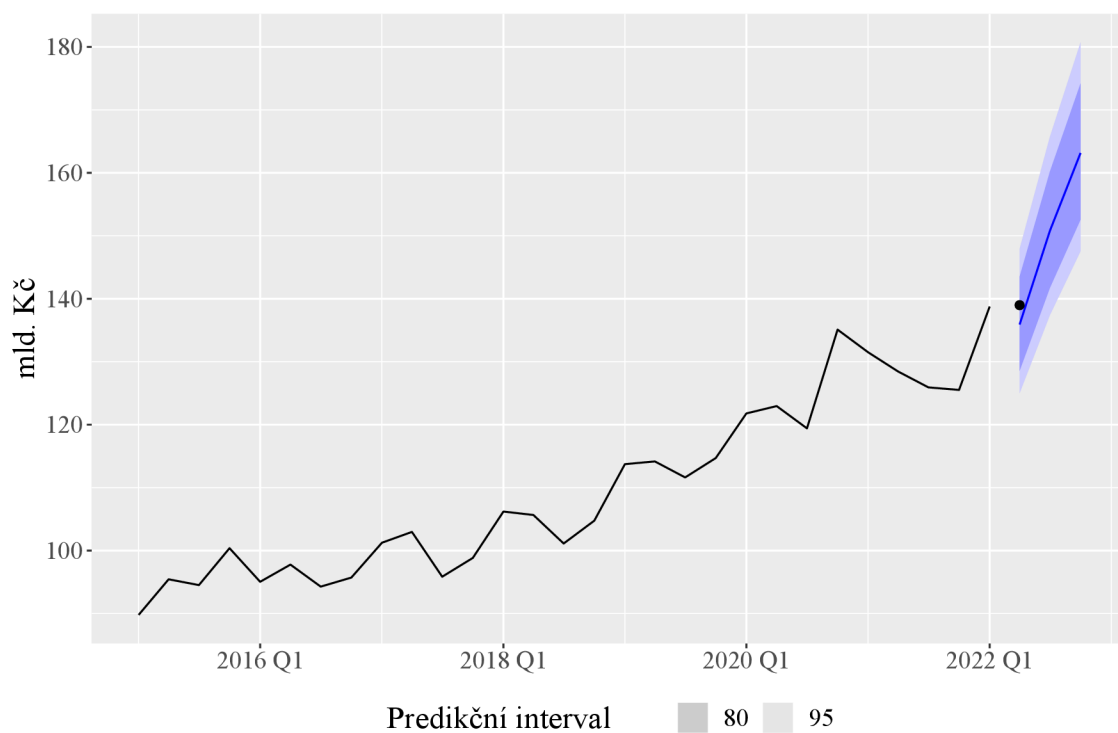
Zdroj: Vlastní zpracování

Obrázek A.17: Předpověď příjmů na DP regr. m. s autokorelovanými rezidui



Zdroj: Vlastní zpracování

Obrázek A.18: Předpověď výdajů na dávky DP regr. m. s autokorelovanými rezidui



Zdroj: Vlastní zpracování