

Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky

Katedra ekonometrie



Bakalářská práce

Ekonometrická analýza nezaměstnanosti v ČR

LS 2016/17

Autor: Anastasiia Melnyk

Vedoucí práce: Ing. Tomáš Formánek, Ph.D.

Čestné prohlášení:

Prohlašuji, že tuto bakalářskou práci na téma „Ekonometrická analýza nezaměstnanosti v ČR“ jsem vypracovala samostatně. Veškeré použité zdroje a podklady uvádím v seznamu literatury a internetových zdrojů.

V Praze dne května 2017

.....

Anastasiia Melnyk

Poděkování:

Chtěla bych poděkovat vedoucímu této bakalářské práce. Ing. Tomáši Formánkovi, Ph.D za ochotu vést moji práci, za trpělivost a pomoc s českou gramatikou, za cenné připomínky, bez kterých by tato práce neměla správný závěr.

Taky bych ráda poděkovala svému kamarádovi Tomáši Pituchovi za kontrolu a opravu práce z hlediska české lexiky a gramatiky.

Abstrakt:

Tato bakalářská práce se zabývá problematikou nezaměstnanosti jako celku, v praktické části konkrétně — ukazatelem obecné míry nezaměstnanosti — v ČR a v Rakousku. Práce nejprve probere teorie obecných pojmů pro seznámení s nezaměstnaností v otevřené ekonomice, pak vysvětlí jednotlivé druhy nezaměstnanosti a faktory které je způsobují. Odděleně popíše Okunův zákon a uvede regresní rovnici, která pak bude ověřena na empirických datech. Dále nás seznámí s vybranými ekonometrickými postupy analýzy nezaměstnanosti, primárně se zaměří na problematiku časových řad, s nimiž pak bude pracovat v praktické části. Konkrétně projde obecné pojmy v ekonometrii, techniky odhadů parametrů regresní rovnice, dekompozici časových řad a jejich charakteristiky, použití dummy proměnných při ekonometrickém modelování, ekonometrické a statistické verifikace modelu. V praktické části se provede empirická analýza a vyhodnocení výsledků – nejprve dekompozice časových řad a řádné očištění o sezónnost, pak test časových řad na jednotkový kořen a následně bude použito diferencování pro očištění o trend. Jako hypotézy budou zvoleny některé faktory popsané v teoretické části, a to: úroveň vzdělání občanů, minimální mzda, globální krize z roku 2008, intervence ČNB, a HDP podle Okunova zákona. Budou modelovány regresní rovnice závislosti míry nezaměstnanosti na jednotlivých faktorech, model bude také postupně rozšiřován o další proměnné, a testován na autokorelaci a heteroskedasticitu. Odhady za použití metody zobecněných nejmenších čtverců budou porovnány s odhady MNČ. Práci shrne závěr o provedené empirické analýze.

Klíčová slova: míra nezaměstnanosti, Okunův zákon, ekonometrická analýza, časová řada.

Abstract:

This bachelor thesis deals with the issue of unemployment in general, and in the practical part concretely – unemployment rate in the Czech Republic and Austria. Firstly, the thesis examines the theory of general concepts for getting acquainted with unemployment in an open economy, then explains different types of unemployment and the factors that cause them. Separately describes Okun's law and introduces the regression equation, which will be verified later on empirical data. Next, it will introduce to us selected econometric methods of unemployment analysis, primarily focusing on the problems of time series, with which it will work in the practical part. Particularly, it will describe general terms in econometrics, techniques of estimation of parameters for regression equation, decomposition of time series and their characteristics, use of dummy variables in econometric modeling, econometric and statistical verification of model. In the practical part empirical analysis and evaluation of the results are going to be performed – firstly, decomposition of time series and proper seasonal adjustment, then test of time series for unit root and subsequently differentiation for detrending. As a hypothesis, some of the factors described in the theoretical part will be selected: citizens' level of education, minimum wage, global crisis of 2008, intervention of CNB, and GDP (based on Okun's law). Regression of unemployment rate on individual factors will be run, the model also will be gradually expanded by other variables and will be tested for the unit root and heteroskedasticity. Estimates using the Generalized Least Squares method will be compared to the OLS estimates. Thesis will be summarized by conclusion of the empirical analysis.

Key words: unemployment rate, Okun's law, econometric analysis, time series.

Obsah

ÚVOD	0
1 FAKTORY OVLIVŇUJÍCÍ NEZAMĚŠTNANOST	1
1.1 OKUNŮV ZÁKON	4
1.2 VÝVOJ NEZAMĚŠTNANOSTI V ČR.....	5
1.3 VÝVOJ NEZAMĚŠTNANOSTI V RAKOUSKU	6
2 VYBRANÉ EKONOMETRICKÉ POSTUPY ANALÝZY NEZAMĚŠTNANOSTI.....	8
2.1 OBECNÉ POJMY EKONOMETRICKÉHO MODELOVÁNÍ	8
2.2 METODA NEJMENŠÍCH ČTVERCŮ, PŘEDPOKLADY JEJÍ POUŽITÍ A VLASTNOSTÍ ZÍSKANÝCH ODHADŮ	9
2.3 DEKOMPOZICE ČASOVÝCH ŘAD.....	11
2.4 STATICKÉ A DYNAMICKÉ MODEL Y	12
2.5 DUMMY PROMĚNNÉ.....	13
2.6 PROBLEMATIKA ČASOVÝCH ŘAD.....	13
2.6.1 Stochastické a stacionární procesy	14
2.6.2 Slabě závislé časové řady	14
2.6.3 AR(1)	14
2.6.4 MA(1).....	15
2.6.5 Integrované řady.....	16
2.6.6 Sériová korelace náhodné složky	16
2.6.7 Heteroskedasticita	17
2.7 VERIFIKACE MODELU	17
2.7.1 F–test.....	17
2.7.2 Dílčí t–testy	18
2.7.3 Vícenásobný upravený koeficient determinace	20
2.7.4 Testy na jednotkový kořen	20
2.7.5 Testování přítomnosti sériové korelace.....	22
2.7.6 Testování heteroskedasticity	24
3 EMPIRICKÁ ANALÝZA A VYHODNOCENÍ VÝSLEDKŮ	25
3.1 DATA	25
3.2 ANALÝZA.....	27
3.2.1 Dekompozice časových řad a sezonní očištění	28
3.2.2 Testování časové řady na jednotkový kořen a diferencování	31

3.2.3	Minimální mzda	32
3.2.4	Úroveň nezaměstnanosti mezi slabě vzdělanými lidmi	34
3.2.5	Dummy proměnná „2008 rok“	36
3.2.6	Dummy proměnná „Intervence“	39
3.2.7	Okunův zákon	40
3.2.8	F test pro sdruženou významnost proměnných “minimální mzda” a “HDP”	42
ZÁVĚR.....		44
LITERATURA		46

Seznam grafů, obrázků a tabulek

Graf č.1 Vývoj míry nezaměstnanosti v ČR 1998–2016.....	5
Graf č.2 Vývoj míry nezaměstnanosti v Rakousku 1999–2016.....	6
Graf č.3 Vývoj míry nezaměstnanosti mezi slabě vzdělanými lidmi v ČR 1999–2016.....	25
Graf č.4 Vývoj změny míry reálného HDP v ČR 1999–2016.....	26
Graf č.5 Vývoj minimální mzdy v ČR 1999–2016	26
Tabulka č.1 Popisné charakteristiky časových řad.....	27
Obrázek č.1 Dekompozice časové řady míry nezaměstnanosti v ČR	28
Obrázek č.2 Dekompozice časové řady míry nezaměstnanosti v Rakousku.....	28
Obrázek č.3 Dekompozice časové řady změny HDP v ČR.....	29
Obrázek č.4 Dekompozice časové řady míry nezaměstnanosti mezi slabě vzdělanými lidmi v ČR	29
Obrázek č.5 Porovnání očištěné a neočištěné o sezonnost časové řady změny HDP v ČR	29
Tabulka č.2b Testování stacionarity AR modelem	30
Tabulka č.2a Testování stacionarity AR modelem.....	30
Tabulka č.3 Výsledky DF testů	31
Tabulka č.4 Výsledky ADF testů	32
Tabulka č.5 Regresní model závislosti míry nezaměstnanosti na minimální mzdě. Porovnání OLS a GLS	33
Tabulka č.6 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelu závislosti míry nezaměstnanosti na minimální mzdě v ČR	34
Tabulka č.7 Regresní model závislosti míry nezaměstnanosti na SVL, rozšířený model. Porovnání OLS a GLS.....	35
Tabulka č.8 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelech A a B	36
Tabulka č.9 Regresní model závislosti míry nezaměstnanosti na dummy proměnné „2008 rok “rozšířený model. Porovnání OLS a GLS	37
Tabulka č.10 Výsledky testování sériové korelace a heteroskedasticity v OLS modelech A a B	38
Tabulka č.11 Regresní model závislosti míry nezaměstnanosti na dummy proměnné „intervence“, rozšířený model. Porovnání OLS a GLS.....	38
Tabulka č.12 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelech A a B	39
Tabulka č.13 Regresní model Okuna, rozšířený model. Porovnání OLS a GLS	40
Tabulka č.14 Výsledky testování sériové korelace a heteroskedasticity v OLS modelech A a B	41
Tabulka č.15 Souhrnný regresní model. Porovnání OLS a GLS	42
Tabulka č.16 Výsledky parciálního F testu.	43

Úvod

V současné době je svět postaven před velké ekonomické a finanční problémy, včetně problémy týkající se nezaměstnanosti a nedostatečného hospodářského růstu. Zvýšení armády nezaměstnaných je velmi diskutabilní záležitost všude po celém světě, nejsou výjimkou ani vyspělé země. Analýza dynamiky počtu nezaměstnaných občanů České republiky, stejně jako ekonometrické modelování dynamiky počtu nezaměstnaných, je důležitým úkolem, neboť nezaměstnanost má velmi vážné ekonomické důsledky, v souvislosti s kterými tento jev vyžaduje důkladnou pozornost. Zkoumání důsledků nezaměstnanosti na národní úrovni ukazuje, že nezaměstnanost snižuje objem hrubého domácího produktu (HDP). Hospodářský růst a míra nezaměstnanosti jsou klíčové ukazatele, které jsou sledovány současně tvůrci politiky a veřejností, protože vytvářejí jasnou představu o ekonomickém rozvoji země. Kromě toho vazba mezi mírou nezaměstnanosti a hospodářským růstem je relevantní makroekonomickou problematikou pokrývající širokou oblast teoretického a empirického výzkumu. Tyto dva úzce související jevy našly svoje vyjádření v "Okunůvě zákonu", který říká, že pokud skutečná míra nezaměstnanosti převyšuje přirozenou míru nezaměstnanosti o jedno procento, pak ztráta HDP je 2-3%. V ekonomické teorii je nezaměstnanost považována za "normální", když dosáhne určité rovnováhy mezi nabídkou práce a poptávkou, tj. počet volných pracovních míst je zhruba stejný jako počet lidí, kteří hledají práci (tento postulát umožňuje kvalitativní nesoulad pracovních míst a pracovníků).

Cílem této práce je vytvořit ekonometrický model popisující dynamiku míry nezaměstnanosti občanů České republiky v závislosti na relevantních faktorech, které nám pomohou pochopit podstatu nezaměstnanosti a povahu procesů, které ji formují.

Osvědčená metoda k analýze nezaměstnanosti (v této práci se zabývám obecnou mírou nezaměstnanosti) je ekonometrická analýza časových řad.

Veškeré výpočty budou provedené v statistických softwarech R a Rstudio.

1 Faktory ovlivňující nezaměstnanost

Nezaměstnaní, s odkazem na normy Mezinárodní organizace práce (ILO), jsou lidé ve věku určeném pro měření ekonomicky aktivního obyvatelstva, v případě ČR starší 15 let, kteří během sledovaného období splňují zároveň následující kritéria:

- byli bez práce,
- hledali práci– apelovali na státní či komerční úřad práce, používali nebo uváděli reklamu v tisku, přímo se obraceli na zaměstnavatele, používali osobní kontakty, nebo dělali kroky k organizování své vlastní podnikatelské činnosti,
- byli schopní nastoupit do práce během určitého referenčního období (obvykle 14 dnů).

Při výpočtu **míry nezaměstnanosti** v ekonomice chceme, aby zahrnovala lidi, kteří chtějí pracovat. Celkový počet lidí, kteří chtějí být v práci se nazývá pracovní síla a může být definována jako součet zaměstnaných a nezaměstnaných:

Pracovní síla = počet zaměstnaných + počet nezaměstnaných (hledají práci)

Potom můžeme definovat míru nezaměstnanosti jako procento pracovní síly, která je bez práce:

Míra nezaměstnanosti = ((Počet nezaměstnaných) / (Pracovní síla)) * 100

Klasifikace typů nezaměstnanosti

Lidé jsou nezaměstnaní z mnoha různých důvodů, a tak můžeme rozdělit nezaměstnanost do tří různých typů:

Frikční nezaměstnanost je krátkodobá nezaměstnanost, která je výsledkem činnosti hledání nové práce. Příkladem je někdo, kdo se stěhuje do Prahy, protože věří, že tam existují lepší pracovní příležitosti, ale je nezaměstnaný po dobu několika týdnů, než si zajistí nové zaměstnání.

Strukturální nezaměstnanost je chronická (trvalá). Nezaměstnanost strukturální je taková forma nezaměstnanosti, která je způsobená zásadními posuny v ekonomice a se zhoršuje vnějšími faktory, jako jsou technologie, konkurence a vládní zásahy. Důvod, proč dochází k strukturální nezaměstnanosti, je kupříkladu nedostatek pracovníků s požadovanými pracovními dovednostmi, nebo že pracovníci bydlí příliš daleko od regionů, kde jsou k dispozici pracovní místa a nemůže se přestěhovat blíž. Pracovní místa jsou k dispozici, ale existuje vážný nesoulad mezi tím, co firmy potřebují a co pracovníci mohou nabídnout.

Cyklická nezaměstnanost je krátkodobá nezaměstnanost, která je spojena s výkyvy ekonomiky. Příkladem toho je, když se ekonomika dostane do recese a pracovníci jsou propuštěni, ale očekávají, že se vrátí zpátky do práce, když se ekonomika zlepší.

Dále projdeme definice a příčiny podrobněji.

Co způsobuje frikční nezaměstnanost?

Frikční nezaměstnanost je nezaměstnanost, která vyplývá z aktivity hledání nového zaměstnání. V podstatě, pracovní trhy spojují firmy s pracovníky. Příkladem může být společnost, která hledá programátora s určitými počítačovými znalostmi a někdo může mít tyto dovednosti. Nicméně, toto může být obtížným úkolem. Společnost musí vyhodnotit velký počet kandidátů, aby našli zaměstnance s nejlepší kvalifikací, a zaměstnanec chce najít společnost, která dá nejlepší nabídku. Všechna vyhledávání trvají dlouho a někdy to znamená trávení času ve stavu nezaměstnaného. Koneckonců, to může být v nejlepším zájmu hledajícího, aby odmítl práci s nízkými platy, i když je nezaměstnaný, pokud si myslí, že může najít vyšší placenou práci v blízké budoucnosti. V tržní ekonomice vždy existují některé firmy, které ukončují svíjí činnost a další nově vznikající společnosti. Tím se spustí proces, ve kterém lidé hledají nová pracovní místa, a firmy hledají nové zaměstnance, z čehož vyplývá, že vždy bude existovat nějaká nezaměstnanost v ekonomice.

Co způsobuje strukturální nezaměstnanost?

K strukturální nezaměstnanosti dochází v důsledku nesouladu dovedností na trhu práce, což může být způsobeno těmito body:

- **Profesní imobilita:** je spojena s obtížemi při studiu nových dovedností použitelných pro nový průmysl, a technologické změny, například nezaměstnaný zemědělec se snaží najít práci v high-tech průmyslu.
- **Geografická imobilita:** týká se obtížnosti při přestěhování s účelem získání práce, např. mohou existovat pracovní místa v Olomouci, ale může být obtížné najít vhodné ubytování nebo vzdělávací instituce pro své děti.
- **Technologická změna:** vede-li vývoj technologií k úsporám pracovní síly v některých odvětvích, pak bude sledovat pokles poptávky po pracovní síle.
- **Strukturální změny v ekonomice:** například pokles ve využití uhelných dolů v důsledku nedostatečné konkurenceschopnosti vede k situaci, kdy mnoho horníků zůstane nezaměstnanými, a následovně zjistí, že je obtížné získat práci v moderních průmyslových odvětvích, jako například v oblasti informačních technologií.

Když se detailně zaměříme na příčinu vzniku strukturální nezaměstnanosti v praxi, můžeme taky říct, že tato forma nezaměstnanosti vznikne, když mzdy nevyvažují nabídku a poptávku po práci. Chceme-li zjistit, proč se strukturální nezaměstnanost vyskytuje, musíme se podívat na důvod, proč se mzda neupravuje. Jako první podstatný důvod můžeme uvést zákony,

kteřé omezují nejnížší úroveň mezd, které firmy mohou nabídnout – minimální mzda. Například, jsem pracovnice, která téměř nemá pracovní dovednosti. Z tohoto důvodu nejsem velmi produktivní, takže maximum, které by zaměstnavatel byl ochotný nabídnout je 50Kč za hodinu pro začátek. Nicméně, minimální mzda je 66Kč za hodinu, takže pro zaměstnavatele je lepší najít pracovníka, jehož produktivita je alespoň ve výši minimální hodinové sazby. Takže budu považována za nezaměstnanou, protože chci pracovat za mzdu ve výši 66Kč za hodinu, ale žádný zaměstnavatel mi nenabídne práci za takovou mzdu. Kvůli tomu budeme očekávat, že nížší průměrná **úroveň kvalifikace** nebo výšší **minimální mzda** vyvolá větší strukturální nezaměstnanost v ekonomice.

Odborové organizace mají podobný vliv na nezaměstnanost. Často takové organizace omezují nabídku práce tím, že omezují počet pracovníků oprávněných pro zaměstnání, s účelem získání výšší mzdy. Tímto způsobem vytváří větší příjem pro ty, kteří mají práci, ale vyvolává strukturální nezaměstnanost pro ty, kteří práce nemají. Takoví nezaměstnaní pracovníci chtějí pracovat za výšší mzdy (nebo dokonce i za nížší mzdy), ale zaměstnavatele je nejsou ochotni zaměstnávat za vysoké mzdy, a zároveň nejsou oprávněni odborovou organizací nabídnout nížší.

V rámci strukturální nezaměstnanosti můžeme také zmínit nezaměstnanost sezonní, která je vlastně její druhem, protože se struktura ekonomiky mění na sezonním základě, poptávka po práci se bude měnit odpovídajícím způsobem. Například v listopadu a prosinci můžeme pozorovat nárůst poptávky po svátečních předmětech a produktech souvisejících s dovolenou, a v červenci a srpnu– nárůst poptávky po zboží a službách souvisejících s prázdninami. Poptávka po těchto výrobcích a službách vytváří poptávku po pracovnících, kteří je mohou dodávat.

Co způsobuje cyklickou nezaměstnanost?

Dohromady strukturální a frikční nezaměstnanost tvoří to, co se jmenuje **přirozená míra nezaměstnanosti**. Slovo „přirozená“ neznamená, že je vítaná (chtěli bychom snížit strukturální nezaměstnanost). V podstatě, je to míra nezaměstnanosti, kterou očekáváme, že bude vždy existovat v ekonomice.

Nicméně, občas ekonomika roste rychleji a někdy pomaleji, a tímto bude ovlivňovat výši nezaměstnanosti. Když ekonomika prudce roste– klesá nezaměstnanost, a když je ekonomika v recesi– stoupá. Nezaměstnanost, vznikající v důsledku dočasných výkyvů ekonomiky, se nazývá cyklická nezaměstnanost. Formálně, cyklická nezaměstnanost je definována jako odchylka nezaměstnanosti od své přirozené míry.

Cyklická nezaměstnanost jde ruku v ruce s hospodářským cyklem. Čím výšší je HDP, tím nížší bude cyklická nezaměstnanost na vrcholu hospodářského cyklu, platí to naopak, když ekonomika je na své nejnížší úrovni hospodářského cyklu. Pokud cyklická míra nezaměstnanosti existuje delší dobu, může to způsobit nenapravitelné škody na pracovní sílu v státu. Uvedenou závislost prokazuje **Okunův zákon**.

1.1 Okunův zákon

Hrubý domácí produkt (HDP) je peněžním vyjádřením celkové hodnoty statků a služeb nově vytvořených v daném období na určitém území; používá se pro stanovení výkonnosti ekonomiky.

Okunův zákon je empirický vztah mezi mírou růstu reálného HDP a změnou nezaměstnanosti. Arthur Okun¹ v poválečném období (1962) vypožoroval, že každé další jedno procento nezaměstnanosti nad 4% je spojeno se snížením reálného důchodu o 3%. Toto empirické pozorování se v dalším desetiletí zdálo být vyhovující, proto se ustálilo pojmenování Okunův zákon. Clark (1983), Gordon (1984) a Prachowny (1993) to podpořili svými teoretickými výzkumy a poté Holloway (1989), Prachowny (1993), Moosa (1997), Attfield a Silverstone (1998) otestovali stabilitu poměru 3:1. Okunův zákon většinou testů prošel, a poměr se později ustálil na 2:1. Závěry se ale začly lišit podle použitých metod a specifikací. Freeman² (2000) navázal na tyto práce a pomocí pásmového filtru pracujícího s metodou centrovaných klouzavých průměrů odhadoval koeficienty neovlivněné hospodářským cyklem. Jako první porovnal tyto koeficienty i regionálně, v osmi různých územích celcích v USA v období 1980–1994, vážený průměr koeficientu 2,22. Spolu s Philipsovou křivkou označil Okunův zákon jako jednu ze základních myšlenek makroekonomie. Lze jej vyjádřit pomocí rovnice:

$$\frac{Y^* - Y}{Y^*} = c(u - u^*) \quad (1.1)$$

Ekonomické důsledky nezaměstnanosti vyjádřené jako rozdíl skutečného a potenciálního produktu, kde Y^* je potenciální produkt, Y je skutečný produkt, u^* přirozená míra nezaměstnanosti, u skutečná míra nezaměstnanosti a c je koeficient lineární závislosti. Pokud je skutečná míra nezaměstnanosti rovna potenciální, produkt je na úrovni potenciálního. Pokud je míra nezaměstnanosti vyšší než přirozená, podíl reálného a potenciálního produktu klesne o podíl vyjádřený konstantou c . Ta obvykle nabývá hodnot v rozmezí 0,2-0,3, což se dá interpretovat jako růst nezaměstnanosti o 0,2-0,3% při neměnném produktu, tak pro pokles míry nezaměstnanosti musí ekonomika růst o více než 4% (30/7). Přirozenou míru nezaměstnanosti a potenciální produkt nemůžeme změřit, tudíž použijeme model se změnami produktu a nezaměstnanosti v čase:

$$\Delta \text{Nezam}_t = \beta_0 + \beta_1 \Delta \text{HDP}_t + \varepsilon_t \quad (1.2)$$

Okun odhaduje tento vztah na čtvrtletních časových řadách jako:

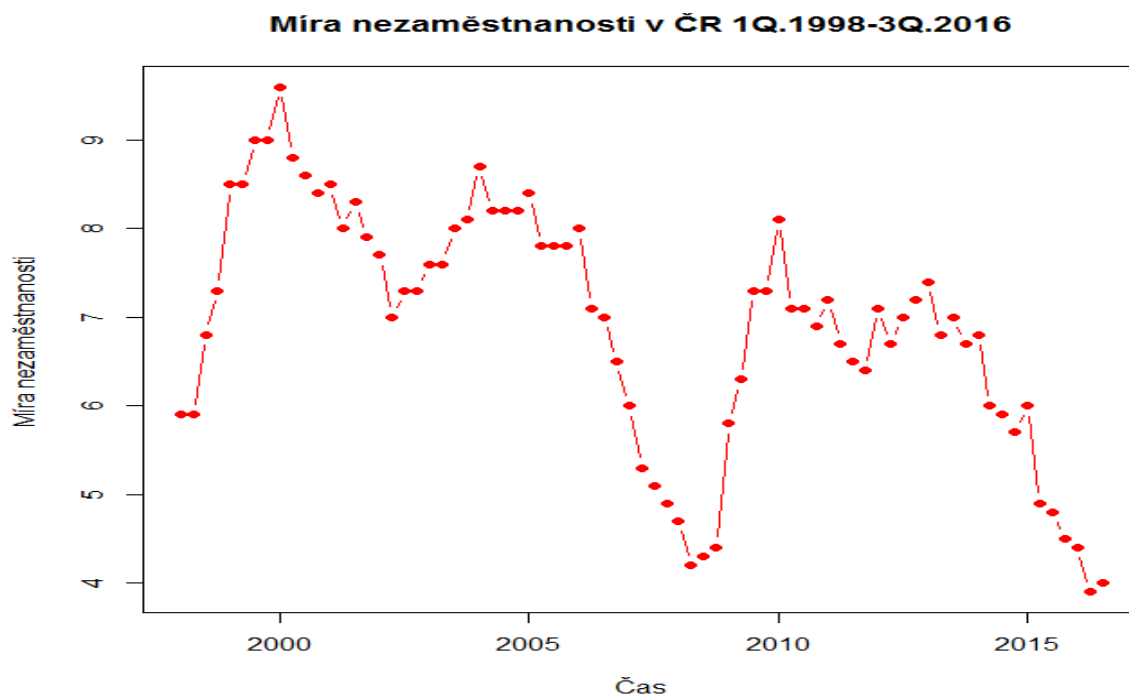
$$\Delta \text{Nezam}_t = 0,3 - 0,07 \Delta \text{HDP}_t + \varepsilon_t \quad (1.3)$$

¹ Arthur M. Okun 1962 - Potential GNP: It's Measurement and Significance

² Donald G. Freeman 2000 - Regional Tests of Okun's Law

1.2 Vývoj nezaměstnanosti v ČR

Na konci 80. let 20. století byla nezaměstnanost v ČR prakticky nulová, což bylo důsledkem komunistického režimu a částečně reformního programu „přestavby“, který probíhal v ČR od konce 70. let. V listopadu 1989 byla provedena v ČSSR demokratická

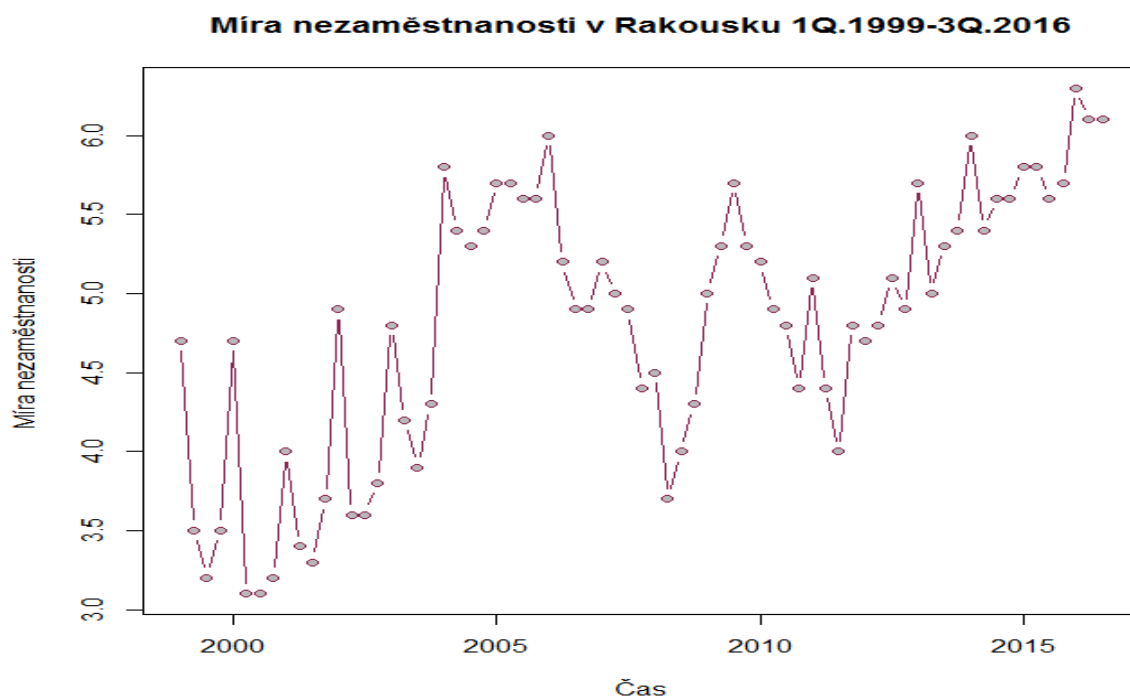


Graf č.1 Vývoj míry nezaměstnanosti v ČR 1998–2016

revoluce, čímž začaly významné změny v struktuře ekonomiky jako celku, a s tím i změny v struktuře nezaměstnanosti. Situace se začala rychle měnit už od roku 1991– růst nezaměstnanosti, pokles počtu zaměstnaných, zejména ve výrobním sektoru.

V této práci máme k dispozici časovou řadu vývoje nezaměstnaných od roku 1998, a první čtvrtletí 1998 roku udává hodnotu nezaměstnanosti 5,9 %. Odtud začíná rychle narůstat a dosahuje svého maxima 9,6 %, a to z důvodu, že se ekonomika nacházela v recesi.

Dne 1. května 2004 roku se česká Republika stává členem Evropské Unie. Z ekonomického hlediska, to vedlo k celkovému zlepšení postavení České republiky na mezinárodním trhu, stejně jako k novým možnostem, které posílili českou ekonomiku, jako je příliv zahraničních investic, přístup na strukturální fondy EU pro rozvoj různých oborů a regionů, nebo příležitost pro české občany, aby se stěhovali do jiných členských států EU za



Graf č.2 Vývoj míry nezaměstnanosti v Rakousku 1999–2016

prací. Všechny limity omezení obchodu byly odstraněny. A to vedlo k výraznému poklesu nezaměstnanosti až do 2. čtvrtletí 2008 roku na úroveň 4,2 %.

Další významné výkyvy byly způsobené krizí 2008 roku. První fáze (srpen 2007–září 2008) neměla na českou ekonomiku výrazné následky, což je docela vidět z hlediska zaměstnanosti na grafu č.1. Druhá fáze zavedla proces růstu nezaměstnanosti, a tím se stav v 2010 roce vrátil na rovinu 8,1 %.

Od roku 2013 se situace neustále zlepšuje a v 4. čtvrtletí 2016 roku úřad evidoval rekordně malou předběžnou hodnotu míry nezaměstnanosti v 3,6 %.

1.3 Vývoj nezaměstnanosti v Rakousku

Pro účely porovnání dynamiky a charakteru nezaměstnanosti mezi dvěma státy se podobnou ekonomikou použijeme časovou řadu míry nezaměstnanosti v Rakousku.

Začátkem 80. let začíná nezaměstnanost v Rakousku nabývat rostoucího dlouhodobého trendu, což je opačné dlouhodobě klesajícímu trendu v ČR.

Situace se zlepšovala od roku 2006, a stejně jako v ČR opak se zhoršila v době krize.

Poslední 5 let míra zaměstnanosti v Rakousku roste, ale stejně zůstává jednou z nejnižších. I když v časové řadě můžeme pozorovat dostatečné velkou variabilitu, zejména

v období do 2004 roku, obecná míra nezaměstnanosti v Rakousku nikdy nepřesahovala 7 %, což můžeme považovat za vlastnost stabilní ekonomiky státu.

Rakousko má nízké procento lidí, kteří jsou nezaměstnaní, ve srovnání s jinými zeměmi v eurozóně a EU, stejně jako v globálním srovnání. Rakousko je v současné době na čtvrtém místě v EU. Pouze v Německu, české Republice a Velké Británii je míra registrované nezaměstnanosti nižší než v Rakousku.

2 Vybrané ekonometrické postupy analýzy nezaměstnanosti

V mojí práci se budu zabývat analýzou časových řad. V této části popíši jednotlivé pojmy, postupy, podklady, metody a testy, které jsou podstatné pro kvalitní ekonometrickou analýzu dat daného typu.

2.1 Obecné pojmy ekonometrického modelování

V analýze ekonomických jevů na základě ekonomických a matematických metod zaujímají zvláštní místo modely, odhalující kvantitativní vztah mezi sledovaným ukazatelem a faktory ho ovlivňující. Vědeckou disciplínou, jejíž předmětem je studium takových kvantitativních aspektů ekonomických jevů a procesů pomocí matematické a statistické analýzy, je ekonometrie, ve které výsledky teoretické analýzy ekonomiky jsou syntetizovány s úsudky matematiky a statistiky. Hlavním cílem ekonometrie je test ekonomických teorií na reálných (empirických) materiálech s využitím metod matematické statistiky.

Hlavním nástrojem ekonometrie je ekonometrický model, tj. ekonomicko-matematický model založený na vícenásobné regresní analýze, jehož parametry jsou odhadovány pomocí matematické statistiky. Tento model slouží jako nástroj pro analýzy a prognózy konkrétních ekonomických procesů založený na skutečné statistické informaci.

Ekonometrické modely mohou být klasifikované podle několika klasifikačních příznaků. Podle matematické funkční formy rozlišujeme modely (rovnice) lineární, nelineární, exponenciální modely a další.

Jednou ze základních klasifikací ekonometrických modelů je klasifikace podle směru a složitosti kauzálních vazeb mezi ukazateli ekonomického systému. Pokud budeme používat termín "proměnné", v každém dostatečně složitém ekonomickém systému dokážeme rozlišit interní proměnné (například zaměstnanost, produkce, produktivita práce) a externí proměnné (např. dodávky zásoby, klimatické podmínky). Pak podle směru a složitostí vztahů mezi vnitřními (endogenními, výstupními) proměnnými a externími (exogenními, vstupními) rozlišujeme následující ekonometrické modely: regresní modely, systémy se zpětnou vazbou mezi proměnnými, rekursivní systémy.

Regresní modely jsou takové modely, které jsou založené na regresní rovnici nebo na soustavě regresních rovnic propojujících mezi sebou hodnoty endogenních a exogenních proměnných. Existují rovnice (modely) jednoduché a vícenásobné regrese. Pokud endogenní (závislou) proměnnou označíme y , a pro označení proměnné exogenní (nezávislé) použijeme x , pak v případě lineárního regresního modelu rovnice pro jednoduchou regresi má tvar:

$$y_t = \beta_0 + \beta_1 x_{t1} + \varepsilon_t \quad (2.1)$$

a rovnice vícenásobné regrese:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, \quad (2.2)$$

kde y_t je hodnota závislé proměnné v čase t , x_{tj} jsou hodnoty nezávislých proměnných v čase t , ε_t je náhodná chyba v čase t , a β_j jsou parametry regrese: β_0 je úrovněová konstanta a β_j pro $j > 0$ jsou koeficienty sklonu křivky a uvádí jaký vliv mají nezávislé proměnné na závislou.

Označíme, že odhady parametrů modelů jednoduché a vícenásobné regrese jsou založeny na metodě nejmenších čtverců.

V moji práci je předmětem analýzy obecná míra nezaměstnanosti, která je jedním z makroekonomických agregátů, takže dále budeme mluvit o takzvaném makroekonomickém modelu (dále jen „makromodel“).

Makroekonomickým modelem se nazývá ekonomicko-matematický model, který odráží výkonnost ekonomiky (regionu, země, na celém světě) jako celku. Makromodely pracují s takovými globálními ukazateli ekonomiky jako jsou národní důchod, hrubý domácí produkt, míra nezaměstnanosti, míra inflace a tak dále. Tyto modely jsou používány pro teoretickou analýzu nejobecnějších zákonů o fungování a rozvoji ekonomiky. Podle povahy závislostí mohou být makromodely deterministické a pravděpodobnostní (stochastické), podle role časového faktoru – statické a dynamické, podle představy proměnných souvisejících s časem – diskrétní a spojitě. V této práci se budu zabývat dynamickým makromodelem ukazatele tržní ekonomiky.

2.2 Metoda nejmenších čtverců, předpoklady její použití a vlastností získaných odhadů

Uvažujeme populační regresní model (2.2).

Metoda nejmenších čtverců je přístup, se kterým přišel Carl Friedrich Gauss (1795) ve věku 18 let. Přiřazuje model k datům pomocí minimalizace součtu druhých mocnin rozdílů vyrovnaných hodnot a vstupních dat, tedy reziduí. Pokud jsou tyto chyby vzájemně nezávislé, tzn. jejich střední hodnota je nulová, je tento odhad zároveň nejlepší existující (Gauss-Markovův teorém). Rovnice pro odhad vypadá následujícím způsobem:

$$\sum_{t=1}^n \left(y_t - \hat{\beta}_0 - \hat{\beta}_1 x_{t1} - \dots - \hat{\beta}_k x_{tk} \right)^2 \rightarrow \min \quad (2.3)$$

Odhad parametrů se dá maticově vyjádřit takto:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y \quad (2.4)$$

kde $\mathbf{X}$ je matice hodnot nezávislých proměnných včetně konstanty (=1), a $\mathbf{y}$ je vektor hodnot závislé proměnné.

Předpoklady pro použití metody nejmenších čtverců pro časové řady:

1. Populační model je lineární v parametrech;
2. Podmíněná střední hodnota náhodné složky v závislosti na vysvětlujících proměnných se rovná nule:

$$E(\varepsilon_t | \mathbf{X}_t) = 0, \quad t = 1, \dots, n$$

Když je splněn tento předpoklad, závislé proměnné jsou striktně exogenní.

3. Není perfektní kolinearita – žádná z vysvětlujících proměnných x_j není konstantní nebo lineární kombinací ostatních závislých proměnných;
4. Podmíněný rozptyl náhodné složky v závislosti na x_j je konstantní v čase a konečný (tzv. homoskedasticita):

$$Var(\varepsilon_t | \mathbf{X}_t) = Var(\varepsilon_t) = \sigma^2, \quad t = 1, \dots, n$$

5. Není autokorelace (sériová korelace) náhodné složky – v závislosti na x_j hodnoty náhodné složky v dvou odlišných obdobích nejsou mezi sebou korelované:

$$Corr(\varepsilon_t, \varepsilon_s | \mathbf{X}_t) = 0, \text{ pro všechna } t \neq s$$

6. Normalita v rozdělení náhodné složky v závislosti na x_j – plyne z předchozích předpokladů:

$$\varepsilon_t \sim N(0, \sigma^2)$$

Všech šest předpokladů spolu tvoří předpoklady klasického lineárního regresního modelu (dále jen KLRM) pro časové řady.

Odhady, získané použitím MNČ, mají následující vlastnosti:

- BLUE – předpoklady 1. až 5. spolu tvoří Gauss-Markovy (dále jen G-M), při jejich splnění odhady MNČ jsou BLUE (Best Linear Unbiased Estimator) ;
- Konzistentní – při zvětšení výběrového vzorku odhad parametrů konverguje k populačnímu parametru:

$$\text{plim}_{n \rightarrow \infty} \hat{\beta} = \beta$$

- Nevychýlené (nezkreslené, nestranné) – platí při splnění G-M předpokladů 1. až 3. a znamená to, že rozdělení odhadu parametru je soustředěné kolem populačního parametru:

$$E(\hat{\beta}_j) = \beta_j, j = 0, 1, \dots, k$$

- Při splnění KLRM předpokladů mají odhady normální rozdělení;
- Vydátne – odhady mají nejmenší rozptyl.

Metoda zobecněných nejmenších čtverců (GLS)

Cochrane-Orcutt nebo **Prais-Winsten** transformace připravuje rovnici pro použití GLS, v případě, že najdeme sériovou korelaci náhodné složky v modelu:

$$\tilde{y}_t = (1 - \rho)\beta_0 + \beta_1\tilde{x}_{t_1} + \dots + \beta_k\tilde{x}_{t_k} + e_t,$$

kde $\tilde{x}_{t_j} = x_{t_j} - \rho x_{t-1,j}$, $\tilde{x}_{t_1} = (1 - \rho^2)^{1/2}x_{t_1}$ a $\tilde{y}_1 = (1 - \rho^2)^{1/2}y_{t_1}$, a ρ je koeficient korelace mezi ε_t a ε_{t-1} z původního regresního modelu.

Však v praxi obvykle neznáme hodnotu ρ takže musíme použít $\hat{\rho}$ pro transformaci, v tomto případě mluvíme o tzn. **realizovatelné metodě nejmenších čtverců (FGLS)**, jejíž odhady mají následující vlastnosti:

- nejsou nevychýlené;
- t a F statistiky mají jenom přibližně t a F rozdělení;
- nejsou BLUE;
- jsou konzistentní.

Metoda vážených nejmenších čtverců (WLS) je druhem metody zobecněných nejmenších čtverců upravená pro případ, že nalezneme heteroskedasticitu (tj. porušení G-M předpokladu o homoskedasticitě). Ke každé hodnotě pozorování přiřadíme váhu (**W**), která je rovna obrácené hodnotě rozptylu její standardní chyby. Poté je tedy u každého pozorování váha úměrná chybě a získáme robustní odhad. Matice **W** je diagonální matice vah. Odhad má tedy po úpravě tento tvar:

$$\hat{\beta}_{WLS} = (X^T W X)^{-1} X^T W Y \quad (2.5)$$

Následujícím způsobem spočteme váhy prakticky:

1. Pustíme regrese $\log(\hat{u}^2)$ na všechny nezávisle proměnné, kde $\hat{u}$ jsou rezidua z původního odhadnutého modelu zatíženého heteroskedasticitou;
2. Uložíme dosažené hodnoty do proměnné $\hat{g}_i$;
3. Spočteme váhy $\hat{h}_i$:

$$\hat{h}_i = e^{\hat{g}_i}$$

Opět odhady, získané metodou WLS při použití vah $\hat{h}_i$ jsou **FGLS** odhady.

2.3 Dekompozice časových řad

Jelikož předmětem zkoumání je časová řada, je třeba rozumět tomu, co ji tvoří.

Časová řada – posloupnost hodnot (y_t) uspořádaná v čase t . (Čas může být diskrétní, $t = 1, 2, 3, \dots$ nebo spojitý $t > 0$). Je typickým představitelem stochastického procesu.

V ekonometrické analýze obvykle identifikujeme následující čtyři typy faktorů, které generují hodnoty časové řady popisující dynamiku procesu, který je předmětem šetření. Mezi ně patří:

1) dlouhodobý faktor (T_t), který tvoří obecný (v dlouhodobém horizontu) směr vývoje v analyzovaných vlastnostech objektu, který je popsán pomocí funkce trendu. Zpravidla je to monotónní funkce. Z grafu č.1 je vidět, že časová řada obecné míry nezaměstnanosti v ČR má dlouhodobě klesající trend;

2) sezónní faktor (S_t), který tvoří opakující se výkyvy analyzovaných charakteristických jevů v určitém ročním období. Na grafu č.1 je zřejmě vidět sezónní výkyvy časové řady, avšak je nutné zjistit jejich sílu;

3) cyklický faktor (C_t), formující změny v analyzovaném jevu způsobené vlivem dlouhodobých cyklů ekonomického, demografického charakteru. Z grafu č.1 vidíme výraznou cyklickou složku (můžeme subjektivně zhodnotit délku cyklu – cca 3 roky);

4) náhodný faktor (ε_t) nebo šum, formující vlastnosti objektu, které nebylo možné identifikovat a systematickým způsobem popsat. Jsou to náhodné, nepravidelné vlivy, jsou zastupované rezidui hodnot časové řady po odstranění vlivů ostatních faktorů.

Dále budou tyto faktory analyzovány ve vztahu k dynamice počtu nezaměstnaných občanů České republiky v období od prvního čtvrtletí 1998 do třetího čtvrtletí 2016 včetně (viz graf č. 1).

V závislosti od charakteru časové řady můžeme použít dva způsoby dekompozice: aditivní nebo multiplikativní:

- Aditivní:

$$y_t = T_t + S_t + C_t + \varepsilon_t \quad (2.6)$$

- Multiplikativní:

$$y_t = T_t \cdot S_t \cdot C_t \cdot \varepsilon_t \quad (2.7)$$

Použijeme model (2.6), pokud je variabilita hodnot nezávislá na čase t . Jinými slovy, velikost rozptylu je konstantní v čase.

Zvolíme model (2.7), když rozptyl časové řady je měnící se v čase.

Pokud charakter dat není zcela zřejmý, a máme potíže s výběrem mezi typy dekompozice, můžeme vyzkoušet oba dva a vybrat si ten, který má větší přesnost.

2.4 Statické a dynamické modely

Statickým modelem se nazývá model:

$$y_t = \beta_0 + \beta_1 z_t + \varepsilon_t, \quad t = 1, \dots, n$$

Ve kterém y_t a z_t jsou datované současně. Taky změna v z v čase t vyvolá okamžitou změnu v y_t .

Dynamickým modelem (taky Finite Distributed Lag model(FDL)) je například model:

$$y_t = \alpha_0 + \beta_0 z_t + \beta_1 z_{t-1} + \beta_2 z_{t-2} + \varepsilon_t, \quad t = 1, \dots, n$$

který se nazývá FDL model druhého řadu.

2.5 Dummy proměnné

Spolu s kardinálními a ordinálními vysvětlujícími proměnnými po převedení na dummy proměnné ve vícenásobné regresní analýze můžeme použít i nominální proměnné, které jsou dichotomické (alternativní), jako je například pohlaví, kuřák a nekuřák, souhlasný a nesouhlasný. Můžeme také použít dichotomické proměnné jako závislé proměnné do regrese, ale v této práci se s tím setkávat nebudu. Když jsou zaznamenávány, jako buď 0 nebo 1, dichotomické proměnné jsou často označovány jako "dummy" proměnné. Označují buď nepřítomnost nebo přítomnost vlastností nebo rysů. Proto fungují jako "dummy" proměnné v otázce. Nejviditelnější použití je, když proměnné buď již je, nebo byla překódována do dvou kategorií. Nicméně, logika dummy proměnných může být také rozšířena, aby nám umožnila zahrnovat nominální proměnné s více než dvěma kategoriemi do vícenásobné regrese, s tím že každou nominální hodnotu proměnné bude představovat jedna dummy proměnná. Příkladem takové proměnné je značka auta, barva, národnost.

Když překódujeme nominální proměnnou, která má více než jednu kategorii do dummy proměnné, získáme o jednu proměnnou méně, než je počet kategorií($k-1$), s tím že jedna kategorie bude představovat bázi pro porovnání s ostatními. Jedním z příkladů použití umělých proměnných je taky představování sezónnosti v časové řadě– denní, měsíční, čtvrtletní, – přidáme 6, 11 nebo 3 dummy proměnné, a budeme porovnávat, jaký budou mít vliv jednotlivá období na časovou řadu v porovnání s bází. Když odhadneme regresní rovnice a získáme své výsledky, odhadovaný koeficient dummy proměnné ukazuje, jak velký rozdíl je, když sledované pozorování patří ke kategorii, pro kterou je dummy proměnná 1.

V mojí práci budu používat daný druh proměnných pro modelování sezónnosti a vlivu krize z roku 2008 na nezaměstnanost v ČR.

2.6 Problematika časových řad

V této části probereme klíčové pojmy, které jsou nutné pro pochopení problematiky časových řad a charakteru jejich analýzy.

2.6.1 Stochastické a stacionární procesy

Stochastický proces je proces, který se vyvíjí v průběhu času v souladu se zákony teorie pravděpodobnosti (např. časová řada). K stochastickým procesům patří třída stacionárních procesů.

Pojem „**stacionární proces**“ hraje důležitou roli v analýze časových řad. Stacionární časová řada je řada, která se řídí takovým procesem, jehož pravděpodobnostní rozdělení je stabilní v čase. Totiž, pokud vezmeme jakýkoli výběr náhodných proměnných v posloupnosti, a pak se přesuneme v čase o h období dopředu, sdružené rozdělení pravděpodobnosti musí zůstat stejné.

Stacionární proces je stochastický proces $\{x_t: t = 1, 2, \dots\}$, ve kterém pro libovolný výběr časových indexů $1 \leq t_1 < t_2 < \dots < t_m$, sdružené rozdělení vysvětlujících proměnných v čase t je stejné jako u vysvětlujících proměnných v čase $t + h$, pro všechny celé $h \geq 1$.

Slabá forma stacionárního procesu: když $\{x_t: t = 1, 2, \dots\}$ má konečný druhý moment, totiž $E(x_t^2) < \infty$ pro všechna t .

Kovarianční stacionární proces: stochastický proces $\{x_t: t = 1, 2, \dots\}$, který má $E(x_t^2) < \infty$ je kovariančním stacionárním procesem, když splňuje následující podmínky:

- x_t má konstantní střední hodnotu: $E(x_t) = \text{const}$;
- x_t má konstantní rozptyl: $\text{Var}(x_t) = \text{const}$;
- pro libovolná $t, h \geq 1$, $\text{cov}(x_t, x_{t+h})$ závisí jenom na h , nikoliv na t .

Když stacionární proces má konečný druhý moment, pak musí být kovariančním stacionárním, ale toto tvrzení neplatí opačně.

2.6.2 Slabě závislé časové řady

Proces $\{x_t: t = 1, 2, \dots\}$ je slabě závislým, pokud při $h \rightarrow \infty$, x_t a x_{t+h} jsou skoro nezávislé (totiž $\text{corr}(x_t, x_{t+h}) \rightarrow 0$). V podstatě jsou to asymptoticky nekorelované kovarianční časové řady. Příkladem jsou AR(1) a MA(1) procesy.

2.6.3 AR(1)

Autoregresní proces prvního řádu, označený AR(1), má formu:

$$y_i = \varphi_0 + \varphi_1 y_{i-1} + \varepsilon_i, \quad (2.8)$$

kde φ_0 je konstanta a φ_1 je koeficient autoregresi.

Uvažujeme indexy i jako zástupci času t , vidíme, že hodnota y v čase $i + 1$ je lineární funkce y v době i plus úrovněová konstanta a náhodná chyba. Podobně obyčejnému lineárnímu regresnímu modelu, předpokládáme, že náhodná složka má normální rozdělení s nulovou střední hodnotou a konstantním rozptylem σ^2 , a taky že chyby jsou nezávislé na hodnotě y :

$$\varepsilon_t \sim N(0, \sigma^2)$$

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0, \quad \text{pro } i \neq j$$

$$\text{cov}(\varepsilon_i, y_j) = 0, \quad \text{pro všechna } i, j$$

Říkáme, že takový proces je stacionární, pokud $|\varphi_1| < 1$, a při analýze časových řad to musíme taky předpokládat. Obecně, proces není stacionární pokud $|\varphi_1| \geq 1$. Pokud $|\varphi_1| = 1$ jde tady o náhodnou procházku, což je diferencovatelně stacionární proces.

Náhodná procházka (Random walk (RW))

$$y_t = y_{t-1} + \varepsilon_t \quad (2.9)$$

Proces, ve kterém se náhodné změny v hodnotách nějaké proměnné v daném okamžiku vyskytují podle schématu, ve kterém tyto hodnoty nejsou závislé na předchozích změnách. Proto, za předpokladu náhodné procházky, proměnná hodnota nemůže nabýt svých předchozích hodnot. Čím déle trvá tento proces, tím širší bude rozptyl rozdělení možných hodnot.

Jak už bylo zmíněno před tím proces náhodné procházky je diferencovatelně stacionárním procesem, což znamená že nestacionární časová řada se stane stacionární po použití metody diferencování na její hodnotách.

Taky rozlišujeme náhodnou procházku s driftem:

$$y_t = a + y_{t-1} + \varepsilon_t \quad (2.10)$$

Která taky je diferencovatelně stacionární.

(Důkaz: $\Delta y_t = y_t - y_{t-1} = a + y_{t-1} + \varepsilon_t - y_{t-1} = a + \varepsilon_t$)

2.6.4 MA(1)

Nejjednodušším procesem, projektovaným na základě procesu bílého šumu, je proces klouzavých průměrů prvního řádu (first-order moving average – MA(1)):

$$x_t = e_t + \alpha_1 e_{t-1}, t = 1, 2, \dots \quad (2.11)$$

$$e_t \sim N(0, \sigma_e^2)$$

$$\text{cov}(x_t, x_{t+1}) = \alpha_1 \text{Var}(e_t) = \alpha_1 \sigma_e^2$$

2.6.5 Integrované řady

Většina z nestacionárních procesů je vysoce perzistentní, totiž jsou "integrované" procesy (obsahují jednotkové kořeny), které se někdy nazývají stochastické trendy. Můžeme považovat tyto posloupnosti za takové, v nichž změna mezi obdobími je stacionární náhodná proměnná. Integrované procesy mohou být převedené na stacionární tím, že použijeme diferencování v čase. Například pro RW označíme:

$$y_{t+h} = e_{t+h} + e_{t+h-1} + \dots + e_{t+1} + y_t \quad (2.12)$$

Proces y_{t+h} má vysoce perzistentní povahu, totiž současná hodnota y_t je signifikantní pro určování hodnoty y v daleké budoucnosti:

$$E(y_{t+h}|y_t) = y_t, \quad \text{pro všechna } h \geq 1$$

Bez ohledu na to, jak daleko do budoucnosti koukáme, naší nejlepší predikcí y_{t+h} je současná hodnota y_t .

Obecně, můžeme dostat z autoregresního procesu obsahujícího q jednotkové kořeny stacionárního procesu tím, že diferencujeme časovou řadu q krát – jinými slovy, q -tá difference takové časové řady je stacionární. Nestacionární časová řada, která může být převedená na stacionární pomocí diferencování q krát se jmenuje integrovaná časová řada řádu q , značíme $I(q)$. Takže časová řada, která neobsahuje v sobě žádný jednotkový kořen, se bude značit $I(0)$.

Termín "integrovaný" pochází z výpočetních vlastností, že integrace a diferenciaci jsou inverzní operace.

Příkladem takových procesů je RW, který obsahuje jeden jednotkový kořen, totiž je $I(1)$.

Testy na jednotkové kořeny jsou popsány v podkapitole 2.7.4.

2.6.6 Sériová korelace náhodné složky

$$u_t = \rho u_{t-1} + e_t, \quad t = 1, \dots, n \quad (2.13)$$

$$e_t \sim N(0, \sigma_e^2), \quad \text{jsou náhodná a nekorelovaná}$$

$$|\rho| < 1$$

Sama po sobě sériová korelace (autokorelace) je porušením G–M předpokladu č. 5, ve kterém je předpokládáno, že chyba je nezávislá na minulosti – nemá „vzpomínky“ na svoji minulost. Takže autokorelovaná chyba je v podstatě korelována sama na sobě a představuje stacionární proces $AR(1)$. Po sobě jdoucí hodnoty časové řady jsou často závislé navzájem, což působí autokorelace náhodné složky.

Sériovou korelaci je třeba vzít v úvahu při testování významnosti, například korelaci mezi dvěma časovými řadami. Mimo jiné, sériová korelace (a trendy) může výrazně snížit efektivní počet stupňů volnosti v časové řadě.

Taky je třeba zmínit vlastnosti odhadů modelu časové řady obsahující v sobě autokorelaci:

- Pokud jsou splněné G–M předpoklady 1 až 3, odhady parametrů regrese jsou nezkreslené, bez ohledu na stupeň autokorelace náhodné složky;
- Za stejných podmínek plus předpoklad slabé závislosti, odhady budou taky konzistentní;
- Odhady už nejsou BLUE:
 - Odhad rozptylu odhadu koeficientu proměnné už není nezkreslený;
 - Z předchozí vlastnosti plyne, že testové statistiky a příslušné testy už nejsou platné, ani asymptoticky;

Autokorelační koeficienty versus lagy se dobře znázorňují v grafu, který je známý jako korelogram.

Testy, které nám pomůžou posoudit přítomnost sériové korelace náhodné složky, jsou popsány v podkapitole 2.7.5

2.6.7 Heteroskedasticita

Heteroskedasticita je porušením 4.G–M předpokladu. Její výsledkem je zkreslený odhad rozptylu odhadnutého parametru a odtud plyne:

- t statistika nemá studentovo t rozdělení, ani asymptoticky;
- F statistika nemá χ^2 rozdělení;
- OLS odhady už nejsou BLUE, ani asymptoticky vydatné.

Testováním heteroskedasticity se zabýváme v podkapitole 2.7.6.

2.7 Verifikace modelu

2.7.1 F–test

Testuje významnost modelu jako celku. F–test (nebo taky Fisherovo kritérium) ověřuje, zda hodnota vysvětlované proměnné závisí na lineární kombinaci vysvětlujících proměnných. Můžeme taky testovat jenom významnost kombinace několika proměnných z modelu.

Testovaná hypotéza:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0,$$

$$H_1: \text{non } H_0$$

"F hodnota" a "P hodnota(F)" statistiky popisují celkovou významnost regresního modelu. Konkrétně to znamená, že testují sdruženou nulovou hypotézu, že všechny regresní koeficienty kromě konstanty jsou rovny nule. Toto testuje neomezený model oproti modelu bez proměnných (nebo taky omezený na několik proměnných, signifikantnost kterých ověřujeme) a s odhadem závislé proměnné jako průměru hodnot závislé proměnné. Hodnota F (v případě testování celkové významnosti) je poměr regresního součtu čtverců dělený součtem čtverců chyb. Její hodnota se bude nacházet v rozmezí od nuly do nekonečna. Vzorec pro výpočet testového kritéria:

$$F = \frac{(SSR_{om} - SSR_{neom}/q)}{SSR_{neom}/(n - k - 1)} > F_{1-\alpha}(q, n - k - 1) \rightarrow \text{zamítám } H_0$$

P hodnota(F) je pravděpodobnost, že nulovou hypotézu pro celý model nezamítáme (tzn., že všechny regresní koeficienty jsou nulové). Například, pokud P hodnota (F) má hodnotu 0.01, pak máme 1 šanci ze 100, že všechny regresní parametry jsou nulové. Tato nízká hodnota by svědčila, že alespoň některé z regresních parametrů jsou nenulové a že regresní model má nějakou validitu při dosazení dat (tj. nezávislá proměnná není úplně náhodná ve vztahu k závislé proměnné).

Co to je signifikantnost?

Jednoduše řečeno, pokud máme za výsledek významnost, tak to znamená, že naše výsledky pravděpodobně nejsou náhodou. Pokud nemáme statistickou významnost za výsledek, můžeme vyhodit testovaná data (protože z nich nedostaneme nic zajímavého); jinými slovy, nelze nulovou hypotézu zamítnout.

2.7.2 Dílčí t-testy

Významnost jednotlivých parametrů testuje t-test (Studentův test), který vyžaduje splnění předpokladů klasického lineárního modelu 1.–6. Jeho podstata leží v tvrzení:

$$\hat{\beta}_j \sim N(\beta_j, \text{var}(\hat{\beta}_j)) \quad j = 1, 2, \dots, k$$

Z čehož vyplývá:

$$\frac{\hat{\beta}_j - \beta_j}{\text{sd}(\hat{\beta}_j)} \sim N(0, 1) \quad j = 1, 2, \dots, k \quad (2.14)$$

Předpoklad normality je obvykle staven na základě Centrální Limitní Věty (CLV), ale v některých případech je to omezující. To znamená, že normalita nemůže být předpokládána vždy. V libovolném příkladu, kde normalita u je předpokládána, plyne opravdu z empirické

zkušenosti. Často se používá metod, kde pomocí transformace, například logaritmování, dostaneme rozdělení, které se blíží k normálnímu, které je snadno ovladatelné z matematického hlediska. Velké vzorky nám umožňují vynechat předpoklad normality, bez významného ovlivnění výsledků.

Takže při splnění KLRM³ předpokladů 1.–6. dostaneme Studentovo t rozdělení:

$$\frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \sim t_{n-k-1} \quad j = 1, 2, \dots, k \quad (2.15)$$

Kde $(k + 1)$ je počet neznámých parametrů v populačním modelu. Výraz (2.15) je důležitý, protože nám to umožňuje testování hypotézy na β_j .

Pokud budeme srovnávat (2.14) s (2.15), vidíme, že Studentovo t rozdělení se odvozuje ze skutečnosti, že parametr σ v $sd(\hat{\beta}_j)$ byl nahrazen jeho odhadem $\hat{\sigma}$, což je náhodnou veličinou. To znamená, že stupně volnosti t jsou $n - k - 1$, které odpovídají stupňům volnosti používaných při odhadu $\hat{\sigma}^2$.

Když jsou stupně volnosti (df) v t rozdělení velké, t rozdělení konverguje k normovanému normálnímu rozdělení:

$$t_n \xrightarrow{n \rightarrow \infty} N(0,1)$$

Testovaná hypotéza:

$$H_0: \beta_j = 0$$

$$H_1: \text{non } H_0$$

Ve vícenásobném regresním modelu β_j měří parciální efekt x_j na y ceteris paribus (při neměnných ostatních faktorech). Takže testujeme hypotézu, že nezávislá proměnná x_j nemá významný vliv na závislou proměnnou y .

Testové kritérium vypadá následovně:

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - \beta_j}{se(\hat{\beta}_j)} \geq T_{1-\frac{\alpha}{2}}(n - k - 1) \rightarrow \text{zamítám } H_0$$

Analogicky F–testu můžeme vypočítat p –hodnotu pro t –test:

$$p - \text{hodnota}_t = P(|T| > |t|)$$

³ Jeffrey M. Wooldridge –Introductory Econometrics a Modern Approach

P–hodnota pro t–test bude mít stejnou interpretaci jako u F–testu ve vztahu k dílčím parametrům.

2.7.3 Vícenásobný upravený koeficient determinace

Vícenásobný koeficient determinace je podíl variability závislé proměnné, který se podařilo vysvětlit odhadnutým regresním modelem, a je mírou dobré shody modelu. Však při porovnání více modelů na jednu (při podmínce, že tyto modely obsahují intercept a mají stejnou funkční závislost), používáme upravený koeficient determinace, který je očištěný o vliv počtu vysvětlujících proměnných. Pohybuje se v rozmezí od 0 do 1, a vypočte se takto:

$$R_u^2 = 1 - \frac{SSR/(n - k - 1)}{SST/(n - 1)} = 1 - \frac{\hat{\sigma}^2}{SST/(n - 1)}$$

Používá se při rozhodování mezi nezahnížděnými modely.

2.7.4 Testy na jednotkový kořen

Jak už bylo probráno v podkapitole 2.6.5 časové řady, obsahující jednotkové kořeny, se v praxi vyskytují často a vyžadují úpravy použitím diferencování. Nejdříve ale musíme zjistit, zda ho ve skutečnosti časová řada obsahuje.

Nejjednodušším způsobem testování na jednotkový kořen je pomocí AR(1) modelu (obecně můžeme testovat AR(q) model):

$$y_t = \alpha + \rho_1 y_{t-1} + e_t, t = 1, 2, \dots, \quad (2.16)$$

kde $\{e_t\}$ má:

$$E(e_t | y_{t-1}, y_{t-2}, \dots, y_0) = 0$$

Taky můžeme označit:

$$\text{corr}(y_t, y_{t-1}) = \hat{\rho}_1$$

Když $\{y_t\}$ je proces tvaru (2.16), pak bude obsahovat jednotkový kořen při $\rho = 1$. Takže testovaná hypotéza bude mít tvar:

$$H_0: \rho = 1 \rightarrow I(1)$$

Alternativní hypotézou je ve většině případů, že $\{y_t\}$ je stacionární proces AR(1), což znamená její slabou závislost a asymptotickou nezávislost :

$$H_1: |\rho| < 1 \rightarrow I(0)$$

Test Dickey-Fullera (DF)

Bereme v úvahu rovnici (2.16). Vhodnou formou testovací rovnice je však:

$$\Delta y_t = \alpha + \theta y_{t-1} + e_t,$$

kde $\Delta y_t = y_t - y_{t-1}$ a $\theta = \rho - 1$. Takže testovaná hypotéza se transformuje na $H_0: \theta = 0$ oproti $H_1: \theta < 0$.

Pro rozhodování o platnosti hypotézy používáme t-test o parametru $\hat{\theta}$, který je popsán v podkapitole 2.7.2. Nicméně, rozdělení používané statistiky se liší od klasického rozdělení t-statistiky (Studentovo rozdělení nebo asymptotické normální rozdělení), i když máme velký výběrový vzorek. Asymptotické rozdělení τ -statistiky při H_0 je známo jako Dickey-Fullerovo rozdělení podle Davida Dickeyho a Wayne Fullera (1979)⁴.

Existují celkově tři druhy testovaných rovnic:

1. Bez konstanty i trendu:

$$\Delta y_t = \theta y_{t-1} + e_t,$$

2. S konstantou, ale bez trendu:

$$\Delta y_t = \alpha + \theta y_{t-1} + e_t,$$

3. S konstantou a s lineárním trendem:

$$\Delta y_t = \alpha_0 + \theta y_{t-1} + \alpha_1 t + e_t, \quad (2.17)$$

Pro každý druh regrese existují vlastní kritické hodnoty DF-statistiky, které jsou tabelovány.

Zamítáme nulovou hypotézu $H_0: \theta = 0$ oproti $H_1: \theta < 0$, když $\tau_{\hat{\theta}} < c$, kde c je zápornou kritickou hodnotou z tabulky značení pro DF-statistiky.

Rozšířený test Dickey-Fullera (ADF)

Když do testovaných regresních rovnic přidáme zpožděné první difference časové řady, pak se rozdělení DF-statistiky (a kritické hodnoty) nezmění. Taková modifikace se jmenuje rozšířený test Dickey-Fullera (Augmented DF, ADF).

Potřebnost přidání zpožděných diferencí je spojená s tím, že proces může být autoregresí ne prvního, ale vyššího řádu.

$$\Delta y_t = \alpha + \theta y_{t-1} + \gamma_1 \Delta y_{t-1} + e_t,$$

⁴ Dickey, D. A.; Fuller, W. A. (1979). „Distribution of the Estimators for Autoregressive Time Series with a Unit Root”.

kde $|\gamma_1| < 1$, takže při nezamítnutí $H_0: \theta = 0$, časová řada je AR(1) proces, při zamítnutí – AR(q) proces (musíme pokračovat v testování jednotkového kořenu vyššího řádu). Obecně můžeme přidat do rovnice p zpoždění Δy_t v závislosti na dynamice sledované časové řady.

2.7.5 Testování přítomnosti sériové korelace

Problematika sériové korelace je popsána v podkapitole 2.6.6. V této části se zaměříme na způsoby testování její přítomnosti.

t-test AR(1) při striktně exogenních regresorech

Požadavek absence korelace mezi u_t a závislými proměnnými ve všech časových okamžicích vylučuje modely obsahující zpožděné proměnné y .

Předpokládáme:

$$\begin{aligned} E(e_t | u_{t-1} u_{t-2}, \dots) &= 0 \\ \text{Var}(e_t | u_{t-1}) &= \text{Var}(e_t) = \sigma_e^2 \end{aligned}$$

Postup:

1. Odhadneme regresní rovnici pro y_t v závislosti na $x_{t_1}, \dots, x_{t_k}$, a uložíme rezidua $\hat{u}_t$.
2. Odhadneme následující model a dostaneme $\hat{\rho}$ a příslušnou testovou statistiku $t_{\hat{\rho}}$:

$$\hat{u}_t = c + \rho \hat{u}_{t-1} + e_t, t = 2, \dots, n \quad (2.18)$$

3. Použijeme $t_{\hat{\rho}}$ pro testování hypotézy $H_0: \rho = 0$ oproti $H_1: \rho \neq 0$.

Durbin-Watson test

Na tento test přišli James Durbin a Geoffrey Watson (1950,1951)⁵⁶. Pro jeho implementaci vyžadujeme splnění předpokladů klasického lineárního modelu (KLM).

Jako v předchozím testu používáme AR(1) model, tj. rovnice (2.), pro testování hypotézy $H_0: \rho = 0$ oproti $H_1: \rho > 0$. Testová statistika vypadá takto:

⁵ Durbin, J.; Watson, G. S. (1950). "Testing for Serial Correlation in Least Squares Regression, I".

⁶ Durbin, J.; Watson, G. S. (1951). "Testing for Serial Correlation in Least Squares Regression, II".

$$DW = \frac{\sum_{t=2}^n (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^n \hat{u}_t^2} \approx 2(1 - \hat{\rho}) \in \langle 0, 2 \rangle \quad (2.19)$$

Z vzorce (2.17) plyne: $\hat{\rho} \approx 0 \rightarrow DW \approx 2$, $\hat{\rho} > 0 \rightarrow DW < 2$, $\hat{\rho} \approx 1 \rightarrow DW \approx 0$.

t-test AR(1) bez striktně exogenních regresorů

V případě, když jedna nebo více závislých proměnných je korelována s u_{t-1} , t a DW statistiky už nejsou validní. To se stane, když například do modelu přidáme zpožděnou závislou proměnnou y_{t-1} (y_{t-1} a u_{t-1} jsou očividně korelované).

V tomto případě používáme buď Durbinovu h statistiku, a nebo alternativně Breusch-Godfrey (BG) test:

1. Odhadneme regresní rovnici pro y_t v závislosti na $x_{t_1}, \dots, x_{t_k}$, a uložíme rezidua $\hat{u}_t$.
2. Odhadneme následující model a dostaneme $\hat{\rho}$ a příslušnou testovou statistiku $t_{\hat{\rho}}$:

$$\hat{u}_t = c + \hat{\rho}\hat{u}_{t-1} + \alpha_1 x_{t_1} + \alpha_2 x_{t_2} + \dots + \alpha_k x_{t_k} + e_t, t = 2, \dots, n \quad (2.20)$$

3. Použijeme $t_{\hat{\rho}}$ pro testování hypotézy $H_0: \rho = 0$ oproti $H_1: \rho \neq 0$ ($\rho > 0$).

Testování AR(q)

Testování hypotézy $H_0: \rho_1 = \dots = \rho_q = 0$ v modelu:

$$\hat{u}_t = c + \rho_1 \hat{u}_{t-1} + \dots + \rho_q \hat{u}_{t-q} + e_t, t = 2, \dots, n \quad (2.21)$$

se uskutečňuje následujícím způsobem:

1. Odhadneme regresní rovnici pro y_t v závislosti na $x_{t_1}, \dots, x_{t_k}$, a uložíme rezidua $\hat{u}_t$.
2. Odhadneme následující model:

$$\hat{u}_t = c + \rho_1 \hat{u}_{t-1} + \dots + \rho_q \hat{u}_{t-q} + \alpha_1 x_{t_1} + \dots + \alpha_k x_{t_k} + e_t, t = (q+1), \dots, n \quad (2.22)$$

3. Použijeme F test pro $\hat{u}_{t-1} + \dots + \hat{u}_{t-q}$. Uvažujeme:
 - a. Když máme striktně exogenní závisle proměnné, můžeme je vyhodit z modelu (2.22);
 - b. Test vyžaduje splnění předpokladu homoskedasticity;

- c. Místo F statistiky můžeme použít LM statistiku, což je jednou z možností Breusch–Godfrey testu pro testování $AR(q)$ sériové korelace, testová statistika– Langreův multiplikátor (dále jen LM):

$$LM = (n - q)R_u^2 \sim \chi^2(q) \quad (2.23)$$

- d. Čtvrtletní a měsíční časové řady, které nebyly očistěné o sezónnosti, musíme otestovat na sezonní formy autokorelace. Například pro čtvrtletní řady místo (2.22) modelu, budeme testovat $\hat{\rho}_4$ v modelu:

$$\hat{u}_t = c + \rho_4 \hat{u}_{t-4} + \alpha_1 x_{t_1} + \dots + \alpha_k x_{t_k} + e_t, t = (q + 1), \dots, n \quad (2.24)$$

2.7.6 Testování heteroskedasticity

V podkapitole 2.6.7 najdeme definice a následky, které působí heteroskedasticita.

Breusch-Pagan-Godfrey test

Testuje nulovou hypotézu $H_0: \delta_1 = \delta_2 = \dots = \delta_k = 0$ z pomocné regrese:

$$\hat{u}_t^2 = \delta_0 + \delta_1 x_{t_1} + \dots + \delta_k x_{t_k} + e_t \quad (2.25)$$

Pokud tato hypotéza platí, rozptyl je konstantní a značí to homoskedasticitu. Pro testování použijeme buď F statistiku, anebo LM statistiku:

$$LM = nR_{\hat{u}^2}^2 \sim \chi^2(k) \quad (2.26)$$

Whiteův test heteroskedarsticity

Testuje nulovou hypotézu $H_0: \delta_1 = \delta_2 = \dots = \delta_{k+h+w} = 0$ (totiž homoskedasticita) oproti $H_1: homoskedasticita$ z pomocné regrese:

$$\hat{u}_t^2 = \delta_0 + \delta_1 x_{t_1} + \dots + \delta_k x_{t_k} + \delta_{k+1} x_{t,k+1}^2 + \dots + \delta_{k+h} x_{t,k+h}^2 + \delta_{k+h+1} x_{t_1} x_{t_2} + \dots + e_t,$$

nebo jednoduše:

$$\hat{u}_t^2 = \delta_0 + \delta_1 \hat{y}_t + \delta_2 \hat{y}_t^2 + e_t, \quad (2.27)$$

kde testujeme hypotézu $H_0: \delta_1 = \delta_2 = 0$.

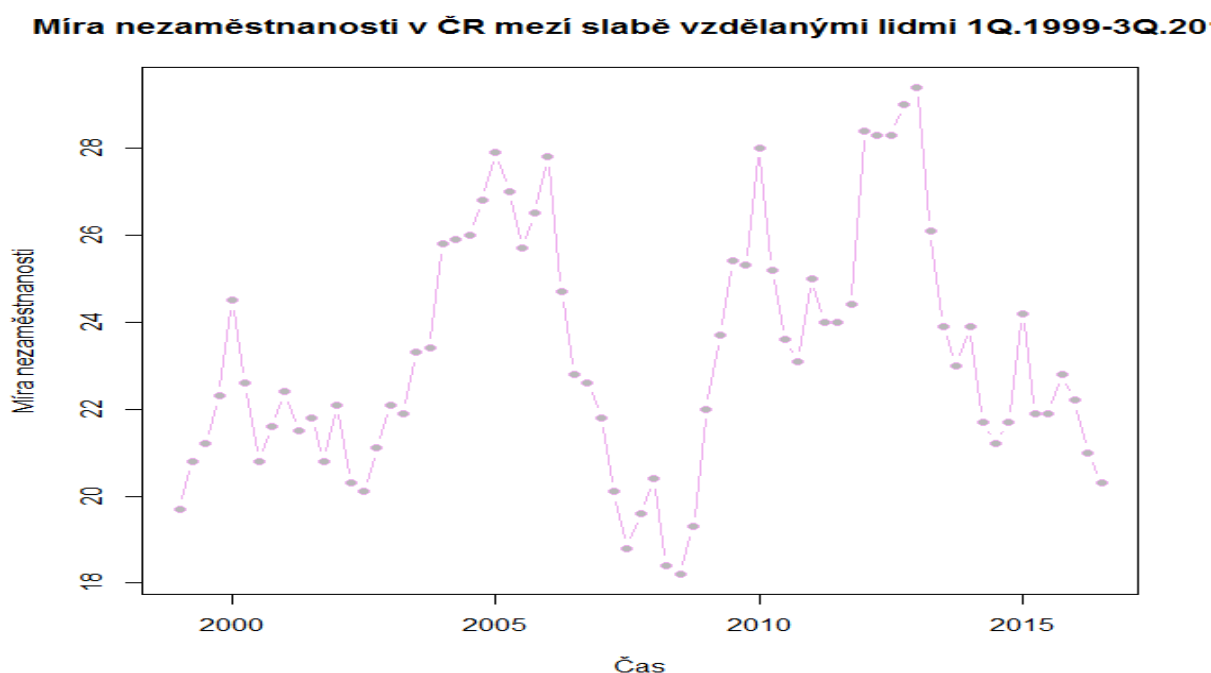
Pro test můžeme použít (v případě (2.27) modelu):

- $F \sim F(2, n - 3)$
- $LM \sim \chi^2(2)$
- p -hodnotu

3 Empirická analýza a vyhodnocení výsledků

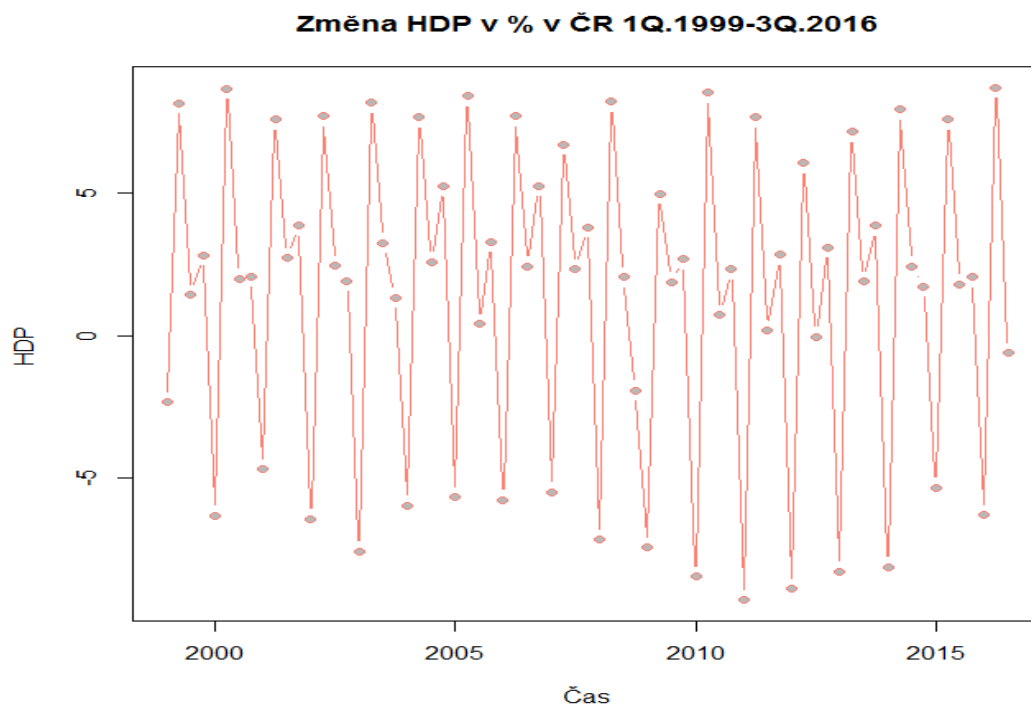
3.1 Data

V mojí bakalářské práci jsem potřebovala časové řady zahrnující v sobě údaje o celkové míře nezaměstnanosti a také podle věkových intervalů, míře růstu reálného hrubého domácího produktu v procentech za jednotlivá období, hodnoty minimální mzdy. Tady je konečná listina použitých dat:



Graf č.3 Vývoj míry nezaměstnanosti mezi slabě vzdělanými lidmi v ČR 1999–2016

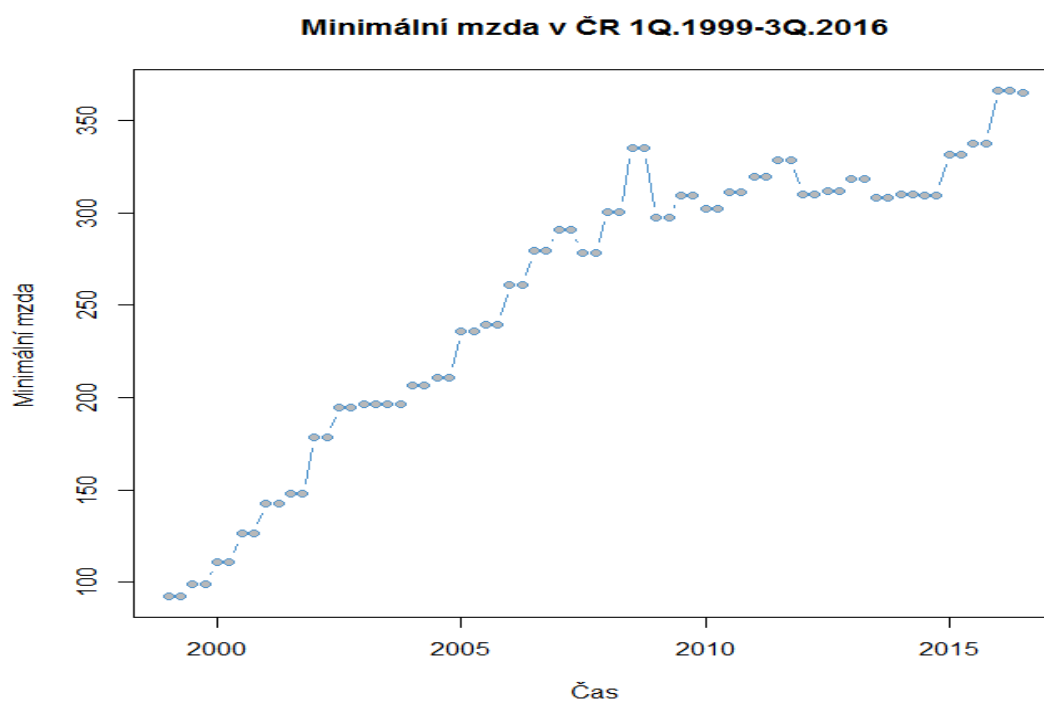
- čtvrtletní časová řada pro celkovou míru nezaměstnanosti (v %) od 1. čtvrtletí 1998 roku za ČR a od 1. čtvrtletí 1999 roku za Rakousko až do 3. čtvrtletí roku 2016 (grafy č.1 a 2);
- čtvrtletní časová řada pro celkovou míru nezaměstnanosti mezi lidmi, kteří mají buď středoškolské, základní anebo nemají vzdělání vůbec (dále jen „slabě vzdělaní lidé“) (v %) od 1. čtvrtletí 1999 roku za ČR do 3. čtvrtletí roku 2016 (graf č.3);
- čtvrtletní časová řada pro reálný HDP (v mil. Kč pro ČR a v mil. Eur pro Rakousko) od 4. čtvrtletí 1998 roku do 3. čtvrtletí roku 2016 za ČR (na grafu č.4



Graf č.4 Vývoj změny míry reálného HDP v ČR 1999–2016

je už upravena časová řada), a od 1. čtvrtletí 1999 roku do 2. čtvrtletí roku 2014 za Rakousko (tentokrát očištěná o sezónnost);

- pololetní časová řada pro minimální mzdu (v eurech) od 1. pololetí 1999 roku do 1. pololetí roku 2017 za ČR (na grafu č.5 je už opravena časová řada);



Graf č.5 Vývoj minimální mzdy v ČR 1999–2016

Data jsou stažena z <https://data.oecd.org> a z webu Českého statistického úřadu a byla vhodně upravena:

- pro celkovou míru nezaměstnanosti pro testování Okunova zákona byly vypočtené rozdíly za dvě po sobě jdoucí období (1. difference), abych dostala mezičtvrtletní změnu v míře nezaměstnanosti. Pro možnost výpočtu regresního modelu, byla časová řada zkrácena na období od 1. čtvrtletí 1999 roku do 3. čtvrtletí roku 2016, aby byla stejně dlouhá jako ostatní;
- pro výpočet mezičtvrtletní změny v reálném HDP v % jsem vydělila hodnoty časové řady jejich předchozími hodnotami, odečetla jsem 1 a vynásobila 100. Takovým způsobem se časová řada vyrovnala s ostatními;
- pololetní časová řada pro minimální mzdu byla převedena na čtvrtletní a zkrácena o poslední období.

Popisné charakteristiky časových řad:

	ČR	Rakousko	Minimální mzda	Změna v HDP
<i>Minimum</i>	3.9	3.1	92.35	-9.26
<i>1. kvantil</i>	6.0	4.3	196.37	-2.12
<i>Medián</i>	7.1	4.9	291.07	2.33
<i>Střední hodnota</i>	6.9	4.8	255.70	1.33
<i>3. kvantil</i>	8.0	5.5	311.39	5.25
<i>Maximum</i>	9.6	6.3	366.35	8.70
<i>Rozptyl</i>	1.97	0.73	6124.92	28.83
<i>Směrodatní odchylka</i>	1.4	0.85	78.26	5.37

Tabulka č.1 Popisné charakteristiky časových řad

3.2 Analýza

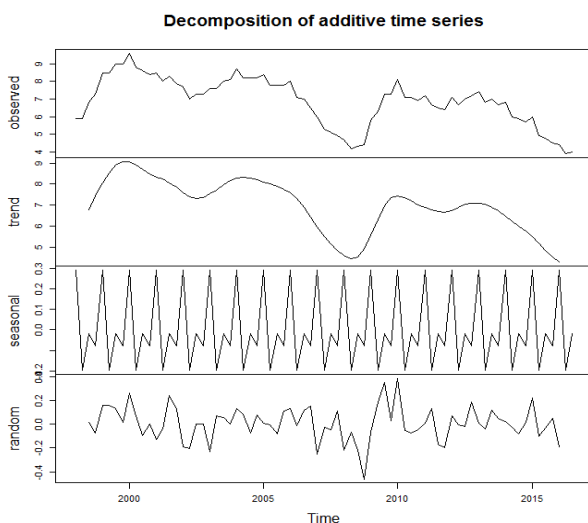
V této části odhadneme regresní modely na našich datech a zhodnotíme je. Konkrétně se tady budeme zabývat:

1. Testováním stacionarity a využití diferencování časové řady;

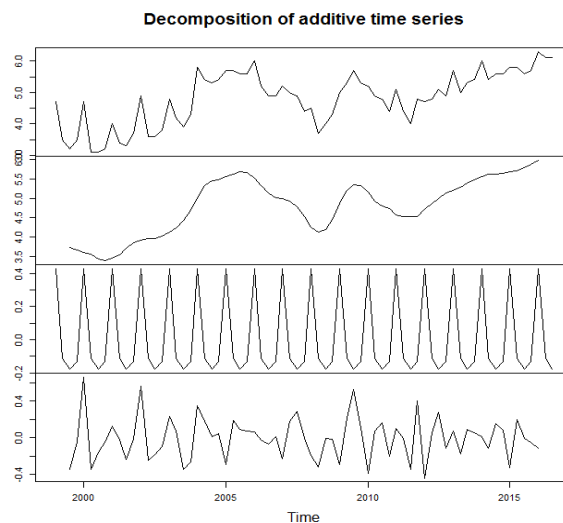
2. Dekompozicí časové řady;
3. Hodnocením vlivu krize 2008 roku;
4. Hodnocením vlivu měnové intervence ČNB;
5. Rozšířením modelu o proměnu pro vzdělání;
6. Rozšířením modelu o proměnu pro minimální mzdu;
7. Klasickým Okunovým modelem;
8. Souhrnným modelem.

3.2.1 Dekompozice časových řad a sezonní očištění

Podklady a metody dekompozice časových řad byly popsány v podkapitole 2.3. A je to důležitý moment v ekonometrické analýze– dekompozice nám pomáhá zkoumat dynamiku časové řady, což v další analýze můžeme použít pro zvolení vhodných hypotéz, testů a předpokladů (například jestli časová řada obsahuje jednotkový kořen a jakého řádu může být,



Obrázek č.1 Dekompozice časové řady míry nezaměstnanosti v ČR



Obrázek č.1 Dekompozice časové řady míry nezaměstnanosti v Rakousku

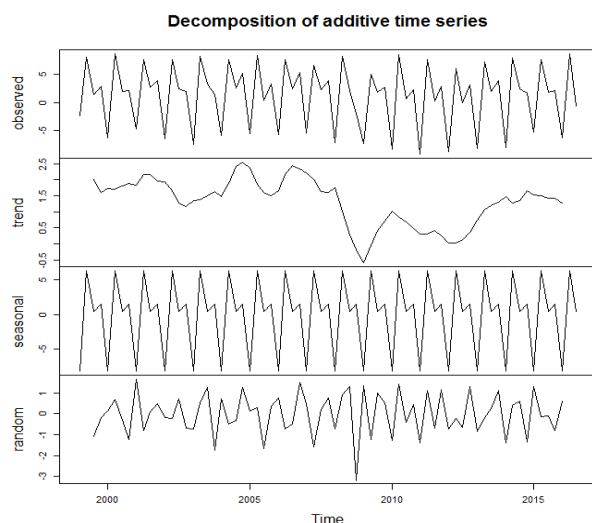
jakého řádu autokorelace se může vyskytnout při regresním modelování, ...).

Dekompozici taky můžeme využít pro očištění časových řad o sezónnost a trend, což pomáhá vyhnout se zdánlivé regresi a špatné interpretaci výsledných testů.

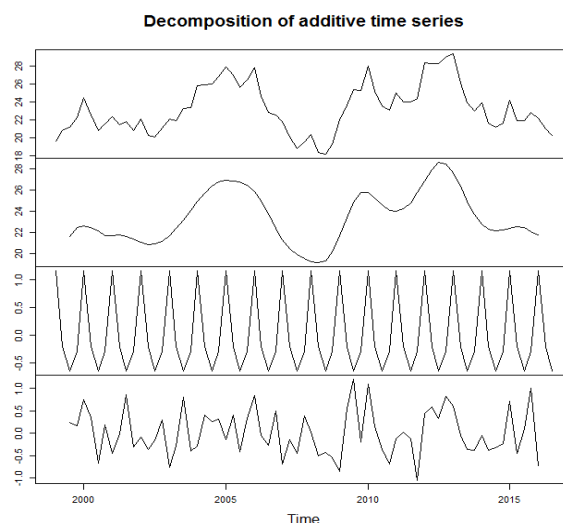
Jak už bylo popsáno v podkapitole 2.3, rozlišujeme dva druhy dekompozice: multiplikativní a aditivní. V této práci se zaměříme na aditivní dekompozici a použijeme k tomu model (2.6). Pro tyto účely nainstalujeme v R balíček *TSA* a použijeme funkci *decompose*⁷,

⁷ <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/decompose.html>

jako vstupní parametry vložíme jednotlivé časové řady, a máme hotovou dekompozici



Obrázek č.3 Dekompozice časové řady změny HDP v ČR



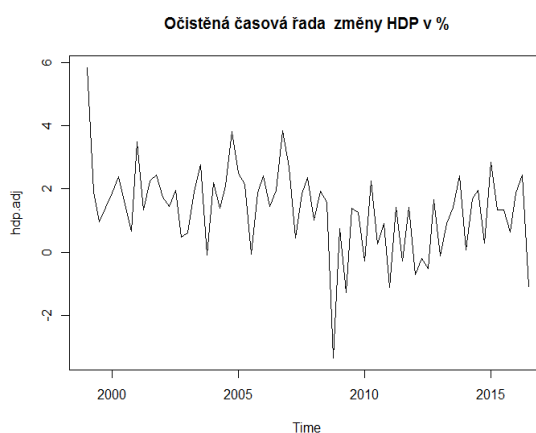
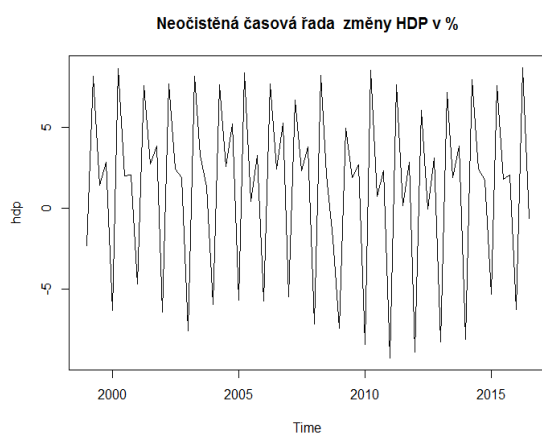
Obrázek č.4 Dekompozice časové řady míry nezaměstnanosti mezi slabě vzdělanými lidmi v ČR

klasickou metodou klouzavých průměrů.

Na obrázku č.1 je zobrazená časová řada míry nezaměstnanosti v ČR podle jednotlivých složek. Nyní je dobře vidět, že časová řada obsahuje v sobě výraznou trendovou a sezonní složku.

Analogicky budeme postupovat s časovými řadami pro Rakousko, HDP, slabě vzdělané lidi a minimální mzdu (obrázky č. 2,3,4).

Jako první při porovnání dynamiky nezaměstnanosti mezi ČR a Rakouskem na obrázcích č.1 a 2 je vidět přímo opačný trend– v ČR na rozdíl od Rakouska je trend klesající. Také



Obrázek č.5 Porovnání očištěné a neočištěné o sezonnost časové řady změny HDP v ČR

	<i>Model:</i>	
	AR(1) pro HDP v ČR	AR(1) pro minimální mzdu
hdp.lag1	0.076 (0.114)	
min.mzd.lag1		0.946*** (0.021)
Constant	1.182*** (0.218)	21.690*** (5.666)
Observations	70	69
R ²	0.007	0.967
Adjusted R ²	-0.008	0.966
Residual Std. Error	1.259 (df = 68)	13.650 (df = 67)
F Statistic	0.450 (df = 1; 68)	1,938.000*** (df = 1; 67)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01

Tabulka č.2b Testování stacionarity AR modelem

můžeme posoudit, že dynamika sezonní složky míry nezaměstnanosti v Rakousku není tak výrazná, jako v ČR.

Pro kvalitnější hodnocení závislostí mezi proměnnými pro další analýzu očistíme časové řady o sezonní složku. Očištěním časových řad o trendovou složku se však zabývat nebudeme, a pro tyto účely je lepší využít diferencování (po provedení testů na jednotkový kořen).

	<i>Model:</i>		
	AR(1) pro ČR	AR(1) pro Rakousko	AR(1) pro SVL
cr.lag1	0.996*** (0.030)		
at.lag1		0.904*** (0.057)	
educ.lag1			0.898*** (0.048)
Constant	-0.032 (0.215)	0.488* (0.279)	2.418** (1.136)
Observations	70	70	70
R ²	0.941	0.784	0.835
Adjusted R ²	0.940	0.781	0.833
Residual Std. Error (df = 68)	0.341	0.384	1.079
F Statistic (df = 1; 68)	1,076.000***	247.400***	344.000***
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01	

Tabulka č.2a Testování stacionarity AR modelem

Sezonní očištění má největší vliv na časovou řadu pro HDP (obrázek č.5), což je výrazným příkladem užitečnosti dané procedury.

3.2.2 Testování časové řady na jednotkový kořen a diferencování

Metody testování časové řady na jednotkový kořen jsou popsány v podkapitole 2.7.4. Tady se pustíme do praktického použití testů a jejich interpretaci. Pro porovnání ukážeme všechny tři způsoby.

Nejprve otestujeme časové řady pomocí AR(1) a jelikož máme časové řady očištěné o sezonní složku, AR(4) modely testovat nebudeme. Takovým způsobem budeme testovat, jak silná je závislost mezi hodnotami v daném období (v čase t) a hodnotami v předchozím období (tj. v minulém čtvrtletí: $t - 1$). Výsledky výchozích AR modelů jsou zobrazené v tabulkách č.2a a č.2b, a můžeme je posoudit: v AR(1) modelu pro míru nezaměstnanosti a také v AR(1) modelu pro minimální mzdu v ČR je $\hat{\rho} \approx 1$, což svědčí pro přítomnost jednotkového kořene, stejný závěr můžeme udělat pro míru nezaměstnanosti v Rakousku a pro míru nezaměstnanosti mezi slabě vzdělanými lidmi, kteří mají $\hat{\rho} \approx 0,9$; model pro změnu HDP má $\hat{\rho} = 0,076$ což nám říká, že časová řada je stacionární, a nemusíme být překvapeni, jelikož na začátku jsme opravovali původní časovou řadu s absolutními hodnotami na relativní změnu.

Pro stejné účely a stejné předpoklady použijeme testy DF a ADF, nastavíme H_0 : jednotkový kořen a pustíme funkce v Rstudio z balíčku *tseries*. Testujeme rovnici (2.17).

Z tabulky č.3 je vidět, že výsledky pro časové řady jsou stejné jako při testování pomocí AR(q).

Výsledky↓, Časová řada→	ČR	Rakousko	HDP	Slabě vzdělané lide	Minimální mzda
DF–testová statistika (kritická hodnota pro $\alpha=0.05$ je -3.5)	-1.2	-2.9	-8.7	-1.8	-1.8
p–hodnota	0.9	0.2	0.01	0.6	0.6
Závěr	nezamítáme H_0	nezamítáme H_0	zamítáme H_0	nezamítáme H_0	nezamítáme H_0

Tabulka č.3 Výsledky DF testů

Výsledky↓, Časová řada→	ČR	Rakousko	HDP	Slabě vzdělané lide	Minimální mzda
Zpoždění	1	1	1	1	1
ADF–testová statistika	-2.6	-2.6	-5.3	-2.4	-1.8
p–hodnota	0.3	0.3	0.01	0.4	0.7
Závěr	nezamítáme H_0	nezamítáme H_0	zamítáme H_0	nezamítáme H_0	nezamítáme H_0

Tabulka č.4 Výsledky ADF testů

V tabulce č.4 máme stejné závěry jako při testování AR(1) modelem a DF testem (kritická hodnota zůstává stejná jako u DF testu): všechny časové řady, kromě časové řady HDP v ČR, obsahují jednotkový kořen 1. řádů.

Závěrem testování časových řad na jednotkový kořen můžeme usoudit: všechny časové řady kromě HDP vyžadují diferenci 1. řádu. Provedeme tyto úpravy v R. Uděláme následovně ADF test, abychom ověřili, zda nějaká časová řada neobsahuje jednotkový kořen vyššího řádu. Výsledky byly ověřené v programu R, kde se můžeme přesvědčit, že už žádná řada neobsahuje jednotkový kořen. Na závěr můžeme říct, že všechny použité řady se řídily I(1) procesem.

Tím pádem, po provedení analýzy dynamiky časových řad dekompozicí, sezonním očištěním, testů na jednotkový kořen a následným diferencováním, jsme dostali očištěné časové řady pro kvalitní ekonometrickou analýzu.

3.2.3 Minimální mzda

Jak už bylo detailně popsáno v kapitole č.1, minimální mzda má velký vliv na nezaměstnanost, jelikož zabraňuje zaměstnavatelům posoudit a zvolit samostatně „cenu“ za konkrétního pracovníka, a tím pádem nutí hledat zaměstnance, který bude stát této minimální mzdy (a to samozřejmě na rozhodnutí zaměstnavatele).

Máme k dispozici pololetní časovou řadu minimální mzdy od 1999 roku Aby nám seděla frekvence v regresním modelu, byla tato časová řada převedena na čtvrtletní jednoduchým kopírováním hodnoty pro 1 pololetí na dva čtvrtletí.

	<i>Dependent variable:</i>	
	cr.fin	
	<i>OLS</i>	<i>GLS</i>
min.mzd4	-0.009*** (0.003)	-0.004* (0.002)
Constant	0.007 (0.044)	-0.015 (0.084)
Observations	70	70
R ²	0.128	
Adjusted R ²	0.116	
Log Likelihood		-15.640
Akaike Inf. Crit.		41.280
Bayesian Inf. Crit.		52.380
Residual Std. Error	0.319 (df = 68)	
F Statistic	10.010*** (df = 1; 68)	
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01	

Tabulka č.5 Regresní model závislosti míry nezaměstnanosti na minimální mzdě. Porovnání OLS a GLS

Diferencované časové rady pro nezaměstnanost (závislá proměnná) a pro minimální mzdu (nezávislá proměnná) v ČR očištěné o sezonnost dosadíme do regresního modelu a odhadneme pomocí MNČ parametry. Dále provedeme verifikaci modelu pomocí t testů, F testu, koeficientu determinace, řádně otestujeme přítomnost sériové korelace náhodné složky a heteroskedasticity.

V tabulce č. 5 najdeme odhadnuté parametry modelu MNČ a výsledky uvedených testů a můžeme se pustit do interpretace: při zvýšení minimální mzdy o 1 euro, klesne nezaměstnanost v průměru o 0,009 procentních bodů a tato změna je signifikantní na méně než 1% hladině významnosti (a to podle t, F testů), koeficient determinace nám říká, že odhadnutým modelem se nám podařilo vysvětlit 12,8% variability závislé proměnné (nezaměstnanosti v ČR).

Zajímavé je, že empirická analýza vyjevila opačný vliv změny minimální mzdy na míru nezaměstnanosti v ČR, a to možná z důvodu, že minimální mzda roste spolu s celkovou cenovou hladinou a ekonomikou, a možná ještě pomaleji, takže nemá takový velký vliv na rozhodnutí zaměstnavatelů.

Jelikož v případě, kdy odhadnutý model je zatížen sériovou korelací náhodné složky nebo heteroskedasticitou, nebude mít vliv na nezkreslenost odhadů parametrů, můžeme jenom pochybovat o správnosti interpretaci testů významnosti, což snadno ověříme. Použijeme pro tyto účely Durbin-Watson a Breusch-Pagan-Godfrey testy z balíčku *lmtest*. Zvolíme $\alpha=0,05$ a najdeme výsledky testů v tabulce č.6. Také připomeňme, že H_0 : *homoskedasticita* pro BPG

test, a H_0 : *absence autokorelace* pro DW test. Z výsledků můžeme usoudit, že model je zatížen autokorelací 2. řádu, takže odhadneme model ještě jednou, ale tentokrát pomocí GLS. Výstupní parametry a testy najdeme v tabulce č.5, kde je můžeme porovnat s odhady OLS.

Z výstupu je vidět, že odhad parametrů proměnné minimální mzdy se změnil z -0,009 na -0,004, konstanta se dokonce změnila z kladné na zápornou, odhady směrodatných chyb odhadnutých parametrů se také změnily a s tím i závěry testů. Při použití metody GLS změna minimální mzdy má signifikantní vliv na míru nezaměstnanosti jenom na 10% hladině významnosti. Střídání znamének mezi odhady OLS a GLS (v tomto případě u konstanty) může být způsobeno i neúplnou specifikací modelu, kdy vzhledem k ignorování dalších vlivů nemusejí být dodrženy podmínky pro *ceteris paribus* interpretaci.

Test	Lag	Autokorelace	Testová statistika	p-hodnota	Závěr
DW	1	0,44	1,077	0,000	zamítáme H_0
	2	0,25	1,421	0,028	zamítáme H_0
	3	0,20	1,516	0,054	nezamítáme H_0
BPG		-	0,7	0,4	nezamítáme H_0

Tabulka č.6 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelu závislosti míry nezaměstnanosti na minimální mzdě v ČR

3.2.4 Úroveň nezaměstnanosti mezi slabě vzdělanými lidmi

Úroveň vzdělání ovlivňuje strukturální nezaměstnanost stejně, jako úroveň kvalifikace pracovníka, takže bude mít velký vliv na nezaměstnanost jako celek. Jelikož sledujeme obecně ukazatele v čase, tj. nemáme k dispozici průřezové jednotky s odlišnou úrovní vzdělání, ukážeme, jak nezaměstnanost mezi slabě vzdělanými lidmi (dále jen SVL) ovlivňuje celkovou nezaměstnanost v ČR.

Budeme porovnávat celkově dva modely mezi sebou: závislost celkové míry nezaměstnanosti na míře nezaměstnanosti SVL (model A) se souhrnným modelem z předchozí podkapitoly (model B) (dál vždy budeme porovnávat jednoduchý regresní model s rozšířeným modelem o další významné proměnné).

	<i>Dependent variable:</i>			
	Celková nezamestnanost v CR			
	<i>OLS</i> (A)	<i>GLS</i> (A)	<i>OLS</i> (B)	<i>GLS</i> (B)
Minimalni mzda			−0.005*** (0.002)	−0.002 (0.002)
SVL	0.233*** (0.024)	0.200*** (0.024)	0.221*** (0.023)	0.197*** (0.022)
Constant	−0.068** (0.026)	−0.064 (0.051)	−0.025 (0.029)	−0.059 (0.045)
Observations	70	70	70	70
R ²	0.579		0.629	
Adjusted R ²	0.573		0.618	
Log Likelihood		9.379		5.880
Akaike Inf. Crit.		−8.757		−1.759
Bayesian Inf. Crit.		2.340		9.264
Residual Std. Error	0.222 (df = 68)		0.210 (df = 67)	
F Statistic	93.640*** (df = 1; 68)		56.790*** (df = 2; 67)	
<i>Note:</i>			*p<0.1; **p<0.05; ***p<0.01	

Tabulka č.7 Regresní model závislosti míry nezaměstnanosti na SVL, rozšířený model. Porovnání OLS a GLS

V tabulce č.7 najdeme výsledky odhadu OLS pro dva modely a můžeme je interpretovat následujícím způsobem: pro jednoduchý model– zvýší-li se nezaměstnanost mezi SVL o 1 procentní bod, vzroste celková míra nezaměstnanosti v průměru o 0,23 procentních bodů a tato změna je významná na hladině významnosti méně než 1%, odhadnutým modelem se podařilo vysvětlit 58% variability závisle proměnné; pro rozšířený model– zvýší-li se nezaměstnanost mezi SVL o 1 procentní bod, vzroste ceteris paribus (nadále všechny proměnné vícenásobné regrese interpretujeme ve smyslu ceteris paribus, ale nebudeme to opakovaně psát) celková míra nezaměstnanosti v průměru o 0,22 procentních bodů a tato změna je významná na hladině významnosti méně než 1%, zvýší-li se minimální mzda o 1 euro, poklesne míra nezaměstnanosti v průměru o 0,005 procentních bodů a tato změna je také významná na hladině významnosti méně než 1%, procento vysvětlené variability vzrostlo na 63%.

Stejně jako v předchozím příkladu musíme modely řádně otestovat na přítomnost sériové korelace a heteroskedasticity, výsledky jsou v tabulce č.8. Teprve zhodnotíme, že model A obsahuje autokorelaci náhodné složky 3. řádu, model B obsahuje autokorelaci náhodné složky 1. řádu a taky navíc k tomu heteroskedasticitu. Jako řešení se opět obrátíme k metodě GLS, kde v modelu B navíc přidáme jako váhy exponenciály dosazených hodnot z pomocné logaritmické regrese čtverců residuí na všechny regresory. Odhady jsou zapsané v tabulce č.7, kde můžeme

vidět nejen změny v odhadech parametrů, ale také změnu v závěru t testu pro proměnnou pro minimální mzdu – proměnná už není signifikantní. Odstraníme proměnnou pro minimální mzdu z modelu a budeme pokračovat v analýze.

<i>Test</i>	<i>Lag</i>	<i>Autokorelace</i>		<i>Testová statistika</i>		<i>p-hodnota</i>		<i>Závěr</i>	
<i>Model</i>		A	B	A	B	A	B	A	B
<i>DW</i>	1	0,38	0,31	1,23	1,38	0,000	0,002	zamítáme H_0	zamítáme H_0
	2	0,27	0,18	1,45	1,63	0,024	0,154	zamítáme H_0	nezamítáme H_0
	3	0,24	0,19	1,50	1,60	0,042	0,140	zamítáme H_0	nezamítáme H_0
<i>BPG</i>		-	-	1,8	6,5	0,2	0,04	nezamítáme H_0	zamítáme H_0

Tabulka č.8 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelech A a B

3.2.5 Dummy proměnná „2008 rok“

Jak jsme popisovali dynamiku časové řady v podkapitole 1.2, globální krize byla uvedena jako jedna z příčin velkého skoku v nezaměstnanosti v ČR na konci 2008 roku.

Pro hodnocení vlivu krize vytvoříme umělou proměnnou „2008 rok“, která nabývá 1 jde-li o období od 4. čtvrtletí 2008 roku až po 1. čtvrtletí 2010 roku, jinak se rovná 0. Pustíme regresi: míra nezaměstnanosti na vytvořenou dummy proměnu pro 2008, abychom zhodnotili samostatný vliv krize (model A). Výsledky najdeme v tabulce č.9: jde-li o období, ve kterém proměnná „2008 rok“ nabývá 1, je nezaměstnanost vyšší v průměru o 0,7 procentních bodů, a tento vliv je významný na hladině významnosti méně než 1%, odhadnutým modelem bylo vysvětleno 34% variability závislé proměnné.

Stejně jako v předchozích empirických úlohách otestujeme model na sériovou korelaci a porušení předpokladu o konstantním rozptylu (tabulka č.10). Závěrem jsme zjistily, že model A má autokorelovanou náhodnou složku 3. řádu, ale není zatížen heteroskedasticitou. V tomto případě nám zase pomůže metoda GLS, kde použijeme vhodnou transformaci podle residuí,

	<i>Dependent variable:</i>			
	Celková nezamestnanost v ČR			
	<i>OLS</i> (A)	<i>GLS</i> (A)	<i>OLS</i> (B)	<i>GLS</i> (B)
SVL			0.194*** (0.022)	0.184*** (0.023)
rok 2008	0.704*** (0.118)	0.717*** (0.158)	0.429*** (0.088)	0.418*** (0.106)
Constant	-0.120*** (0.035)	-0.107* (0.060)	-0.103*** (0.024)	-0.101*** (0.032)
Observations	70	70	70	70
R ²	0.343		0.690	
Adjusted R ²	0.333		0.680	
Log Likelihood		-6.229		13.250
Akaike Inf. Crit.		22.460		-16.500
Bayesian Inf. Crit.		33.550		-5.478
Residual Std. Error	0.277 (df = 68)		0.192 (df = 67)	
F Statistic	35.430*** (df = 1; 68)		74.450*** (df = 2; 67)	
<i>Note:</i>			*p<0.1; **p<0.05; ***p<0.01	

Tabulka č.9 Regresní model závislosti míry nezaměstnanosti na dummy proměnné „2008 rok“
“rozšířený model. Porovnání OLS a GLS

které se řídí procesem AR(3). V tabulce č.9 jsou odhady, které se příliš neliší od odhadů získaných OLS, a závěry testů zůstávají stejné.

Přidáme “2008 rok“ do předchozího modelu, ze kterého jsme už odstranili proměnnou pro minimální mzdu. Odhadnutý model obsahuje autokorelaci 1. řádu (tabulka č.10), z čeho opět plyne potřeba použití GLS. V tabulce č.9 jsou výsledné odhady, které můžeme interpretovat: za stejné míry SVL je nezaměstnanost v ČR vyšší v průměru o 0,42 procentních bodů během krize 2008 roku, je výše této změny signifikantní na hladině významnosti méně než 1%, a tímto modelem jsme vysvětlili 69% rozptylu míry nezaměstnanosti v ČR. Všechny proměnné v daném modelu jsou signifikantní, a taky sdruženě signifikantní na hladině významnosti méně než 1%.

V Rakousku měla krize špatný dopad na míru nezaměstnanosti, ale o něco menší: odhadnutý koeficient u dummy proměnné „2008 rok“ se rovná 0,414 oproti 0,717 v ČR, a je signifikantní na hladině významnosti méně než 5%.

Test	Lag	Autokorelace		Testová statistika		p-hodnota		Závěr	
Model		A	B	A	B	A	B	A	B
DW	1	0,35	0,25	1,23	1,50	0,000	0,018	zamítáme H_0	zamítáme H_0
	2	0,18	0,11	1,53	1,74	0,044	0,260	zamítáme H_0	nezamítáme H_0
	3	0,19	0,15	1,49	1,66	0,050	0,178	zamítáme H_0	nezamítáme H_0
BPG		-	-	3,6	4,4	0,06	0,1	nezamítáme H_0	nezamítáme H_0

Tabulka č.10 Výsledky testování sériové korelace a heteroskedasticity v OLS modelech A a B

Dependent variable:				
Celkova nezamestnanost v CR				
	OLS (A)	GLS (A)	OLS (B)	GLS (B)
SVL			0.186*** (0.022)	0.180*** (0.022)
rok 2008			0.408*** (0.085)	0.397*** (0.101)
Intervence	-0.272*** (0.094)	-0.309*** (0.101)	-0.132** (0.055)	-0.144* (0.073)
Constant	-0.002 (0.043)	0.027 (0.090)	-0.073*** (0.026)	-0.068** (0.034)
Observations	70	70	70	70
R ²	0.110		0.714	
Adjusted R ²	0.097		0.701	
Log Likelihood		-10.150		13.000
Akaike Inf. Crit.		30.310		-13.990
Bayesian Inf. Crit.		41.410		-0.855
Residual Std. Error	0.322 (df = 68)		0.185 (df = 66)	
F Statistic	8.425*** (df = 1; 68)		55.040*** (df = 3; 66)	

Note:

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Tabulka č.11 Regresní model závislosti míry nezaměstnanosti na dummy proměnné „intervence“, rozšířený model. Porovnání OLS a GLS

3.2.6 Dummy proměnná „Intervence“

Jelikož se v současné době hodně mluví o ukončení intervencí ČNB, je zajímavé zkoumat, jestli měla nějaký pozitivní vliv na míru nezaměstnanosti.

Devizové intervence jsou devizové obchody centrální banky, jejichž cílem je ovlivnění kurzu domácí měny (v se ČR měna oslabovala). V ČR bylo jejich cílem dosažení cíle inflace v rozmezí 1-3%. Pro dosažení inflačního cíle udržovala ČNB kurz koruny nad hranici 27 korun za euro, a to pomocí devizových intervencí– vytváření nových korun, za které se nakupují eura.

V naší časové řadě můžeme pozorovat dvě období intervencí: 3. čtvrtletí 2001–2. čtvrtletí 2002 roku a 4. čtvrtletí 2013 roku až po 4. čtvrtletí 2016 roku. Takže vytvoříme příslušnou umělou proměnnou, která nabývá 1 v uvedených čtvrtletích, a pustíme na ni regresi (model A). Výsledný model najdeme v tabulce č.11. Po provedení DW a BPG testů (tabulka č.12), jsme zjistili, že v modelu je přítomná autokorelace náhodné složky 2. řádu a heteroskedasticita, takže zase musíme využít kombinovanou GLS pro případ nalezení autokorelované náhodné složky a heteroskedasticity současně. Interpretovat už budeme rovnou GLS odhady: v době intervencí klesá nezaměstnanost v průměru o 0,31 procentního bodu, tento vliv je významný na hladině významnosti méně než 1%, odhadnutým modelem bylo vysvětleno 11% variability závislé proměnné. Takže můžeme mluvit o užitečnosti provádění devizových intervencí pro

Test	Lag	Autokorelace		Testová statistika		p-hodnota		Závěr	
Model		A	B	A	B	A	B	A	B
DW	1	0,50	0,21	0,959	1,56	0,000	0,034	zamítáme H_0	zamítáme H_0
	2	0,35	0,12	1,234	1,70	0,000	0,182	zamítáme H_0	nezamítáme H_0
	3	0,19	0,16	1,562	1,62	0,112	0,154	nezamítáme H_0	nezamítáme H_0
BPG		-	-	5,1	7,6	0,02	0,05	zamítáme H_0	zamítáme H_0

Tabulka č.12 Výsledky testování sériové korelaci a heteroskedasticity v OLS modelech A a B

zaměstnanost.

Následujícím krokem rozšíříme předchozí model o proměnnou „intervence“ (model B), abychom zkoumali sdruženou významnost vysvětlujících proměnných a zároveň jaké procento variability závislé proměnné už jsme schopni vysvětlit. Po testování autokorelace a heteroskedasticity (tabulka č.12) víme, že model je zatížený autokorelací náhodné složky 1. řádu a heteroskedasticitou, zase se obrátíme na GLS a odhady interpretujeme následujícím

způsobem: po využití GLS proměnná „intervence“ je významná na hladině významnosti méně než 10%, a při neměnných ostatních faktorech, mají intervence ČNB pozitivní vliv na míru nezaměstnanosti – snižuje v průměru o 0,14 procentního bodu; ostatní proměnné zůstávají významné na hladině významnosti méně než 1%, stejné výsledky má i test na sdruženou signifikantnost. Na daný okamžik můžeme vysvětlit 71,4% rozptylu závisle proměnné.

3.2.7 Okunův zákon

Modelujeme Okunův zákon rovnicí (1.2), a zkusíme si prokázat vztah (1.3) na

	<i>Dependent variable:</i>			
	Celková nezamestnanost v ČR			
	<i>OLS</i> (A)	<i>GLS</i> (A)	<i>OLS</i> (B)	<i>GLS</i> (B)
SVL			0.186*** (0.022)	0.178*** (0.022)
rok 2008			0.407*** (0.091)	0.434*** (0.113)
HDP	-0.074** (0.032)	-0.005 (0.026)	-0.001 (0.019)	0.018 (0.018)
Intervence			-0.132** (0.056)	-0.143** (0.072)
Constant	0.036 (0.056)	-0.039 (0.094)	-0.072* (0.039)	-0.094** (0.045)
Observations	70	70	70	70
R ²	0.076		0.714	
Adjusted R ²	0.062		0.697	
Log Likelihood		-14.710		10.950
Akaike Inf. Crit.		41.410		-7.905
Bayesian Inf. Crit.		54.730		7.316
Residual Std. Error	0.328 (df = 68)		0.187 (df = 65)	
F Statistic	5.564** (df = 1; 68)		40.660*** (df = 4; 65)	
<i>Note:</i>			*p<0.1; **p<0.05; ***p<0.01	

Tabulka č.13 Regresní model Okuna, rozšířený model. Porovnání OLS a GLS

empirických datech (model A).

Při použití OLS pro odhady parametrů vztahu závislosti změny míry nezaměstnanosti v ČR na změně HDP se nám podařilo přiblížit jak ke konstantě, tak i ke sklonu modelu Okuna (tabulka č.14): aby došlo k poklesu nezaměstnanosti, musí HDP v průměru růst rychleji než o 0,49 (0,036/0,074) procentního bodů čtvrtletně, a při růstu HDP o 1 procentní bod dojde k poklesu nezaměstnanosti v průměru o 0,038 (0,036-0,074) procentního bodu (taková

interpretace plyne z toho, že používáme tedy 1. differenci časových řad, tzn. mezičtvrtletní změnu v míře nezaměstnanosti a v HDP), a je tato změna významná na hladině významnosti méně než 5%, odhadnutý model vysvětluje 7,6% variability závisle proměnné.

Po testování modelu usoudíme autokorelaci náhodné složky 2. řádu, ale absenci heteroskedasticity (tabulka č.13). Při použití GLS pro odhad parametrů, proměnná změny HDP už není významná a vyvola pokles v nezaměstnanosti při libovolné pozitivní změně v HDP (díky záporné úroňové konstantě), ale připomenme, že kvůli tomu, že používáme FGLS, nemají odhady vlastnost nezkreslenosti, ale jsou konzistentní (máme jenom 70 pozorování takže to nejde ověřit) a můžeme spolehnout na testy významnosti, které tentokrát svědčí o nevýznamnosti vlivu HDP na vývoj míry nezaměstnanosti.

Při odhadu rozšířeného modelu GLS (při odhadu OLS se objevila autokorelace náhodné složky, ale tentokrát 1. řádu, viz. tabulka č.13) proměnná HDP nejmon přestala být

Test	Lag	Autokorelace		Testová statistika		p-hodnota		Závěr	
Model		A	B	A	B	A	B	A	B
DW	1	0,39	0,21	1,159	1,56	0,002	0,036	zamítáme H_0	zamítáme H_0
	2	0,29	0,12	1,343	1,70	0,006	0,180	zamítáme H_0	nezamítáme H_0
	3	0,20	0,16	1,504	1,62	0,068	0,094	nezamítáme H_0	nezamítáme H_0
BPG		-	-	0,67	7,7	0,4	0,1	nezamítáme	nezamítáme

Tabulka č.14 Výsledky testování sériové korelace a heteroskedasticity v OLS modelech A a B

významnou, ale i změnila znaménko na kladné, což také můžeme vysvětlit možností zkreslení odhadů FGLS. Ostatní proměnné modelu zůstávají významné, a také sdruženě významné.

V Rakousku odhadnutá rovnice udává významnější závislost změny v míře nezaměstnanosti na změně HDP: odhadnutý koeficient u HDP se rovná -0,205 a je významný na hladině významnosti méně než 1%, konstanta je rovna 0,2025. Koeficienty můžeme interpretovat takto: aby klesala míra nezaměstnanosti v Rakousku musí HDP v průměru růst alespoň o 1 procentní bod každé další čtvrtletí. Přiblížili jsme se odhadnutým modelem „Okunuvské“ konstantě, což nemůžeme říct pro koeficient sklonu, však zůstává HDP významnou vysvětlující proměnnou pro míru nezaměstnanosti v Rakousku a vysvětluje 11,8% variability míry nezaměstnanosti v Rakousku.

3.2.8 F test pro sdruženou významnost proměnných “minimální mzda” a “HDP”

Ačkoliv proměnné “minimální mzda” a “HDP” nejsou významné podle t testů (viz.

	<i>Dependent variable:</i>	
	Celková nezamestnanost v ČR	
	<i>OLS</i>	<i>GLS</i>
SVL	0.187*** (0.022)	0.180*** (0.022)
rok 2008	0.372*** (0.094)	0.392*** (0.117)
HDP	0.009 (0.021)	0.022 (0.019)
Minimalni mzda	-0.003 (0.002)	-0.002 (0.002)
Intervence	-0.122** (0.056)	-0.134* (0.070)
Constant	-0.064 (0.039)	-0.082* (0.046)
Observations	70	70
R ²	0.722	
Adjusted R ²	0.701	
Log Likelihood		6.102
Akaike Inf. Crit.		3.796
Bayesian Inf. Crit.		21.070
Residual Std. Error	0.186 (df = 64)	
F Statistic	33.300*** (df = 5; 64)	
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01		

Tabulka č.15 Souhrnný regresní model. Porovnání OLS a GLS

tabulka č.15), a to ani metodou GLS (má autokorelační náhodnou složku), mohou být sdruženě významné. Pro ověření dané hypotézy použijeme parciální F test, kde $H_0: \beta_3 = \beta_4 = 0$, tzn. sdružená nesignifikantnost. Výsledné statistiky a závěr najdeme v tabulce č. 16. Podařilo se nám zamítnout H_0 a to znamená, že minimální mzda a HDP mají sdruženě významný vliv na míru nezaměstnanosti v ČR– přispívají významnými informacemi k míře nezaměstnanosti, když proměnné SVL, „rok 2008“ a „intervence“ byly přijaty v úvahu.

Podobný závěr by mohl být způsobený tím, že máme v modelu proměnnou SVL, která je hodně korelována se závislou proměnnou ($corr = 0,76$), a tím může zvyšovat rozptyly odhadů,

což má za následek nízké hodnoty testovacích kritérií pro t testy– regresní koeficienty se jeví statisticky nesignifikantní.

<i>Testovací statistika</i>	<i>p-hodnota</i>	<i>Závěr</i>
13,79	0,001	zamítáme H_0

Tabulka č.16 Výsledky parciálního F testu.

Závěr

Cílem práce byla analýza míry nezaměstnanosti v ČR, její závislosti na různých faktorech, empirické ověření platností předpokladů a nakládání s problematikou časových řad v praxi, a také jednodušší porovnání její dynamiky s dynamikou míry nezaměstnanosti v Rakousku, ověření platnosti Okunova zákona v ČR a v Rakousku.

Ohledně porovnání dynamiky nezaměstnanosti v ČR a Rakousku můžeme udělat následující závěr: obě časové řady jsou závislé na sezonní složce – výrazné zvýšení míry nezaměstnanosti v prvních čtvrtletí a postupný pokles v následujících; pro ČR však můžeme pozorovat větší sezonní změny v porovnání z Rakouskem. Obě časové řady v sobě obsahují stochastický trend (jednotkový kořen), ale mají různou povahu: v případě ČR má časová řada dlouhodobě klesající trend ($T=-0,04$), v případě Rakouska – dlouhodobě rostoucí ($T=0,03$). Globální krize měla špatný významný vliv jak na míru nezaměstnanosti v ČR, tak i na míru nezaměstnanosti v Rakousku, v Rakousku ale o něco menší.

Prakticky jsme zkoumali vliv takových faktorů jako jsou: míra nezaměstnanosti mezi lidmi, kteří mají buď středoškolské, základní, anebo nemají vzdělání vůbec, krize 2008 roku a období po ní, minimální mzda, devizové intervence ČNB, změna v HDP. Předpoklady o vlivu těchto faktorů jsou postaveny na teoriích, které byly popsány v 1. kapitole. V rámci empirické analýzy závislosti míry nezaměstnanosti na minimální mzdě samostatně se nám podařilo prokázat významnost pozitivního vlivu růstu minimální mzdy na míře nezaměstnanosti v ČR – snížení míry nezaměstnanosti, což jde v rozporu s teoretickým tvrzením negativního vlivu, totiž zvýšení míry nezaměstnanosti. Další hypotézou jsme zvolili významnost vlivu míry nezaměstnanosti mezi slabě vzdělanými lidmi na celkovou nezaměstnanost a došli jsme k závěru, že tato hypotéza platí a nezávislá proměnná má výrazný negativní vliv na celkovou míru nezaměstnanosti, což není překvapivé (jsou mezi sebou silně korelované, korelační koeficient=0,76). Krize 2008 roku jako faktor vysvětluje až 33% dynamiky nezaměstnanosti v ČR, což je velké procento z ohledem na to, že měla vliv jen na 6 čtvrtletí z 70 zkoumaných; a měla signifikantně špatný vliv na celkovou míru nezaměstnanosti. Otázkou také bylo, jestli má nějaký vliv intervence ČNB, a empirickou analýzou jsme zjistili, že z hlediska nezaměstnanosti splňují intervence svůj cíl a mají na ni pozitivní signifikantní vliv.

Dalším bodem práce bylo ověřit platnost Okunova zákona na empirických datech. Pomocí jednoduchého modelu závislosti změny v míře nezaměstnanosti na změně HDP na použitých datech jsme zjistili, že v současné době se platnost Okunova zákonu nepotvrzuje, což nám říká, že Okunův zákon v ČR není dominantní pro dynamiku nezaměstnanosti. V ČR nemá HDP významný vliv na míru nezaměstnanosti a tímto modelem se nám podařilo vysvětlit jen 7,6% variability vysvětlované proměnné. Asi převazují jiné vlivy jako například strukturální změny při vstupu do EU, globalizace, nástup IT oborů... V Rakousku ale podle skutečných výsledku regresní analýzy závěrem je významnost vlivu HDP na nezaměstnanost, i když

odhadnuté koeficienty neodpovídají teoretickým. Odhadnutý model vysvětluje 11,8% variability závislé proměnné, což je lepším výsledkem v porovnání s ČR.

Nejlepší výsledky z hlediska vysvětlení variability (až 72%) se ale podařilo získat pomocí rozšířeného modelu o všechny vysvětlující proměnné, které jsme před tím zkoumali zvlášť. Po potřebném použití metod robustních vůči heteroskedasticitě jsme získali model lépe popisující vysvětlovanou proměnnou míry nezaměstnanosti. Ale v takovém rozšířeném modelu proměnné pro minimální mzdu a změnu HDP se ukázaly být nesignifikantní zvlášť, ale signifikantní sdruženě. Ostatní proměnné zůstaly významné. Z tohoto pohledu se projekt jeví jako úspěšný. Na našich datech se nám podařilo potvrdit vliv některých faktorů, uvedených v teorii.

Literatura

- [1] WOOLDRIDGE, J M. *Introductory econometrics: a modern approach*. Mason: South-Western Cengage Learning, 2009. ISBN 978-0-324-58162-1.
- [2] HUŠEK, R. *Ekonometrická analýza*. Praha: Oeconomica, 2007. ISBN 978-80-245-1300-3.
- [3] Mankiw, NG. *Macroeconomics*, 7th Edition. Worth Publishers; 2010.
- [4] Arthur M. Okun 1962 *Potential GNP: It's Measurement and Significance*
- [5] Donald G. Freeman 2000 *Regional Tests of Okun's Law*

Internetové zdroje:

- [6] <https://data.oecd.org/> : Datový zdroj.
- [7] <https://CRAN.R-project.org/view=TimeSeries> : Návod na výběr balíčků v R pro analýzu časových řad podle úlohy.
- [8] <https://socserv.socsci.mcmaster.ca/jfox/Books/Companion/appendix/Appendix-Timeseries-Regression.pdf> : Regrese s časovými řadami pomocí GLS v R.
- [9] <https://cran.r-project.org/web/packages/lmtest/lmtest.pdf> : Popis funkcí balíčku *lmtest* v R.
- [10] <https://cran.r-project.org/web/packages/stargazer/stargazer.pdf> : Instrukce pro balíček *Stargazer* v R.
- [11] <https://anomaly.io/seasonal-trend-decomposition-in-r/> : Tutorial pro dekompozici v R.
- [12] <https://www.jalobe.com/doc/uroot.pdf> : další informace o testování jednotkového kořenu.
- [13] <https://cran.r-project.org/web/packages/tseries/tseries.pdf> : Instrukce pro balíček *tseries* v R.
- [14] <http://www.stat.columbia.edu/~martin/W2024/R6.pdf> : Tutorial výpočtu parciálního F testu v R.
- [15] <https://www.novinky.cz/ekonomika/428225-cesko-ma-nejnizsi-nezamestnanost-v-historii.html>
- [16] http://www.historickygis.cz/sites/default/files/svoboda_p._nemeskal_j._2015_nezamestnanost_v_cesku_v_historickem_pohledu_geograficke_rozhledy_24_c._4_s._28-29.pdf

- [17] <http://www.e-polis.cz/clanek/vyvoj-nezamestnanosti.html> : Tento a předchozí dva články jsem použila pro seznámení s vývojem míry nezaměstnanosti v ČR.
- [18] http://www.wifo.ac.at/jart/prj3/wifo/resources/person_dokument/person_dokument.jart?publikationsid=25832&mime_type=application/pdf
- [19] <http://www.tradingeconomics.com/austria/unemployment-rate>
- [20] <http://countrystudies.us/austria/79.htm>
- [21] <https://www.socialeurope.eu/2014/03/youth-unemployment-austrians/> : Tento a předchozí tři články jsem použila pro seznámení s vývojem míry nezaměstnanosti v Rakousku.
- [22] <https://www.vlada.cz/assets/media-centrum/aktualne/VLIV-MINIMALNI-MZDY-NA-NEZAMESTNANOST-V-CR.pdf> : Teorie vlivu minimální mzdy na míru nezaměstnanosti.
- [23] <http://www.frbsf.org/economic-research/publications/economic-letter/2005/july/age-and-education-effects-on-the-unemployment-rate/> : Teorie vlivu vzdělání na míru nezaměstnanosti.
- [24] http://ekonomika.idnes.cz/financni-krize-zvysila-pocet-nezamestnanych-fnr-/ekonomika.aspx?c=A081208_090147_ekonomika_pin
- [25] https://www.cnb.cz/miranda2/export/sites/www.cnb.cz/cs/verejnost/pro_media/konference_projevy/vystoupeni_projevy/download/Zamrazilova_20081202_brno.pdf : Tento a předchozí internetový zdroj jsem použila pro seznámení s globální krizí a její důsledky pro českou ekonomiku.
- [26] <http://zpravy.e15.cz/intervence-cnb-wiki>
- [27] https://www.cnb.cz/cs/faq/menovy_kurz_jako_nastroj_menove_politiky.html
- [28] <https://www.web4trader.cz/prehledna-historie-intervenci-cnb-46464/>
- [29] http://wiki.rvp.cz/Kabinet/Ucebni_texty/%C4%8Clov%C4%9Bk_a_spole%C4%8Dnost_m%C4%9Bnov%C3%A1_politika/12._Devizov%C3%A9_rezervy_a_devizov%C3%A9_intervence : Tento a předchozí tři články jsem použila pro seznámení s pojmem „intervence“ a zjištění období, ve kterých byly prováděny v ČR.
- [30] <http://www.investopedia.com/dictionary/> : Jako zdroj definicí řady pojmů v statistice a ekonometrii.