

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS

Kursinis darbas

**Japonijos jaunimo nedarbo lygio modeliavimas ir
prognozavimas**

Modeling and forecasting youth unemployment rate in Japan

Gabrielė Ruminavičiūtė

VILNIUS 2019

**MATEMATIKOS IR INFORMATIKOS FAKULTETAS
TAIKOMOSIOS MATEMATIKOS INSTITUTAS
STATISTINĖS ANALIZĖS KATEDRA**

Darbo vadovas: j. asist. Andrius Buteikis

Darbas apgintas 2019 m. _____ mėn. ____ d.

Turinys

1	Įvadas	3
2	Ekonominis kontekstas	3
2.1	Nedarbo lygio ir užimtumo sąvokos	3
2.2	Jaunimo nedarbo lygis pasaulyje	4
2.3	Jaunimo nedarbo lygio įtaka	4
2.4	Japonijos jaunimo nedarbo lygis	5
2.5	Literatūros apžvalga	5
3	Duomenys ir metodologija	6
3.1	Duomenys	6
3.2	Metodologija	6
3.2.1	Stacionarumo tikrinimas	6
3.2.2	Vienmačių laiko eilučių modelis	7
3.2.3	Daugiamatčių laiko eilučių modelis	8
3.2.4	Modelio tinkamumo tikrinimas	10
3.2.5	Prognozavimas	10
4	Praktinė dalis	12
4.1	Pirminė statistinė duomenų analizė	12
4.1.1	Aprašomoji analizė	12
4.1.2	Stacionarumo testo analizė	13
4.2	Ekonometrinių modelių specifikacija	14
4.2.1	Sezoninio autoregresinio integruoto slenkamųjų vidurkių proceso modeliavimas	14
4.2.2	Vektorinės autoregresijos proceso modeliavimas	15
4.2.3	Vektorinis paklaidų korekcijos modelis	18
4.2.4	Modelių prognozavimo rezultatai	19
5	Išvados bei rekomendacijos	22
6	Pirmas priedas	23

Japonijos jaunimo nedarbo lygio modeliavimas ir prognozavimas

Santrauka

Jaunimo nedarbas yra globali problema, turinti didelį neigiamą poveikį ekonominiam ir socialiniam gyvenimui. Tačiau jaunimo nedarbo lygio temos aktualumui augant, tampa vis prasmingiau kalbėti, kodėl Japonija išlaiko vieną žemiausių jaunuolių nedarbo lygio reitingų. Todėl šio tyrimo tikslas yra ištirti Japonijos jaunimo nedarbo lygio laiko eilutę, jai sumodeliuoti atitinkamus ekonometrinius modelius pasitelkiant makroekonominis veiksniai bei palyginti šiuos modelius, sudarant Japonijos nedarbo lygio laiko eilutės prognozę. Tyrime sudaryti SARIMA, VAR ir VECM modeliai, naudojant OECD Japonijos ketvirtinius jaunimo nedarbo lygio duomenis nuo 1968 iki 2018 metų. Taip pat naudojami BVP pokyčių duomenys bei infliacija. Visi sudaryti modeliai yra tinkami ir adekvatūs. Tačiau visų modelių liekanų dispersija yra nepastovi. Nustatyta, kad SARIMA modelis yra tinkamiausias Japonijos jaunimo nedarbo lygiui prognozuoti. Vis dėlto egzistuoja ryšis tarp jaunimo nedarbo ir BVP pokyčių bei nedarbo ir infliacijos. Bet koks impulsas ateinantis iš BVP pokyčių ar infliacijos, veikia Japonijos jaunimo nedarbo lygį. Taip pat tyrime yra nustatyta, kad trumpojo laikotarpio prognozės yra geresnės, negu ilgojo, dėl to tolimesni tyrimai reikalingi, kad pagerinti modelių prognozavimą ilgalaikiam jaunimo nedarbo lygiui nustatymui.

Raktiniai žodžiai : Jaunimo nedarbo lygis; Modeliavimas; Prognozavimas; Laiko eilutės

Modeling and forecasting youth unemployment rate in Japan

Abstract

Youth unemployment is a global problem that has a major negative impact on economic and social life. However, as the topic of youth unemployment rises, it becomes more and more meaningful to talk about the reasons of why Japan maintains one of the lowest youth unemployment rates. Therefore, the main objective of this study is to investigate the time series of Japanese unemployment rate, specify appropriate econometric models using macroeconomic factors, and to compare these models by forecasting Japanese unemployment rate. The study compiled SARIMA, VAR, and VECM models using OECD Japan quarterly youth unemployment data from 1968 to 2018. Also using data on GDP changes and inflation. All models are appropriate and adequate. However, the SARIMA model is the most appropriate for predicting youth unemployment in Japan. On the other hand, there is a link between youth unemployment and changes in GDP and between unemployment and inflation. Any impulse coming from GDP changes or inflation affects the youth unemployment rate in Japan. The study also found that short-term forecasts are better than long-term ones, and further research is needed to improve model prediction for long-term youth unemployment.

Key words : Youth unemployment rate; Modeling; Forecasting, Time series

1 Įvadas

Jaunimas laikomas visuomenės atrama. Jų judrumas, inovatyvumas, veržlumas ir galia veda visą visuomenę į priekį. Jauni žmonės yra labiau atviri, produktyvūs ir besikeičiantys, kurių potencialas kurti ateities ekonominę gerovę yra didžiulis. Atskirti šiuos žmones nuo ekonomikos, visuomenės ir darbo rinkos yra neįmanoma. Tačiau vis dėlto šiuo metu jauni žmonės susiduria su didžiulio nedarbo problema. Todėl jaunimo nedarbo lygio aktualumui augant, tampa vis prasmingiau kalbėti, kodėl Japonija išlaiko vieną žemiausių jaunuolių nedarbo lygio reitingus.

Šio tyrimo tikslas yra ištirti Japonijos jaunimo nedarbo lygio laiko eilutę, jai sumodeliuoti atitinkamus ekonometrinius modelius pasitelkiant makroekonominis veiksniai bei palyginti šiuos modelius, sudarant Japonijos nedarbo lygio laiko eilutės prognozę. Tyrimui išsikėliau tokius uždavinius:

- Patikrinti nagrinėjamų laiko eilučių stacionarumą.
- Sumodeliuoti geriausią vienamati modelį jaunimo nedarbo lygiui.
- Sumodeliuoti daugiamačių laiko eilučių modelius.
- Atlikti liekanų analizę.
- Ištirti sąryšį tarp jaunimo nedarbo lygio ir makroekonominių veiksnių.
- Sudaryti modelių prognozes ir jas palyginti.

Šį darbą sudaro dvi dalys: teorinė ir praktinė. Teorinėje dalyje aprašomas temos ekonominis kontekstas, duomenys ir tyrimo metodologija. Ekonominiame kontekste pristatomos svarbiausios sąvokos, aprašoma jaunimo nedarbo lygio situacija pasaulyje bei daroma įtaka ekonominiam ir socialiniam gyvenimui. Tuomet pereinama prie jaunimo nedarbo lygio Japonijoje bei literatūros apžvalgos, kurioje pristatomi kiti tyrimai, nagrinėjantys nedarbo lygio laiko eilutės bei joms sudaryti modeliai. Duomenų ir metodologijos skyrelyje yra aprašomi tyrime naudojami duomenys bei metodologija.

Praktinė dalis susideda iš pirminės statistinės duomenų analizės, kurioje pateikiama aprašomoji analizė bei stacionarumo testo rezultatai. Taip pat aprašomos ekonometrinių modelių specifikacijos ir jų prognozių rezultatai. Darbo pabaigoje pateiktos išvados ir rekomendacijos, priedai bei literatūros sąrašas.

2 Ekonominis kontekstas

2.1 Nedarbo lygio ir užimtumo sąvokos

Nedarbo sąvoka yra glaudžiai susijusi su jaunimo nedarbo sąvoka. Pagal J.Tartilą "Plačiąja prasme į nedarbą galima pažvelgti kaip į vieną svarbiausių neefektyvaus gamybos išteklių (visų pirma darbo jėgos) panaudojimo priežasčių." [12] Todėl nedarbą galima laikyti kaip gamybos išteklių švaistymą. Tuomet jaunimo nedarbo sąvoką galime suprasti intuityviai - gamybos išteklių švaistymą, kurių savininkai yra jauni žmonės. Šiame tyrime jaunimas laikomas nuo penkiolikos iki dvidešimt keturių metų.

Visas nedarbo lygis gali būti išreiškiamas formule:

$$\text{nedarbas} = \frac{\text{bedarbių skaičius}}{\text{visa darbo jėga}} \times 100\%.$$

Tuomet, jaunimo nedarbui apskaičiuojant yra atrenkami jauni žmonės, ne studentai ir mokiniai, neturintys darbo ir padalinami iš visos jaunimo darbo jėgos.

Užimtumas yra priešinga sąvoka nedarbui. Šis rodiklis taip gali būti apskaičiuojamas formule:

$$\text{užimtumas} = \frac{\text{dirbantys asmenys}}{\text{visa darbo jėga}} \times 100\%.$$

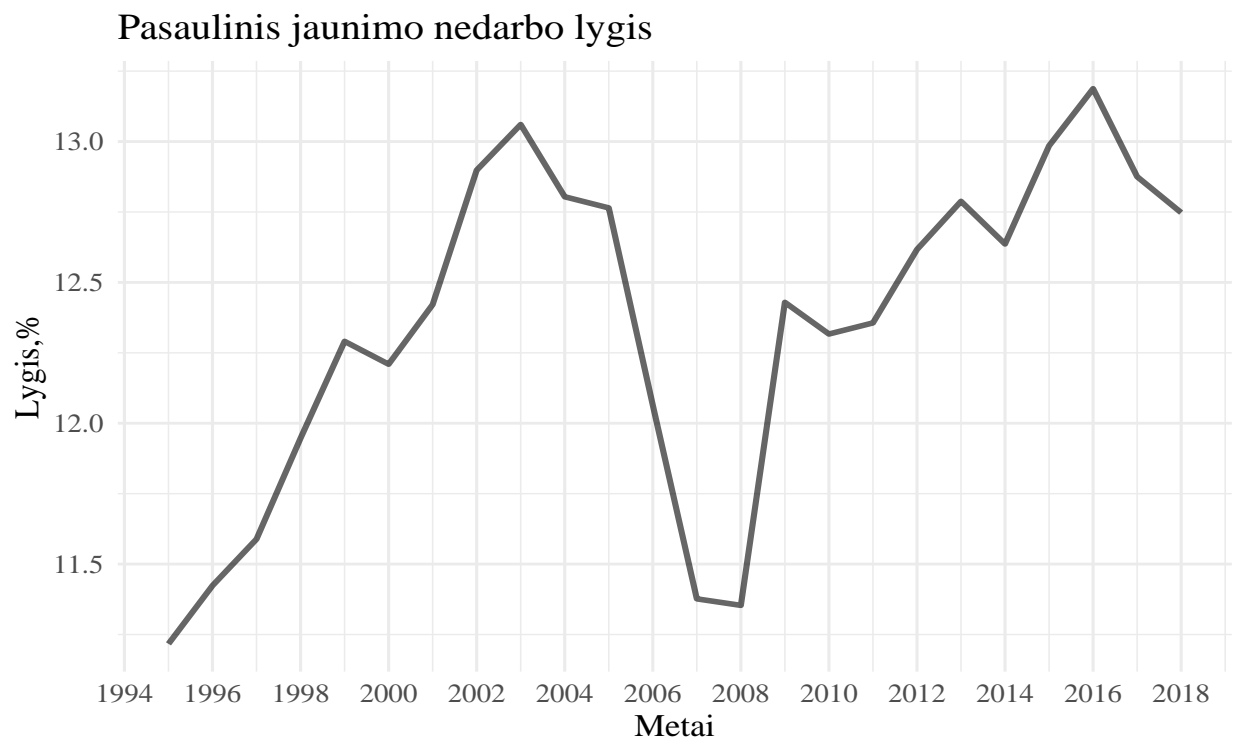
Visuomenėje yra neįmanomas visiškas užimtumas, tačiau kyla klausimas – kokio dydžio nedarbo lygis yra optimalus. Natūralus nedarbo lygis yra toks nedarbas, kuris susidaro pasiekus potencialųjį bendrąjį nacionalinį produktą arba, kai numatomi infliacijos tempai sutampa su faktiniais infliacijos tempais. Jungtinių Amerikos Valstijų Federalinė rezervų sistema teigia, kad natūralus nedarbo lygis yra 4,5%-5%.

2.2 Jaunimo nedarbo lygis pasaulyje

Jaunimo nedarbas - globali problema. Tarptautinė darbo organizacija teigia, kad jaunuolių nuo 15 iki 24 metų nedarbo lygis pasaulyje yra tris kartus didesnis nei suaugusiųjų. 2018 metais jaunimo nedarbo lygis siekė 13%, kuomet suaugusiųjų - 4.3% [4]. Pastaraisiais dešimt metų žemiausias jaunimo nedarbo lygis buvo per ekonominę pasaulinę krizę (11.3%), tačiau ekonomikai augant - jaunimo nedarbo lygis taip pat didėja (žr.1 pav.).

Labiausiai šis rodiklis kilo Pietų Afrikoje, arabų šalyse ir Lotynų Amerikoje. Nors daugiausia jaunų žmonių gyvena besivystančiose šalyse, Europoje taip pat stebimas drastiškas jaunimo užimtumo kritimas. Graikija(17,2%), Ispanija(13.9%) ir Italija(9,9%) patiria jaunimo nedarbo krizę.

Šiek tiek kitokia situacija yra Azijoje. Čia jaunimo bedarbių skaičius išlieka gana stabilus. Tačiau Pietų Azijoje ir Pietryčių Azijoje yra aukštesnis jaunimo nedarbo lygis palyginus su Rytų Azija.



1 pav.: Nedarbo lygis pasaulyje

2.3 Jaunimo nedarbo lygio įtaka

Nors yra žinoma, kad ekonomikos nuosmukis gali iššaukti nedarbo didėjimą, tačiau taip pat ir pats nedarbas gali atvirkščiai neigiamai veikti ekonomiką [5]. Bedarbiai nebemoka tam tikrų mokesčių, kuriuos asmenys mokėdavo su gautu atlyginimu - gyventojų pajamų, privalomojo sveikatos ir socialinio draudimo mokesčius. Tai reiškia valstybės biudžeto mažėjimą, skolos augimą ir tai dar labiau skatina ekonomikos lėtėjimą. Taip pat dėl mažėjančių pajamų, bedarbiai vis mažiau išleidžia pinigų - taupydami, jie dar labiau stabdo ekonomiką. Jaunimo nedarbas lygiai taip pat neigiamai veikia ekonominį gyvenimą - vis mažiau pinigų yra išleidžiamas, mažesnis kiekis pinigų išleidžiamas į ekonomiką, todėl taip netgi yra skatinamas ekonomikos lėtėjimas.

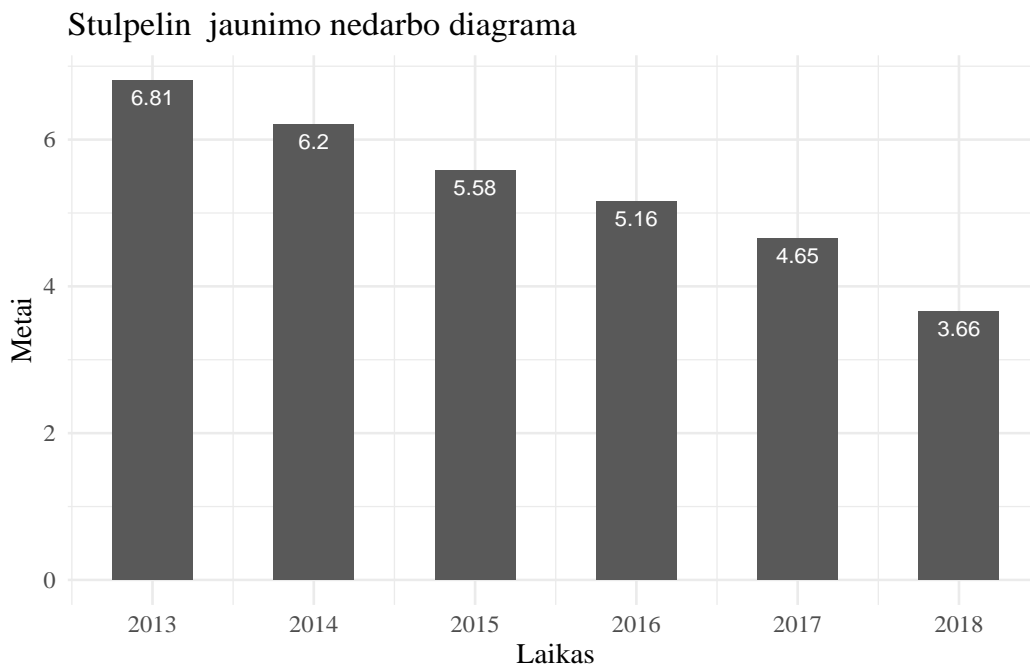
Tačiau jaunimo nedarbas labiausiai veikia socialinį žmonių gyvenimą. A.Hammarström nustatė, kad jaunimo nedarbas neigiamai veikia jaunuolių sveikatą, ypačiai pažeidžiama tampa jų psichinė būklė [3]. Nedarbantis jaunimas daro įtaką ir kriminaliniam pasauliui - auga narkotikų vartojimas, apiplėšimų skaičius, didėja gaujos. Taip pat yra nustatyta, kad jaunimo nedarbas labai stipriai veikia asmeninį gyvenimą: didesnis šansas, kad vėlesnėse stadijose asmuo bus taip pat nedarbantis, jo alga bus mažesnė, truks įvairių įgūdžių, kamuos didelis nepasitenkinimas gyvenimu [8].

2.4 Japonijos jaunimo nedarbo lygis

Japonija viena iš šalių, kurioje jaunuolių užimtumas yra didžiausias, tai yra nedarbo lygis siekia tik 3,69%. Lyginant su visu užimtumo lygiu šalyje - Japonijoje nedarbas siekia 2,5%. Todėl ir čia pasireiškia suaugusiųjų didesnis užimtumo lygis, kaip ir visame pasaulyje. Tačiau lyginant su viso pasaulio duomenimis, Japoniją galima laikyti mažo jaunimo nedarbo fenomenu.

Kita vertus, Japonija ne visada turėjo aukštą jaunimo darbo lygį. 1990 metais įvykusi Japonijos krizė, stipriai paveikė šalies jaunuolių gyvenimą. Greitai senėjančiai japonų visuomenei tuo metu negrėse nedarbas, tačiau jaunimui drastiškai sumažėjo darbo galimybės [1].

Vidutinis jaunimo nedarbo lygis Japonijoje nuo 1968 iki 2018 metų siekia 5.61%. Šiame laiko intervale aukščiausio nedarbo lygį pasiekta 2010 metais birželio mėnesį - 10,9%, o mažiausias 1968 metų lapkritį - 1,2%. Diagramoje (žr. 2 pav.) matoma kaip keičiasi jaunuolių, nuo 15 iki 24 metų nedarbo lygis pastaraisiais šešiais metais (2013-2018). Grafikas atskleidžia, kad nuo 2013 metų Japonijoje jaunuolių nedarbo lygis mažėjo, pasiekdamas 3,66 procentus 2018 metais, tai yra kiekvienais metais krisdavo maždaug po 1,2%.



2 pav.: Japonijos nedarbo lygis 2013-2018

2.5 Literatūros apžvalga

Makroekonominį reiškinį modeliavimas ir prognozavimas padeda lengviau įvertinti problemas, kylančias iš šalių ekonominės padėties. Jaunimo nedarbas yra vienas iš makroekonominį kintamųjų, kuris yra tyrėjų modeliuojamas ir prognozuojamas tam, kad geriau įvertinti situaciją ne tik konkrečioje šalyje, bet ir regionuose (pvz. Europos Sąjunga). Šis situacijos įvertinimas padeda valstybėms geriau įgyvendinti tam tikrus ekonominius sprendimus, pavyzdžiui padėti jaunimui lengviau pereiti iš švietimo į darbo rinką. Kaip ir kiti makroekonominiai kintamieji, jaunimo nedarbo lygis yra analizuojamas pasitelkiant įvairius ekonometrinius modelius.

Vienas iš ekonometrinių modelių, plačiai naudojamas prognozuojant makroekonominis duomenis yra autoregresinis integruotas slenkamųjų vidurkių procesas (ARIMA). Autoriai M. González-Fernández ir C. González-Velasco naudoja būtent ARIMA(0,1,1) Italijos jaunimo nedarbo lygiui prognozuoti [2]. Korėjos jaunimo nedarbo lygiui prognozuoti buvo pasirinkta SARIMA(1,0,0)(0,1,0)₁₂ [6].

Vektorinės autoregresijos modelis (VAR) taip pat labai naudingas modeliuojant ir prognozuojant makroekonominis duomenis, jį naudoti pasiūlė C. Sims [10]. Šio modelio privalumas yra tas, kad jį įvertinti galima mažiausių kvadratų metodu (MKM). Nors konkrečios šalies jaunimo nedarbas nėra plačiai nagrinėjamas VAR modeliu, tačiau yra atlikta įvairių tyrimų, kur VAR modelis pritaikomas nagrinėjant visą nedarbo lygį. Pavyzdžiui, M. Simionescu panaudojo būtent šį modelį Rumunijos nedarbui įvertinti kartu su infliacijos duomenimis [9]. H. Ohlsson vertindamas darbo rinkos politiką su nedarbu ir atlyginimais Švedijoje, naudodamas vektorinės autoregresijos modelį nustatė, kad vykstant nedarbo lygio šokams, daugėja darbo jėgos kūrimo programų [7]. Tačiau VAR yra sudėtinga naudoti tada, kuomet norime įvertinti efektą, padarytą naujos politinės taisyklės [11]. Šiame darbe bus taip pat naudojama ARIMA ir VAR modeliai įvertinti jaunimo nedarbo lygiui.

3 Duomenys ir metodologija

3.1 Duomenys

Tyrimui naudojami Japonijos jaunimo nedarbo lygio ketvirtiniai duomenys nuo 1968 metų pirmojo iki 2018 ketvirtojo ketvirčio, surinkti iš OECD. Viso 204 stebėjimai, iš kurių 180 bus naudojami modelių sudarymui, o likusieji 24 stebėjimų modelio prognozavimui.

Kadangi tyrimas susideda tiek iš vienmatės, tiek daugiamatės laiko eilučių modelių sudarymo, bus naudojama papildomai kiti ekonominiai kintamieji. Ekonominėje teorijoje yra įrodytas sąryšis tarp bendrojo vidaus produkto ir nedarbo lygio: jeigu krenta BVP - kyla nedarbo lygis ir atvirkščiai. Taip pat yra netiesioginis sąryšis tarp infliacijos ir nedarbo lygio - šis sąryšis nusakomas Filipso kreive - augant infliacijai, krenta nedarbo lygis ir atvirkščiai. Kadangi yra daroma prielaida, kad bendrojo vidaus produkto pokyčiai daro įtaką nedarbo lygiui, darbe taip pat bus naudojami sezonškai pakoreguoti procentiniai Japonijos praėjusio periodo BVP pokyčiai bei infliacijos ketvirtiniai duomenys nuo 1968 metų iki 2018 iš OECD.

3.2 Metodologija

Laiko eilutė gali būti apibrėžta kaip "to paties kintamojo reikšmės skirtingais laikotarpiais" ("Ekonomikos terminų žodynas"). Stebint tik vieno kintamojo reikšmių kitimą, yra gaunama vienmatė laiko eilutė, o į daugialypę laiko eilutę yra įtraukiama daugiau stebimo objekto rodiklių reikšmių, todėl daugialypė laiko eilutė yra labiau naudojama, kuomet norima paaiškinti efektus ir sąryšį tarp kintamųjų.

3.2.1 Stacionarumo tikrinimas

Prieš sudarant modelį laiko eilutėms, reikia patikrinti laiko eilutės stacionarumą. Stacionarumas atsiranda tuomet, kai laiko eilutės vidurkis ir autokovariacija išlieka tokia pati, tai yra šios charakteristikos laike nekinta. Tai reiškia, kad bet koks bendrasis laiko eilučių kintamųjų bendras statistinis skirstinys nepriklauso nuo laiko. Todėl laiko eilutę $\{Y_t, t \in \mathbb{Z}\}$ vadinama stacionaria jeigu:

1. $\forall t \in \mathbb{Z} \quad EY_t^2 < \infty,$
2. $\forall t \in \mathbb{Z} \quad EY_t = EY_0,$
3. $Cov(s, t) = Cov(s + h, t + h), \quad \forall s, t, h \in \mathbb{Z}.$

Praplėstasis Dikio-Fulerio (ADF) testas

Praplėstasis (angl. augmented) Dikio-Fulerio testas (ADF), vienas iš populiariausių testų stacionarumo tikrinimui. Jis skirtas vienetinių šaknų egzistavimo nustatymui. Jeigu laiko eilutė Y_t neturi vienetinės šaknies, sakoma, kad ji - stacionari.

ADF testas atliekamas modeliui:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \dots + \delta_p \Delta Y_{t-p} + \varepsilon_t,$$

čia

- α - konstanta,
- β tendencijos koeficientas,
- p - ankstinių skaičius.

Hipotezės gali būti užrašomos taip:

$$\begin{cases} H_0 : & \gamma = 0 \text{ (laiko eilutė turi vienetinę šaknį)} \\ H_1 : & \gamma < 0. \end{cases}$$

Testo statistika

$$DF_t = \frac{\hat{\gamma}}{SE(\hat{\gamma})},$$

čia SE - standartinė koeficiento paklaida. Testo statistika lyginama su atitinkama Dikio-Fulerio testo, taikomo autoregresiniam modeliui, kritine reikšme.

3.2.2 Vienmačių laiko eilučių modelis

Sezoninio autoregresinio integruoto slenkamųjų vidurkių proceso modelis (SARIMA)

Sezoninio autoregresinio integruoto slenkamųjų vidurkių proceso modelis (SARIMA) yra bendresnė ARIMA modelio forma, kuri įtraukia tiek sezoniškumo, tiek nesezoniškumo charakteristikas. Panašiai kaip ARIMA modelyje, prognozių reikšmės yra laikomos tiesine praeities reikšmių ir praeities paklaidų kombinacija. SARIMA modelis žymimas kaip $ARIMA(p, d, q)(P, D, Q)_S$. Modelio forma yra:

$$\phi(L)\Phi(L^S)(1-L)^d(1-L^S)^D Y_t = \theta(L)\Theta(L^S)\epsilon_t.$$

čia

- Y_t - laiko eilutė;
- ϵ_t - baltojo triukšmo paklaidų išraiška;
- p, d, q - nesezoninio AR modelio eilė, integracijos laipsnis ir nesezoninio MA modelio eilė atitinkamai;
- P, D, Q - sezoninio AR modelio eilė, integracijos laipsnis, ir sezoninio MA modelio eilė;
- S - sezoniškumo eilė, šiuo atveju $S=4$ ketvirtiniams duomenims;
- L - ankstinio operatorius $L^k Y_t = Y_{t-k}$.

Modelyje yra įtraukti atitinkami eilių p ir q AR ir MA procesai:

$$\begin{aligned} \phi(L) &= 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_{p-1} L^{p-1} - \phi_p L^p, \\ \theta(L) &= 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_{q-1} L^{q-1} + \theta_q L^q. \end{aligned}$$

Taip pat sezoninės eilių P ir Q funkcijos:

$$\begin{aligned} \Phi(L^S) &= 1 - \Phi_1 L^S - \Phi_2 L^{2S} - \dots - \Phi_{P-1} L^{(P-1)S} - \Phi_P L^{PS}, \\ \Theta(L^S) &= 1 + \Theta_1 L^S + \Theta_2 L^{2S} + \dots + \Theta_{Q-1} L^{(Q-1)S} + \Theta_Q L^{QS}. \end{aligned}$$

SARIMA eilės pasirinkimas

Sezoninio ir nesezoninio autoregresinio ir slenkamojo vidurkio ankstinių komponentų p , P , q ir Q yra nustatomi remiantis autokoreliacijos (ACF) ir dalinės autokoreliacijos grafikais (PACF). Šie grafikai suteikia informacijos apie vidinę koreliaciją tarp laiko eilučių stebėjimų skirtingu laiku tam, kad būtų lengviau susiorientuoti tarp sezoninių ir nesezoninių ankstinių. Tiek ACF, tiek PACF turi išsišovimus (angliškai *spikes*) ir nukirtimus (angliškai *cut off*) ankstiniuose k ir k_S . Modelio eilė yra nustatoma pagal statistiškai reikšmingų išsišovimų skaičių.

Geriausios SARIMA eilės pasirinkimas

Pagal ACF ir PACF grafikus galimos įvairios SARIMA(p, d, q)(P, D, Q) $_S$ eilės, su skirtingais p , P , q ir Q reikšmingais ankstiniais. Dėl to, reikia atrinkti optimalų sezoninio ir nesezoninio komponentų ankstinių ilgį naudojantis tam tikru kriterijumi. Tyrimė naudojamas Akaike informacijos kriterijus (angl. *Akaike information criterion* (AIC)), kuris suskaičiuojamas taip:

$$AIC = -2 \ln L + 2m,$$

kur $m = p + P + q + Q$, o L - tikėtinas. Geriausias SARIMA(p, d, q)(P, D, Q) $_S$ modelis yra laikomas su mažiausia AIC reikšme.

3.2.3 Daugiamačių laiko eilučių modelis

Vektorinės autoregresijos proceso modeliavimas

VAR modelis yra vienas lanksčiausių, lengviausiai naudojamų modelių daugiamačių laiko eilučių analizei. Jis yra naudingas greitai besikeičiančioms ekonominių ir finansinių laiko eilučių aprašymui ir prognozavimui. Šių modelių prognozės yra gana lanksčios, nes jas galima susieti su tam tikrais modelių kintamųjų ateities galimybėmis. Vektorinės autoregresijos modelio eilės p , VAR(p) modelis gali būti užrašomas taip:

$$\mathbf{Y}_t = c + \pi_1 \mathbf{Y}_{t-1} + \pi_2 \mathbf{Y}_{t-2} + \dots + \pi_p \mathbf{Y}_{t-p} + \epsilon_t,$$

čia

- c - k vektorių stulpelis, susidedantis iš konstantų,
- $t = 1, \dots, T$ ir \mathbf{Y}_t yra procesas lygus $(k \times 1)$ vektorių stulpeliui $(Y_{1t}, Y_{2t}, \dots, Y_{kt})'$,
- π_i yra $(k \times k)$ koeficientų matrica,
- ϵ_t yra $(k \times 1)$ baltojo triukšmo procesas, kur $E\epsilon_t = 0$.

VAR eilės parinkimas

Kaip ir kiti laiko eilučių modeliai, taip ir VAR modelis turi turėtų geriausią ankstinio ilgį. Bendru atveju, reikia pasirinkti tokias eiles $m = 0, \dots, pmax$, kurios minimizuotų tam tikrus modelio pasirinkimo kriterijus. Modelio pasirinkimo kriterijus turi bendrinę išraišką:

$$C(m) = \ln \left| \sum_m \hat{\epsilon}_t \hat{\epsilon}_t' \right| + c_T \cdot \varphi(m, k),$$

kur

- $\sum_m = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t \hat{\epsilon}_t'$ - liekanų įverčių covariacijų matrica eilės m modeliui,
- $\varphi(m, k)$ yra eilės m funkcija, kuri apriboja aukštos VAR modelio eiles,
- c_T yra seka, priklausanti nuo imties dydžio ir nustato konkretų kriterijų.

$\ln \left| \hat{\Sigma}_m \right|$ yra nedidėjanti eilės m funkcija, kuomet $\varphi(m, k)$ didėja kartu su m . Tinkamo ankstinio pasirinkimas turi būti toks, kad šias abi funkcijas optimizuotų.

Trys dažniausiai naudojami informacijos kriterijai yra Akaike, Schwarz (SC) ir Hannan-Quin (HQ) informacijos kriterijai:

$$\begin{aligned} AIC &= \ln \left| \hat{\Sigma}_m \right| + \frac{2}{T}mk^2, \\ SC(m) &= \ln \left| \hat{\Sigma}_m \right| + \frac{\ln T}{T}mk^2, \\ SC(m) &= \ln \left| \hat{\Sigma}_m \right| + \frac{2 \ln(\ln T)}{T}mk^2. \end{aligned}$$

Visais atvejais $\varphi(m, k) = mk^2$ yra VAR parametų skaičius modelyje su eile m ir kintamųjų skaičiumi k .

Kointegracija

Kointegracijos vektorius yra žinomas kaip ilgalaikiu pusiausvyros santykiu tarp VAR modelio kintamųjų. Kuomet egzistuoja reikšmingas kointegracijos vektorius, VAR modelis turėtų būti praplėstas paklaidų korekcija. Kitaip tariant, grynas VAR modelis gali būti naudojamas tuomet, kai kointegracijos sąryšio tarp kintamųjų nėra. Prieš sudarant VAR modelį, būtina atlikti kointegracijos testą.

Kointegracija tikslas yra susieti integruotas nestacionarias laiko eilutes su ilgalaikiu sąryšiu. Kointegracijos lygties egzistavimas yra interpretuojamas kaip ilgalaikio ryšio pusiausvyros taškas tarp kintamųjų. Jeigu yra k laiko eilučių, kurių integruotumo eilė yra pirma - $I(1)$ - taip pat gali egzistuoti iki $k - 1$ nepriklausomų tiesinių sąryšių, kurie yra $I(0)$.

Johansen testas kointegracijai patikrinti

Tegul r yra π rangas. Čia π vektorių α ir kointegravimo parametų β matrica, $\pi = \alpha\beta$. H_0 : sistemoje nėra nei vieno kointegravimo vektoriaus ($r = 0$). Jei ši hipotezė atmetama, tikrinama hipotezė, kad sistemoje yra ne daugiau vieno kointegravimo vektoriaus ($r \leq 1$). Taip tikrinama tol, kol nulinės hipotezės atmesti negalima.

Vektorinis paklaidų korekcijos modelis

Vektorinis paklaidų korekcijos modelis (VECM) yra naudojamas tuomet, kai laiko eilutės yra nestacionarios ir tarp jų egzistuoja bent vienas ar daugiau ilgojo laikotarpio sąryšis. VECM yra vienas iš modelių, kuris gali tiesiogiai nustatyti lygį po kurio kintamasis gali atsistatyti į pusiausvyrą po tam tikro šoko iš kitų kintamųjų. Taip pat jis yra naudingas nustatyti trumpojo ir ilgojo laikotarpio efektus kiekvienam kintamajam. VECM(p) modelis su kointegracijos rangu $r \leq k$ apibūdinamas taip:

$$\Delta \mathbf{Y}_t = c + \Pi \mathbf{Y}_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta \mathbf{Y}_{t-1} + \epsilon_t,$$

čia

- $\Delta \mathbf{Y}_t = \mathbf{Y}_t - \mathbf{Y}_{t-1}$, kintamųjų skirtumų vektorius,
- c - konstantų vektorius,

- Π - kointegracijos matricos koeficientas ($\Pi = \alpha\beta'$, kur α - paklaidų korekcijos komponentės koeficientas, β - kointegracijos vektorius),
- Γ_i - matrica, kurios koeficientas yra $(k \times k)$,
- ϵ_t - paklaidų vektorius.

3.2.4 Modelio tinkamumo tikrinimas

Modelio adekvatumo vertinimas yra svarbi laiko eilučių analizės dalis, skirta liekanų analizei. Sudaryto modelio liekanos turėtų turėti baltojo triukšmo savybės: nekoreliuotos (nėra likusios jokios informacijos, kuri neįtraukta į modelį), su vidurkiu nuliu.

Liekanų analizė

Kuomet liekanos nėra baltasis triukšmas, apskaičiuoti skirtumai tarp parametrų tampa ne-nuoseklūs ir nepatikimi, taip atlikti tam tikri testai tampa negalimi modelių įvertinimui. Be to, modelių prognozės gali būti neefektyvios dėl didelių prognozių paklaidų skirtumų. Dėl to, labai svarbu atlikti liekanų analizę prieš naudojant atitinkamą modelį.

Liekanų autokoreliacijos testas

Liekanų autokoreliacijai nustatyti galima tiek grafiniu būdu, tiek statistiniu testu. Tačiau dėl grafinio būdo neformalumo ir netikslumo, liekanų autokoreliacijai nustatyti bus naudojami statistiniai testai.

Breusch-Godfrey testas

Tegul VAR modelio liekanos u_t yra išreiškiamos forma:

$$u_t = D_1 u_{t-1} + \dots + D_h u_{t-h} + v_t.$$

v_t aprašo baltojo triukšmo procesą. Dėl to, hipotezės autokoreliacijos testui u_t yra:

$$\begin{cases} H_0 : D_1 = \dots = D_h = 0 & (\text{autokoreliacijos nėra}) \\ H_1 : D_j \neq 0 & \text{bent vienam } j < h \quad (\text{autokoreliacija yra}). \end{cases}$$

Testo statistika yra apskaičiuojama:

$$BG = \frac{n - p - k - 1}{p} \cdot \frac{R^2}{1 - R^2} \sim F(p, n - p - k - 1),$$

čia R^2 - determinacijos koeficientas, k - kintamųjų skaičius.

3.2.5 Prognozavimas

Prognozavimas yra paprasčiausias ateities spėjimo būdas remiantis praeities ir dabartiniais duomenimis bei tendencijų analize.

Prognozavimo atlikimo metodas

Modeliai turėtų būti atrenkami ne tik pagal tai, kaip jie puikiai atitinka duomenis, bet ir pagal tai, kaip tiksliai modelio prognozės atitinka tikrąsias reikšmes. Kadangi vienas iš tyrimo tikslų yra modelių palyginimas sudarant prognozę, geriausias modelis bus atrinktas pagal tikrųjų ir prognozuotų reikšmių atitikimą. Tačiau kartais atsitinka taip, kad modelis yra geriausias tik toms reikšmėms, kurioms yra sudaromas modelis, bet prognozuojamoms reikšmėms sudarytas modelis visai netinka. Dėl to, kad išvengti šios problemos, tyrimui naudojamas prognozavimo būdas, kur

turimi duomenys atskirti į dvi dalis: didesnioji dalis naudojama modelių sudarymui, o likusioji, mažesnioji dalis, prognozavimui. Todėl, modelis, kurio prognozės labiausiai atitiks tikrąsias reikšmes, bus laikomas geriausiu. Todėl duomenys nuo 1968Q1 iki 2012Q4 (180 stebėjimų) bus naudojami modelio sudarymui, o likusieji nuo 2013Q1 - 2018Q4 prognozavimui.

Prognozavimo tikslumas

Prognozavimo tikslumas yra kriterijus, skirtas įvertinti modelių prognozavimo efektyvumą. Jis nustato kaip gerai modelis gali nustatyti ateities reikšmes. Tyrime naudojami dvi paklaidų statistikos: vidutinė absoliuti paklaida (MAE) ir vidutinis kvadratinis nuokrypis (RMSE). Kuo mažesnės šios reikšmės, tuo tikslesnės prognozės.

Tarkime $Y_{\frac{t+h}{t}}$ yra h žingsnių Y_{t+h} prognozė, o $\varepsilon_{\frac{t+h}{t}} = Y_{t+h} - Y_{\frac{t+h}{t}}$ - atitinkama prognozės paklaida. Tuomet prognozės įvertinimo statistikos pagal N h -žingsnių gali būti apibrėžtos taip:

$$MAE = \frac{1}{N} \sum_{j=t+1}^{t+N} \left| \varepsilon_{\frac{j+h}{j}} \right|,$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=t+1}^{t+N} \left(\varepsilon_{\frac{j+h}{j}} \right)^2}.$$

Vidutinis kvadratinis paklaidos nuokrypis (RMSE) yra naudojamas matuoti laiko eilučių skirtumą tarp numatytų ir tikrųjų reikšmių. Tai prognozių galios matas. Vidutinė absoliuti paklaida yra skirta išmatuoti kiek prognozės yra artimos tikrosioms reikšmėms. Dėl to, tas modelis, kurio RMSE ir MAE reikšmės yra mažiausios, yra laikomas geriausiu prognozių modeliu.

4 Praktinė dalis

4.1 Pirminė statistinė duomenų analizė

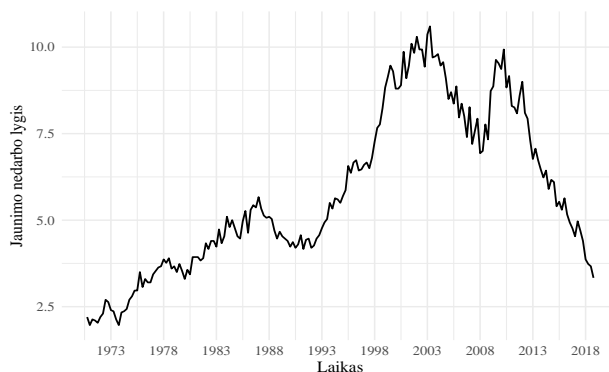
4.1.1 Aprašomoji analizė

Aprašomoji statistika, kurioje yra vidurkis, standartinis nuokrypis, minimali ir maksimali reikšmė, aprašyta 1-oje lentelėje. Matoma, kad visos reikšmės yra gana skirtingos.

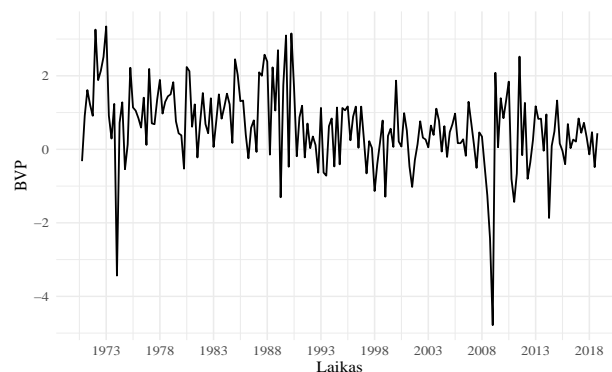
Kintamasis	Stebėjimai	Vidurkis	Standartinis nuokrypis	Min.	Max.
Jaunimo nedarbo lygis	204	5,575	2.436	1.333	10.6
Procentiniai praėjusio periodo BVP pokyčiai	204	0.723	1.199	-4.783	5.698
Infliacija	204	0.656	1.234	1.249	9.302

1 lentelė: Aprašomoji kintamųjų statistika

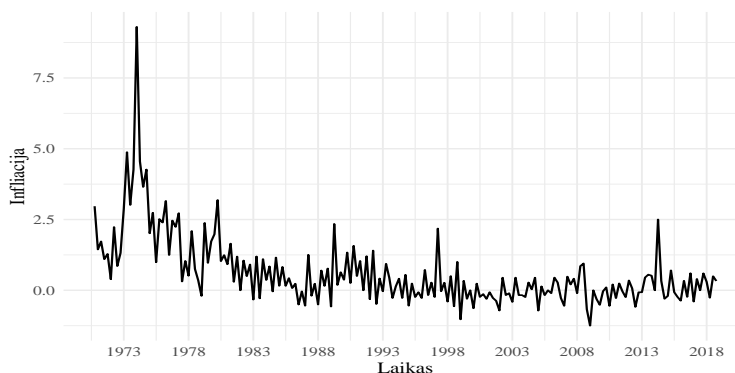
Jaunimo nedarbo lygio sklaida apie vidurkį yra didesnė negu BVP pokyčių ar infliacijos.



(a) Jaunimo nedarbo lygis Japonijoje



(b) Bendrojo vidaus produkto pokyčiai



(c) Infliacija Japonijoje

3 pav.: Laiko eilutės

Iš Japonijos jaunimo nedarbo lygio grafiko galima pastebėti (3a pav.), kad ilgą laiką nuo 1968 iki 2003 metų, jaunimo nedarbo lygis augo. Po 2003-ųjų - nedidelis kritimas, bet 2009-ųjų metų ekonominė krizė vėl neigiamai atsiliepė šiam rodikliui, aukštas jaunimo nedarbas laikėsi maždaug iki 2010-tųjų metų. Šiuo metu - Japonijos nedarbo lygio rodiklis - žemumose.

BVP pokyčių grafikas (3b pav.) atskleidžia, kad BVP pokyčiai yra išsidėstę apie vidurkį, su didesniais nuokrypiais 1974-tais metais, kuomet Japonija išgyveno ekonomikos nuopolį ir per pasaulinę ekonominę krizę 2009-ais metais.

Inflacijos grafikas (3c pav.) atskleidžia, kad 1974-tųjų metų Japonijos ekonomikos krizė neigiamai paveikė ir infliacijos lygį - jis čia aukščiausias visoje 1968-2018 metų istorijoje. Tačiau po jų -

infliacija laikosi daugmaž tame pačiame lygyje su keliais didesniais nuokrypiais.

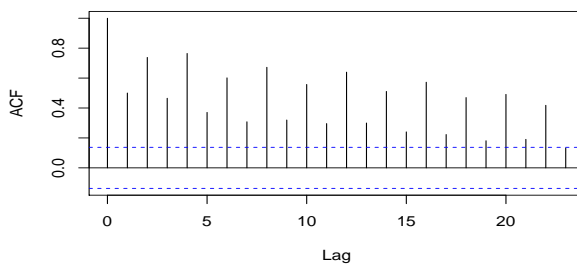
4.1.2 Stacionarumo testo analizė

Prieš bandant sudaryti tinkamą modelį, duomenims reikia atlikti stacionarumo savybės testą. Tai galima padaryti tiek vizualiai, tiek taikant praplėstąjį (angl. *augmented*) Dikio-Fulerio testą - ADF. 3 paveikslėlyje yra pavaizduoti visų ekonominių kintamųjų laiko eilučių grafikai. Šiame grafike pastebima, kad nestacionari yra jaunimo nedarbo laiko eilutė (3a), o infliacija (3c), atrodo, turi sezoninę komponentę. Bendrojo vidaus produkto pokyčių laiko eilutė (3b) vizualiai atrodo stacionarūs.

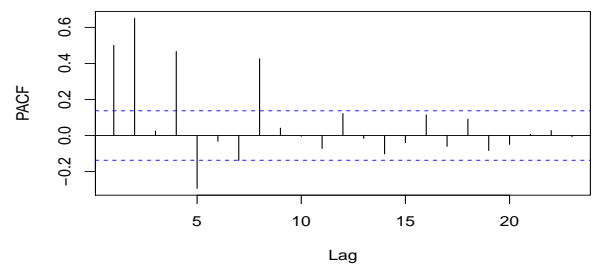
Atlikus ADF testą, rezultatai (2 lentelė) atskleidžia, kad Japonijos jaunimo nedarbo lygio laiko eilutė turi vienetinę šaknį, kadangi testo reikšmė didesnė už kritinę reikšmę. BVP pokyčiai pagal testo statistiką yra stacionarūs. Inflacijai taip pat atmetama vienetinės šaknies nulinė hipotezė su prielaida, kad laiko eilutė turi driftą su trendu. Tačiau autokoreliacijos ir dalinės autokoreliacijos (ACF ir PACF atitinkamai) grafikai atskleidžia, kad ji turi sezoninę komponentę (žr. 4a ir 4b), dėl šios priežasties, inflacijos laiko eilutės sezoniškumą reikia pašalinti. Sezoniškumo pašalinimui imame sezoninius inflacijos skirtumus. Tokį pat testą vėl atliekame tiek diferencijuotiems jaunimo nedarbo duomenims, tiek inflacijos sezoniniams skirtumams.

Kintamasis	Laiko eilutė su driftu		Laiko eilutė su driftu ir trendu	
	Testo statistika	p reikšmė	Testo statistika	p reikšmė
Jaunimo nedarbo lygis	-0.320	0.552	0.463	0.99
Procentiniai praėjusio periodo BVP pokyčiai	-4.74	0.01	-6.04	0.01
Infliacija	-2.96	0.043	-4.15	0.01
Kritinė 5% reikšmė	-2.89		-3.45	

2 lentelė: Vienetinės šaknies testo rezultatai



(a) Autokoreliacijos grafikas

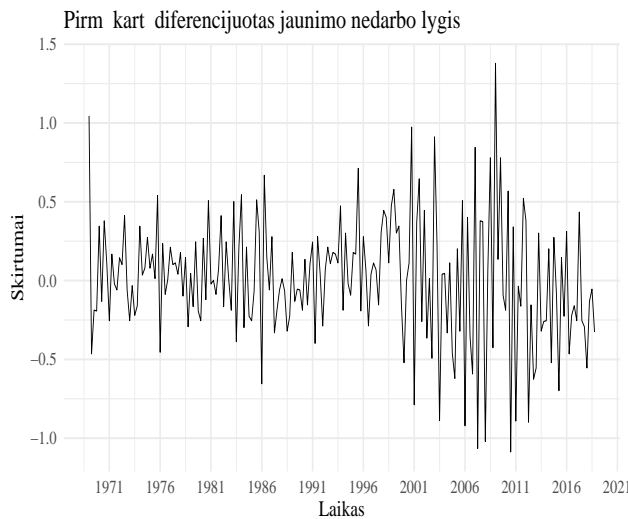


(b) Dalinės autokoreliacijos grafikas

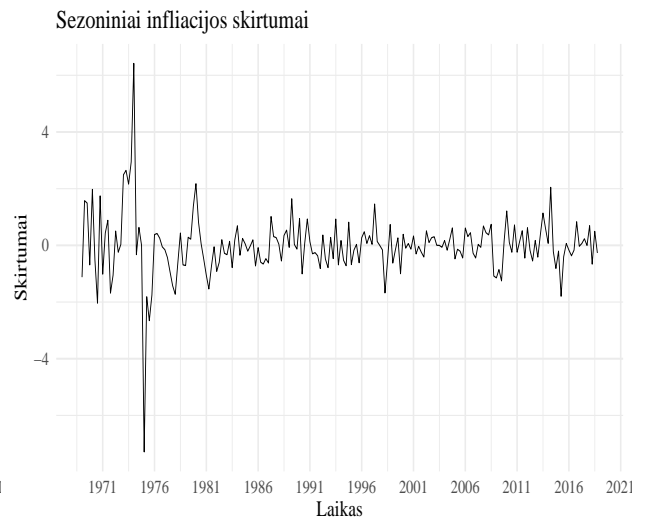
4 pav.: Inflacijos autokoreliacijos ir dalinės autokoreliacijos grafikai

Šį kartą vizualiai galima pamatyti, kad jaunimo nedarbo ir inflacijos skirtumų laiko eilutės atrodo stacionarios (5 pav.). Atliktas ADF testas parodo (3 lentelė), kad abiejų laiko eilučių skirtumai yra stacionarūs, nes kritinės reikšmės yra didesnės už testo statistiką, todėl hipotezę apie vienetinę šaknį yra atmetama. Tačiau 5a grafikas atskleidžia ir jaunimo nedarbo lygio sezoniškumą, todėl vietoje ARIMA modelio, bus sudaromas SARIMA modelis.

Taigi, Japonijos jaunimo nedarbo eilutės integruotumo eilė yra pirma, todėl jaunimo nedarbo lygis bus modeliuojamas pirmiesiems laiko eilutės skirtumams, naudojant sezoninio autoregresinio integruoto slenkamųjų vidurkių proceso (SARIMA) ir vektorinės autoregresijos (VAR) modelius.



(a) Diferencijuota Japonijos jaunimo nedarbo eilutė



(b) Diferencijuota infliacijos laiko eilutė

5 pav.: Laiko eilutės po diferencijavimo

Kintamasis	Laiko eilutė su driftu		Laiko eilutė su driftu ir trendu	
	Testo statistika	p reikšmė	Testo statistika	p reikšmė
Jaunimo nedarbo lygio skirtumai	-5.08	0.01	-5.51	0.01
Infliacijos skirtumai	-8.21	0.01	-8.19	0.01
Kritinė 5% reikšmė	-2.89		-3.45	

3 lentelė: Vienetinės šaknies šaknies rezultatai po diferencijavimo

4.2 Ekonometrinių modelių specifikacija

4.2.1 Sezoninio autoregresinio integruoto slenkamųjų vidurkių proceso modeliavimas

Pradiniams jaunimo nedarbo lygio duomenims yra nustatyta pirma laiko eilutės integruotumo eilė, t.y. $d = 1$. Žiūrint į laiko eilutės ACF ir PACF grafikus (6 pav.) matoma, kad pirmas išskirtinis ankstinys yra pirmas, todėl galimos MA(1) ir AR(1) nesezoninės komponentinės. Tačiau parinkti SARIMA modeliui eiles, žiūrint tik į ACF ir PACF grafikus yra nepakankama. Todėl remiantis AIC ir BIC kriterijais, pasirinktas SARIMA(1,1,0)(1,1,0)₄ modelis:

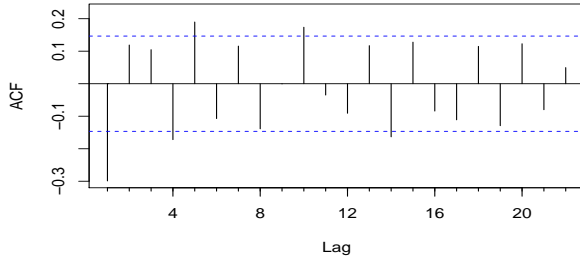
$$(1 - \phi L)(1 - L)(1 - \Phi L^4)(1 - L^4)Y_t = \epsilon_t,$$

kur

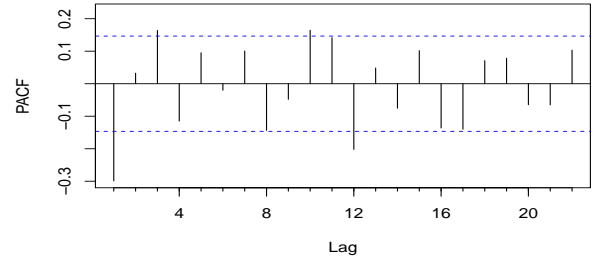
- $\hat{\phi} = -0.3592$,
- $\hat{\Phi} = -0.4784$,
- $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$, čia $\hat{\sigma}^2 = 0.2352$.

Norint įsitikinti, ar modelis yra adekvatus, reikia atlikti liekanų analizę. Modelis yra adekvatus tuomet, kai jo liekanos yra baltojo triukšmo procesas. Grafike jos yra išsidėsčiusios apie nulinę tiesę (7 paveikslėlis), kas indikuoja modelio tinkamumą.

Breusch-Godfrey testas atskleidžia liekanų nekoreliuotumą. SARIMA(1,1,0)(1,1,0)₄ modelio atveju p reikšmė = 0.8118 > 0.05, todėl hipotezė apie liekanų nekoreliuotumą neatmetama, t.y. liekanos yra nekoreliuotos. Liekanų vidurkis yra -0.008, artimas nuliui. Taigi, modelio liekanos yra baltasis triukšmas, o tai reiškia, kad modelis yra adekvatus.

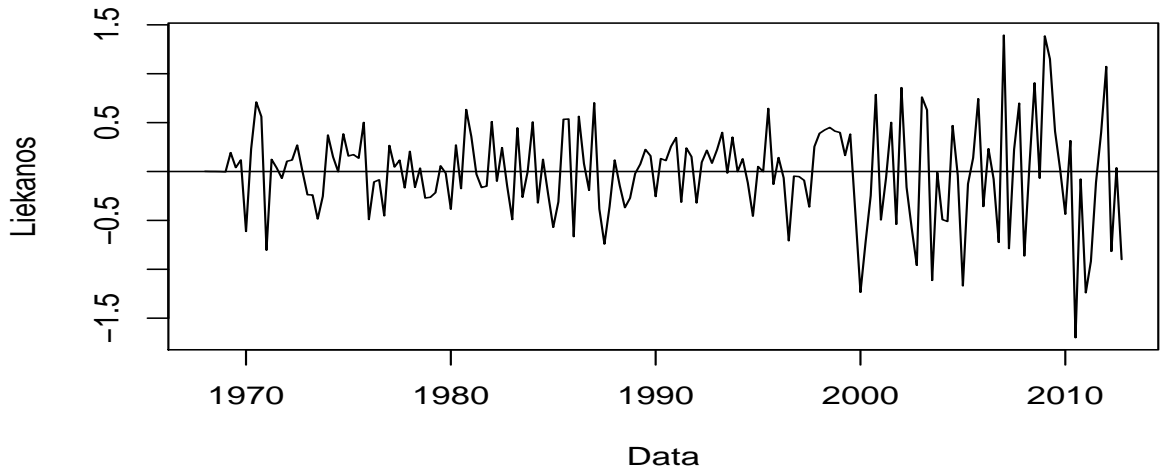


(a) Autokoreliacijos grafikas



(b) Dalinės koreliacijos grafikas

6 pav.: Diferencijuoto jaunimo nedarbo lygio autokoreliacijos grafikai



7 pav.: SARIMA modelio liekanų grafikas

Tačiau matome, kad nuo 2000-ųjų metų dispersija yra didesnė. Jeigu jaunimo nedarbo lygis susijęs su infliacija ir BVP – VAR modeliai galėtų padėti tiksliau įvertinti nuo 2010 metų prasidėjusią jaunimo nedarbo lygio kritimo tendenciją.

4.2.2 Vektorinės autoregresijos proceso modeliavimas

Bus sudaryti du atskiri VAR modeliai - diferencijuotiems jaunimo nedarbo lygio duomenims su bendrojo vidaus produkto pokyčiams, žymėsime VAR_{bvp} bei sezoniskai diferencijuotiems jaunimo nedarbo lygio duomenims su infliacijos sezoniniais skirtumais, žymėsime VAR_{inf} .

VAR_{bvp}

Remiantis AIC kriterijumi, pasirinktas $\text{VAR}_{bvp}(13)$ modelis:

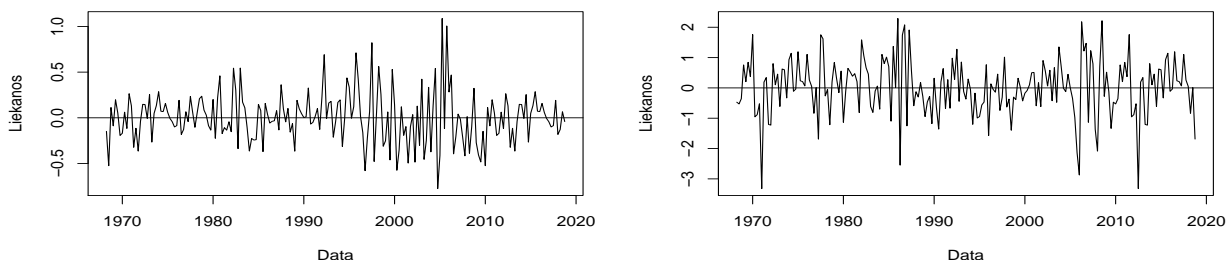
$$\Delta \mathbf{Y}_t = \sum_{i=1}^{13} \Theta_i \Delta \mathbf{Y}_{t-i}, \quad (4.1)$$

čia

- $\Delta \mathbf{Y}_t = (\Delta y_{1t}, \Delta y_{2t})'$ - laiko eilučių vektorius, kur Δy_{1t} Japonijos nedarbo lygio skirtumai, Δy_{2t} - BVP pokyčiai (kadangi mano originalūs duomenys yra BVP pokyčiai, laiko eilutė papildomai nėra diferencijuota),

- Θ_i - koeficientų matricos (nurodytos 23 psl.).

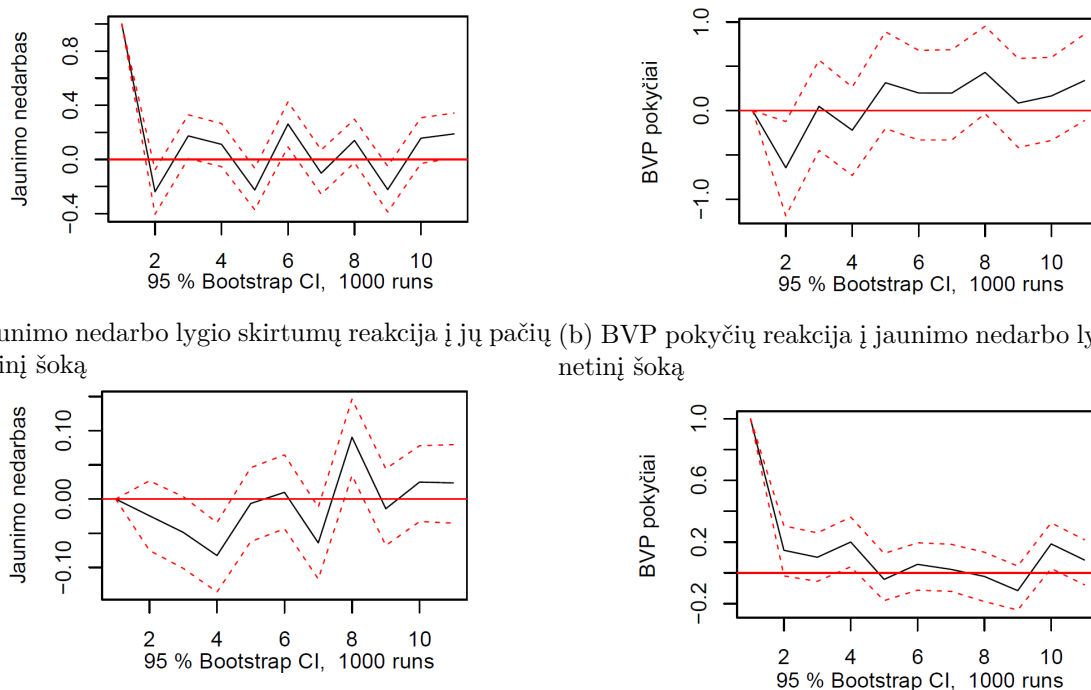
Atlikus vizualią liekanų analizę, galima pastebėti, kad grafiškai jos yra išsidėsčiusios apie nulį (8 pav.). Atlikus Breusch-Godfrey testą liekanų autokoreliacijai nustatyti, matoma, kad liekanos yra nekoreliuotos (p reikšmė - 0.1462). Liekanų vidurkis yra 0.014. Todėl modelio liekanos yra baltasis triukšmas, kas indikuoja, kad $\text{VAR}_{bvp}(13)$ modelis adekvatus.



(a) Jaunimo nedarbo skirtumų $\text{VAR}_{bvp}(13)$ liekanos (b) BVP pokyčių $\text{VAR}_{bvp}(13)$ modelio liekanos

8 pav.: $\text{VAR}_{bvp}(13)$ modelio liekanos

Iš $\text{VAR}_{bvp}(13)$ modelio impulso-atsako grafiko matyti (9 pav.), kad šokai iš abiejų laiko eilučių nemenkai daro įtaką kintamųjų reikšmėms. Impulsai, ateinantys iš jaunimo nedarbo lygio skirtumų taip pat keičiasi laike ir daro įtaką BVP pokyčiams - iš pat pradžių impulsai veikia neigiamai, tačiau laikui bėgant - poveikis pasidaro teigiamas (9b pav.). BVP pokyčių šokas taip pat turi įtakos jaunimo nedarbo lygio skirtumams. Jeigu BVP pokyčių impulsas teigiamas, tai jaunimo nedarbą veikia neigiamai. Todėl tiksliau galime įvertinti BVP pokyčių sąryšį su jaunimo nedarbo lygiu nenagrinėjant individualių $\text{VAR}(13)$ modelio parametrų.



(a) Jaunimo nedarbo lygio skirtumų reakcija į jų pačių vienetinį šoką (b) BVP pokyčių reakcija į jaunimo nedarbo lygio vienetinį šoką

(c) Jaunimo nedarbo lygio reakcija į BVP pokyčių vienetinį šoką (d) BVP pokyčių reakcija į jų pačių vienetinį šoką

9 pav.: Impulso-atsako grafikai

VAR_{inf}

Remiantis AIC kriterijumi, buvo pasirinktas penktos eilės VAR_{inf} modelis:

$$\Delta \mathbf{Y}_s = \sum_{i=1}^5 \Theta_i \Delta \mathbf{Y}_{s-i} + \mu_s, \quad (4.2)$$

čia

- μ_s - konstanta (nurodyta 23 psl);
- $\Delta \mathbf{Y}_s = (\Delta y_{1s}, \Delta y_{2s})'$ - laiko eilučių vektorius, kur Δy_{1s} sezoniniai Japonijos nedarbo lygio skirtumai, Δy_{2s} - sezoniniai infliacijos skirtumai,
- Θ_i - koeficientų matricos (nurodytos 23 psl.).

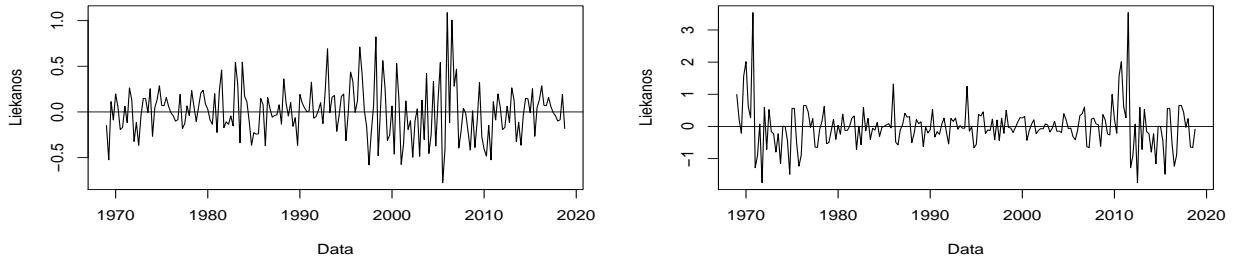
Kadangi jaunimo nedarbo lygis ir infliacija turi vienodos eilės integruotumą, reikia patikrinti, ar kintamieji yra kointegruoti. Kointegruotumo patikrinimui naudotas Johansen testas. Rezultatai matomi 4 lentelėje.

Maksimalus rangas	Testo statistika	5% kritinė reikšmė
0	7.92	14.90
1	2.58	8.18

4 lentelė: Johansen testas kointegracijai patikrinti

Taigi, matoma, kad tarp jaunimo nedarbo lygio ir infliacijos ilgojo laikotarpio sąryšio nėra. Todėl VAR_{inf} modelį sudarys pirmieji jaunimo nedarbo lygio ir infliacijos sezoniniai skirtumai.

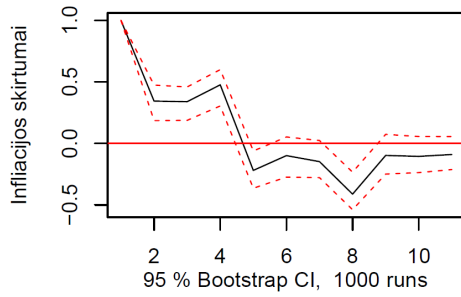
Atlikus tokią pat liekanų analizę, matoma, kad liekanos yra išsidėsčiusios apie nulinę tiesę (10 pav.). Breusch-Godfrey testo p-reikšmė yra 0.05364, todėl griežtai paneigti liekanų autokoreliatumo 0.05 lygmeniu negalime, tačiau nulinės hipotezės atmesti tikrai negalime su 0.1 reikšmingumo lygmeniu. Liekanų vidurkis $(-1.17) \times 10^{-17}$ - arti nulio. Taigi, liekanos sudaro baltojo triukšmo procesą.



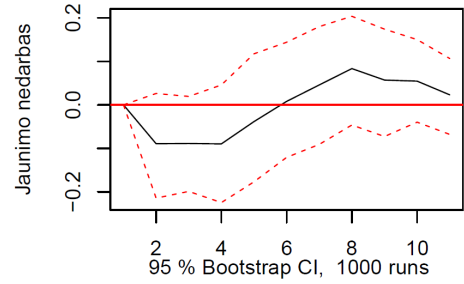
(a) Sezoninių jaunimo nedarbo skirtumų $\text{VAR}_{inf}(5)$ (b) Sezoninių infliacijos pokyčių $\text{VAR}_{inf}(5)$ modelio liekanos

10 pav.: $\text{VAR}_{inf}(5)$ modelio liekanos

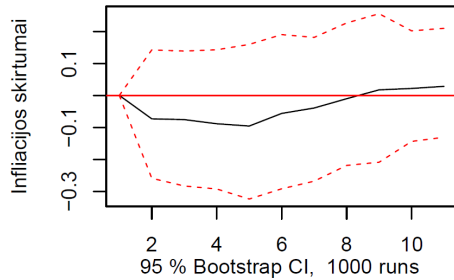
Šį kartą, nagrinėjant $\text{VAR}_{inf}(5)$ modelio impulso-atsako grafikus matome, kad infliacijos teigiamieji pokyčiai neigiamai veikia jaunimo nedarbo lygį (11a pav. ir 11b pav.) Tačiau sunku pasakyti, ar įvykę jaunimo nedarbo pokyčiai irgi paveikia infliaciją, ar jie keičiasi tik dėl pirminio šoko. Bet tuo pačiu metu žiūrint į jaunimo nedarbo šokus matome (11d pav.), kad jie daro įtaką infliacijos skirtumams (11c pav.), bet jų pasikliautiniai intervalai didesni, lyginant su infliacijos šokų įtaka jaunimo nedarbui (11b pav.).



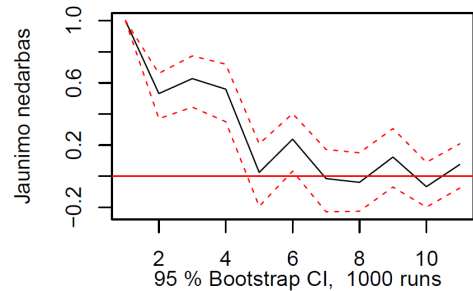
(a) Inflacijos reakcija į jų pačių vienetinį šoką



(b) Jaunimo nedarbo lygio reakcija į infliacijos vienetinį šoką



(c) Inflacijos reakcija į jaunimo nedarbo lygio vienetinį šoką



(d) Jaunimo nedarbo lygio skirtumų reakcija į jų pačių vienetinį šoką

11 pav.: Impulso-atsako grafikai

4.2.3 Vektorinis paklaidų korekcijos modelis

Kadangi du kintamieji - jaunimo nedarbas ir infliacija yra pirmosios integruotumo eilės - $I(1)$, norint sudaryti bendrą visų trijų kintamųjų modelį, reikia patikrinti jų kointegruotumą. Taip pat svarbu dar kartą pabrėžti, kad Johansen testas yra atliktas originaliems duomenims, kur jaunimo nedarbo lygis ir infliacija yra nediferencijuoti, o BVP laiko eilutė yra pokyčių laiko eilutė.

Imant VAR(8), Johansen testo rezultatai matomi 5 lentelėje.

Maksimalus rangas	Testo statistika	5% kritinė reikšmė
0	26.07	21.07
1	6.50	14.90
2	2.53	8.18

5 lentelė: Johansen testo rezultatai

Kadangi yra vienas kointegravimo ryšys, laiko eilutėms bus sudarytas VECM(7) modelis:

$$\Delta \mathbf{Y}_t = \alpha \beta' \mathbf{Y}_{t-1} + \sum_{i=1}^7 \Gamma_i \Delta \mathbf{Y}_{t-i}, \quad (4.3)$$

čia

- $\Delta \mathbf{Y}_t = (\Delta y_{1t}, \Delta y_{2t}, \Delta y_{3t})'$ - laiko eilučių vektorius, kur Δy_{1t} Japonijos nedarbo lygio skirtumai, Δy_{2t} - BVP pokyčių pokyčiai bei Δy_{3t} infliacijos skirtumai,
- $\beta' \mathbf{Y}_{t-1}$ - paklaidų korekcijos komponentė ECT_{t-1} ,
- Γ_i - koeficientų matricos (nurodytos 23 psl.),
- α - paklaidų korekcijos komponentės koeficientas (nurodytas 23 psl.)

4.3 lygtį galime užrašyti ir taip:

$$\Delta \mathbf{Y}_t = \alpha ECT_{t-1} + \sum_{i=1}^7 \Gamma_i \Delta \mathbf{Y}_{t-i}. \quad (4.4)$$

Kointegracijos sąryšis:

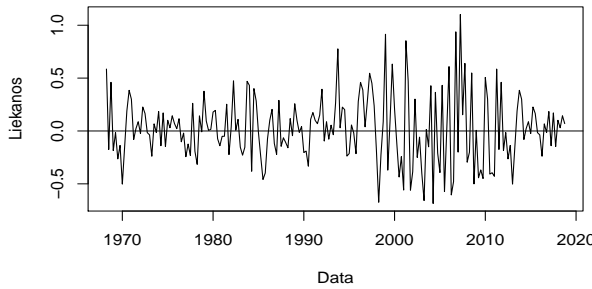
$$\text{jaunimo nedarbas}_{t-1} = -1.244685 \times \text{BVP pokyčiai}_{t-1} - 0.5830744 \times \text{inflacija}_{t-1} \quad (4.5)$$

Iš šios lygties pastebima, kad visoms kitoms sąlygoms esant lygioms, mažėjant BVP pokyčiui, jaunimo nedarbo lygis auga 1.245 procentus. Inflacijai mažėjant, jaunimo nedarbas taip pat auga, tačiau mažesniu procentu - 0.583.

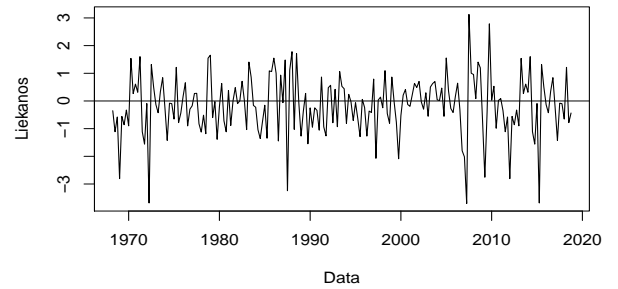
VECM(7) modelio liekanos vėlgi gavosi baltasis triukšmas:

- Liekanos išsidėsčiusios apie nulinę tiesę (12 pav.),
- Breusch-Godfrey testas liekanoms H_0 hipotezės neatmetė (p reikšmė - 0.06828),
- Liekanų vidurkis yra arti nulio: $(-1.48) \times 10^{-14}$.

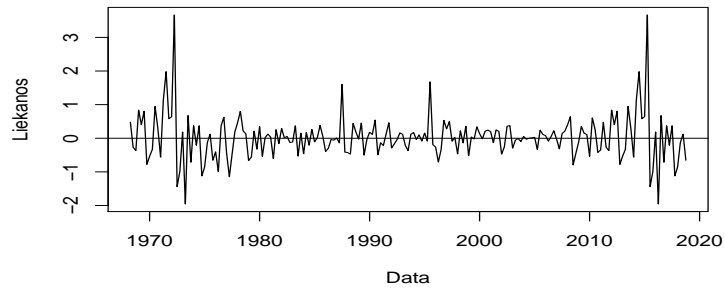
Tai indikuoja modelio adekvatumą.



(a) VECM(7) jaunimo nedarbo liekanos



(b) VECM(7) BVP liekanos



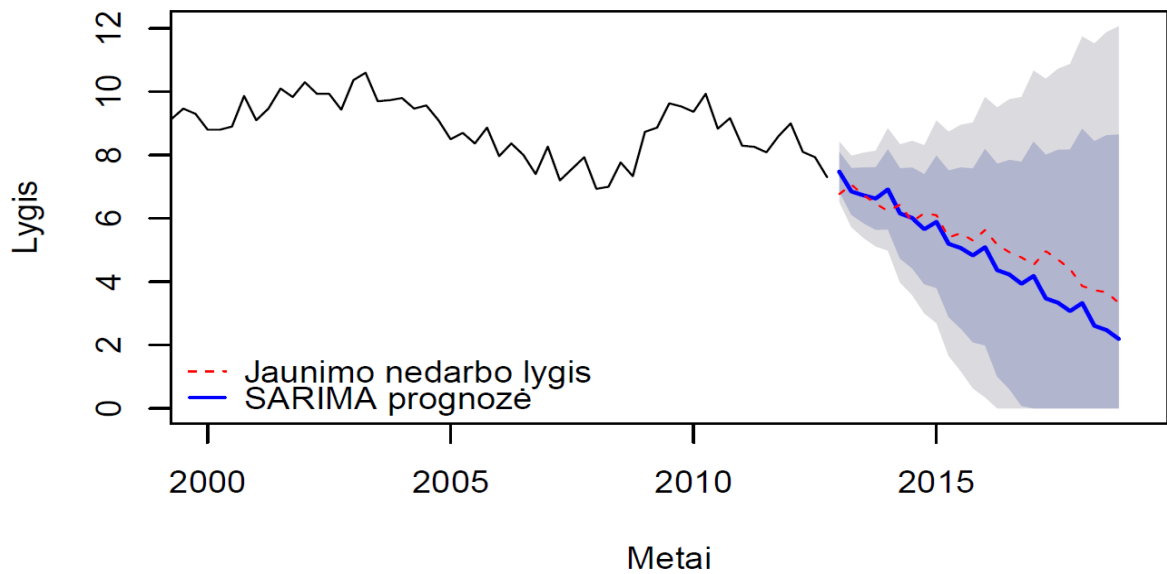
(c) VECM(7) infliacijos liekanos

12 pav.: VECM(7) modelio liekanos

4.2.4 Modelių prognozavimo rezultatai

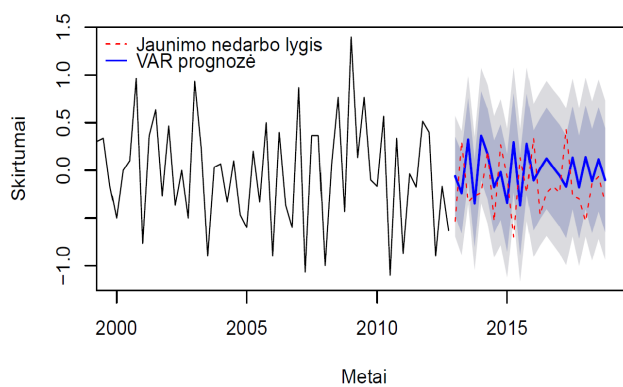
SARIMA(1,1,0)(1,1,0)₄ modelio prognozės

Iš SARIMA(1,1,0)(1,1,0)₄ modelio prognozės matyti (13 pav.), kad tikrosios ir prognozuotos reikšmės šiek tiek skiriasi. Tačiau tas skirtumas didėja didėjant ir metams, todėl trumpuoju laikotarpiu prognozė yra geresnė, negu ilguoju laikotarpiu. Tą indikuoja pasiklivimo intervalų plėtimasis. Čia MAE yra 0.3509, o RMSE - 0.475.

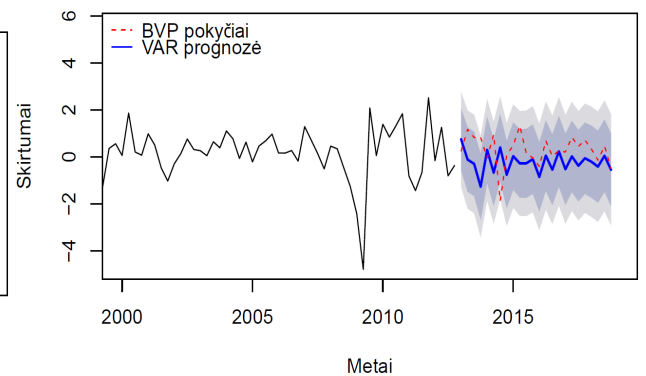


13 pav.: Nedarbo lygio SARIMA prognozė ir tikrosios reikšmės

VAR_{bvp} modelio prognozės



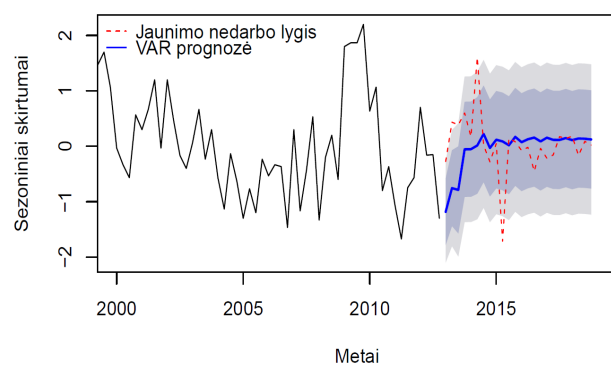
(a) Nedarbo skirtumų prognozės ir tikrosios reikšmės



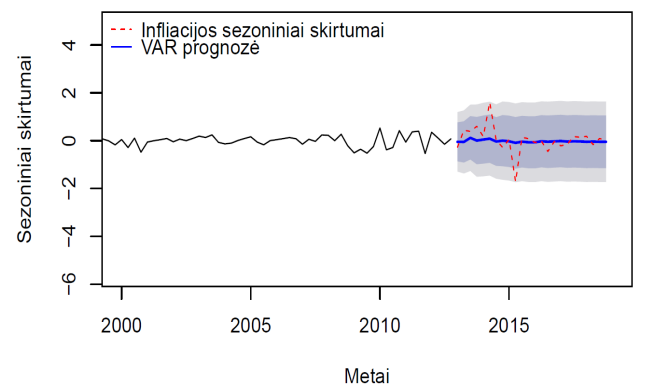
(b) Bendrojo vidaus produkto pokyčių prognozė ir tikrosios reikšmės

14 pav.: VAR_{bvp} prognozės ir tikrosios reikšmės

VAR_{inf} modelio prognozės



(a) Nedarbo skirtumų prognozės ir tikrosios reikšmės

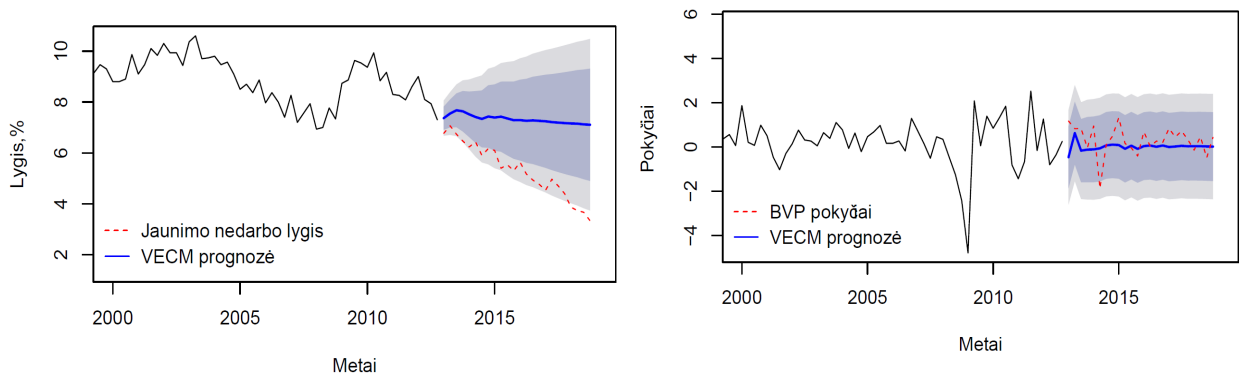


(b) Inflacijos skirtumų prognozė ir tikrosios reikšmės

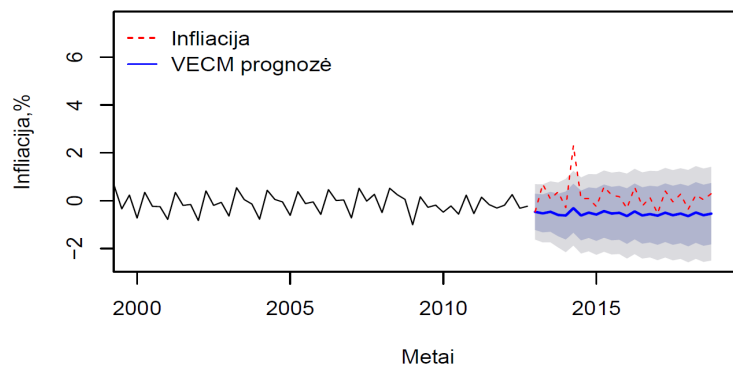
15 pav.: VAR_{inf} prognozės ir tikrosios reikšmės

Iš 14a ir 15a paveikslėlių galima pastebėti, kad VAR_{bvp} modelio prognozės panašesnės į tikrąsias reikšmes nei VAR_{inf} modelio prognozės. Tą patvirtina ir modelių vidutinė absoliuti paklaida (MAE) bei vidutinis kvadratinis paklaidos nuokrypis (RMSE): jaunimo nedarbo lygio skirtumų MAE ir RMSE yra 0.3825 ir 0.4463 atitinkamai, o sezoniniai jaunimo lygio skirtumai MAE ir RMSE reikšmės yra 0.7077 ir 0.7919 atitinkamai. Tačiau, BVP pokyčių prognozė yra labiausiai atitolusi nuo tikrųjų reikšmių. Įvertinus modelio MAE ir RMSE reikšmės - 0.779 ir 0.9799 - jos yra didesnės už sezonines infliacijos reikšmės - 0.3029 (MAE) ir 0.3086 (RMSE), todėl šiuo atveju VAR_{inf} penktos eilės modelis adekviau atliko infliacijos sezoninių skirtumų prognozę, nei VAR_{bvp} modelio tryliktos eilės BVP pokyčių prognozę.

VECM(7) modelio prognozės



(a) VECM prognozė Japonijos jaunimo nedarbo lygiui (b) Bendrojo vidaus produkto pokyčių VECM prognozė



(c) Infliacijos Japonijoje VECM prognozė

16 pav.: VECM(7) prognozė

VECM(7) prognozės jaunimo nedarbo lygiui ($MAE = 2$ ir $RMSE = 2,197$) ir infliacijai ($MAE = 0.7$ ir $RMSE = 0.865$) yra labiausiai nutolusios nuo tikrųjų reikšmių (16 pav.). Tačiau, vektorinis paklaidų korekcijos modelis BVP prognozes įvertino tiksliau, negu VAR_{bvp} modelio prognozės - $MAE = 0.5$ ir $RMSE = 0.722$.

5 Išvados bei rekomendacijos

Japonijos jaunimo nedarbo lygiui modeliuoti naudojami tiek vienmačių, tiek daugiamačių laiko eilučių modeliai SARIMA, VAR ir VECM. Visi sudaryti modeliai yra adekvatūs ir tinkami, tačiau modelių liekanų dispersija yra vietomis kiek didesnė. Sudarytų modelių prognozių rezultatai lyginami su tikrosiomis reikšmėmis bei prognozių tikslumas nustatomas naudojantis tinkamumo rodikliais (MAE ir RMSE). Nustatyta, kad SARIMA modelis Japonijos jaunimo nedarbui prognozuoti yra geriausias. Taip pat yra nustatytas ryšis tarp jaunimo nedarbo ir BVP pokyčių bei nedarbo ir infliacijos. Bet koks impulsas ateinantis iš BVP pokyčių ar infliacijos, veikia Japonijos jaunimo nedarbo lygį. Todėl galima patvirtinti, kad ekonominėje teorijoje įrodytas sąryšis tarp BVP pokyčių ir nedarbo galioja ir jaunimo nedarbo lygiui: jeigu krenta BVP - kyla jaunimo nedarbo lygis ir atvirkščiai. Taip pat šiame tyrime parodyta, egzistuoja ryšis tarp jaunimo nedarbo lygio ir infliacijos. Šis sąryšis atitinka Filipso kreivės idėją, kad infliacija susijusi su jaunimo nedarbo lygiu. Tačiau tyrimas indikuoja, kad prognozavimas trumpuoju laikotarpiu yra geresnis už ilgojo laikotarpio prognozes.

Vis dėlto, dėl modelių liekanų dispersijos nepastovumo, vienas iš uždavinių galėtų būti gerinti modelius tam, kad liekanų dispersija būtų maždaug tokia pačia lygyje. Taip pat, kadangi trumpojo laikotarpio prognozės yra geresnės, prasminga būtų toliau nagrinėti tinkamiausią modelį Japonijos jaunimo nedarbo lygio ateities reikšmėms prognozuoti. Naudinga atlikti tolesnius tiek paprastų, tiek sudėtingų ekonometrinių modelių tyrimus, tam, kad būtų galima numatyti ir ilgalaikį jaunimo nedarbo lygį. Tolimesniems tyrimams gali būti naudinga modeliuoti jaunimo nedarbą naudojant struktūrinį VAR modelį.

6 Pirmas priedas

Ivertintų VAR_{bvp} , VAR_{inf} ir VECM modelių parametrų matricos

$$\begin{aligned}
 &4.1 \text{ modelio matricos: } \hat{\Theta}_1 = \begin{pmatrix} -0.238817 & -0.024433 \\ -0.64486 & 0.14656 \end{pmatrix}, \hat{\Theta}_2 = \begin{pmatrix} 0.100670 & -0.050950 \\ -0.01124 & 0.06398 \end{pmatrix}, \\
 &\hat{\Theta}_3 = \begin{pmatrix} 0.145184 & -0.081834 \\ -0.07792 & 0.14445 \end{pmatrix}, \hat{\Theta}_4 = \begin{pmatrix} -0.237785 & 0.004263 \\ 0.49247 & -0.15508 \end{pmatrix}, \hat{\Theta}_5 = \begin{pmatrix} 0.117606 & 0.034741 \\ 0.04713 & 0.05984 \end{pmatrix}, \\
 &\hat{\Theta}_6 = \begin{pmatrix} 0.061798 & -0.047740 \\ 0.32794 & 0.02031 \end{pmatrix}, \hat{\Theta}_7 = \begin{pmatrix} 0.152506 & 0.065243 \\ 0.25481 & 0.00733 \end{pmatrix}, \hat{\Theta}_8 = \begin{pmatrix} -0.170663 & 0.020998 \\ 0.26562 & -0.05569 \end{pmatrix}, \\
 &\hat{\Theta}_9 = \begin{pmatrix} 0.146957 & 0.032135 \\ -0.20777 & 0.24589 \end{pmatrix}, \hat{\Theta}_{10} = \begin{pmatrix} 0.319950 & -0.017439 \\ 0.52024 & 0.12067 \end{pmatrix}, \hat{\Theta}_{11} = \begin{pmatrix} 0.122221 & -0.014783 \\ 0.45790 & 0.03895 \end{pmatrix}, \\
 &\hat{\Theta}_{12} = \begin{pmatrix} -0.245466 & 0.042218 \\ 0.18299 & -0.12352 \end{pmatrix}, \hat{\Theta}_{13} = \begin{pmatrix} 0.090596 & 0.030885 \\ 0.03302 & 0.23423 \end{pmatrix}.
 \end{aligned}$$

$$\begin{aligned}
 &4.2 \text{ modelio matricos: } \hat{\Theta}_1 = \begin{pmatrix} 0.53117 & -0.08926 \\ -0.073301 & 0.344245 \end{pmatrix}, \hat{\Theta}_2 = \begin{pmatrix} 0.33869 & -0.01060 \\ -0.011521 & 0.214395 \end{pmatrix}, \\
 &\hat{\Theta}_3 = \begin{pmatrix} 0.03883 & 0.02174 \\ 0.005259 & 0.278389 \end{pmatrix}, \hat{\Theta}_4 = \begin{pmatrix} -0.51270 & 0.08135 \\ 0.016773 & -0.558857 \end{pmatrix}, \hat{\Theta}_5 = \begin{pmatrix} 0.28150 & -0.033431 \\ -0.028202 & -0.029584 \end{pmatrix}, \\
 &\hat{\mu}_s = \begin{pmatrix} 0.04133 \\ -0.017864 \end{pmatrix}.
 \end{aligned}$$

$$\begin{aligned}
 &4.3 \text{ modelio matricos: } \hat{\alpha} = \begin{pmatrix} -0.0009 \\ -0.0039 \\ -0.0037 \end{pmatrix}, \hat{\Gamma}_1 = \begin{pmatrix} -0.3160 & 0.2193 & -0.1220 \\ -0.0282 & -0.8317 & 0.0467 \\ -0.0512 & -0.2483 & -0.6018 \end{pmatrix}, \\
 &\hat{\Gamma}_2 = \begin{pmatrix} 0.0106 & -0.0128 & -0.0329 \\ -0.0660 & -0.6339 & 0.0496 \\ -0.0905 & -0.1555 & -0.2947 \end{pmatrix}, \hat{\Gamma}_3 = \begin{pmatrix} 0.1441 & 0.5298 & 0.0606 \\ -0.1106 & -0.4305 & 0.0465 \\ -0.0824 & -0.3188 & -0.0570 \end{pmatrix}, \\
 &\hat{\Gamma}_4 = \begin{pmatrix} -0.0749 & 0.3165 & -0.0428 \\ -0.1638 & -0.4558 & 0.1353 \\ -0.0789 & -0.3878 & 0.2803 \end{pmatrix}, \hat{\Gamma}_6 = \begin{pmatrix} 0.1587 & 0.2918 & -0.2216 \\ -0.1398 & -0.3247 & 0.1015 \\ -0.0598 & -0.0442 & -0.0617 \end{pmatrix}, \\
 &\hat{\Gamma}_7 = \begin{pmatrix} 0.0992 & 0.1342 & -0.0435 \\ -0.0969 & -0.2257 & 0.0736 \\ -0.0333 & -0.1659 & -0.2060 \end{pmatrix}, \hat{\Gamma}_8 = \begin{pmatrix} 0.1790 & 0.1459 & 0.0331 \\ -0.0817 & -0.0113 & -0.0040 \\ -0.0052 & -0.0678 & -0.4609 \end{pmatrix}.
 \end{aligned}$$

Literatūra

- [1] Y. Genda. Who Really Lost Jobs in Japan? Youth Employment in an Aging Japanese Society. In *Labor Markets and Firm Benefit Policies in Japan and the United States*, pages 103–134. National Bureau of Economic Research, Inc, 2003.
- [2] M. González-Fernández and C. González-Velasco. Can Google econometrics predict unemployment? Evidence from Spain. *Economics Letters*, 170(C):42–45, 2018.
- [3] A. Hammarström. Health consequences of youth unemployment. *Public Health*, pages 403–412, 1994.
- [4] International Labour Office. *Global Employment Trends for Youth 2017*. ILO, 2017.
- [5] F. M. Kreishan. Economic growth and unemployment: An empirical analysis. *Journal of Social Sciences* 7, pages 228–231, 2011.
- [6] C. Kwon and J. Jung. Forecasting youth unemployment in Korea with web search queries. In O. Galinina, S. Balandin, and Y. Koucheryavy, editors, *Proceedings of the the 9th Conference on Internet of Things, Smart Spaces, and Next Generation Networks and Systems*, page 3–14.
- [7] H. Ohlsson. Labor market policy, unemployment and wages—a VAR model for Sweden 1969–1990.
- [8] Putun, M. and Karatas, A. S. and Akyildiz I. E. The economic consequences of the youth unemployment case in EU countries: a critical analysis. *International Journal of Economics and Finance Studies*, 2017.
- [9] M. Simionescu. Kalman Filter or VAR Models to Predict Unemployment Rate in Romania? *Naše gospodarstvo/Our economy*, 61(3):3–21, 2015.
- [10] C. Sims. Macroeconomics and reality. *Econometrica*, 48(1):1–48, 1980.
- [11] J. H. Stock and M. W. Watson. Vector autoregressions. *Journal of Economic Perspectives*, 15(4):101–115, 2001.
- [12] J. Tartilas. Nedarbo problema. *Jurisprudencija*, 2005.