

VILNIAUS UNIVERSITETAS

LAURYNAS NARUŠEVIČIUS

EKONOMETRINIS BANKŲ STABILUMO VERTINIMAS

Daktaro disertacijos santrauka  
Fiziniai mokslai, matematika (01 P)

Vilnius, 2017 metai

Disertacija rengta 2012–2016 metais Vilniaus universitete.

**Mokslinis vadovas:**

prof. habil. dr. Alfredas Račkauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P).

Disertacija ginama viešame gynimo tarybos posėdyje:

**Pirmininkas** – prof. habil. dr. Remigijus Leipus (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P).

**Nariai:**

prof. dr. Audronė Jakaitienė (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P);

prof. habil. dr. Kęstutis Kubilius (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P);

prof. dr. Marijus Radavičius (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P);

prof. habil. dr. Anne Philippe (Nanto universitetas, Prancūzija, fiziniai mokslai, matematika – 01 P).

Disertaciją bus ginama viešame gynimo tarybos posėdyje 2017 m. liepos mėn. 4 d. 15 val. Vilniaus universiteto Matematikos ir informatikos fakulteto prof. Jono Kubiliaus (102) auditorijoje.

Adresas: Naugarduko g. 24, LT-03225 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2017 m. birželio mėn. 5 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir Vilniaus universiteto interneto svetainėje adresu: [www.vu.lt/naujienos/ivykiu-kalendorius](http://www.vu.lt/naujienos/ivykiu-kalendorius)

VILNIUS UNIVERSITY

LAURYNAS NARUŠEVIČIUS

ECONOMETRIC ASSESSMENT OF BANK STABILITY

Summary of doctoral dissertation  
Physical sciences, mathematics (01 P)

Vilnius, 2017

Doctoral dissertation was written in 2012–2016 at Vilnius University.

**Scientific supervisor:**

prof. habil. dr. Alfredas Račkauskas (Vilnius university, physical sciences, mathematics – 01 P).

The dissertation will be defended at the public meeting of the council:

**Chairman** – Prof. Habil. Dr. Remigijus Leipus (Vilnius University, Physical Sciences, Mathematics – 01 P).

**Members:**

Prof. Dr. Audronė Jakaitienė (Vilnius University, Physical Sciences, Informatics – 09 P);

Prof. Habil. Dr. Kęstutis Kubilius (Vilnius University, Physical Sciences, Mathematics – 01 P);

Prof. Dr. Marijus Radavičius (Vilnius University, Physical Sciences, Mathematics – 01 P);

Prof. Habil. Dr. Anne Philippe (Nantes University, France, Physical Sciences, Mathematics – 01 P).

The dissertation will be defended at the public meeting of the council in prof. Jonas Kubilius lecture room (room 102), Vilnius University, Faculty of Mathematics and Informatics at 3 p.m. on 4 July, 2017.

Address: Naugarduko st. 24, LT-03225 Vilnius, Lithuania.

The summary of the dissertation was distributed on 5 June, 2017.

The dissertation is available at the library of Vilnius University and VU website:  
[www.vu.lt/naujienos/ivykiu-kalendorius](http://www.vu.lt/naujienos/ivykiu-kalendorius)

# *Disertacijos aprašymas*

## **1 Įžanga**

Bankų sektoriaus ir visos finansų sistemos stabilumo tema pastaruoju metu sulaukia ypatingo dėmesio tiek centriniuose bankuose, tiek tarp kitų finansų rinkų dalyvių, tiek iš visuomenės. Nors finansų krizės nėra naujas reiškinys, tačiau finansų sistemos stabilumu centriniai bankai pradėjo rūpintis tik pastaraisiais dešimtmečiais. Finansų krizės yra retos, bet po jų dažniausiai vyksta ekonomikos nuosmukis, todėl centriniai bankai ir priežiūros institucijos turi imtis veiksmų, kuriais būtų siekiama užtikrinti finansų sistemos stabilumą.

Makroekonominis testavimas nepalankiausiomis sąlygomis yra viena iš tokių priemonių. Tai būdas iš anksto įvertinti bankų sistemos atsparumą įvairių rūšių rizikai, galinčiai pasitvirtinti ateityje. 2007-2008 metais įvykusi pasaulinė finansų krizė ypač paspartino testavimo nepalankiausiomis sąlygomis metodų plėtotę ir taikymą. Ap-skritai testavimas nepalankiausiomis sąlygomis yra atliekamas komercinių bankų, priežiūros institucijų ir centrinių bankų, siekiant įvertinti viso finansų sektoriaus ar konkrečios finansų įstaigos atsparumą nepalankiems ekonomikos pokyčiams. Skiriamas testavimas iš apačios į viršų (angl. *bottom-up*) ir testavimas iš viršaus į apačią (angl. *top-down*). Testavimą iš apačios į viršų atlieka komerciniai bankai, naudodami turimus duomenis ir savus modelius. Centrinis bankas ar kita priežiūros institucija gali nustatyti tam tikrus modeliavimo apribojimus (pvz., pateikti bendrus scenarijus, nurodyti privalomas taikyti prielaidas, nustatyti metodologijos principus) ir, taikydami savas analizės priemones, tikrinti bei vertinti komercinių bankų gautus rezultatus. Tokiu testavimo metodu siekiama įvertinti konkretaus banko atsparumą ekonomikos sukrėtimams, taigi, testavimas iš apačios į viršų yra viena iš mikroprudencinės bankų priežiūros priemonių. Testavimas iš viršaus į apačią yra atliekamas centrinių bankų, dažniausiai komercinių bankų į vertinimo procesą neįtraukiant. Atliekant jį, taikomos taisyklės, scenarijai ir modeliavimo prielaidos, vienodos visiems bankams. Testavimas iš viršaus į apačią dažnai būna priemonė, padedanti palyginti rezultatus, gautus atlikus testavimą iš apačios į viršų. Taip galima aptikti komercinių bankų atlikto testavimo prieštaras. Testavimas iš viršaus į apačią yra viena iš makroprudencinės priežiūros priemonių, nes šia procedūra siekiama įvertinti visos bankų sistemos atsparumą nepalankiems ekonomikos sukrėtimams.

Pagrindinis šioje disertacijoje pristatomos testavimo nepalankiausiomis sąlygomis procedūros tikslas – kiekybiškai įvertinti visos Lietuvos bankų sistemos ir ją sudarančių institucijų atsparumą nepalankiems ekonomikos sukrėtimams. Atliekant mokymo testavimą nepalankiausiomis sąlygomis, vertinamas bankų kapitalo pakan-

kamumas esant nepalankiam makroekonominiam scenarijui. Testuojamas laikotarpis apima dvejus metus, nuosekliai modeliuojamos ketvirtinės bankų pelno (nuostolio) ataskaitos eilutės. Pasiūlyta procedūra atitanka testavime iš viršaus į apačią taikomus principus.

Kaip vienas iš pagrindinių ekonominės veiklos finansavimo šaltinių, bankai gali daryti įtaką verslo ciklas. Kita vertus, bankų pajamos kinta laike, kadangi jos priklauso nuo bendro ekonomikos aktyvumo. Banko pelningumas yra vienas iš svarbiausių veiksnių, lemiančių banko stabilumą, skolinimo pajėgumą ir yra svarbi dalis atliekant testavimo nepalankiausiomis sąlygomis, nes pelnas gali kompensuoti didelę dalį patirtų kredito nuostolių. Stabilus bankų sektorius gali stimuliuoti ekonomiką ir sugeba atlaikyti ekonominius sukrėtimus. Dėl šios priežasties svarbu suprasti, koks yra ryšis tarp banko pajamų ir makroekonominių rodiklių. Be to, tai gali padėti įvertinti bankų sektoriaus stabilumą.

Vienas iš disertacijoje nagrinėjamų tikslų yra ištirti ryšį tarp Lietuvos bankų sektoriaus pelningumo ir jį lemiančių veiksnių. Šio ryšio žinojimas yra naudingas tiek komerciniams bankams, tiek ir priežiūros institucijoms, kurios yra atsakingos už finansų sektoriaus stabilumą. Šioje disertacijoje pritaikomas panelinis paklaidų korekcijos modelis (angl. *panel error correction model*), kuris padeda įvertinti ilgalaikius ir trumpalaikius vidinius ir išorinius veiksnius, susijusius su banko pelno nuostolio ataskaitos eilutėmis (grynosiomis palūkanų pajamomis, grynosiomis paslaugų ir komisinių pajamomis bei operacinėmis išlaidomis). Sugrupuoto vertinimo metodas (angl. *pooled mean group*), kuris buvo pasiūlytas Pesaran ir kt. (1997, 1999), nustato homogeniškus ilgalaikius koeficientus ir leidžia heterogeniškumą tarp trumpo laikotarpio koeficientų. Taigi šis tyrimas prisideda prie negausios mokslinės literatūros nagrinėjančios Lietuvos bankų sektorių. Be to, pritaikomas panelinio modelio vertinimo metodas, kuris nebuvo naudojamas šioje tyrimų srityje.

Kitų autorių tyrimais nustatyta, kad bankų pelningumas priklauso nuo bankų vidinių, bankų sektoriaus ir makroekonominių veiksnių. Tyrimuose naudoti kintamieji ir jų poveikis pelningumui yra skirtingas, kadangi naudojami skirtingi duomenų rinkiniai. Keletas tyrimų (Andersen ir kt. 2008; Albertazzi ir Gambacorta 2009) nagrinėja atskiras bankų pajamų ir išlaidų eilutes. Tačiau nei vienas iš šių tyrimų neįtraukia Lietuvos bankų sektoriaus į savo analizę. Šiame darbe tiriame, ar galima rasti panašius ryšius tarp bankų pelningumo ir jį lemiančių veiksnių besivystančioje Lietuvos ekonomikoje.

Antroje disertacijos dalyje iš Lietuvos bankinio sektoriaus analizavimo pereinama į Europos bankų sektoriaus analizę, visų pirma į klasterinę analizę. Pastaraisiais metais klasterinė analizė, kuria siekiama iš daugelio stebimų duomenų išskirti panašias grupes, išpopuliarėjo mokslinėje literatūroje. Viena iš klasterinės anaizės

sričių yra laiko eilučių klasterizavimas, kurioje laiko eilučių duomenų grupavimas padeda aptikti charakteringas trajektorijas, prognozuoti būsimą raidą ir kt. Klasterinėje analizėje taikomi metodai gali būti išskirstyti į tris kategorijas: metodai kurie remiasi (1) pradinių duomenų panašumu; (2) tam tikromis savybėmis, gautomis iš pradinių duomenų ir (3) iš pradinių duomenų gautais modeliais. Trumpai priminsime, kad panašumo/nepanašumo matas  $D$  tarp dviejų objektų  $X$  ir  $Y$  yra simetriškas ( $D(X, Y) = D(Y, X)$ ), ne neigiamas ( $D(X, Y) > 0$ ) ir tenkiną lygybę  $- D(X, X) = 0$ . Panašumo matas gali būti, bet nebūtinai yra metrika, t.y. tenkina nelygybę  $D(X, Z) \leq D(X, Y) + D(Y, Z)$ . Vienas iš disertacijos tikslų yra nagrinėti įvairius panašumo matus ir juos pritaikyti turimai Europos bankų duomenų imčiai. Kadangi bankų sektorius yra heterogeniškas todėl būtų naudinga surasti bankų grupes, pasižyminčias panašiomis savybėmis, ir pritaikyti ar kalibruoti tam tikrus makroprudencinius instrumentus, kurie tiktų atitinkamai grupei. Disertacijoje nagrinėjame galimybes suskirstyti bankus į panašias grupes.

Šioje disertacijoje nagrinėjami panašumo matai, kurie remiasi laiko eilučių arba funkcinių duomenų savybėmis. Vienu atveju bankai grupuojami į klasterius pagal vieną konkretų jų veiklos rodiklį, kitu atveju bankai klasterizuojami pagal kelis rodiklius. Kadangi atliekant klasterinę analizę iš anksto nėra žinoma, kiek yra klasterių, kitas svarbus uždavinys yra surasti klasterių skaičių, kuris geriausiai tinka turimai duomenų imčiai. Be to, svarbu, kad gauti klasteriai ne tik pasižymėtų geromis statistinėmis savybėmis, bet ir būtų ekonomiškai paaiškinami.

Paskutinė disertacijos dalis skirta įvertinti, kokie funkcinės duomenų analizės metodai būtų naudingi vertinant ir prognozuojant tam tikrą klasterį. Klausimas, kas yra geriau, prognozuoti agreguotą procesą tiesiogiai, o paskui išskaidyti ar prognozuoti atskirus komponentus ir tada sujungti juos, formuojant bendro proceso prognozę, yra svarbus daugelyje empirinių tyrimų. Tai taip pat žinoma, kaip prognozavimo iš viršaus į apačią ar iš apačios į viršų problema (žiūrėti pvz. Montgomery ir kt. 2007). Dažniausiai yra sunku teoriškai pagrįsti, kuris būdas geriau tinka konkrečiu atveju. Dėl to šis klausimas dažnai vertinamas empiriškai išbandant abu būdus.

Šioje disertacijoje nagrinėjama bankinių duomenų vertinimo ir prognozavimo problema. Analizuojamas Europos bankų kapitalo pakankamumo rodiklis, kuris parodo bankų galimybes atlaikyti galimus nuostolius, kylančius iš kredito rizikos, rinkos rizikos, operacinės rizikos ir kitų rizikų. Be to, šio rodiklio reikalavimas užtikrina, kad bankai neplės savo veiklos neturėdami pakankamai kapitalo.

Kitas dalykas, kuris yra nagrinėjamas disertacijoje, tai privalumai, kuriuos suteikia funkcinių duomenų analizės metodai. Tariant, kad kiekvienam  $j = 1, \dots, N$

turima atsitiktinė funkcija  $X_j = (X_j(t), 0 \leq t \leq T)$ , nagrinėjame agreguoto proceso

$$\hat{X}_N = \frac{1}{N} \sum_{j=1}^N X_j$$

prognozavimą taikant abu principus, t.y. iš viršaus į apačią ir iš apačios į viršų. Pirmu atveju įvertiname agreguoto proceso modelį  $\hat{X}_t$ , tada prognozuojame  $\hat{X}_N(T+h)$  tam tikram laiko periodui  $h$ . Antru anveju, įvertinamas modelis kiekvienam elementui  $X_j$  ir iš tų elementų prognozių  $X_j(T+h)$  sudaroma agreguoto dydžio  $\hat{X}_N(T+h)$  prognozė. Vertindami modelius ir atlikdami prognozes pasinaudojame funkcinių duomenų privalumu, t.y. galime diskretizuoti funkciją į tiek taškų, kiek reikia modelio vertinimui.

Galiausiai, disertacijoje pasiūlomas naujas funkcinių duomenų regresijos modelis bei jo įvertinimo procedūra. Šis modelis išsiskiria iš kitų funkcinės regresijos modelių tuo, kad jis skirtas įvertinti sąryšiui tarp vienos stochastinio proceso realizacijos ir kitų funkcinių kintamųjų.

## 2 Mokslinė problema ir tyrimo objektas

Pasaulinė finansų krizė, įvykusi 2007-2008 metais, atskleidė bankų sektoriaus ir visos finansų sistemos stabilumo svarbą ekonomikai ir jos raidai. Nors finansų krizės nėra naujas reiškinys, tačiau finansų sistemos stabilumu susirūpinta tik paskutiniu metu. Disertacijoje siekiama ekonometriškai įvertinti sąryšius tarp makroekonomikos kintamųjų raidos ir bankų veiklos rodiklių kaitos. Šie sąryšiai yra svarbūs, norint iš anksto įvertinti bankų sistemos atsparumą įvairių rūšių rizikai, galinčiai pasitvirtinti ateityje.

Klasterinė laiko eilučių analizė ir funkcinių duomenų analizė pastaruoju metu išpopuliarėjo mokslinėje literatūroje. Tačiau šie metodai vis dar yra mažai pritaikyti, analizuojant bankų duomenis ir jų raidą. Disertacijoje analizuojami teoriniai funkcinių duomenų modeliai ir jų panaudojimas bankinių duomenų prognozavimui. Funkcinių duomenų analizės metodų pritaikymas galėtų suteikti naujų galimybių tyrinėti sąryšius, kurie yra svarbūs bankiniame sektoriuje ir gali lemti jų stabilumą.

## 3 Tikslas ir pagrindiniai uždaviniai

Disertacija sudaryta iš dviejų pagrindinių dalių, todėl yra du pagrindiniai tikslai, kurie buvo pasiekti. Pirmasis tikslas – išanalizuoti Lietuvos bankų sistemą ir joje veikiančius bankus bei sukurti naują metodiką, kuri padėtų įvertinti jų atsparumą



esant neigiamam ekonomikos vystymuisi. Antrasis tikslas – nagrinėti turimos Europos bankų imties grupavimą į klasterius, remiantis jų veiklos rodikliais, ir palyginti panašumo matus, kurie remiasi laiko eilučių ir funkcinių duomenų savybėmis. Be to, siekiama svarstyti klasterio vidutinio proceso prognozavimo iš viršaus į apačią ir iš apačios į viršų problemą.

Norint pasiekti nustatytus tikslus, buvo suformuluoti tokie uždaviniai:

1. Sukurti metodiką, kurią būtų galima reguliariai naudoti, vertinant Lietuvos bankų sistemos stabilumą
2. Sudaryti ekonometrinius modelius, kurie nustatytų sąryšius tarp bankų kredito nuostolių ir makroekonominių kintamųjų dinamikos.
3. Prisidėti prie mokslinės literatūros, nagrinėjančios Lietuvos bankinį sektorių ir pritaikyti panelinį paklaidų korekcijos modelį, nustatant vidinius ir išorinius veiksnius, kurie daro įtaką bankų pajamoms ir išlaidoms ilgu ir trumpu laikotarpiu.
4. Pasiūlyti naujus panašumo matus ir palyginti su kai kuriais esamais panašumo matais, kurie remiasi laiko eilučių ar funkcinių duomenų savybėmis.
5. Išplėsti kelis panašumo matus iš vienmačio atvejo į daugiamatį.
6. Pasiūlyti teorinius funkcinių duomenų modelius ir juos pritaikyti, analizuojant funkcinius bankų duomenis.

## 4 Tyrimų metodika

Pirmoje darbo dalyje daugiausia remiamasi tiesinės regresijos ir dinaminiais panelinių duomenų modeliais. Kredito nuostoliai vertinami pasitelkiant tiesinės regresijos modelį, o bankų veiklos pelnas modeliuojamas, taikant dinaminį panelinių duomenų modelį arba panelinį paklaidų korekcijos modelį. Antroje darbo dalyje naudojami įvairūs laiko eilučių ar funkcinių duomenų panašumo matai. Bankai į klasterius sugrupuoti pasitelkiant hierarchinį jungimo klasterizavimo algoritmą. Klasterizavimo kokybė vertinama naudojant vidutinio silueto pločio, Dunn ir Caliński and Harabasz indeksus. Vidutinio proceso raida prognozuojama, naudojant autoregresinį ir funkcinių duomenų regresijos modelius.

## 5 Ginamieji teiginiai

1. Pasiūlyta makroekonominio testavimo nepalankiausiomis sąlygomis procedūra yra naudinga, vertinant Lietuvos bankų sistemos ir ją sudarančių bankų atsparumą nepalankiems makroekonominės aplinkos pokyčiams. Siūloma testavimo procedūra remiasi tarptautinėmis rekomendacijomis ir geriausia praktika.
2. Panelinis paklaidų korekcijos modelis gali būti naudojamas, nustatant vidinius ir išorinius veiksnius, kurie daro įtaką bankų pajamoms ir išlaidoms ilgu ir trumpu laikotarpiu.
3. Parodyta, kad funkcinių duomenų analizės metodai gali būti naudojami bankinių duomenų klasterinėje analizėje. Pasiūlyti nauji panašumo matai, kurie remiasi funkcinių duomenų savybėmis, labiau tiko bankinių duomenų klasterizavimui nei panašumo matai, kurie remiasi laiko eilučių savybėmis.
4. Parodyta, kad teorinių modelių, pristatytų disertacijoje, pritaikymas leidžia gauti vienodai gerus iš apačios į viršų ir iš viršaus į apačią bankinių duomenų prognozavimo rezultatus.

## 6 Darbo mokslinis naujumas ir aktualumas

Šioje disertacijoje pristatoma nauja Lietuvos bankų sistemos makroekonominio testavimo nepalankiausiomis sąlygomis procedūra, kurios pagrindinis tikslas – įvertinti bankų sistemos atsparumą. Pasiūlyta metodika remiasi tarptautine praktika ir patarimais ir yra naudojama, siekiant kiekybiškai įvertinti, ar bankų sukauptos kapitalo atsargos yra pakankamos, kad galėtų atlaikyti nepalankius makroekonominės aplinkos pokyčius. Sudarant testavimo nepalankiausiomis sąlygomis procedūrą daugiasia dėmesio skiriama ekonometriniais modeliams, leidžiantiems susieti makroekonominės aplinkos rodiklius su bankų veiklos rodikliais. Kadangi bankų pelnas yra svarbus veiksnys darantis įtaką jų stabilumui ir skolinimo pajėgumui, disertacijoje detaliau nagrinėjami bankų pajamas ir išlaidas lemiantis veiksniai. Panaudojamas panelinis paklaidų korekcijos modelis, kuris padeda įvertinti ilgalaikius ir trumpalaikius veiksnius, lemiančius bankų pajamas ir išlaidas. Tokiu būdu prisidedama prie negausios mokslinės literatūros, kuri nagrinėja Lietuvos bankų sektorių, bei pritaikomas vertinimo būdas, kuris yra naujas šioje tyrimų srityje.

Antroje disertacijos dalyje nagrinėjama Europos bankų klasterizavimo problema ir lyginami panašumo matai, kurie remiasi laiko eilučių ir funkcinių duomenų savybėmis. Atliekant klasterinę analizę buvo pasiūlyti du nauji panašumo matai, kurie remiasi funkcinių duomenų savybėmis. Pasiūlyti metodai buvo vieni iš labiausiai

tinkančių bankinių duomenų klasterizavimui. Be to, darbe du panašumo matai buvo praplėsti iš vienmačio atvejo į daugiamatį. Paskutinėje disertacijos dalyje funkcinių duomenų analizės metodai ir jų teikiami privalumai pritaikyti nagrinėjant bankų klasterius. Buvo pasiūlyti teoriniai funkcinių duomenų modeliai ir pristatytas naujas funkcinės regresijos modelis ir jo vertinimo būdas. Šie modeliai pritaikyti atliekant iš viršaus į apačią ir iš apačios į viršų bankinių duomenų prognozavimą. Funkcinių duomenų analizės metodų pritaikymas suteikia naujų galimybių tyrinėti sąryšius, kurie yra svarbūs bankiniame sektoriuje ir gali lemti jų stabilumą.

## **7 Darbo struktūra ir apimtis**

Disertacija parašyta anglų kalba. Disertaciją sudaro 5 skyriai: įvadas, 2 moksliniams tyrimams paskirti skyriai, išvados ir literatūros sąrašas. Kiekvienas ekonometrinis tyrimas yra nagrinėjamas atskirame skyriuje – glaustai pristatoma problema, susiję moksliniai rezultatai ir jų reikšmė, pristatoma tyrimų metodika ir pagrindiniai rezultatai. Bendra darbo apimtis yra 154 puslapiai.

## **8 Moksliniai rezultatai**

### **8.1 Makroekonominis testavimas nepalankiausiomis sąlygomis ir bankų pelningumas**

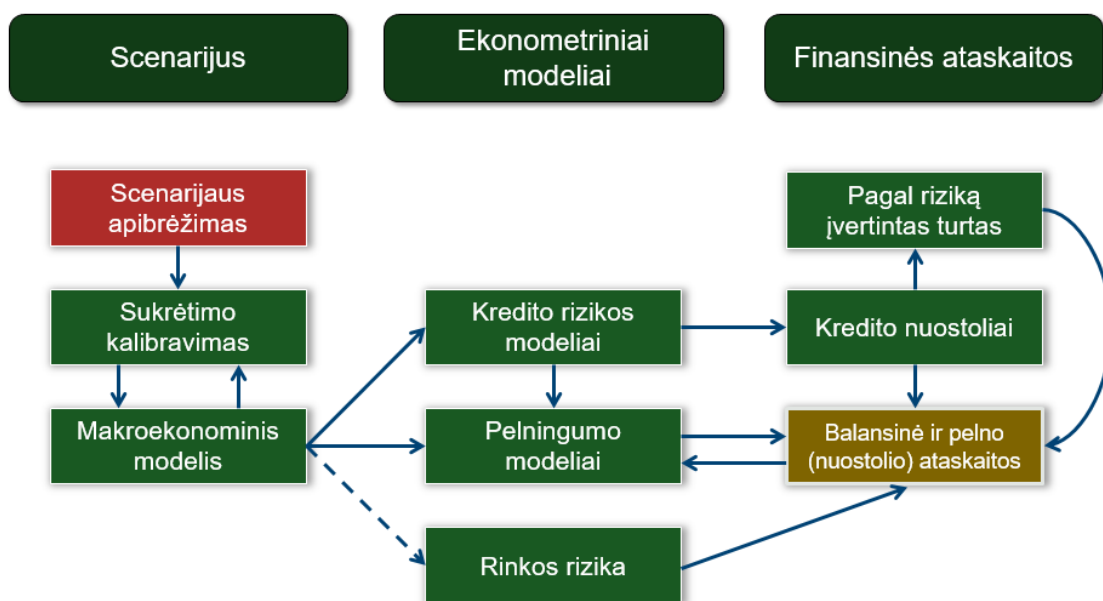
#### **8.1.1 Makroekonominis testavimas nepalankiausiomis sąlygomis**

Disertacijoje pristatoma testavimo nepalankiausiomis sąlygomis (TNS) procedūra, padeda įvertinti, ar bankų sukauptos kapitalo atsargos būtų pakankamos padengti paskolų portfelio nuvertėjimo nuostoliams, kurių galėtų susidaryti įvykus nepalankiems makroekonominės aplinkos pokyčiams. Pasiūlytos makroekonominio TNS procedūros trys pagrindiniai etapai (jų sudėtinės dalys ir tarpusavio sąryšiai pavaizduoti 1 pav.) yra šie:

- Pirmajame etape sudaromas makroekonominis scenarijus. Jo pagrindu toliau vertinamas bankų sistemos atsparumas. Scenarijus sudaromas naudojantis Lietuvos struktūriniu makroekonominiu modeliu, oficialiųjų makroekonominių rodiklių statistinėmis savybėmis, pasitelkiamas ir ekspertinis vertinimas.
- Kitame TNS procedūros etape taikomi ekonometriniai modeliai, kurie padeda susieti makroekonominių kintamųjų dinamiką su banko kredito rizikos ir pelningumo kaita. Šie modeliai yra dviejų tipų: kredito nuostolių modeliai ir pel-

ningumo modeliai; taikant pastaruosius modeliuojamos kitos pelno (nuostolio) ataskaitos eilutės.

- Galiausiai modeliavimo rezultatai iš skirtingų blokų susiejami į bendrą pelno (nuostolio) ataskaitą. Kartu įvertinama kapitalo kaita ir pagal riziką įvertinto turto kaita. Šie kintamieji apibrėžia tikslinį modeliavimo kintamąjį – kapitalo pakankamumo rodiklį, kuriuo remiantis daromas pagrindinės išvados apie banko atsparumą. Kiti modeliuojami rodikliai taip pat gali suteikti papildomos informacijos apie banko veiklos specifiką.



1 pav.. Testavimo nepalankiausiomis sąlygomis procedūros schema

Šaltinis: sudaryta autorių.

### *Testavimo prielaidos*

Taikant TNS gaunami rezultatai nėra prognozė. Priešingai, tai mažai tikėtinų įvykių analizė, pateikiamos išvados yra sąlyginės. Gautus rezultatus reikia vertinti, atsižvelgiant į daromas prielaidas. Atliekant TNS, taikomos šios nekintančio balanso (angl. *static balance sheet*) prielaidos, leidžiančios sukonkretinti skaičiavimus ir atsiriboti nuo neprognozuojamų aspektų:

- Per testuojamą laikotarpį bankų paskolų portfelio struktūra lieka nepakitusi.
- Natūralią paskolų portfelio amortizaciją kompensuoja suteiktos naujos paskolos, todėl paskolų portfelis, skaičiuojamas bendrąja verte, nekinta.
- Jei per testuojamą laikotarpį buvo uždirbta pelno, tai jis naudojamas kapitalui didinti.

- Bankai nemoka dividendų ir nedidina kapitalo jokiais kitomis priemonėmis.
- Pagal riziką įvertinto turto pokyčius lemia tik paskolų portfelio kokybės pokyčiai.
- Tariama, kad bankų priežiūros ir valdžios institucijos nesiima veiksmų ekonomikos sukrėtimo padariniams švelninti.
- Neįtraukiami galimi pačių bankų strateginiai sprendimai ir jų įtaka kapitalo pakankamumo rodikliui.

### *Ekonometrinis modeliavimas*

Komercinių bankų patiriami nuostoliai, susiję su kredito rizika, daro didžiausią įtaką jų turto kokybei ir atitinkamai — kapitalo pakankamumo rodikliui. Todėl atliekant makroekonominį TNS kredito rizikos modeliavimas laikomas vienu iš svarbiausių elementų, leidžiančių įvertinti galimą bankų mokumą ir stabilumą. Modeliuojami galimi konkretaus banko kredito nuostoliai, atsižvelgiant į sudarytą hipotetinį makroekonomikos scenarijų, t. y. analizuojami sąryšiai tarp kredito rizikos ir makroekonomikos kintamųjų. Visi makroekonominiai kintamieji taikomi kaip egzogeniniai modelio kintamieji, todėl jie ir nulemia testavimo nepalankiausiomis sąlygomis rezultatus.

Atliekant lygčių vertinimą, taikomas mažiausių kvadratų metodas. Įverčių reikšmingumas tikrinamas pasitelkiant *Newey-West* kovariacinės matricos įvertį, kuriuo atsižvelgiama į modelio paklaidų heteroskedastiškumą ir autokoreliuotumą. Egzogeniniai kintamieji parenkami taip, kad atitiktų ekonominę logiką ir gerai paašškintų duomenis, t. y. siekiama, kad įvertinti koeficientai būtų atitinkamo ženklo ir statistiškai reikšmingi.

Bankų veiklos pelnas yra itin svarbi bendro vertinimo dalis, nes pelnas gali kompensuoti didelę dalį patiriamų kredito nuostolių ir taip daryti reikšmingą poveikį galutiniam modeliavimo rezultatui. Dėl šios priežasties bankų pelningumo modeliavimas įtraukiamas į mokumo TNS. Bankų veiklos pelnas išskaidomas į šešias sudedamąsias dalis: 1) grynosios palūkanų pajamos; 2) grynosios paslaugų ir komisinių pajamos; 3) grynosios investicinės veiklos pajamos, 4) kitos veiklos pajamos, 5) operacinės išlaidos; 6) amortizacija. Vertinant šias eilutes atskirai, o ne visą veiklos pelną bendrai, galima tiksliau nustatyti ryšius su realiąja ekonomika, be to, matyti, kurios vertinamos eilutės pokyčiai daro didžiausią įtaką bankų pelningumui.

Pelningumo modeliavimui pasirinktas dinaminis panelinių duomenų (angl. *dynamic panel data*) modelis, nes jis leidžia įvertinti ryšius, kurie vienodai veikia visą bankų sistemą. Tai svarbu, nes atliekamas TNS iš viršaus į apačią, t. y. reikia

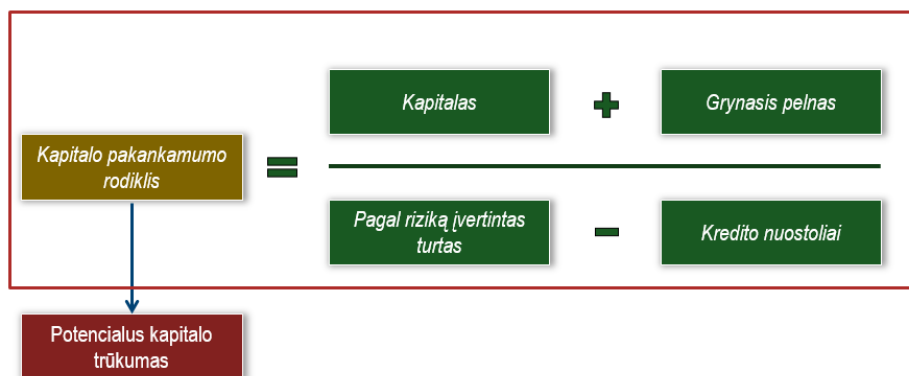
palyginti skirtingų bankų testavimo rezultatus. Be to, į modelį įtraukti banko specifiniai kintamieji ir nestebimas banko individualus efektas leidžia išskirti rezultatus neprarandant palyginamumo galimybes. Pažymėjus banką indeksu  $i$ , o ketvirtį — indeksu  $t$ , bendrai veiklos pelno ir išlaidų modeliavimas aprašomas šia lygtimi:

$$Y_{b,t} = \alpha + \eta_b + \beta Y_{b,t-1} + \sum_{j=1}^k \gamma_j M_{j,t} + \sum_{s=1}^l \delta_s B_{s,b,t} + \varepsilon_{b,t}, \quad (1)$$

čia  $Y_{b,t}$  – modeliuojama pelno (nuostolio) ataskaitos eilutė (grynosios palūkanų pajamos, grynosios paslaugų ir komisinių pajamos arba operacinės išlaidos),  $\eta_b$  – nestebimas banko individualus efektas;  $M_{j,t}$  – makroekonominiai kintamieji;  $B_{s,b,t}$  – banko specifiniai kintamieji.

#### Modeliavimo rezultatai

Visi ekonometriškai modeliuojami dydžiai yra įtraukiami į supaprastintą pelno (nuostolio) ataskaitą. Pelno (nuostolio) ataskaitos galutinis kintamasis yra grynasis pelnas. Jis vėliau nulemia banko kapitalo dinamiką, o ši savo ruožtu apibrėžia kapitalo pakankamumo rodiklio reikšmę.



2 pav.. Kapitalo pakankamumo vertinimas

Šaltinis: sudaryta autorių.

Kapitalo pakankamumo rodiklis yra esminis kintamasis, kuriuo remiantis apibendrinami TNS rezultatai ir nustatomas potencialus kapitalo trūkumas (žr. 2 pav.). Kapitalo pakankamumo rodikliui įtakos gali turėti ir pagal riziką įvertinto turto kaita. Ši kaita nėra tiesiogiai modeliuojama – atsižvelgiant į prielaidas, pagal riziką įvertinto turto kaita yra apibrėžta paskolų portfelio kokybės pokyčiais.

Disertacijoje pasiūlyta testavimo procedūra leidžia įvertinti ne konkretaus banko, o viso bankų sektoriaus būklę, tiksliau, jo mokumą, t.y. nustatyti, ar bankų sukauptos kapitalo atsargos yra pakankamos padengti paskolų portfelių nuvertėjimo nuostoliams, kurių galėtų susidaryti įvykus nepalankiems makroekonominės aplinkos

pokyčiams. Todėl sudarant testavimo procedūra daugiausia dėmesio skiriama ekonometriniais modeliams, leidžiantiems susieti makroekonominės aplinkos rodiklius su bankų veiklos rodikliais.

### 8.1.2 Bankų pelningumas

Šioje disertacijos dalyje detaliau nagrinėjamos Lietuvoje veikiančių bankų pelno (nuostolio) ataskaitos eilutės, t.y. grynosios palūkanų pajamos, paslaugų ir komisinių pajamos ir operacinės išlaidos, siekiant nustatyti vidinius ir išorinius veiksnius, kurie daro įtaką šių eilučių pasikeitimams ilgu ir trumpu laikotarpiu. Tyrime nagrinėjami duomenys nuo 2004 iki 2013, taigi apimamas prieš krizinis laikotarpis bei taip pat ir po krizinis laikotarpis.

Remiantis kitais tyrimais, kurie nagrinėjo kitų šalių bankų sektorius, pradiniam etape buvo atrinkta vienuolika paaiškinančiųjų kintamųjų. Dalis kintamųjų yra išoriniai, t.y. jie parodo bendrą ekonomikos situaciją šalyje. Į šią grupę patenka realusis Lietuvos BVP, eksportas, infliacija, trumpo laikotarpio palūkanų norma (3 mėnesių VILIBOR), nedarbo lygis. Kita grupė kintamųjų atspindi pačių bankų charakteristikas: banko turtas, paskolų portfelio dydis, kredito nuostolių ir paskolų portfelio santykio rodiklis, sudaryti atidėjiniai paskoloms. Be to, įtraukiamas Herfindalio-Hirschmano indeksas (HHI), kuris parodo Lietuvos bankų rinkos struktūrą, o tiksliau jos koncentraciją.

#### *Ekonometrinis modelis*

Kadangi bankų veiklos pelningumas ar tam tikros pelno (nuostolio) ataskaitos eilutės pasižymi pastovumu laike, todėl daugelis tyrimų šioje srityje taiko dinaminis panelinius modelius. Šiame tyrime taip pat taikomas dinaminis panelinis modelis, tačiau atsižvelgdami į tai, kad Lietuvos bankų sektorius yra heterogeniškas, norime įvertinti ilgalaikius sąryšius, kurie būtų bendri visiems bankams, o trumpo laikotarpio sąryšiai priklauso nuo kiekvieno banko atskirai. Taigi, šiam tyrimui geriausiai tinka Sugrupuoto vertinimo būdas (PMG), kuris buvo pasiūlytas Pesaran ir kt. (1997, 1999). PMG vertinimo būdas apriboja ilgo laikotarpio koeficientus, tačiau laisvasis narys, trumpo laikotarpio koeficientai ir paklaidų korekcijos koeficientas skiriasi kiekvienam bankui. Kaip teigė Haque (1999), nepaisant trumpo laikotarpio heterogeniškumo, galima gauti klaidingas išvadas apie ilgo laikotarpio sąryšius.

Tarkime, kad turime panelinį autoregresinį pasiskirsčiusių vėlinių (angl. *panel autoregressive distributed lag*) modelį  $ARDL(p, q_1, \dots, q_n)$ , pagal kurį priklausomi kintamieji yra paaiškinami savo pačių vėlyniais bei banko specifinių ar makroekonominų kintamųjų vėlyniais. Pažymėjus skerspjūvio duomenis  $i = 1, 2, \dots, N$  ir laiko periodą  $t = 1, 2, \dots, T$ , panelinis  $ARDL(p, q_1, \dots, q_n)$  modelis gali būti pateikiamas

tokia lygtimi:

$$y_{it} = \mu_i + \sum_{j=1}^p \lambda_{ij} y_{i,t-j} + \sum_{j=0}^{q_i} \delta'_{ij} B_{i,t-j} + \sum_{j=0}^{q_i} \gamma'_{ij} M_{t-j} + \varepsilon_{it}, \quad (2)$$

čia:  $y_{it}$  yra paaiškinamasis kintamasis (grynosios palūkanų pajamos, paslaugų ir komisinių pajamos, operacinės išlaidos),  $B_{it}$  - banko specifiniai kintamieji,  $M_t$  - makroekonominiai kintamieji,  $\mu_i$  - nestebimi individualus efektai.

Paėmus pirmus skirtumus  $ARDL(p, q_1, \dots, q_n)$  specifikacijos (2), galime perparametrizuoti modelį į:

$$\begin{aligned} \Delta y_{it} = & \mu_i + \phi_i y_{i,t-1} + \beta'_i B_{it} + \eta'_i M_t + \\ & + \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{q_i-1} \delta^*_{ij} \Delta B_{i,t-j} + \sum_{j=0}^{q_i-1} \gamma^*_{ij} \Delta M_{t-j} + \varepsilon_{it}, \end{aligned} \quad (3)$$

čia:  $\phi_i = -(1 - \sum_{j=1}^p \lambda_{ij})$ ;  $\beta_i = \sum_{j=0}^{q_i} \delta_{ij}$ ;  $\eta_i = \sum_{j=0}^{q_i} \gamma_{ij}$ ;  $\lambda^*_{ij} = -\sum_{m=j+1}^p \lambda_{im}$ ,  $j = 1, 2, \dots, p-1$ ;  $\delta^*_{ij} = -\sum_{m=j+1}^{q_i} \delta_{im}$ ,  $j = 1, 2, \dots, q_i-1$ ;  $\gamma^*_{ij} = -\sum_{m=j+1}^{q_i} \gamma_{im}$ ,  $j = 1, 2, \dots, q_i-1$ , t.y. koeficientai yra pradinių (2) lygties koeficientų funkcijos.

Toliau (3) lygtis gali būti pertvarkyta į panelinę paklaidų korekcijos lygtį, t.y. priklausomojo kintamojo pokyčiai yra paaiškinami skirtumais nuo ilgo laikotarpio pusiausvyros ir kitų kintamųjų trumpo laikotarpio pokyčiais:

$$\begin{aligned} \Delta y_{it} = & \phi_i (y_{i,t-1} - \alpha_i - \beta_i^* B_{it} - \eta_i^* M_t) - \\ & - \sum_{j=1}^{p-1} \lambda^*_{ij} \Delta y_{i,t-j} - \sum_{j=0}^{q_i-1} \delta^*_{ij} \Delta B_{i,t-j} - \sum_{j=0}^{q_i-1} \gamma^*_{ij} \Delta M_{t-j} + \varepsilon_{it}, \end{aligned} \quad (4)$$

čia:  $\phi_i$  - žymi paklaidų korekcijos koeficientą arba koregavimo į pusiausvyros reikšmę greitį; ilgo laikotarpio koeficientai:  $\alpha_i = -\frac{\mu_i}{\phi_i}$ ;  $\beta_i^* = -\frac{\beta_i}{\phi_i}$ ;  $\eta_i^* = -\frac{\eta_i}{\phi_i}$ .

Pesaran ir kt. (1999) teigė, kad PMG įvertinys yra suderintas ir efektyvus, jei galioja ilgo laikotarpio homogeniškumas. Norėdami patikrinti ilgo laikotarpio homogeniškumo hipotezę, atliekamas Hausmano testas, kur lyginami PMG ir grupės vidurkio (angl. *mean group*) MG įvertiniai. MG įvertinį pasiūlė Pesaran ir Smith (1999), o jame nėra ilgo laikotarpio koeficientų apribojimo. Pesaran ir Smith (1999) parodė, kad MG įvertinys pateikia suderintus ilgo laikotarpio koeficientus, tačiau yra neefektyvus, jei galioja ilgo laikotarpio koeficientų homogeniškumas.

#### *Modeliavimo rezultatai*

Empiriniai tyrimo rezultatai rodo, kad banko dydis yra svarbus ilgo laikotarpio lemiantys veiksnys tiek pajamos, tiek ir išlaidoms. Lietuvos bankų sektorius vis dar vystosi, todėl bankai dar nėra pakankamai dideli, kad susidurtų su masto neveiks-



mingumu, kuris buvo rastas kituose tyrimuose (pvz. Pasiouras ir Kosmidou, 2007; Coffinet ir Lin, 2010). Lietuvoje veikiantis bankai galėtų būti priskirti prie mažo ir vidutinio dyžio bankų, dėl to jie gali išnaudoti masto ekonomiją. Kaip ir tikėtasi, ekonomikos aktyvumas yra svarbus makroekonominis veiksnys, darantis įtaką banko pajamoms. Didėjantis BVP rodiklis skatina daugiau skolintis ir investuoti, kas lemia didesnes grynąsias palūkanų pajamas. Iš kitos pusės, didesnis ekonomikos aktyvas verčia bankus tenkinti didesnius transakcijų kiekius ir patirti didesnius paskolų priežiūros kaštus, kas didina jų operacines išlaidas. Šie sąryšiai atitinka Albertazzi ir Gambacorta (2009), Dietrich ir Wanzenried (2011) ir daugelio kitų tyrimų išvadas, kurie taip pat rado bankų pajamų procikliškumo savybę.

Darbe atliktas vertinimas rodo, kad palūkanų normos pokyčiai ir kredito nuostoliai daro trumpo laikotarpio įtaką grynosioms palūkanų pajamoms. Teigiamą palūkanų normos įtaką atspindi faktą, kad bankai turi rinkos galią greitai padidinti skolinimo normą. Kaip ir Athanasoglou ir kt. (2008) bei Coffinet ir Lin (2010), šiame tyrime randama, kad prastėjanti paskolų kokybė mažina bankų galimybes generuoti pajamas, todėl kredito nuostoliai neigiamai veikia grynąsias palūkanų pajamas. Empiriniai rezultatai taip pat rodo, kad realusis eksportas yra svarbus veiksnys, lemiantis grynąsias paslaugų ir komisinių pajamas. Taigi, valiutos keitimo paklausos ir bankinių operacijų paklausos pokyčiai daro įtaką bankų pajamoms ir paslugų ir komisinių. Trumpo laikotarpio sąryšis tarp operacinių išlaidų ir vidutinio darbo užmokesčio atspindi faktą, kad darbuotojų atlyginimai sudaro reikšmingą operacinių išlaidų dalį. Visi trumpalaikiai sąryšiai atitinka išankstinius teorinius lūkesčius.

Kitų tyrime naudotų bankų specifinių arba makroekonominių kintamųjų nustatyta įtaka buvo mažesnė arba jie iš viso nedarė įtakos bankų pelningumui. Pavyzdžiui, buvo nustatyta, kad HHI yra nereikšmingas veiksnys visoms trimis nagrinėtoms pajamų (nuostolio) ataskaitos eilutėms, todėl šis kintamasis nėra įtrauktas į galutinius lygtis. Šis rezultatas atitinka Claeys ir Vennet (2004) išvadas, kurie nustatė, kad koncentracija nėra svarbus veiksnys rytų Europos bankams.

Šiame tyrime naudotas metodas leidžia mums išskirti ilgalaikius ir trumpalaikius veiksnius, kurie daro įtaką bankų pajamoms ir išlaidoms. Šie rezultatai gali būti panaudoti, analizuojant Lietuvos bankų sektoriaus veiklos stabilumą. Vis dėlto, PMG vertinimo metodas reikalauja turėti pakankamai didelę duomenų imtį. Todėl didesnė Lietuvos bankų duomenų imtis leistų į modelį įtraukti daugiau kintamųjų bei turėti geresnį supratimą apie ilgalaikius ir trumpalaikius sąryšius. Šis klausimas galėtų būti sprendžiamas ateities tyrimuose.

## 8.2 Klasterinė analizė ir prognozavimas

### 8.2.1 Klasterinė analizė

Antroje disertacijos dalyje analizuojami Europos sąjungos šalių bankų duomenys. Pirmojo tyrimo tikslas susideda iš dviejų dalių. Pirmiausia siekiama nagrinėti turimos bankų imties grupavimą į klasterius, remiantis jų veiklos rezultatais 1999 – 2013 metai. Antrasis tikslas – palyginti skirtingus panašumo matus ir nustatyti, kuris iš jų geriausiai tinka bankinių duomenų klasterizavimui.

Šioje klasterinėje analizėje naudoti šeši bankų veiklos rodikliai. Buvo įtraukti trys pelningumo rodikliai: vidutinio turto graža (ROAA) (angl. *return on average assets*), vidutinės nuosavybės graža (ROAE) (angl. *return on average equity*) ir grynoji palūkanų normų marža (NII) (angl. *net interest margin*); veiklos efektyvumo rodiklis – išlaidų ir pajamų santykis (CIR) (angl. *cost to income ratio*); kredito kokybės rodiklis – atidėjinių paskolų nuostoliams ir paskolų portfelio grynąja verte santykis (LLP); banko rizikingumo/stabilumo rodiklis – kapitalo pakankamumo rodiklis (CAR) (angl. *capital adequacy ratio*). Šie bankų veiklos rodikliai yra pagrindiniai indikatoriai, kurie parodo, kokia yra situacija bankų sektoriuje.

Tyrimė naudojami metiniai nekonsoliduoti bankų finansinių ataskaitų duomenys, apimantis laikotarpį nuo 1999 iki 2013 metų. Duomenų imtis gauta iš Bureau van Dijk *Bankscope* duomenų bazės ir įtraukia komercinius, taupymo ir kooperatinius bankus, kurie veikia Europos sąjungos šalyje. Atlikus pradinių duomenų valyvą, toliau analizėje naudota duomenų imtis svyravo nuo 260 bankų (kapitalo pakankamumo rodiklis) iki 1332 bankų (vidutinio turto graža).

#### *Laiko eilučių klasterizavimo metodai*

Iš pradžių tariame, kad turima duomenų imtis yra daugiamatės laiko eilutės. Taigi, analizuojami duomenys yra tokios formos:

$$\mathbf{x}_t^{(i)} = (x_{1,t}^{(i)}, \dots, x_{d,t}^{(i)}), \quad t = 1, \dots, T; \quad i = 1, \dots, N.$$

čia indeksas  $i = 1, \dots, N$  žymi bankus, indeksas  $t$  žymi laiką (šiuo atveju metus), o  $d$  žymi tam tikrą bankų veiklos rodiklį. Kadangi duomenų imtyje esantys bankai turi ne visus rodiklius, todėl daugiausia taikysime klasterizavimą, remdamiesi vienu rodikliu. Be to, naudodami pelningumo ir veiklos efektyvumo rodiklius taikysime daugiamačio klasterizavimo metodus, t.y. klasterizuosime pagal keturis bankų veiklos rodiklius.

Kaip teigė Liao (2005) bei Batista ir kt. (2014), atliekant laiko eilučių klasterizavimą panašumo mato tarp dviejų laiko eilučių pasirinkimas yra vienas iš svarbiausių pasirinkimų, kurį reikia atlikti. Šioje analizėje naudojome šešis panašumo matus, skirtus laiko eilučių klasterizavimui.

Tyrime naudojamas Euklidinis atstumas kaip pradinis metodas skirtas laiko eilučių klasterizavimui. Tarkime, kad  $x = (x_1, \dots, x_T)$  yra  $i$ -tojo banko tam tikro rodiklio reikšmės, o  $y = (y_1, \dots, y_T)$  žymi  $j$ -tojo banko ( $i, j = 1, \dots, N$  ir  $i \neq j$ ) atitinkamo rodiklio reikšmės. Tada Euklidinis atstumas aprašomas taip:

$$D_{\lambda, \text{EUCL}}(x, y) = \left( \sum_{t=1}^T ((x_t - \lambda x_{t-1}) - (y_t - \lambda y_{t-1}))^2 \right)^{1/2}, \quad (5)$$

čia  $\lambda$  yra svertinis parametras. Darbe naudojami du variantai:  $\lambda = 0$  ir  $\lambda = 1$ . Tiksliau, antru atveju naudojama  $D_{\Delta, \text{EUCL}} = D_{0, \text{EUCL}}(x, y) + D_{1, \text{EUCL}}(x, y)$ .

Kitas naudojamas panašumo matas yra Chouakria and Nagabhushan (2007) pristatytas adaptyvus panašumo indeksas (angl. *adaptive dissimilarity index*), kuris remiasi adaptyvia reguliavimo funkcija ir atsižvelgia tiek į laiko eilučių kitimą laike tiek į laiko eilučių reikšmių artumą. Batista ir kt. (2011, 2014) pristatė nuo sudėtingumo nekintanti panašumo matą, kuris įvertina sudėtingumo skirtumus tarp dviejų laiko eilučių  $x$  ir  $y$ . Bohte ir kt. (1980), Geleano ir Peña (2000) bei keletas kitų autorių naudojo įvertintą autokoreliacijos funkciją siekdami išmatuoti skirtumą tarp dviejų laiko eilučių. Galiausiai taikomas dinaminis laiko iškreipimas (angl. *dynamic time warping*), kuris buvo pasiūlytas Berndt ir Clifford (1994). Šis atstumo matas yra populiarus ir dažnai taikomas laiko eilučių klasterizavime. Taigi, šie atstumo matai buvo taikyti vertinant dviejų laiko eilučių  $x$  ir  $y$  panašumą.

#### *Funkcinių duomenų klasterizavimo metodai*

Toliau tyrime laikoma, kad turimi duomenys yra stebimi kaip kreivės  $\mathbf{x}^{(i)}(t) = (x_1^{(i)}(t), \dots, x_d^{(i)}(t))$ ,  $t \in [0, T]$ ,  $i = 1, \dots, N$ , t.y. atsitiktiniai dydžiai aprašantys duomenis yra tolydaus laiko atsitiktinis procesas:

$$\mathbf{X}^{(i)}(t) = (x_1^{(i)}(t), \dots, x_d^{(i)}(t)), \quad t \in [0, T], \quad i = 1, \dots, N.$$

Tačiau realiai kreivės yra stebimos diskretaus laiko momentais:

$$\mathbf{x}_j^{(i)} = \mathbf{X}^{(i)}(j/T) + \varepsilon^i(j/T), \quad j = 1, \dots, T.$$

Funkcijas  $\mathbf{x}^{(i)}(t)$ ,  $t \in [0, T]$  atstatome, naudodami įvairias glodinimo technikas (pvz., žiūrėti Ramsey ir Silverman, 2005), tokiu būdu gaudami funkcinis duomenis:

$$\widehat{\mathbf{x}}^{(i)}(t), \quad t \in [0, T], \quad i = 1, \dots, N,$$

kurie yra naudojami funkcinų duomenų klasterinėje analizėje.

Tyrime taikome šešis atstumo tarp funkcinų duomenų matavimo variantus. Pir-

mu atveju, atstumas tarp dviejų kreivių vertinamas Hausdorff atstumu. Šis atstumo matas vertina didžiausią atstumą nuo taško vienoje kreivėje iki artimiausio taško kitoje kreivėje. Kiti du panašumo matai remiasi Ferraty and Vieu (2006), kurie pasiūlė dviejų etapų klasterizavimo algoritmą. Pirmame etape naudojami B-splainai, siekiant aproksimuoti funkcinius duomenys. Tarkime, kad turime B-splainų bazinių funkcijų aibę  $B = b_1, \dots, b_N$ . Tada B-splainais aproksimuotos kreivės  $q$  eilės išvestinė gali būti užrašoma taip:  $\hat{x}^{(q)} = \sum_{n=1}^N c_n B_n^{(q)}$ . Antrame etape vertinamas dviejų kreivių panašumas formule:

$$D_B(x, y) = \sqrt{\frac{1}{T} \int_T (\hat{x}^{(q)}(t) - \hat{y}^{(q)}(t))^2 dt}. \quad (6)$$

Šiame tyrime taikome du atvejus, t.y. vertiname kreivių panašumą, kada  $q = 0$  ( $D_{\text{BASIS}}$ ) ir kada  $q = 1$  ( $D_{\text{DERIV}}$ ).

Toliau tyrime pristatomi du nauji panašumo matai. Šie matai, mūsų žiniomis, iki šiol nebuvo taikyti funkcinių duomenų klasterizavime. Vienas iš matų remiasi funkcijų Hölderio savybe ir yra sudarytas iš dviejų dalių. Pirmoje dalyje vertinama, kaip arti viena kitos yra dvi funkcijos. Vertindami dviejų funkcijų artumą skaičiuojame supremumą tarp jų. Antroji mato dalis parodo, kaip panašiai funkcijos kinta. Taigi Hölder atstumo matas aprašomas lygtimi:

$$D_{\text{Hölder}}(x, y) = \sup_t |x(t) - y(t)| + \sup_{t \neq s} \frac{|(x(t) - y(t)) - (x(s) - y(s))|}{|t - s|^\alpha}, \quad (7)$$

čia  $\alpha \in (0, 1]$  vadinamas Hölderio eksponentė.

Antasis pasiūlytas matas yra konstruojamas panašiai. Pirmoje dalyje vertiname, kaip arti B-splainais aproksimuotos funkcijos yra arti viena kitos. Antroje dalyje, naudodami pirmos eilės išvestines vertinami, kaip panašiai funkcijos kinta. Šis panašumo matas aprašomas tokia taisykle:

$$D_{\text{SUP}}(x, y) = \sup_t |x(t) - y(t)| + \sup_t |x^{(q)}(t) - y^{(q)}(t)|. \quad (8)$$

Taigi pasiūlyti panašumo matai vertina tiek artumą tarp dviejų funkcijų, tiek ir jų kitimo panašumą.

Galiausiai naudojamas atstumo matas, kuris remiasi funkcijų išskaidymu į funkcines pagrindines komponentes. Taigi, kaip ir laiko eilučių atveju, naudojame šešis skirtingus atstumo tarp dviejų funkcijų matus.

#### *Daugiamačių duomenų klasterizavimo metodai*

Iki šiol pristatyti klasterizavimo metodai buvo taikoma vienmatėms laiko eilutėms

arba vienmatėms funkcijoms. Tačiau tyrime taip pat taikyti atstumo tarp daugiamatųjų duomenų skaičiavimo metodai. Pavyzdžiui, daugiamatis Euklidinis atstumas aprašomas formule:

$$D_{\lambda, \text{EUCL}}(\mathbf{x}, \mathbf{y}) = \left( \sum_{j=1}^d \sum_{t=1}^T ((x_{jt} - \lambda x_{j,t-1}) - (y_{jt} - \lambda y_{j,t-1}))^2 \right)^{1/2}. \quad (9)$$

Disertacijoje praplečiami Chouakria ir Nagabhushah (2007) (adaptyvus panašumo indeksas) bei Batista ir kiti (2011,2014) (nuo sudėtingumo nekintantis panašumo matas) pristatyti panašumo matai iš vienmačio atvejo į daugiamatį. Adaptyvus panašumo indeksas tarp daugiamatųjų laiko eilučių  $\mathbf{x}$  ir  $\mathbf{y}$  užrašomas taip:

$$\text{CORT}(\mathbf{x}, \mathbf{y}) = Q_{\mathbf{x}}^{-1/2} Q_{\mathbf{x}, \mathbf{y}} Q_{\mathbf{x}}^{-1/2},$$

čia

$$Q_{\mathbf{x}} = \sum_{t=1}^{T-1} (\mathbf{x}_{t+1} - \mathbf{x}_t)^{\tau} (\mathbf{x}_{t+1} - \mathbf{x}_t), \quad Q_{\mathbf{x}, \mathbf{y}} = \sum_{t=1}^{T-1} (\mathbf{x}_{t+1} - \mathbf{x}_t)^{\tau} (\mathbf{y}_{t+1} - \mathbf{y}_t).$$

Jei  $\lambda_{\max}(\text{CORT}(\mathbf{x}, \mathbf{y}))$  žymi didžiausią tikrinę matricos  $\text{CORT}(\mathbf{x}, \mathbf{y})$  reikšmę, tada daugiamatis adaptyvus panašumo matas aprašomas lygtimi:

$$D_{\text{CORT}}(\mathbf{x}, \mathbf{y}) = \phi_k[\lambda_{\max}(\text{CORT}(\mathbf{x}, \mathbf{y}))] \cdot D_{\text{EUCL}}(\mathbf{x}, \mathbf{y}), \quad (10)$$

čia  $\phi(u)$  yra eksponentinė adaptyvi reguliavimo (angl. *tuning*) funkcija.

Sudėtingumo įvertis  $CE(\mathbf{x})$  daugiamatį atveju aprašomas taip:

$$CE(\mathbf{x}) = \left( \sum_{j=1}^d \sum_{t=1}^{T-1} (x_{jt+1} - x_{jt})^2 \right)^{1/2}.$$

Tada nuo sudėtingumo nekintantis panašumo matas užrašomas lygtimi:

$$D_{\text{CID}}(\mathbf{x}, \mathbf{y}) = CF(\mathbf{x}, \mathbf{y}) \cdot D_{\text{EUCL}}(\mathbf{x}, \mathbf{y}), \quad (11)$$

čia  $CF(\mathbf{x}, \mathbf{y})$  yra sudėtingumo koregavimo faktorius.

Kiti panašumo matai ( $D_{\text{ACF}}$ ,  $D_{\text{DTW}}$ ,  $D_{\text{Hausdorff}}$ ,  $D_{\text{Hölder}}$ ,  $D_{\text{BASIS}}$ ,  $D_{\text{DERIV}}$ ,  $D_{\text{SUP}}$ ,  $D_{\text{FPCA}}$ ) daugiamatį atveju apskaičiuoti remiantis formule:

$$D_{\text{M}}(\mathbf{x}, \mathbf{y}) = \left( \sum_{j=1}^d [D_j(x, y)]^2 \right)^{1/2}, \quad (12)$$

čia  $D_j(x, y)$  yra pakoordinačiui apskaičiuoti panašumo matai.

### *Klasterinės analizės rezultatai*

Klasterinės analizės rezultatai parodė, kad panašumo mato pasirinkimas daro reikšmingą įtaką bankų grupavimui. Klasterių skaičiaus pasirinkimas taip pat priklauso nuo naudojamo panašumo mato ir nuo vertinimo metodo. Kaip teigė Batista ir kt. (2014) panašumo matas yra svarbiausias laiko eilučių klasterizavimo komponentas. Dėl to, naudinga turėti kelis atstumo tarp laiko eilučių ar funkcijų skaičiavimo matus ir palyginti gautus rezultatus. Remdamiesi vidutiniu silueto pločiu galime daryti išvadą, kad nėra vieno panašumo mato, kuris būtų geriausias visais tirtais atvejais. Vienu atveju, taikant panašumo matą, kuris remiasi pirma išvestine, ar taikant mūsų pasiūlytą panašumo matą, kur taikoma Hölderio sąlyga, galima gauti didžiausią vidutinį silueto plotį, t.y. gauname geresnius klasterizavimo rezultatus. Kitu atveju, naudingiau taikyti funkcinės pagrindines komponentes ir jomis paremtą panašumo matą. Vis dėlto, klasterizavimo metodai, kurie remiasi funkciniais duomenimis ir jų savybėmis dažniausia duoda geresnius klasterizavimo rezultatus nei metodai, kurie remiasi laiko eilučių savybėmis. Šiame tyrime naudojant laiko eilučių panašumo matus gavome santykinai prasčiausius klasterizavimo rezultatus. Kita išvada galėtų būti tai, kad paprastas Euklidinis atstumas yra geras atstumo matas atliekant bankų grupavimą į klasterius. Trečia išvada – remiantis vidutiniu silueto pločiu, mūsų pasiūlyti du panašumo matai buvo tarp labiausiai tinkančių bankinių duomenų klasterizavimui.

Be to, darbe nustatyta, kad nėra vienareikšmiško atsakymo, koks klasterių skaičiaus turėtų būti grupuojant bankus. Jei nagrinėjame CAR, LLP ar CIR rodiklius, tada skirtingi klasterių tikrinimo metodai siūlo suskirstyti bankų nuo 2 iki 4 grupių. Tačiau, jei nagrinėjame pelningumo rodiklius, tada rezultatai yra įvairūs ir klasterių skaičius galėtų būti pasirinktas remiantis ekspertiniu vertinimu.

Galima pastebėti, kad jei skaidytume bankus į 20 grupių, tokiu atveju gautume kelis klasterius sudarytus iš daug bankų, o kiti klasteriai būtų sudaryti iš vieno ar keleto bankų. Vėlgi, skirtingiems bankų rodikliams gauname nuo 6 iki 12 klasterių. Šie didesni klasteriai galėtų būti toliau analizuojami ir naudojami tam tikrų makroprudencinių įrankių sudarymui.

Daugiamatis klasterizavimas parodė, kad yra tikslinga grupuoti bankus į klasterius, remiantis jų veiklos pelningumo ir efektyvumo rodikliais. Gauname panašius vidutinio silueto pločio rodiklio rezultatus kaip ir vienmačio klasterizavimo atveju. Žinoma, jei analizuotume viename klasteryje esančių bankų atskirus rodiklius, tai matytume, kad tam tikrais atvejais bankai nebūtų sugrupuoti pagal vienmatį atvejį. Tačiau, tam tikrais atvejais yra naudinga taikyti daugiamačius klasterizavimo metodus, kai norima išskirstyti turimus laiko eilučių duomenis pagal kelias savybes.

## 8.2.2 Prognozavimas

Ketvirtojoje disertacijos dalyje nagrinėjama prognozavimo iš viršaus į apačią (angl. *top-down*) ir iš apačios į viršų (angl. *bottom-up*) palyginimo problema. Dažniausiai yra sunku iš anksto teoriškai pagrįsti, kuris variantas geriau tinka konkrečioms duomenims. Todėl šiame darbe pritaikomi abu prognozavimo variantai ir lyginamas jų gerumas prognozuojant bankų klasterio vidutinio kapitalo pakankamumo rodiklio reikšmes.

Iš pradžių nagrinėjamas teorinis bankų rodiklių raidos modelis. Tarkime, kad kiekvienam  $j = 1, \dots, N$  turime atsitiktinę funkciją  $X_j = (X_j(t), 0 \leq t \leq T)$ . Pasirinkus atskaitos taškus  $0 = \tau_0 < \tau_1 < \dots < \tau_n = T$ , kiekvienam  $j$  turima atsitiktinė imtis  $(X_j(\tau_k), k = 0, 1, \dots, n)$  ir nagrinėjama toks modelis:

$$X_j(\tau_k) = \beta_j X_j(\tau_{k-1}) + \boldsymbol{\gamma}'_j \mathbf{Z}_k + \varepsilon_{jk}, \quad k = 1, \dots, n, \quad (13)$$

čia  $\mathbf{Z}_k$  yra bendras paaiškinamasis  $p$ -dimensijos vektorius nepriklausomas nuo balto triukšmo proceso  $(\varepsilon_{jk}, j, k \geq 0)$ :  $\mathbf{E} \varepsilon_{jk} = 0$  ir  $\mathbf{E} \varepsilon_{ik} \varepsilon_{j\ell} = \sigma^2 \delta_{ij} \delta_{k\ell}$ , kur  $\delta_{st} = 0$  jei  $s \neq t$  ir  $\delta_{ss} = 1$ . Tariame, kad  $\beta_j$  yra nepriklausomi vienodai pasiskirstę atsitiktiniai dydžiai su bendru pasiskirstymu intervale  $(-a, a)$ ,  $\{\boldsymbol{\gamma}_j\}$  yra nepriklausomi vienodai pasiskirstę  $p$ -mačiai atsitiktiniai vektoriai su baigtiniu vidurkiu. Tariame, kad yra bendras nepriklausomumas tarp  $\beta_j$ ,  $\{\boldsymbol{\gamma}_j\}$  ir  $\{\varepsilon_j\}$ .

Nežinomi parametrai gali būti randami, naudojant mažiausių kvadratų metodą:

$$(\widehat{\beta}_j, \widehat{\boldsymbol{\gamma}}_j) := \underset{(b, \mathbf{c})}{\operatorname{argmin}} \sum_{k=1}^n \left[ X_j(\tau_k) - b X_j(\tau_{k-1}) - \mathbf{c}' \mathbf{Z}_k \right]^2. \quad (14)$$

Apibrėžiamas agreguotas procesas:

$$\overline{X}_N(t) = N^{-1} \sum_{j=1}^N X_j(t), \quad t \in [0, T].$$

Toliau pateikti rezultatai apibūdina asimptotinę proceso  $\overline{X}_N(\tau_k)$  elgesį, kada  $N \rightarrow \infty$ .

**1 Teorema.** *Tarkime, kad  $X_j(0) = 0$ . Tada kiekvienam  $1 \leq k \leq n$  galioja*

$$\lim_{N \rightarrow \infty} \mathbf{E} \left[ |\overline{X}_N(\tau_k) - Y_k|^2 \right] = 0,$$

čia

$$Y_k = \sum_{i=0}^{k-1} \mathbf{E}(\beta_1^i) \mathbf{E}(\boldsymbol{\gamma}'_1) \mathbf{Z}_{k-i}.$$

Remiantis Teorema 1 galime pastebėti, kad asimptotini agreguotų atsitiktinių funkcijų  $(\bar{X}_N(t), t \in [0, T])$  baigtinės dimensijos pasiskirstymo elgesį kontroliuoja paaiškinamieji kintamieji  $Z_1, \dots, Z_p$  ir koeficientais  $a_1, \dots, a_p$ , kai  $N$  yra didelis pagal tikimybę

$$(\bar{X}_N(\tau_k))_{1 \leq k \leq n} \approx \left( \sum_{i=0}^{k-1} a'_i \mathbf{Z}_{k-i} \right)_{1 \leq k \leq n}.$$

Tuo atveju, kai  $(X_j(\tau_k), k = 0, 1, \dots, n)$  atitinka atsitiktinių koeficientų regresijos modelį:

$$X_j(\tau_k) = \boldsymbol{\gamma}'_j \mathbf{Z}_k + \varepsilon_{jk}, \quad k = 1, \dots, n,$$

remiantis didžiųjų skaičių dėsnio turime:

$$\bar{X}_N(\tau_k) \xrightarrow[n \rightarrow \infty]{P} \mathbf{E}(\boldsymbol{\gamma}'_1) \mathbf{Z}_k$$

Agreguotam procesui  $X = P - \lim_{N \rightarrow \infty} \bar{X}_N$ , remdamiesi 1 teorema galime nagrinėti funkcinį regresijos modelį:

$$X(t) = \sum_{j=1}^p \int_0^t Z_j(t-s) \beta_j(s) ds + \varepsilon(t), \quad t \in [0, T], \quad (15)$$

čia funkcijos  $\beta_1, \dots, \beta_p$  yra nežinomi modelio parametrai, o  $Z_1, \dots, Z_p$  yra paaiškinamosios funkcijos, kurios gali būti deterministinės arba atsitiktinės. Pagrindinis skirtumas tarp šio modelio ir klasikinių funkcinės regresijos modelių yra tai, kad neturime daugiau informacijos išskyrus vieną atsako funkcijos  $X$  ir kiekvieno paaiškinamojo kintamojo  $Z_1, \dots, Z_p$  realizaciją. Iš šio vieno stebėjimo mums reikia įvertinti modelio parametrus ir padaryti statistines išvadas.

Siekiant įvertinti parametrus  $\beta_j$  yaikoma tokia procedūra. Visų pirma tariama, kad

$$\beta_j(t) = \sum_{k=1}^d \beta_{jk} u_k(t),$$

čia  $u_1(t), \dots, u_d(t)$  yra tam tikrų funkcijų aibė, tokiu būdu (15) modelis perrašomas į:

$$X(t) = \sum_{j=1}^p \sum_{k=1}^d \beta_{jk} y_{jk}(t) + \varepsilon(t), \quad t \in [0, 1], \quad (16)$$

čia

$$y_{jk}(t) = \int_0^t Z_j(t-s) u_k(s) ds.$$



Pažymint

$$Y(t) = (y_{11}(t), \dots, y_{pd}(t))', \quad B = (\beta_{11}, \dots, \beta_{pd})'$$

Perrašome (16) modelį į kompaktiškesnę formą:

$$X(t) = Y'(t)B + \varepsilon(t), \quad t \in [0, T]. \quad (17)$$

Norint įvertinti  $B$  imami atskaitos taškai  $\tau_k, k = 1, \dots, n$  ir turima:

$$X(\tau_k) = Y'(\tau_k)B + \varepsilon(\tau_k), \quad k = 1, \dots, n. \quad (18)$$

Tada  $B$  mažiausių kvadratų įvertinys yra

$$\hat{B} = \left( \sum_{k=1}^n Y(\tau_k)Y'(\tau_k) \right)^{-1} \sum_{k=1}^n Y(\tau_k)X(\tau_k). \quad (19)$$

#### *Atvejo analizė*

Atlikdami empirinį tyrimą, naudojame metinius nekonsoliduotus bankų kapitalo pakankamumo rodiklio (KPR) duomenis nuo 1999 iki 2013 metų. Duomenų imtis gauta iš Bureau van Dijk *Bankscope* duomenų bazės. Egzogeniniai kintamieji: Europos sąjungos BVP ir Eonia vienos nakties palūkanų norma yra gauti iš Eurostato ir Europos centrinio banko duomenų bazių. Iš pradžių turimi duomenis yra suglodinami, naudojant B-spline funkcijas, o tada iš bankų, pasinaudojant klasterine analize, sudaromos grupės, kurios pasižymi panašia kapitalo pakankamumo rodiklio raida. Toliau kiekvienas klasteris analizuojamas atskirai.

Modelis, kuris skirtas iš apačios į viršų prognozavimo atveju užrašomas lygtimi:

$$X_j(\tau_k) = \alpha + \beta_j X_j(\tau_{k-1}) + \sum_{i=1}^2 \gamma_i Z(\tau_k) + \varepsilon_j(\tau_k), \quad (20)$$

čia  $X_j(\tau_k)$  – bankų kapitalo pakankamumo rodiklis,  $Z(\tau)$  – egzogeniniai paaiškinamieji kintamieji (šiuo atveju EU BVP augimas ir Eonia palūkanų norma).

Turėdami kiekvieno banko įvertį, sudarome konkretaus klasterio vidurkio raidos įvertinį:

$$\hat{\bar{X}}_{BU} = \frac{1}{N} \sum_{j=1}^N \hat{X}_j \quad (21)$$

Panašiai kaip ir iš apačios į viršų atveju, taip ir iš viršaus į apačią atveju nagrinėjama lygti:

$$\bar{X}(\tau_k) = \alpha + \beta \bar{X}(\tau_{k-1}) + \sum_{i=1}^2 \gamma_i Z(\tau_k) + \varepsilon(\tau_k), \quad (22)$$

čia  $\bar{X}(\tau_k)$  yra vidutinės klasterio KRP reikšmės, t.y. tiesiogiai vertinamas vidurkio procesas.

Trečiasis modelis, kuris yra vertinamas darbe, yra šioje disertacijoje pasiūlytas funkcinės regresijos modelis:

$$X(t) = \sum_{j=1}^p \int_0^t Z_j(t-s)\beta_j(s)ds + \varepsilon(t), \quad t \in [0, T], \quad (23)$$

čia funkcijos  $\beta_1, \dots, \beta_p$  yra nežinomi modelio parametrai, o  $Z_1, \dots, Z_p$  yra regresoriai.

Atliekant modelių prognozavimo vertinimą, turimi duomenys skaidomi į mokymo imtį (duomenis iki 2011 metų pabaigos) ir testavimo imtį (duomenis nuo pirmo 2012 metų duomenis iki 2013 metų pabaigos). Tada įvertiname iš apačios į viršų modelį ir iš viršaus į apačią modelį imdami tiek taškų ( $\tau_k$ ), kad imtis atitiktų metinius, pusmetinius, ketvirtinius arba mėnesinius duomenis. Tokiu būdu atitinkamai turime prognozuoti 2, 4, 8 ir 24 stebėjimus. Funkcinės regresijos modelis taip pat vertinamas, turint duomenis iki 2011 metų pabaigos. Tada sudaromos prognozuojamos funkcinio proceso reikšmės iki duomenų imties pabaigos. Vertinant prognozės tikslumą imami 2, 4, 8 ir 24 taškai, kurie atitiktų iš apačios į viršų ir iš viršaus į apačią modeliais prognozuojamus stebėjimus.

#### *Prognozavimo dalies išvados*

Vertindami pirmąjį modelį pasinaudojama vienu iš privalumu, kurios teikia funkciniai duomenis, t.y. galimybė tolydžią kreivę diskredituoti į bet kokį kiekį taškų. Vertinant modelių tikslumą, abu modeliai (iš apačios į viršų ir iš viršaus į apačią) duoda panašius rezultatus. Be to, abiem atvejais autoregresinis koeficientas artėja prie 1, didinant stebėjimų skaičių. Vertinant modelių prognozes, matome, kad abiejų modelių tikslumas yra panašus. Trim atvejais iš šešių šiek tiek tikslesnes prognozes duoda iš viršaus į apačią modelis, tačiau kitais trim atvejais iš apačios į viršų modeliu gautos prognozės buvo tikslesnės.

Funkcinis regresijos modelis ir pasiūlytas jo vertinimo metodas yra naudingas aprašant bankinių duomenų pokyčius. Pasiūlytas modelis gana tiksliai aprašo realius duomenis imties pradžioje, tačiau vėliau skirtumas tarp pradinių duomenų ir įvertintų reikšmių šiek tiek padidėja. Bet kurio atveju, modelis gana gerai aprašo pagrindinius stochastinio proceso raidos aspektus. Pasiūlytas funkcinės regresijos modelis išsiskiria tuo, kad galime įvertinti priklausomybę tarp funkcinų duomenų, turėdami tik vieną stochastinio proceso realizaciją. Šiuo modeliu gautų prognozių tikslumas yra palyginamas, t.y. panašus, su prognozėmis, gautomis labiau tradiciniais vertinimo būdais.

Apibendrinant, šiame darbe pateikti įvairių modelių vertinimo rezultatai rodo,

kad analizuojant bankinius duomenis gali būti naudinga pasitelkti funkcinių duomenų analizės metodus bei pasinaudoti tokių duomenų tam tikromis savybėmis. Tokie metodai gali suteikti papildomų įžvalgų apie sąryšius, kurie yra svarbūs bankiniame sektoriuje.

## 9 Publikacijos

Pristatomos disertacijos rezultatai publikuoti 5 moksliniuose straipsniuose. Keturi iš jų jau yra publikuoti, o vienas pateiktas publikavimui specializuotame recenzuojamame moksliniame žurnale.

Publikuoti straipsniai:

1. Naruševičius L. (2013). *Modelling banks' profitability using dynamic panel data estimation*. Social technologies, Vol. 3(2), p. 278-287. ISSN 2029-7564 (Online).
2. Butkus V. and L. Naruševičius (2015). *Lietuvos bankų sistemos makroekonominis testavimas nepalankiausiomis sąlygomis*. Pinigų studijos, No. 1, p. 74-93. ISSN 1648-8970 (Online).
3. Naruševičius L. and A. Račkauskas (2016). *Comparing dissimilarity measures: a case of banking ratios*. Informatica, Vol. 27(3), p. 649-672. ISSN 1822-8844 (Online), (ISI WoS).
4. Naruševičius L. (2017). *Bank profitability and macroeconomy: evidence from Lithuania*. Technological and Economic development of Economy. ISSN 2029-4921 (Online), (ISI WoS).

Straipsniai pateikti publikavimui:

1. Naruševičius L. and A. Račkauskas (2017). *Forecasting with functional data: case study*.

Disertacijoje gauti rezultatai buvo pristatyti šiose mokslinėse konferencijose ir seminaruose:

1. Naruševičius L. *Modelling banks' profitability using dynamic panel data estimation*. International Conference Social Technologies'13. October 10-11, 2013, Vilnius, Lithuania.
2. Naruševičius L. and A. Račkauskas. *Comparing dissimilarity measures: a case of banking ratios*. X<sup>th</sup> Tartu Conference on Multivariate Statistics. 27 June - 1 July, 2016, Tartu, Estonia.
3. A. Račkauskas and L. Naruševičius. *Forecasting with functional data: case study*. 17<sup>th</sup> Applied Stochastic Models and Data Analysis International Conference, 6-9 June, 2017, London, United Kingdom.

Dalyvavimas konferencijose, kurios yra susijusios su šios disertacijos tema ir buvo naudingos didinant supratimą apie šios srities mokslinių tyrimų klausimus:

1. *3<sup>rd</sup> Annual European Stress testing Forum for Banks*, September 13-14, 2014, Amsterdam, The Netherlands.
2. *Conference on Financial Stability and Stress Testing*, June 3-5, 2015, Basel, Switzerland.

## 10 Apie autorių

Laurynas Naruševičius gimė 1985 m. kovo mėn. 8 d., Kėdainiuose.

### *Išsilavinimas*

**2012 – 2016 m.:** Vilniaus Universitetas, Matematikos ir informatikos fakultetas, matematikos doktorantūros studijos.

**2008 – 2010 m.:** Vilniaus Universitetas, Matematikos ir informatikos fakultetas, magistras. Ekonometrijos studijų programa

**2004 – 2008 m.:** Vilniaus Universitetas, Matematikos ir informatikos fakultetas, bakalauras. Ekonometrijos studijų programa

**2007 rudens semestras:** Atėnų Ekonomikos ir Verslo Universitetas, ERASMUS programa

**2000 – 2004 m.:** Kėdainių šviesioji gimnazija (baigta su pagyrimu)

## 11 Summary

After the Global financial crisis of 2007-2008 there has been a strong interest of research that is related with financial system or research that is dealing with the banking system. One of the main topics is how to estimate the stability of the banks and the whole banking system. In the first part of this thesis we were working with the Lithuanian banking data and tried to address the riskiness and profitability issues of the system. In the second part, we considered the data from European banks and proposed clustering and forecasting methods that could be applied to these data. The problems, considered in this thesis, help to get some new insights into this stream of research and into the field of functional data analysis.

In the thesis, we presented a macroeconomic top-down stress testing methodology, which is used to assess the resilience of the Lithuanian banking sector. In this work, we briefly introduced a stress testing methodology and described the main components of it. Then we reviewed the macroeconomic scenario design. The main focus, however, was on the development and estimation of the so-called satellite models, which help link the dynamics of macroeconomic variables to credit risk and profitability of a bank. It is worth noting that the focus was on testing bank solvency, i.e. assessing the potential insolvency of the bank due to regulatory capital shortages in the adverse macroeconomic scenario.

The presented methodology is used to run a top-down stress test. It means that the models and assumptions used are the same for all banks, which makes the results comparable across different banks. On the other hand, such an approach does not fully capture the specific aspects of individual banks. The proposed stress testing methodology is in line with the literature on stress testing and is capable of producing robust and meaningful results.

In the thesis we further analyzed the Lithuanian banking data and pay closer attention to the profitability of our banks. We examined the long-term and short-term relationship between bank profitability and bank-specific or macroeconomic variables. In the analysis, we applied a pooled mean group estimator, developed by Pesaran et al. (1997, 1999), which constrains long-term coefficient across cross-sectional units and at the same time allows intercept, short-term coefficients and adjustment to the equilibrium relationship to differ. The empirical results supported the idea that the Lithuanian banking sector is still developing and banks could be attributed to small and medium size banks and, therefore, they can exploit the economies of scale and scope. Further, in line with other studies, we found a procyclical behaviour of banks revenues and expenses, i.e. GDP development is an important determinant which makes the impact on banks' profits. However, a larger

data set for Lithuanian banks would help us to include more determinants in the model and have a better understanding of the long-term and short-term relationship.

In the second part of the thesis, we considered clustering of the banks according to their performance ratios. We approached the problem from two perspectives. One way was to look at the data as time series and apply dissimilarity measures that are based on raw data or some properties of time series. Another approach taken here was to consider that our data is continuous functions (i.e. functional data) and then we used dissimilarity measures, based on the functional data properties. In the thesis we have introduced two dissimilarity measures that are based on functions and take into account not only how close the two curves are, but also how similarly they change over time. In addition to the univariate clustering, where banks are grouped into clusters according to one bank performance ratio, a multivariate clustering is applied, where banks are clustered based on several ratios. For that reason, we extended several dissimilarity measures from the univariate case into a multivariate case.

The results of cluster analysis show that there is no clear answer which dissimilarity measure is best for banking data and how many clusters we get. However, the clustering, based on distance measures that used data as functions mostly outperformed distance measures, based on time series properties. Since in the cluster analysis data are unlabelled, the number on clusters is not known a priori. Division of banks into 20 clusters has revealed that there are from 6 to 12 clusters (depending on the banking ratio) that are formed by a larger number of banks and other clusters are formed by one or several banks. Those larger clusters could be taken as the basis for further research.

In the last part of the thesis, we considered the top-down versus bottom-up forecasting problem. We explored the advantages of the functional data and estimated top-down and bottom-up models to forecast a future development of the European banks' capital adequacy ratio. Furthermore, in this work, we proposed a novel functional regression model and introduce its estimation method. This model could be used to estimate the functional relationship between one realization of the stochastic process and other functional covariates.

The estimation of the models showed that, in this case, there are no significant difference between the top-down and bottom-up approaches while forecasting a future development of the mean process. The autoregressive models with exogenous variables and functional regression models provided similar results in terms of RMSE and MAE. Moreover, the proposed functional regression model was able to capture the main characteristic behaviour of the banking data and the result of out-of-sample forecasting performance was comparable with more traditional estimation techniques.

es. Nevertheless, a further analysis of the model and more empirical applications is still necessary.

To sum up, in the thesis, various aspects of the banking data analysis were considered. The stress testing, profitability, clustering and forecasting problems were discussed and analyzed. The results that were received here help to add some new insights into the stream of research that is dealing with banking data, in particular concerning stability of the banking system, as well as into the field of functional data analysis. Of course, future work and improvements are still needed. For example, it would be useful to apply quantile regression in stress testing procedure, allowing the estimation of nonlinear relationships between macroeconomic variables and bank indicators. Since various banking ratios are overlapped, it would be also interesting to apply fuzzy clustering algorithms to the banking data. These and other issues could be addressed in the future studies.