

TALLINNA TEHNIKAÜLIKOOL
Majandusteaduskond
Majandusanalüüsi ja rahanduse instituut

Artur Kohv

**EESTI TÖÖSTUSETTEVÕTETE KAPITALI STRUKTUURI
SEOSSED KASUMLIKKUSEGA**

Magistritöö

Õppekava ärirahandus ja majandusarvestus, peaeriala ärirahandus

Juhendaja: Kaido Kepp, MA

Tallinn 2021

Deklareerin, et olen koostanud lõputöö iseseisvalt ja olen viidanud kõikidele töö koostamisel kasutatud teiste autorite töödele, olulistele seisukohtadele ja andmetele, ning ei ole esitanud sama tööd varasemalt ainepunktide saamiseks. Töö pikkuseks on 14 699 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Artur Kohv

(allkiri, kuupäev)

Üliõpilase kood: 176682TARM

Üliõpilase e-posti aadress: artur.kohv@gmail.com

Juhendaja: Kaido Kepp, MA:

Töö vastab kehtivatele nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(nimi, allkiri, kuupäev)

SISUKORD

LÜHIKOKKUVÕTE	5
SISSEJUHATUS	6
1. KAPITALI STRUKTUURI JA KASUMLIKKUSE SEOS	9
1.1. Teoreetilised lähtepunktid	9
1.1.1. Modigliani ja Milleri kapitali struktuuri irrelevantsuse teooria	9
1.1.2. Kompromissiteooria	11
1.1.3. Agendi- ja vaba rahavoo agendikulude teooriad	13
1.1.4. Finantshierarhia teooria	14
1.1.5. Kapitali struktuuri teooriate kokkuvõte	15
1.2. Empiirilised uurimused	16
1.2.1. Uurimused Eestis	16
1.2.2. Uurimused teistes lääneriikides	17
2. ANDMED JA METOODIKA	22
2.1 Andmed	22
2.1.1. Uurimisobjekt	22
2.1.2. Valim ja valimi korrigeerimine	22
2.1.3. Kasutatud näitajad	24
2.1.4. Tunnuste kirjeldav statistika	28
2.2. Kasutatav meetoodika	30
2.2.1. Mitmene lineaarne regressioon	32
2.2.2. Otsustusmetsa algoritm	34
2.2.3. Otsustusmetsa mudeldamise protsess	38
3. TULEMUSED JA JÄRELDUSED	44
3.1. Lineaarse regressiooni mudelite tulemused	44
3.2. Otsustusmetsa mudelite tulemused	47
3.2.1. Koguvõla ja kasumlikkuse seosed	49
3.2.2. Pikaajalise võla ja kasumlikkuse seosed	51
3.2.3. Lühiajalise võla ja kasumlikkuse seosed	53
3.3. Järeldused ja ettepanekud	55
KOKKUVÕTE	65
SUMMARY	68
KASUTATUD ALLIKATE LOETELU	71
LISAD	80

Lisa 1. Valimisse lülitatud NACE Rev. 2 koodid ja nende tähendus	80
Lisa 2. ROE andmestiku korrelatsioonimaatriks	81
Lisa 3. ROA andmestiku korrelatsioonimaatriks	82
Lisa 4. Lineaarse regressiooni testide tulemused	83
Lisa 5. Otsustusmetsa ROE andmestiku treening- ja testandmestiku kirjeldav statistika	84
Lisa 6. Otsustusmetsa ROA andmestiku treening- ja testandmestiku kirjeldav statistika	85
Lisa 7. Otsustusmetsa ROE treening- ja testandmete ROE suhtelise sageduse graafik	86
Lisa 8. Otsustusmetsa ROA treening- ja testandmete ROA suhtelise sageduse graafik	87
Lisa 9. R skript otsustusmetsa I mudeli näitel	88
Lisa 10. Otsustusmetsa I mudeli jääkliikmete histogramm ja hajususgraafik	92
Lisa 11. Otsustusmetsa II mudeli jääkliikmete histogramm ja hajususgraafik	94
Lisa 12. Otsustusmetsa III mudeli jääkliikmete histogramm ja hajususgraafik	96
Lisa 13. Otsustusmetsa IV mudeli jääkliikmete histogramm ja hajususgraafik	98
Lisa 14. Otsustusmetsa V mudeli jääkliikmete histogramm ja hajususgraafik	100
Lisa 15. Otsustusmetsa VI mudeli jääkliikmete histogramm ja hajususgraafik	102
Lisa 16. I regressioonimudeli jääkliikmete histogramm	104
Lisa 17. II regressioonimudeli jääkliikmete histogramm	105
Lisa 18. III regressioonimudeli jääkliikmete histogramm	106
Lisa 19. IV regressioonimudeli jääkliikmete histogramm	107
Lisa 20. V regressioonimudeli jääkliikmete histogramm	108
Lisa 21. VI regressioonimudeli jääkliikmete histogramm	109
Lisa 22. Hüperparameeter mtry ristvalideerimisega optimeerimise tulemused	110
Lisa 23. Otsustusmetsa kontroll- ja põhimudelite treening- ja testvea täpsus	111
Lisa 24. Otsustusmetsa I mudeli (ROE-VK) tunnuste olulisus	112
Lisa 25. Otsustusmetsa II mudeli (ROE-PVK) tunnuste olulisus	113
Lisa 26. Otsustusmetsa III mudeli (ROE-LVK) tunnuste olulisus	114
Lisa 27. Otsustusmetsa IV mudeli (ROA-VK) tunnuste olulisus	115
Lisa 28. Otsustusmetsa V mudeli (ROA-PVK) tunnuste olulisus	116
Lisa 29. Otsustusmetsa VI mudeli (ROA-LVK) tunnuste olulisus	117
Lisa 30. Otsustusmetsa I mudeli kontrollnäitajate osalise sõltuvuse joonised	118
Lisa 31. Otsustusmetsa IV mudeli kontrollnäitajate osalise sõltuvuse joonised	119
Lisa 32. Otsustusmetsa II mudeli kontrollnäitajate osalise sõltuvuse joonised	121
Lisa 33. Otsustusmetsa V mudeli kontrollnäitajate osalise sõltuvuse joonised	122
Lisa 34. Otsustusmetsa III mudeli kontrollnäitajate osalise sõltuvuse joonised	124
Lisa 35. Otsustusmetsa VI mudeli kontrollnäitajate osalise sõltuvuse joonised	125
Lisa 36. Lihtlitsents	127

LÜHIKOKKUVÕTE

Magistritöö eesmärk on hinnata Eesti töötleva tööstuse ettevõtete kasumlikkuse ja võlataseme vahelise seose olulisust, suunda ja tugevust. Sealhulgas käsitletakse eraldi lühiajalise-, pikaajalise- ja koguvõla taseme seost kasumlikkusega. Valimi moodustasid üle 2000 unikaalse ettevõtte ja üle 12 500 koguvaatluse perioodil 2011-2019.

Seoste hindamiseks kasutati lineaarse regressiooni ühendatud mudelit ja valdkonnas vähe kasutatud otsustusmetsa (ingl *random forest*). Kahe erineva keerukusega meetodika rakendamine võimaldas eristada seoste piirkondi, mille osas võib järeldusi usaldusväärseks pidada, nendest piirkondadest, mille kohta magistritöö raames lõplikke järeldusi teha ei saanud.

Mõlema meetodi tulemusi arvesse võttes järeldus, et pikaajalise võla taseme ja kasumlikkuse vahel esineb positiivne seos, kui pikaajaline võlg moodustab varadest vähem kui 14%. Sama seos kehtib koguvõla taseme ja kasumlikkuse vahel, kuid seos on nõrgem. Lühiajalise võla taseme ja kasumlikkuse vahel on negatiivne seos, kui lühiajaline võlg moodustab varadest rohkem kui 6%.

Väiksema koguvõla ja pikaajalise võla taseme ning kasumlikkuse vahel leitud positiivne seos toetab agendikulude teooria argumenti, et teatud maani muudab võlg ettevõtte efektiivsemaks. Samuti toetab koguvõla ja pikaajalise võla tase kuni 14%-ni staatilise kompromissiteooria seisukohta, et väiksema võlaga kaasnevate riskidega ettevõtetel on rohkem võlga. Lühiajalise võla ja kasumlikkuse vahel leitud negatiivne seos toetab finantshierarhia teooriat, mille järgi eelistavad ettevõtted oma vahendite kasutamist.

Metoodikad andsid vastupidised tulemused piirkonnas, kus koguvõla või pikaajalise võla osakaal varadest ületab 14% või kui lühiajaline võlg moodustab varadest vähem kui 6%. Seetõttu ei saa nendel juhtudel tunnuste vahelise seose osas lõplikku järeldust teha.

Võtmesõnad: kapitali struktuur, kasumlikkus, töötlev tööstus, lineaarne regressioon, otsustusmets

SISSEJUHATUS

Kapitali struktuuri kui kasumlikkuse ühe võimaliku mõjuri juhtimine on oluline nii ettevõtte omanike kui ka juhtkonna jaoks. Samas on kapitali struktuuri teooriad eri seisukohtadel, milline on kasumlikkuse ja võlataseme vaheline seos ning ka empiiriliste uurimuste tulemused on selles küsimuses vastuolulised.

Staatilise kompromissiteooria kohaselt on seos võlataseme ja kasumlikkuse vahel positiivne, sest suurema kasumlikkusega ettevõtetel on väiksemad võlga kaasnevad riskid, mistõttu on võlakapital nende jaoks kättesaadavam ja seda kasutatakse rohkem (Frank, Goyal 2009). Finantshierarhia teooria kohaselt on aga seos negatiivne, kuna ettevõtted eelistavad kasutada esmalt jaotamata kasumit, seejärel võlga ja viimaseks raha kaasamist väliste investorite kaudu. Kuna kasumlikel ettevõtetel on rohkem jaotamata kasumit, on neil vähem võlga. (Myers, Majluf 1984)

Empiirilised uurimused on tuvastanud Euroopa ettevõtete kasumlikkuse ja võlataseme vahel nii negatiivse seose (nt Goddard *et al.* 2005), positiivse seose (nt Margaritis, Psillaki 2010) kui ka seose puudumise (nt eraettevõtetes Hanousek, Shamshur 2011). Samuti on leitud kasumlikkuse ja võlataseme vahel mitte-lineaarne seos (nt Pattitoni *et al.* 2014). Magistritöö panustab valdkonda võttes fookusse Eesti töötleva tööstuse sektori.

Uurimisobjekt on kitsendatud Eesti töötleva tööstuse ettevõtetele kahel põhjusel:

- 1) Ettevõtete tegevusala mõjutab nende kapitali struktuuri (Talberg *et al.* 2008; Michaelas *et al.* 1999) ja kasumlikkust (Pattitoni *et al.* 2014);
- 2) Eesti maksusüsteemi ja finantseerimistava unikaalsus.

Töötlev tööstus on üks Eesti olulisemaid sektoreid nii tööhõive, toodangu- kui ka ekspordimahu vaatenurgast. Samuti on töö autoril laialdased valdkonnaalased teadmised, mis võimaldavad näha ja tõlgendada uuritavaid seoseid asjatundlikult.

Magistritöö eesmärgiks on hinnata Eesti töötleva tööstuse ettevõtete kasumlikkuse ja võlataseme vahelise seose olulisust, suunda ja tugevust. Eesmärgi saavutamiseks on autor püstitanud järgnevad uurimisküsimused:

- 1) Milline on kogu võlakohustuste ja kasumlikkuse seos?

- 2) Milline on pikaajaliste võlakohustuste ja kasumlikkuse seos?
- 3) Milline on lühiajaliste võlakohustuste ja kasumlikkuse seos?

Töö empiirilises uurimuses on sõltuv muutuja kasumlikkus, mille mõõdikuna kasutab autor omakapitali tootlust (ROE) ja varade tootlust (ROA). Finantsvõimendust kirjeldavad muutujad on võlakohustuste ja varade jagatis ehk võlakordaja ja selle komponendid – pikaajaliste võlakohustuste kordaja ja lühiajaliste võlakohustuste kordaja. Kontrollnäitajad on ettevõtte vanus, omakapital määra, käibe, käibe aastane kasvumäär, põhivarade osakaal koguvaredest ja lühiajaliste kohustuste kattekordaja.

Empiirilise uurimuse läbiviimiseks kasutab autor kahte analüüsimeetodit: lineaarset regressiooni ning otsustusmetsa ehk juhuslikku metsa (ingl *random forest*, *random decision forest*). Lineaarne regressioon on kapitali struktuuri ja kasumlikkuse seoste uurimustes laialt kasutatud meetodika. Otsustusmets on masinõppe meetod, mis on võrreldes lineaarse regressiooniga uuem ja vähem levinud. Finantsvõimendust eri meetoditega mudeldanud Amini *et al.* (2021) leidsid, et otsustusmets saavutas võrreldes teiste masinõppe ja lineaarsete mudelitega, sh lineaarse regressiooniga, parema ennustustäpsuse. Otsustusmetsa üks eelis on suutlikkus tuvastada keerulisemaid mitte-lineaarseid seoseid (*Ibid*). Autorile teadaolevalt pole meetodit Eestis kapitali struktuuri ja kasumlikkuse seoste uurimiseks varem kasutatud.

Töö empiirilises osas kasutatakse Eesti töötleva tööstuse ettevõtete finantsandmeid aastatel 2011-2019. Andmete allikas on Orbis Europe andmebaas. Autor kasutab mõlema meetodika puhul sama andmestikku, kuid lineaarses regressioonis on andmed struktureeritud paneelandmetena ning otsustusmetsas ristanandmetena – sama ettevõtte erinevate aastate andmeid vaadeldakse eraldiseisvate vaatlustena. ROE-d sõltuva tunnuseks kasutavas lõplikus valimis on 12 568 vaatlust ja ROA-d sõltuva tunnuseks kasutavas andmestikus 12 703 vaatlust. ROE valimi moodustasid 2064 ja ROA valimi 2063 unikaalset ettevõtet 9-aastase perioodi jooksul.

Empiiriliseks analüüsiks otsustusmetsaga kasutab autor R programmeerimiskeelt (R Core Team 2021), kasutajaliidest RStudio (RStudio Team 2020) ning „randomForest“ teeki (Liaw, Wiener 2002), mis on Breimani (2001a) otsustusmetsa implementatsioon. Lineaarset regressiooni teostab autor kasutades vabavaraalist tarkvara Gretl.

Töö esimeses peatükis on antud ülevaade kapitali struktuuri teooriatest ja empiiriliste uurimuste järeldustest. Välja on toodud teooriates esitatud seisukohtade võrdlused ja vastuolud. Empiirilise kirjanduse ülevaates on asetatud rõhk uurimustele Euroopas ja tööstussektoris.

Töö teises peatükis on käsitletud andmeid ja metoodikat. Peatükis kirjeldatakse muu hulgas uurimisobjekti, andmestiku korrigeerimist, kasutatud tunnuseid ning põhjendatakse kontrollnäitajate ja meetodite valikut. Peatükis avatakse lineaarse regressiooni ja otsustusmetsa meetodeid ning mudeldamise protsessi.

Töö kolmandas peatükis on välja toodud lineaarse regressiooni ja otsustusmetsaga teostatud empiirilise uurimuse tulemused ja järeldused ning kahe meetodi tulemuste võrdlus. Samuti on kirjeldatud analüüsi kitsaskohad ning ettepanekud aspektide kohta, mis võivad olla aineks edasisteks uurimusteks.

Magistritöö tulemused on praktiliselt kasulikud Eesti tööstusettevõtete juhtidele ja omanikele kapitali struktuuri puudutavate juhtimisotsuste tegemiseks ning informatiivsed välistele osapooltele (finantseerijad, tarnijad, töötajad), kes saavad otsuseid tehes magistritöö tulemusi arvesse võtta.

Autor tänab juhendajat Kaido Keppi asjatundlike nõuannete ja perekonda igakülgse toetuse eest.

1. KAPITALI STRUKTUURI JA KASUMLIKKUSE SEOSED

1.1 Teoreetilised lähtepunktid

Kapitali struktuuri all peetakse silmas proportsiooni oma- ja võõrkapitali kasutamise vahel, mis moodustab ettevõtte kogukapitali. Neli peamist kapitali juhtimise ja ettevõtte äritegevuse finantseerimise võimalust on kasutada olemasolevat omakapitali (so eelmiste perioodide jaotamata kasum, lihtaktsiad), emiteerida uut omakapitali, kasutada erinevaid võlainstrumente (laenud, liisingud, võlakirjad jm) või hübriidkapitali (mezzanine, eelisaktsiad, vahetusvõlakirjad jm). Optimaalne kapitali struktuur viib ettevõtte kapitalikulud miinimumini ja maksimeerib seeläbi ettevõtte väärtust (Horne, Wachowicz 2008).

Valdkonnaalased teooriad püüavad selgitada, milline on sobilik kapitali struktuur ning millistest kaalutlustest lähtuvalt ettevõtte enda kapitali struktuuri valivad. Kapitali struktuuri teooriate kohaselt mõjutavad kapitali struktuuri valimist erinevad tegurid, nende seas maksud, ettevõtte varade struktuur, kasvutempo, ainulaadsus, valdkond, suurus, tulu volatiilsus ja kasumlikkus (Titman, Wessels 1988). Brealey *et al.* (2019) hinnangul ei ole olemas ühte kapitali struktuuri teooriat, mis hõlmaks kõikide ettevõtete kapitali struktuuri valikuid. On olemas aga mitmeid teooriaid, mis on kapitali struktuuri valimisel abiks olenevalt ettevõtte varast, ärist ja olukorrast (*Ibid*). Käesoleva töö aspektist on olulised eeskätt kapitali struktuuri teooriate argumendid kasumlikkuse ja kapitali struktuuri seoste kohta.

1.1.1. Modigliani ja Milleri kapitali struktuuri irrelevantsuse teooria

Optimaalset kapitali struktuuri puudutavale akadeemilisele diskussioonile pani aluse Modigliani ja Milleri 1958.a nn kapitali struktuuri irrelevantsuse teooria (ingl *capital structure irrelevancy theory*). Teooria esimese ja magistritöö kontekstis olulisima postulaadi kohaselt ei mõjuta ettevõtte kapitali struktuur selle väärtust. See tähendab, et täielikult omakapitaliga finantseeritud ettevõtte väärtus on võrdne ettevõttega, mis finantseerib enda tegevust kasutades nii võõrfinantseeringut kui omakapitali, *ceteris paribus*. Teooria kohaselt sõltub ettevõtte väärtus vaid selle oodatavatest tuleviku rahavoogude nüüdisväärtusest (ettevõtte j näitel) (Modigliani, Miller 1958):

$$V_j \equiv (S_j + D_j) = \frac{X_j}{\rho_k} \quad (1)$$

kus

V_j – ettevõtte j turuväärtus,

S_j – omakapitali (noteeritud ettevõtte puhul lihtaktsiate) turuväärtus,

D_j – võõrkapitali turuväärtus,

X_j – oodatav kasum enne intressikulu,

ρ_k – diskonteerimismäär, mis vastab k -klassi kuuluvate ettevõtete riskiastmele.

Teooria teiseks postulaadiks on ettevõtte kapitali hinna sõltumatus tema kapitali struktuurist – ettevõtte keskmine kapitali hind on teooria kohaselt võrdne (konkreetsesse klassi kuuluva) ettevõtte kapitaliseerimismääraga (*Ibid*):

$$\frac{X_j}{(S_j + D_j)} \equiv \frac{X_j}{V_j} = \rho_k \quad (2)$$

Teooria autorid näitavad, et olukordades, mil eeltoodud seosed ei kehti (st erinevate kapitali struktuuridega ettevõtete turuväärtused erinevad), tekivad arbitraaživõimalused, mille ekspluateerimise tulemusel ülehinnatud kapitali hind langeks ja alahinnatud kapitali hind tõuseks, taastades seeläbi ettevõtete turuväärtuste võrdsuse ja leitud seoste kehtivuse. (*Ibid*)

Brealey *et al.* (2019, 436) on võtnud Modigliani ja Milleri teooria kokku järgnevalt: finantsvõimendus ei ole võlutrikk, mis ettevõtte kapitali hinda vähendaks. Nende hinnangul kiputakse võlakapitali kasutamisel unustama, et võlaga kaasneb lisaks intressikulule ka kõrgema tootluse kulu, mida investorid kõrgema riski tõttu soovivad. Nende kulude tõttu ei teki finantsvõimenduse kasutamisest kapitalikulu vähenemist ning ettevõtte väärtus ei muutu. (*Ibid*)

Modigliani ja Milleri teooria oli valdkonnas teedrajavaks, pakkudes raamistiku, millest lähtuvalt järgmisi uuringuid läbi viia. Teooria kasutas probleemi lahendamiseks mitmeid olulisi lihtsustusi, mis tegelikkuses reeglina ei kehti ja mida võib pidada teooria peamiseks puuduseks. Nendeks on maksude, tehingu-, pankroti- ja agendikulude arvestamata jätmine ning turuosaliste seas info asümmeetria, täielikult efektiivsete finantsturude olemasolu ning üle maailma sarnaste ettevõtete samasse riskiklassi kuulumise eeldamine.

1963.a täiendasid Modigliani ja Miller enda varasemat teooriat, võttes arvesse maksude olemasolu ja sellega kaasneva „maksukilbi” efekti. Viimane seisneb selles, et klassikalist ehk täielikult mitte-integreeritud maksusüsteemi rakendatavates riikides (nt USA-s ja mitmetes teistes OECD riikides)

arvestatakse maksustatavast kasumist maha laenukohustustega kaasnevad intressikulud. Kuna väiksemal maksukoormusel on positiivne mõju ettevõtte rahavoogudele, järeldasid Modigliani ja Miller, et ettevõtte väärtus suureneb proportsionaalselt võlakapitali kasutamise määraga. Kuigi sellest järeldub, et optimaalne, so ettevõtte väärtust maksimeeriv kapitali struktuur koosneb täielikult võõrkapitalist, nendivad teooria autorid, et tegelikkuses ei pruugi ettevõtted sellist olukorda soovida. Näiteks võib piisava võimenduse korral olla täiendava võlakapitali hind ka tekkivat maksusoodustust arvesse võttes kallim kui ettevõtte jaotamata kasumil, mis tuleneb kreditoride piirangutest või investorite endi tulumaksukohustusest. Samuti võib juhatas soovida säilitada tulevikus suuremat paindlikkust. (Modigliani, Miller 1963)

Mitmed hilisemad tööd on aga välja toonud selle, et üksikisiku tasandi maksud loovad maksukilbist saadavale eelisele piirid ja võivad teha laenu võtmise isegi ebaatraktiivseks (Miller 1977; Farrar, Selwyn 1967; Stiglitz 1972). Jong *et al.* (2008) on leidnud, et kapitali struktuuri mõjurid on eri riikides erinevad, sealhulgas ka maksustamine ja selle mõju. Jensen ja Meckling (1976) on Modigliani ja Milleri maksusoodustuse argumenti kritiseerinud öeldes, et see pole ainus põhjus, miks ettevõtted võlga kasutavad. Autorid toovad välja, et ettevõtted kasutasid võlakapitali ka enne maksukilbi teket ning ettevõtte kasutab võlga ka siis, kui omakapital pole piisav, et võimalikke kasumlikke investeerimisvõimalusi kasutada.

Laenukapitali maksueelist Eesti maksusüsteemi kontekstis on uurinud Sander (2005), kelle järgi sõltub maksueelise olemasolu ja selle suurus ettevõtte omanike juriidilisest staatusest ja ettevõtte dividendipoliitikast. Sama autori 2014.a Konkurentsiametile koostatud raportis märgitakse, et Eestis maksustatakse ühekordselt nii oma- kui ka võõrkapitali väljamakseid, mistõttu maksukilpi ei teki (Sander 2014). Kokkuvõtlikult sõltub maksukilbi olemasolu riikide maksupoliitikast.

1.1.2. Kompromissiteooria

Modiglianit ja Millerit võib pidada nende 1963.a maksukilbiga täiendatud mudeli tõttu kompromissiteooria (ingl *trade-off theory*) pioneerideks. Modigliani ja Milleri alustatud kapitali struktuuri valiku diskursust arendasid pankrotikulude aspektist Robichek ja Myers (1966), Stiglitz (1972), Kraus ja Litzenberger (1973) ning Scott (1976). Kompromissiteooria järgi püüavad ettevõtted maksimeerida enda väärtust otsides tasakaalupunkti maksukilbist saadava tulu ning võimaliku pankrotikulu ja muude võlga kaasnevate kulude vahel (Myers 2015).

Otsesed pankrotikulud on näiteks pankrotti mineva ettevõtte administratiivsed- ja õiguskulud. Kaudne pankrotikulu on ettevõtte väärtuse vähenemine, kui ettevõtte ei suuda võlakohustusi täita. (Baxter 1967) Muud võlaga kaasnevad kulud võivad tekkida, kui kõrge võlakoormusega ettevõtte on sunnitud tegema negatiivse nüüdisväärtusega finantstehinguid (Myers 2015). Solomoni (1963) hinnangul suureneb võlakapitali suurenedes kapitali hind, kuna kõrge võlatasemega ettevõtte suhtes nõuab kreditor kõrgema riski tõttu kõrgemat tootlust. Seetõttu eelistavad ettevõtted Solomoni arvates optimaalset võörkapitali-omakapitali suhet.

Brealey *et al.* (2019, 482) on võtnud kompromissiteooria kokku järgnevalt: ettevõtte väärtus võrdub väärtusega, kui ettevõtte oleks täielikult omakapitaliga finantseeritud pluss võlast tulenev maksukilp miinus võlaga kaasnevad kulud. Ettevõtete maksukilbist tulenevate eeliste ja võlaga kaasnevate kulude varieeruvuse tõttu on ettevõtete võlataseme eesmärgid erinevad. Ettevõtted, millel on palju mitteriskantset materiaalselt põhivara ja maksustatavat tulu, peaksid valima kõrgema võlakapitali taseme. Vähem kasumlikud ettevõtted, millel on palju riskantset immateriaalselt põhivara, peaksid aga peamiselt toetuma omakapitali kasutamisele ning võtma vähem võlakapitali. (*Ibid*)

Võlaga kaasnevate kulude aspektist on Titman (1984) ning Maksimovic ja Titman (1991) argumenteerinud, et kõrge finantsvõimenduse tase võib vähendada nõudlust ettevõtte toodete järgi ja suurendada klientide, töötajate ja tarnijatega seotud kulusid, kuna kõrge võla tase vähendab nende usku ettevõtte elujõulisusse. Bolton ja Scharfstein (1990) on argumenteerinud, et kõrge võlatasemega ettevõtted võivad kannatada madalama võlatasemega konkurentide agressiivse strateogia tõttu.

Kompromissiteoorial on staatiline ja dünaamiline versioon. Staatilise kompromissiteooria järgi kaaluvad ettevõtted võlast tulenevaid maksueeliseid ja äriske. Kui tänu võlale maksude arvelt kokkuhoitud tulu ja võlast tulenev kulu on sama suured, on tegu optimaalse võlatasemega. Staatilise kompromissiteooria järgi liiguvad ettevõtted seega nende jaoks optimaalse konkreetse võlataseme poole. (Brennan, Schwartz 1978; DeAngelo, Masulis 1980; Leland 1994) Staatilise kompromissiteooria järgi on ettevõtete kasumlikkuse ja võlataseme vahel positiivne seos, kuna kasumlikel ettevõtetel on rohkem võimalik kasu lõigata maksukilbist ja väiksemad võlakapitali kasutamise seotud riskid ehk pankrotioht (González, González 2011; Miglo 2010; Jõeveer 2018). Uuematel kiirel kasvuga ettevõtetel kasumlikkuse ja võlataseme vahel negatiivne seos, kuna kiire kasvuga ettevõtete jaoks on võlaga kaasnevad riskid suuremad (Miglo 2010).

Staatiline kompromissiteooria ei võta arvesse väärtpaperite emiteerimise ja tagasiostmisega seotud tehingukulud ning varasemalt kogunenud likviidseid vahendeid. Neid detaile võtab arvesse dünaamiline kompromissiteooria. Dünaamilise kompromissiteooria järgi on ettevõtetel optimaalne võlataseme vahemik, mille raames ettevõtte oma võlataset muudab. Ettevõtte liigub soovitud oma- ja võõrkapitali suhte suunas juhul, kui optimaalse suhte suunas liikumine ei ole kulukam kui olemasoleva suhte säilitamine. Näiteks võib kõrge kasumlikkusega ettevõtte eelistada investeerimist jaotamata kasumi arvelt, kui see võimaldab tulevikus võlakapitali kaasamisega seotud tehingukulud vähendada. (Fisher *et al.* 1989; Miglo 2010) Dünaamiline kompromissiteooria räägib võlakirjade väljaandmisega seonduvatest kuludest. Kuna Eestis on kapitaliturud USA ja Lääne-Euroopaga võrreldes suhteliselt vähe arenenud, on autori hinnangul Eesti kontekstis paslikum rääkida võlakapitali võtmisega seonduvatest kuludest (lepingutasud, intressikulu, kohustistasud jmt).

1.1.3. Agendi- ja vaba rahavoo agendikulude teooriad

Agenditeooria järgi mõjutab kapitali struktuuri valikuid ettevõtte omandistruktuur ja juhtimine. Agendikulud tulenevad ettevõttes tegutseva agendi (juhatuse liige, kes pole omanik) ja omanike erinevatest huvidest, mille tõttu viimane kulusid kannab. Agent ei pruugi alati tegutseda viisil, mis on omaniku vaatest kõige optimaalsem. Agendikulud moodustuvad saamata jäänud tulust, mille omanik agendi ebaoptimaalsete otsuste tõttu kaotab ja kuludest agendi üle järelevalve teostamiseks ja agendikonflikti vähendamiseks. Lisaks on agendisuhe ka omanike ja võlausaldajate vahel. Näiteks ei pruugi pankrotiohus ettevõtte omanikel olla motivatsiooni teha ettevõtte käekäigu jaoks parimaid valikuid, kuna nendest saavad kasu vaid võlausaldajad, mitte nemad ise. (Jensen, Meckling 1976)

Vaba rahavoo agendikulude teooria (ingl *theory of agency cost of free cash flows*) järgi võib esialgu olla võlal kasumlikkusele positiivne mõju, kuna võlg muudab ettevõtte efektiivsemaks. Agendid käituvad võla tõttu omanikele kasulikumal viisil, eriti ettevõtetes, millel on suur vaba rahavoog. Võla võtmiseta on suurem risk, et agendid kulutaksid raha ebaefektiivselt või projektidele, mis end ei õigusta. Võla võtmisel on investeerimisotsused rohkem läbi mõeldud ja distsipliini suurendab laenuandja järelevalve. Kiiresti kasvavates ettevõtetes, mille vaba rahavoog on piiratud, on võla distsiplineeriv efekt väiksem. (Jensen 1986)

Samas suurenevad võlaga omanike ja võlausaldajate vahelised agendikulud, mistõttu on teatud võlataseme ületamisel ettevõtte väärtusele negatiivne mõju. (Jensen 1986) Võlaga seotud

peamised agendikulud on agendi järelevalve ja agendikonflikti vähendamise kulu, võla mõjul kasumlike investeerimisvõimaluste kasutamata jätmise alternatiivkulu ning pankroti või restruktureerimise kulud. (Jensen, Meckling 1976) Optimaalne võõr- ja omakapitali osakaal on seega võla tulu ja kulu tasakaalupunktis. (Jensen 1986)

Kokkuvõtlikult on seega agendikulude teooria kohaselt kasumlikkuse ja võlataseme vahel tagurpidi U-kujuline suhe. Teatud tasemeni suurendab võlg ettevõtte efektiivsust vähendades ettevõtte omanike ja ettevõtte siseste agendite vahelisi agendikulusid. Teisest küljest suurenevad võlaga seotud kulud, sh omanike ja võlausaldajate vahelised agendikulud, mistõttu teatud tasemest võib võla mõju olla kasumlikkusele negatiivne. Autori hinnangul kattuvad enamike Eesti tööstusettevõtete omanike ja juhtkonna ringid vähemalt osaliselt. Seetõttu ei pruugi keskmine võla efektiivistav mõju olla sektoris tugev.

1.1.4. Finantshierarhia teooria

Myers ja Majluf 1984.a püstitatud finantshierarhia teooria (ingl *pecking order theory*) järgi eelistavad ettevõtted info asümmeetria tingimustes kasutada esmalt eelmiste perioodide jaotamata kasumit, seejärel väliseid laene ja viimasena omakapitali suurendamist uute omanike kaasamise teel. Laenude võtmine ja uute omanike kaasamine on kulukam, kuna viimastel on ettevõtte käekäigu kohta vähem infot, mistõttu soovitakse laenule või investeringule suuremat tootlust. (Myers, Majluf 1984)

Erinevalt kompromissiteooriast ei arvesta finantshierarhia teooria maksudega ega laenust tulenevate kuludega (Myers 2015) Finantshierarhia teooriast võib järeldada, et kõrge kasumlikkusega ettevõtted kasutavad vähem laenukapitali kui vähese kasumlikkuse ja seega väiksemate likviidsete vahenditega ettevõtted (Abor 2007). Seega peaks finantshierarhia teooria kohaselt olema ettevõtte kasumlikkuse ja võlakapitali taseme vahel negatiivne seos.

Finantshierarhia teooria järgi on küpse ettevõtte võlataseme kasv võrdne ettevõtte väliskapitali vajadusega. Teooria järgi ei ole ettevõtetel konkreetset võlataseme eesmärki. (Myers 1984; Myers 2015) Finantshierarhia teooria ei selgita ettevõtete võlakapitali kasutamise poliitikat täielikult. Teooria ei suuda lõplikult selgitada eri suuruse, vanuse ja edukusega ettevõtete kapitali struktuuri valikuid. (Myers 2015)

Info assümmeerial põhinevad ka signaalide edastamise teooria (ingl *signalling theory*) ja aktsiaturu ajastamise teooria (ingl *market timing theory*). Signaalide edastamise teooria järgi eelistavad ettevõtted võlakapitali, kui nende arvates on ettevõtte alahinnatud ning aktsiate emiteerimist, kui nende arvates on ettevõtte ülehinnatud. Ettevõtte võlakapitali suurenemine indikeerib, et ettevõtte on kindel oma tuleviku rahavoos. (Ross 1977; Leland, Pyle 1977; Harris, Raviv 1985) Signaalide teooria kohaselt on ettevõtte kasumlikkuse ja võlakapitali osakaalu vahel positiivne seos. Aktsiaturu ajastamise teooria eitab optimaalse võõr- ja omakapitali suhte olemasolu. Teooria põhipostulaadi kohaselt oleneb kapitali struktuur aktsiaturu olukorrast ja aktsiate hinnast. Teooria väidab, et ettevõtte kapitali struktuur kujuneb välja aktsiaturul ettevõttele kasulike emiteerimiste ja tagasiostmistele tulemusena. (Lucas, McDonald 1990; Kroajczyk *et al.* 1992; Baker, Wurgler 2002) Signaalide edastamise ja aktsiaturu ajastamise teooriad on käesoleva töö kontekstis vähem relevantset, kuna valimis on vaid üksikud börsiettevõtted.

1.1.5. Kapitali struktuuri teooriate kokkuvõtte

Põhiliste valdkonnaalaste teooriate kokkuvõtteks võib öelda, et kuigi täiusliku ja täielikult efektiivse turu tingimustes ei ole kapitali struktuur ja ettevõtte kasumlikkus omavahel seotud ning kapitali kulu eri kapitali struktuuride puhul ei muutu, on tegelikus elus mitmeid aspekte, mis siiski loovad kapitali struktuuri ja ettevõtte kasumlikkuse vahele seose. Nende aspektide seas on maksud, võlaga seotud riskid, agendikulud, tehingukulud, info asümmeetria ettevõtte siseste ja väliste isikute seas, majanduse tsüklilisus ja aktsiahinna kõikumine. Teooriad on eri seisukohtadel, milline on ettevõtte kasumlikkuse ja võlataseme vaheline seos.

Staatilise kompromissiteooria järgi on ettevõtte võlatase ja kasumlikkus positiivses seoses, kuna kasumlikematel ettevõtetel on väiksemad võlaga kaasnevad riskid. Dünaamilise kompromissiteooria järgi võib olla seos kasumlikkuse ja võlataseme vahel negatiivne, kui ettevõtte jaoks on tehingukulud arvestades mõistlikum finantseerida projekte jaotamata kasumi arvelt. Agenditeooria järgi on võlakapitali kasutamine kasumlikkusega esialgu positiivses seoses kasumlikkusega, kuna vähendab agendikulud omanike ja ettevõtte siseste agentide vahel. Samas suurendab võlg pankrotiriski ja agendikulud omanike ja võlausaldajate vahel, mistõttu võib teatud võlataseme ületamisel olla kasumlikkusele negatiivne mõju. Finantshierarhia teooria järgi on ettevõtte võlatase ja kasumlikkus negatiivses seoses, kuna ettevõtted eelistavad esmalt kasutada jaotamata kasumit.

1.2. Empiirilised uurimused

Kasumlikkuse ja kapitali struktuuri seoseid on uuritud erinevates arengutasemega riikides ja erinevates sektorites, keskendudes nii ettevõtte-, valdkonna- kui ka riigispetsiifilistele faktoritele. Samuti on uuritud kapitali struktuuri valikuid eri suuruse, vanuse ning nii börsil noteeritud ja ka noteerimata ettevõtete seas. Uurimuste fookus on erinev – ühest küljest uuritakse mõjureid, mille baasil ettevõtted oma kapitali struktuuri valivad (kapitali struktuuri determinandid) ning teisest küljest uuritakse, milline on võla kasutamise mõju kasumlikkusele (kasumlikkuse determinandid). Lisaks on ettevõtete seas läbi viidud kvalitatiivseid küsitlusi eesmärgiga paremini mõista ettevõtete võlapoliitika rolli mängivaid aspekte ning nende koosõla kapitali struktuuri teooriatega. Magistritöös keskendutakse uurimustele Eestis ja tööstussektoris.

1.2.1. Uurimused Eestis

Eesti on olnud hõlmatud Mokhova ja Zineckeri (2013) Euroopa Liidu riikide tööstusettevõtete, Jõeveeri (2018) ning Haasi ja Peetersi (2006) Kesk- ja Ida-Euroopa ettevõtete, Hanouseki ja Shamshuri (2011) Ida-Euroopa ettevõtete, Škuláňová (2019) kaheksa Euroopa riigi taristuehituse ettevõtete, Nivorozhkini (2005) viie Euroopa riigi uurimuses ning Norvaišiene (2012) Balti börsiettevõtete kapitali struktuuri uurimuses. Ainult Eesti mitte-finantsettevõtteid on uurinud Seppa (2008).

Uuritud on eri perioode: aastaid 1993-2001 (Haas, Peeters 2006), 1997–2001 (Nivorozhkin 2005), 1996-2006 (Hanousek, Shamshur 2011), 2002-2003 (Seppa 2008), 2002-2011 (Norvaišiene 2012), 2006-2011 (Mokhova, Zinecker 2013), 2006-2014 (Jõeveer 2018), 2009-2018 (Škuláňová 2019).

Eesti ettevõtete kapitali struktuuri on uuritud peamiselt lineaarsete mudelitega, sh mitmese lineaarse regressiooni (Mokhova, Zinecker 2013; Jõeveer 2018; Hanousek, Shamshur 2011; Seppa 2008) ja mitmemõõtmelise lineaarse regressiooniga (Norvaišiene 2012). Kasutatud on ka lineaarse regressiooni dünaamilisi mudeleid (Jõeveer 2018; Haas, Peeters 2006; Nivorozhkin 2005). Škuláňová (2019) on kasutanud üldistatud momentide meetodit.

Kasumlikkuse mõõdikutena on kasutatud ROI-d ja EBIT-it (Seppa 2008), EBITDA suhet koguvaradesse (Jõeveer 2018), ROE-d ja ROA-d (Norvaišiene 2012). Finantsvõimenduse mõõdikutena on sarnaselt käesoleva magistritöoga kasutatud võlakordajat ja pikajaliste ning

lühiajaliste võlakohustuste kordajaid (Mokhova, Zineckeri 2013; Norvaišiene 2012) või võla ja omakapitali suhet (Jõeveer 2018).

Peamiselt on Eesti ettevõtetes tuvastatud kasumlikkuse ja finantsvõimenduse vahel statistiliselt oluline negatiivne seos (nt Mokhova, Zinecker 2013; Jõeveer 2018; Norvaišiene 2012; Nivorozhkin 2005, Seppa 2008). Seppa (2008) leidis, et suurtes Eesti ettevõtetes oli negatiivne seos võlataseme ja kasumlikkuse vahel tugevam kui keskmise ja väikese suurusega ettevõtetes. Sama trendi on tuvastanud ka Škuláňová (2019).

Kolmekümne kahe Euroopa Liidu liikmesriigi, sh Eesti, tööstusettevõtete uurimuses perioodil 2006-2011 leidsid Mokhova ja Zineckeri (2013), et kasumlikkusel oli enamikes uuritud riikides võlaga negatiivne seos, va Islandil ja Sloveenias, kus oli seos positiivne. Eestis oli kasumlikkusel oluline negatiivne seos nii lühiajalise-, pikaajalise- kui ka koguvõlaga.

Norvaišiene (2012) leidis perioodidel 2002-2004, 2005-2007 ja 2008-2011 Eesti 14 börsiettevõtete koguvõla, pikaajalise võla ja lühiajalise võla taseme ning kasumlikkuse (ROE ja ROA) vahel valdavalt olulise negatiivse seose, kuigi leidis ka statistiliselt ebaolulisi tulemusi. Näiteks ROE ja pikaajalise võla taseme seos oli oluline kõikidel perioodidel, ROA puhul oli oluline ainult perioodil 2002-2004.

Škuláňová (2019) leidis perioodil 2009-2018 Eesti taristuehituse ettevõtete võlataseme ja lühiajalise võla taseme ning ROA vahel negatiivse seose. Keskmise suurusega Eesti ettevõtetes, mida oli uurimuses 133, olid koefitsendid vastavalt -1,070 ja -0,881. Teiste riikide puhul olid statistiliselt olulised negatiivsed seosed ROA ja finantsvõimenduse näitajate vahel pigem tugevad ning tunnuste vahelised positiivsed seosed nõrgad.

Leidub ka uurimusi, mis ei tuvastanud Eesti ettevõtete kasumlikkuse ja võlataseme vahel statistiliselt olulist seost. Hanousek ja Shamshur (2011) leidsid statistiliselt olulise nõrga negatiivse seose vaid börsiettevõtete puhul, kuid mitte eraettevõtetes.

1.2.2. Uurimused teistes lääneriikides

Enamik empiirilistest uurimustest on tuvastanud kasumlikkuse ja finantsvõimenduse vahel statistiliselt olulise seose. Euroopa Liidus on laiahaardeliselt kasumlikkuse determinante uurinud mitmete riikide võrdluses Goddard *et al.* (2005) ning Pattitoni *et al.* (2014), kes mõlemad

tuvastasid kasumlikkuse olulise determinandina võlataseme, kuigi nad kasutasid erinevaid võlataseme mõõdikuid.

Belgia, Prantsusmaa, Itaalia ja Ühendkuningriikide tööstussektori kasumlikkuse determinante perioodil 1993–2001 üldistatud momentide meetodiga uurinud Goddard *et al.* (2005) leidsid, et võlakohustuste ja omakapitali suhe on ROA-ga negatiivselt seotud. Lisaks leidsid nad negatiivse seose ettevõtte suuruse ja ROA vahel, positiivse seose turuosa ja ROA vahel ning positiivse seose likviidsuse ja ROA vahel.

Pattitoni *et al.* (2014) käsitlesid 15-st eri Euroopa riigist üle 30 tuhande ettevõtte, millest enamus olid tööstusettevõtted. Valimis olid hõlmatud Austria, Belgia, Taani, Soome, Prantsusmaa, Saksamaa, Kreeka, Iirimaa, Itaalia, Luksemburgi, Hollandi, Portugali, Hispaania, Rootsi ja Ühendkuningriigid. Uuritud periood oli 2004–2011. Kaheksa lineaarse mudeliga leidsid nad võlataseme ja ROA vahel olulise negatiivse suhte. Võimalikuks põhjuseks tõid nad, et kõrge võlatasemega ettevõtetel jääb kehvematel aegadel rahast puudu, mistõttu on nad sunnitud loobuma kasumlikest investeeringutest. Täiendava analüüsimisega tuvastasid nad võlataseme ja ROA vahel mittelineaarse suhte. Mittelineaarse mudeliga, mis baseerus võlakordaja ruutfunktsioonil, leidsid nad, et kui võlakordaja on madal, on koguvõla taseme ja ROA suhe positiivne. Alates teatud tasemest on suhe aga negatiivne (nende uurimuses alates 32% koguvõla tasemest). Samas ületasid nende uurimuses enamik ettevõtetest selle künnise, mistõttu nende hinnangul andsid ka lineaarsed mudelid küllaltki õige tulemuse.

Lääneriikide ettevõtete kapitali struktuuri determinantide laiahaardelised uuringud (Rajan, Zingales 1995; Jong *et al.* 2008; Mokhova, Zinecker 2013) on tuvastanud kasumlikkuse ja võlataseme vahel enamjaolt negatiivse seose. Rajan ja Zingales (1995) uurisid arenenud riikide ettevõtete kapitali struktuuri lineaarse regressiooni ja Tobit regressiooni meetoditega. Nende valim hõlmas väljaspool finantssektorit tegutsevaid ettevõtteid G7 riikides: Ameerika Ühendriigid, Jaapan, Saksamaa, Prantsusmaa, Itaalia, Ühendkuningriigi ja Kanada. Nad leidsid, et peale Saksamaa oli uuritavates riikides ettevõtte kasumlikkuse (mõõdetuna EBITDA suhe varade bilansilisse väärtusesse) ja võlataseme vahel negatiivne seos. Bilansiväärtuses omakapitali mudelis oli oluline negatiivne seos Ameerika Ühendriikide, Kanada ja Jaapani ettevõtetel. Turuväärtuses omakapitaliga mudelis oli lisaks eelmainitud kolmele riigile oluline negatiivne seos ka Ühendkuningriigi ettevõtetel. Saksamaal, Itaalias ja Prantsusmaal olulist seost ei tuvastatud.

Täiendavalt leidsid Rajan ja Zingales, et negatiivne seos kasumlikkuse ja koguvõla (suhtena varadesse) vahel tugevnes USA-s, Jaapanis, Itaalias ja Kanadas ettevõtte suurenemisega, kuid UK-s jällegi nõrgenes. Autorid nentisid, et seose tagamaad pole neile täielikult arusaadavad, erinevaid põhjuseid võib olla mitmed ja põhjuste mõjud erisuunalised. Ühe võimaliku põhjusena tõid nad välja, et kasumlikel ettevõtetel on nii rohkem vaba raha kui ka kvaliteetseid investeerimisvõimalusi, millest üks vähendab vajadust võlakapitali järele ja teine jällegi suurendab. (*Ibid*)

Jong *et al.* (2008) kritiseerisid varasemate empiiriliste uurimuste eeldust, et ettevõtte spetsiifilised kapitali struktuuri mõjurid on riikides samad. Nad uurisid hariliku ja kaalutud vähimruutude meetoditega perioodil 1997-2001 nii ettevõttest kui ka riigist tulenevaid kapitali struktuuri mõjureid 42 riigis, millest pooled olid arenenud riigid ja pooled vähem arenenud. Nende valimis oli kokku 12 000 ettevõtet, mille seas polnud finants- ega kommunaal-ettevõtteid. Nad leidsid, et ettevõtte spetsiifilised tegurid ei ole riikides sama mõjuga, kuid enamik valitud mõjuritest (materiaalne vara, äririsk, suurus, maksumäär, kasv, kasumlikkus, likviidsus), suutsid selgitada suure osa finantsvõimendusest. Kasumlikkuse ja finantsvõimenduse vahel tuvastasid nad negatiivse suhte 25 riigis, millest enamik olid arenenud riigid. Finantsvõimenduse määratlesid nad pikaajalise võla suhtena koguvara turuväärtusesse ning kasumlikkuse põhitegevuse tulu suhtena vara bilansilisse väärtusse. Nad tõdesid, et riigipõhised tegurid, sh sisemajanduse kogutoodangu kasv, võlakirjaturu arengutase, korra- ja riskitegevuse efektiivsus ning võlausaldajate ja aktsionäride õiguste kaitse, võivad mõjuda ka kasumlikkuse ja finantsvõimenduse suhtele. Näiteks korra- ja riskitegevuse efektiivsus vähendab agendikulude rolli, mistõttu efektiivsema korra- ja riskitegevusega riikides on kasumlikkuse ja finantsvõimenduse suhe tugevamalt negatiivne.

Kesk- ja Ida-Euroopa ettevõtete perioodi 1993-2001 andmeid analüüsinud Haas ja Peeters (2006) leidsid kuues riigis kümnest kasumlikkuse ja finantsvõimenduse vahel statistiliselt olulise negatiivse seose. Eestis olulist seost ei tuvastatud. Valdava olulise negatiivse seose põhjusena tõid nad välja info assümmeetria ja kõrge laenuandjate riskipremia arvestades, et uuritava perioodil oli tegu üleminekumajandustega. Samas märkisid nad ära, et ka arenenud riikides on kõrgema kasumlikkusega ettevõtetel keskmiselt väiksem laenuandjate riskipremia, viidates muu hulgas Rajani ja Zingalesi (1995) uurimusele.

Ühendkuningriikides on teostatud mitmeid uurimusi, milles on ettevõtete kasumlikkuse ja võlataseme vahel tuvastatud peamiselt negatiivne seos. Michaelas *et al.* (1999) leidsid fikseeritud

efektidega mudeliga perioodi 1986-1995 Ühendkuningriigi 3500 väikse ja keskmise suurusega ettevõtete uurimuses, et lühiajalise võla tase tõusis majanduslanguse ajal ning vähenes majandusseisu paranedes. Nad leidsid, et nii koguvõla, lühiajalise võla kui ka pikaajalise võla tase on maksude eelse kasumi marginaaliga negatiivselt seotud. Hall *et al.* (2000) leidsid Ühendkuningriikide 3500 väikese ja keskmise suurusega ettevõtte regressioonanalüüsis lühiajalise võla taseme ja maksude eelse kasumi marginaali vahel negatiivse seose. Pikaajalise võla taseme puhul nad olulist seost ei tuvastanud.

Eelnevatele vastukaaluks on leidnud Prantsuse tekstiili- ja kemikaalitööstuse ettevõtetes perioodil 2002-2005 kasumlikkuse ja koguvõla taseme vahel positiivse suhte Margaritis ja Psillaki (2010), kes kasutas kasutades hariliku vähimruutude ja kvantiilregressiooni meetodeid. Nende hinnangul kinnitab tulemus agendikulude teooria argumenti, et võlg tõstab ettevõtete efektiivsust. Samuti positiivse seose on Itaalia keskmise suurusega tööstusettevõtetes leidnud de Luca (2014). Ta tuvastas positiivse seose ROE ning koguvõla, pikaajalise ja lühiajalise võla taseme vahel. Samuti positiivse seose leidis ta ROA ning pikaajalise ja lühiajalise võla taseme vahel. Meetodina kasutas de Luca (2014) lineaarset regressiooni.

Lisaks on positiivse seose kasumlikkuse ja koguvõla taseme leidnud Poola kõrge kasumlikkusega väljaspool finantssektorit tegutsevates börsiettevõtetes Mazur (2007). Kõrgelt kasumlikud ettevõtted olid uurimuses need, mille kasumlikkus oli valimis kõrgem kui 67. pretsentiilil. Madala kasumlikkusega ettevõtete puhul (madalam kasumlikkus kui 33. pretsentiilil) oli seos negatiivne. Mudelis, kus ettevõtted kasumlikkuse alusel jaotatud polnud, oli seos negatiivne. Uuritav periood oli 2000-2004.

Seega võib öelda, et Euroopa Liidu ettevõtete kapitali struktuuri uurimustes on valdavalt lineaarseid mudeleid kasutades jõutud järeldusele, et kasumlikkuse ja finantsvõimenduse vahel on oluline negatiivne seos. Samas on Euroopa Liidu ettevõtete uurimustes tuvastatud ka positiivset seost, lisaks jõutud ka järeldusele, et seos puudub. Samuti on tuvastatud mitte-lineaarset seost.

Lääneriikides väljaspool Euroopa Liitu on tuvastatud tööstussektori ettevõtete finantsvõimenduse ja kasumlikkuse vahel samuti nii positiivseid kui ka negatiivseid seoseid. Gill *et al.* (2011) on leidnud lineaarse regressiooniga New Yorgi börsi teenindus- ja tööstusettevõtetes perioodil 2005-2007 positiivse seose koguvõla ja pikaajalise võla taseme ning ROE vahel. Tööstusettevõtetes leidis lisaks positiivne suhe ka lühiajalise võla taseme ja ROE vahel. Autorite hinnangul võis

positiivne seos pikaajalise võla ja kasumlikkuse vahel tuleneda USA-s toimunud majanduslangusest ning madalatest intressimääradest. Roden ja Lewellen (1995) on leidnud multinomiaalse logistilise regressiooniga USA-s ROA ja võlataseme vahel statistiliselt olulise positiivse seose. Samale tulemusele, so positiivne seos võla osakaalu ja kasumlikkuse vahel, jõudis varem USA ettevõtete puhul Taub (1975). Negatiivse seose on tuvastatud finantsvõimenduse ja kasumlikkuse vahel LISREL meetodiga Ameerika tööstussektori uurimuses perioodil 1974-1982 Titman ja Wessels (1988).

Lääneriikides on tehtud ettevõtete võlapoliitika küsitlusi, mille järelduste kohaselt on ettevõtete peamine motivaator võla kasutamiseks finantsiline paindlikkus. Võla võtmisel peetakse olulisteks aspektideks ka tulu volatiivsust ja krediidireitingut. Küsitlustest on leitud eelkõige toetust kompromissiteooriale – kõrgema pankrotiriskiga ettevõtted võtavad vähem laenu. (Brounen *et al.* 2006; Bancel, Mittoo 2004; Graham, Harvey 2001)

Kokkuvõtlikult võib empiirilistest uurimustest järeldada, et nii Eestis, teistes lääneriikides kui ka mujal on kapitali struktuuri uurimustes leitud kasumlikkuse ja võlataseme vahel valdavalt negatiivne seos – nt Rajan ja Zingales (1995), Goddard *et al.* (2005), Mokhova ja Zinecker (2013), Jõeveer (2018). Teisalt leidub ka positiivse seose tuvastanud töid – nt Margaritis ja Psillaki (2010), da Luca (2014). Peamiselt on kapitali struktuuri uurimustes kasutatud lineaarseid mudeleid, ent leidub töid, mis on tuvastanud kasumlikkuse ja võlataseme vahel mittelineaarse seose – Pattitoni *et al.* (2014), Dalci (2018).

2. ANDMED JA METOODIKA

2.1 Andmed

2.1.1. Uurimisobjekt

Töö uurimisobjekt on Eesti töötleva tööstuse ettevõtted. Uurimisobjekt on kitsendatud Eesti töötleva tööstuse ettevõtetele kahel põhjusel:

- 1) Ettevõtete tegevusala mõjutab nende kapitali struktuuri (Talberg *et al.* 2008; Michaelas *et al.* 1999) ja kasumlikkust (Pattitoni *et al.* 2014);
- 2) Eesti maksusüsteemi ja finantseerimistava unikaalsus.

Töötlev tööstus on üks Eesti suurimaid sektoreid, mis moodustas aastatel 2011-2019 Eestis loodud lisandväärtusest 14,7-16,5% ja on sama perioodi igal aastal panustanud majanduskasvu 0,2-2,0 protsendipunkti (Eesti Statistikaamet, tabel RAA0042). Sektori ettevõtete otsene eksport moodustab üle poole Eesti koguekspordist ja ca 70% kogu kaupade ekspordist (*Ibid*, tabel VK02). Töötleva tööstuse suuremad alamsektorid koondkäibe alusel olid 2018.a seisuga puidusektor, toiduainetööstus ja elektroonikatööstus (*Ibid*, tabel EM001).

Statistikaameti andmetel kuulus 2018.a seisuga töötleva tööstuse sektorisse 7363 ettevõtet (8,1% kõigist Eesti ettevõtetest) ja sektor andis tööd ligi 108 tuhandele inimesele (24,3% tööjõust). Töötleva tööstuse ettevõtete summaarne käive oli 2018.a 13,2 miljardit eurot ja puhaskasum 1,6 miljardit eurot (keskmine puhaskasumi marginaal 12.1%), moodustades 19,4% Eesti ettevõtete kogukäibest ja 31,9% kasumist. Sektori ettevõtteid iseloomustab suhteliselt suur varade maht – ettevõtete koguarade summa oli sama aasta seisuga 10,4 miljardit eurot, millest omakapital oli 6,0 miljardit eurot ja omakapitali osakaal 58%. (*Ibid*, tabelid EM001, EM001A, EM009)

Seega on töötlev tööstus Eesti majanduse struktuuris oluline eksportöör ja tööandja. Samuti on sektor oluline lisandväärtuse looja ja panustanud riigi majanduskasvu.

2.1.2. Valim ja valimi korrigeerimine

Magistritöös kasutatakse andmeid Eesti tööstusettevõtete kohta perioodil 2011-2019. Andmete allikas on Orbis Europe andmebaas, mis sisaldas magistritöö kirjutamise hetkel ettevõtete andmeid aastatel 2011-2020 (kümme aastat). Kuna enamik ettevõtteid ei olnud andmete kasutamise ajaks

enda 2020.a tulemusi avaldanud, otsustas autor 2020.a andmed valimist välja jätta. Kuna valitud ajaperiood ei sisalda globaalse majanduskriisi aastaid, võivad saadud tulemused erineda uuringutest, mis hõlmavad tervet majandustsüklit.

Esialgseteks otsingukriteeriumiteks olid ettevõtte äriiline aktiivsus, asumine Eestis, tegutsemine erasektoris ja vastavus töötleva tööstussektori NACE Rev. 2 koodidele 10-31 (vt täpsemalt lisast 1). Lisaks eelmainitud tingimustele pidi ettevõtte kohta olema saadaval andmeid vähemalt ühes järgnevatest tunnustest: käive, varad, töötajate arv, lühiajaliste kohustuste kattekordaja (CR), omakapitali määr, materiaalsed põhivarad, käibevarad, lühiajalised võlakohustused, pikaajalised võlakohustused, ROE ja ROA. Lisaks otsingukriteeriumitele saadi andmebaasist ettevõtte nimi, registrikood ja asutamise kuupäev. Otsing tehti iga aasta kohta eraldi ja nimetatud kriteeriumitele vastavaid ettevõtteid oli iga aasta kohta 9221. Algandmete baasilt arvatati mitmed uued finantssuhtarvud ja bilansikirjed, sh koguvõlg, vanus, aastane käibekasv, põhivarade osakaal varadest, võlakordaja, pikaajaliste võlakohustuste kordaja ja lühiajaliste võlakohustuste kordaja.

Algsest valimist elimineeriti ettevõtted, mille aastane käive oli vähem kui 50 tuhat eurot ja bilansimaht vähem kui 39 tuhat. Valimist jäeti välja väiksemad kui 50 tuhande eurose käibega ettevõtted, sest autori hinnangu kohaselt ei oma nii väikesed tööstusettevõtted üldjuhul head ligipääsu võõrfinantseerimisele. Eeldust toetab fakt, et sellise grupi ettevõtete võlakohustuste mediaan nimiväärtus oli vaid kaks tuhat eurot. Lisaks on suur osa grupi kuuluvatest ettevõtetest kas mitteaktiivsed (38% elimineeritud vaatlustest) või valdusettevõtted, mida iseloomustab väike käive, kuid suur varade maht (19% elimineeritud vaatlustest). Kokkuvõtlikult ei ole sellised ettevõtted magistr töö fookuses. 50 tuhande euro piir on ühtlasi mikroettevõtte määratluse üheks tingimuseks (RPS § 3). Samadel põhjustel on valimist välja jäetud väiksema kui 39 tuhande eurose bilansimahuga ettevõtted, kusjuures kasutatud piirmäär on korrutis mikroettevõtte definitsiooniga seotud käibepiirmäärast ning töötleva tööstuse sektori kaalutud keskmisest varade ja käibe suhtest (78%) perioodil 2011-2018. 2019.a andmeid ei olnud töö kirjutamise hetkel avaldatud.

Töö eesmärgist lähtuvalt on kriitilise tähtsusega võlakohustuste ja ROE ning ROA info olemasolu. Seetõttu elimineeriti need vaatlused, kus puudus pikaajaliste või lühiajaliste võlakohustuste ning ROE või ROA suurus. Samuti jäid välja vaatused, mille puhul oli ROE ja ROA sisendiks kasutatav kasumlikkuse mõõdik EBT suurem kui käive. EBT saab käivet ületada ainult juhul kui ettevõttel esineb suhteliselt suures mahus muid äritulusid, mida käibe alla ei arvestata. Sellisel juhul on aga ROE või ROA suurus moonutatud ja ei ole teiste vaatlustega üheselt võrreldav. Samuti jäid

lõplikust valmist välja väärtused, mis ei saa tegelikkuses eksisteerida. Sealhulgas vaatlused, kus põhivara oli negatiivne, põhivarade suurus ületas bilansimahtu, võlakohustuste suurus oli negatiivne, võlakohustused ületasid kogukohustuste mahtu.

Otsustusmetsa kasutamise üheks eelduseks on puuduvate väärtusteta andmestik. Puuduvate väärtuste asendamiseks on erinevaid meetodeid, nende seas puuduvate väärtuste asendamine mediaaniga. Kui puuduvaid väärtusi on palju, võib puuduvate väärtuste asendamine moonutada tunnuste permitatsiooniolulisuse tulemusi (Hapfelmeier *et al.* 2014), mistõttu eemaldati andmestikust kõik puuduvate väärtustega vaatlused. Ühtegi puuduvat väärtust ei esinenud ainult tunnustel vanus ja põhivarade suurus. Kapitali struktuuri determinantide uurimuses eemaldasid puuduvate väärtustega vaatlused ka Amini *et al.* (2021), kes kasutasid samuti ühe meetodina otsustusmetsa.

Viimaseks välistuseks oli ekstreemsete vaatluste eemaldamine. Ekstreemsetena käsitleti iga tunnuse neid vaatlusi, mis asusid tunnuse mediaanväärtusest rohkem kui kolme standardhälve kaugusel. Tunnused, mille puhul erindeid ei ilmnenu, olid omakapitali määr ja põhivarade osakaal koguaradest. Kokku välistati sõltuva muutujana ROE-d kasutavast andmestikust 2363 ja ROA-d kasutavast andmestikust 2228 ekstreemset vaatlust.

ROE-d sõltuva tunnuse kasutava valimi lõplikuks suuruseks kujunes 12 568 vaatlust ja ROA-d kasutava andmestiku lõplikuks suuruseks 12 703 vaatlust. ROE valimi moodustasid 2064 ja ROA valimi 2063 unikaalset ettevõtet 9-aastase perioodi jooksul.

2.1.3. Kasutatud näitajad

Töös on sõltuv muutuja kasumlikkus, mille näidikuna kasutab autor ROE-d ja ROA-d. Samu suhtarve on kasumlikkuse näidikutena uurimustes kasutanud ka nt Norvaišiene (2012).

Esimene sõltuv muutuja on ROE ehk omakapitali tulukus. ROE on traditsiooniline ettevõtete kasumlikkuse mõõdik, mis iseloomustab ettevõtte võimekust luua investeeritud kapitalist kasumit. Omanikke rahuldavaks ROE tasemeks võib pidada sellist, mis ületab teiste sama riskantsusega investeeringute tulumäära. Magistritöö kontekstis on suhtarv leitud EBT-i ehk maksueelse kasumi ja omakapitali jagatisena. Traditsioonilises ROE arvutuses kasutatakse EBT-i asemel puhaskasumit, kuid autori hinnangul on Eesti ettevõtete puhul EBT-i baasilt arvatud ROE parem kasumlikkuse mõõdik, sest dividendide väljamaksetest tulenevad dividendimaksud on ettevõtete

lõikes ebaühtlased ja sõltuvad üksnes omanike otsustest väljamakseid teha. Seetõttu võib põhitegevuselt kasumlikuma ettevõtte puhaskasum olla dividendimaksude tõttu väiksem kui vähemkasumliku, dividende vähem või mitte-jaotava ettevõtte puhaskasum. Töö fookuses on just eelkõige ettevõtete põhitegevuslik kasumlikkus ja seetõttu on EBT-i kasutamine põhjendatud.

Teine sõltuv muutuja on ROA ehk varade tulukus. Sarnaselt ROE-ga on tegemist laialt kasutatava ettevõtte kasumlikkust iseloomustava suhtarvuna. ROA iseloomustab ettevõtte võimekust enda varasid kasutades kasumit teenida. Eelmises lõigus toodud põhjustel eelistab autor ka ROA puhul arvutuses jagatavana kasutada EBT-i.

Võlakohutusi iseloomustatavateks selgitavateks muutujateks on autor valinud võlakordaja ja selle komponendid lühiajalise võla kordaja ja pikaajalise võla kordaja. Tegemist on uurimisküsimustele vastamiseks vajalike selgitavate muutujatega, millest igauhega loodi regressioonimudel ROE ja ROA kirjeldamiseks (kokku kuus regressioonimudelit). Võlakordaja (VK) on võlakohustuste ja varade jagatis, lühiajaliste võlakohustuste kordaja (LVK) lühiajaliste võlakohustuste ja varade jagatis ja pikaajaliste võlakohustuste kordaja (PVK) pikaajaliste ja varade jagatis. Võlataset kirjeldavate tunnustena on VK-d, PVK-d ja LVK-d kasutanud varasemates uurimustes nt Michaelas *et al.* (1999) ja Hall *et al.* (2000).

Võlakohustuste alla kuuluvad bilansis kajastatavad laenukohustised, liisingud, lühiajaliste krediidilimiitide kasutus, väljastatud võlakirjad, mõnedel juhtudel regressiga faktooringuvõlg ja muud. Võlakohustuste alla ei kuulu tuletisinstrumendid, tarnijatele tasumata arved, viitvõlad ja bilansis mitte kajastatavad krediiditooted, nt garantiikirjad ja kasutusrendikohustused. Alates 2019. aastast kajastuvad ka kasutusrendikohustused IFRS alusel raamatupidamist korraldavates ettevõtetes. Autori hinnangul on selliseid ettevõtteid Eestis siiski vähe ja raamatupidamislik muudatus ei moonuta oluliselt võlakohustuste andmestikku.

Lühiajalisteks loetakse sellised võlakohustused, mille tasumise tähtaeg on bilansipäevast järgneva 12 kuu jooksul. Selle alla kuuluvad üldjuhul lühiajalised krediidilimiidid ja käibekapitali laenud. Samuti kajastatakse lühiajalise kohustusena pikaajalise võlakohustuse seda osa, mis on kavas tasuda bilansipäevast alates 12 kuu jooksul (lühiajaline osa).

Viimaseks, lühiajalise kohustusena kajastatakse ka enda olemuselt pikaajalisi võlakohutusi (nt suure *bullet*-maksega laenud), kui kohustuse tähtaeg on bilansipäevast alates 12 kuu jooksul ja

ettevõttel ei ole kokkuleppeid kohustuste refinantseerimiseks. Autori hinnangul võib selline raamatupidamislik kajastus andmeid moonutada, näidates lühiajalisi kohustusi suuremana kui tegelikkus. Kahjuks ei ole pikaajaliste kohustuste lühiajalist osa ja pikaajalisi lõppevaid, kuid refinantseeritavaid võlakohustusi võimalik muudest lühiajalistest võlakohustustest hõlpsasti eristada. Seetõttu jätkatakse võlakohustuste andmetega muutmata kujul. Pikaajaliseks loetakse neid võlakohustusi, mis ei ole lühiajalised.

Kasumlikkuse ning VK, PVK ja LVK vaheliste seoste eeldatavad suunad on negatiivsed, kuna empiirilise kirjanduse ülevaatest nähtus (alapeatükk 1.2), et peamiselt on Eestis ja teistes lääneriikides teostatud kapitali struktuuri uurimustes tuvastatud negatiivne seos. Lisaks eelnimetatutele on regressiooni ja otsustusmetsa lülitatud muud empiirilises kirjanduses laialt kasutust leidnud ja Orbis Europe andmebaasist kättesaadavad kasumlikkust selgitavad kontrollnäitajad.

Esimene kontrollnäitaja on ettevõtte vanus (VANUS). Empiirikas on ettevõtte vanuse ja kasumlikkuse vahel tuvastatud nii negatiivset kui ka positiivset suhet. Positiivse suhte on tuvastanud näiteks Papadogonas (2007) Kreeka tööstussektoris ja negatiivse suhte Istanbuli börsil noteeritud ettevõtetes Doğan (2013). Ühest küljest koguvad ettevõtted vanusega kogemust, mis võib aidata neil efektiivsemalt tegutseda ja olla seega kasumlikud (Balik, Gort 1993). Teisest küljest võib vanuse mõju kasumlikkusele olla negatiivne, kui ettevõtte on keskkonnas toimunud muutustele reageerimisel aeglane – vanusega ettevõtte struktuurne inerts suureneb (Hannan, Freeman 1984). Tuginedes viimasele argumentile on ootuspärane suund vanuse ja kasumlikkuse vahel negatiivne.

Teine kontrollnäitaja on käibe aastane kasvumäär (YoY). Oodatav YoY ja kasumlikkuse seos on positiivne – positiivse seose on tuvastanud näiteks Euroopa Liidu riikides Pattitoni *et al.* (2014). Positiivse seose põhjenduseks on nad pakkunud selle, et ettevõtte kasv mõjub töötajatele motiveerivalt.

Kolmas kontrollnäitaja on lühiajaliste kohustuste kattekordaja (CR). Likviidussuhtarv on leitud käibevarade ja lühiajaliste kohustuste jagatisena. CR-i ja kasumlikkuse vahel on positiivse seose tuvastanud näiteks Istanbuli börsiettevõtetes Doğan (2013) ning Belgia, Prantsusmaa, Itaalia ja Ühendkuningriikide tööstussektoris Goddard *et al.* (2005). Viimase hinnangul vähendab kõrgem

likviidsuse tase riski, et ettevõtte ei suuda täita oma lühiajalisi kohustusi. CR-i ja kasumlikkuse oodatav seos on positiivne.

Neljas kontrollnäitaja on omakapitali määr (OKm). Suhtarvu leidmiseks jagatakse omakapital varadega. OKm-i kasutatakse laialdaselt ettevõtte riskantsuse hindamiseks. OKm-i ja kasumlikkuse oodatav seos on positiivne. Positiivse seose on leidnud OKm-i ja ROA vahel Bangladeshis tööstussektoris Rahman *et al.* (2019).

Viies kontrollnäitaja on põhivarade osakaal varadest (POV). Statistikaameti andmetel moodustas töötleva tööstuse ettevõtete põhivarast 70% materiaalne põhivara, 22% pikaajalised finantsinvesteeringud, 6% immateriaalne põhivara ja 2% kinnisvarainvesteeringud (Eesti Statistikaamet, tabel EM009). Pattitoni *et al.* (2014) tuvastasid Euroopa Liidu ettevõtete uurimuses käibevara osakaalul varadest ja kasumlikkuse vahel positiivse suhte. Kuna käibevara osakaalu tõustes põhivarade osakaal väheneb, siis POV-i oodatav seos kasumlikkusega negatiivne.

Autor hõlmas esialgsesse andmestikku käibe, vara ja töötajad, et välja selgitada, milline suuruse mõõdik on magistritöö kontekstis kõige suurema relevantsem. Kuna tunnuste vaheline korrelatsioon osutus väga tugevaks, so korrelatsioonikordaja vahemikus 74-91%, otsustati regressioonmudelites võimaliku multikollineaarsuse ja sellega seonduvate moonutuste vältimiseks valida lõplikesse regressioonimudelitesse kolmest üks parima kirjeldusvõimega ettevõtte suurust kirjeldav tunnus. Kõige suurema kirjeldusvõimega ja regressioonimudeli infokriteeriume parandava mõjuga oli käive, mis jäi lõplikesse mudelitesse sisse.

Ka otsustusmetsa mudelites ei kasutata tunnuseid, mille omavaheline korrelatsioon on enam kui 70%, kuna otsustusmets on tugevalt korreleerivate tunnuste suhtes kallutatud, hinnates nende olulisust tegelikust suuremaks (Strobl *et al* 2007). Võrreldavuse säilitamiseks jäetakse ka otsustusmetsa mudelisse sisse käive ja selgitavad tunnused varad ning töötajate arv jäetakse välja.

Kuuendaks kontrollnäitajaks valiti käive (KAIVE). Logaritmitud käivet on ettevõtte suurust iseloomustava tunnusena Kreeka börsiettevõtete kasumlikkuse determinantide uurimisel hõlmanud Lazaridis ja Tryfonidis (2006) ning Istanbuli börsil noteeritud ettevõtetes Doğan (2013), kes leidsid mõlemad käibe logaritmi ja kasumlikkuse vahel statistiliselt olulise positiivse seose. Doğan (2013) argumenteeris, et suuremad ettevõtted, st ettevõtted, kellel on rohkem töötajaid, vara

ja käivet, on efektiivsemad, kuna löikavad kasu massiefektist. Seega on käibe ja kasumlikkuse vahel oodatav seos positiivne.

Käibe logaritmine vähendab tunnuse jaotuse positiivset asümmeetriat ja tagab mudeli lineaarsuse tunnuse parameetri suhtes. Samuti on logaritmisel positiivne mõju jääkliikmete heteroskedastiivsusele. Selgitava tunnuse transformeerimine ei mõjuta otsustusmetsa tulemusi, kuna otsustuspuud ei ole selgitavate tunnuste monotoonse transformeerimise suhtes tundlikud (Hastie *et al.* 2017). Seetõttu on eri meetodikate abil saadud tulemused regressioonimudelid käibe logaritmisest hoolimata võrreldavad. Selgitavate tunnuste arvutusmeetodid ja nende oodatavad seosed kasumlikkusega on toodud kokkuvõtlikult tabelis 1.

Tabel 1. Selgitavate tunnuste oodatavate seoste suunad kasumlikkusega.

Lühend	Tunnus	Arvutusmeetod	Oodatav seos
VK	Võlakordaja	Võlakohustused / varad	-
PVK	Pikaajal. võlakoh. kordaja	Pikaajalised võlakohustused / varad	-
LVK	Lühiajal. võlakoh. kordaja	Lühiajalised võlakohustused / varad	-
KAIVE	Käive	Naturaallogarim käibest	+
YoY	Käibe aastane kasvumäär	(Käive aastal 1 / käive aastal 0) - 1	+
VANUS	Ettevõtte vanus	Vaatlusaasta - asutamisaasta	-
OKm	Omakapitali määr	Omakapital / varad	+
POV	Põhivarade osakaal varadest	Põhivarad / varad	-
CR	Lüh. kohustuste kattekordaja	Käibevarad / lühiajalised kohustused	+

Allikas: autori koostatud

2.1.4. Tunnuste kirjeldav statistika

Magistritöös on sõltuvaks tunnuseks valitud kaks erinevat kasumlikkuse mõõdikut - ROE ja ROA, millel on sarnased, kuid mitte identsed andmestikud. Andmestikud on koostatud lähtudes samasugustest valimi korrigeerimise põhimõtetest, kuid ROE andmestikus puudub teine sõltuv tunnus ROA ja vastupidi; kõik ülejäänud tunnused ühtivad. Lisaks on ROE andmestik 135 vaatluse võrra väiksem võrreldes ROA andmestikuga, sest tunnusel ROE oli 135 erindit rohkem. Alljärgnevalt on toodud kirjeldav statistika ROE andmestiku kohta tabelis 2 ja ROA andmestiku kohta tabelis 3.

Tabel 2. ROE mudelites kasutatud tunnuste kirjeldav statistika (n=12 568).

Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROE (%)	14%	12%	33%	-205%	224%
VANUS (aastat)	13,7	13,0	7,1	0,0	39,0
KAIVE (tuh eur)	2 738	739	5 987	50	82 013
YoY (%)	10%	6%	27%	-96%	120%
CR (ühikuta)	2,6	1,8	2,4	0,0	16,6
OKm (%)	55%	56%	23%	0%	100%
POV (%)	48%	48%	24%	0%	100%
VK (%)	21%	17%	18%	0%	85%
PVK (%)	13%	8%	15%	0%	65%
LVK (%)	8%	5%	8%	0%	41%

Allikas: autori koostatud.

Tabel 3. ROA mudelites kasutatud tunnuste kirjeldav statistika (n=12 703).

Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROA (%)	7%	6%	13%	-44%	55%
VANUS (aastat)	13,6	13,0	7,1	0,0	39,0
KAIVE (tuh eur)	2 749	749	5 975	50	82 013
YoY (%)	10%	6%	27%	-96%	120%
CR (ühikuta)	2,5	1,7	2,4	0,0	16,6
OKm (%)	53%	55%	25%	-74%	100%
POV (%)	48%	48%	24%	0%	100%
VK (%)	21%	17%	18%	0%	85%
PVK (%)	13%	8%	15%	0%	65%
LVK (%)	8%	5%	9%	0%	41%

Allikas: autori koostatud.

Tabelitest nähtub andmestike sarnasus. Kõige olulisemad erinevused on sõltuvate tunnuste vahel. ROE väärtuste hajuvus on oluliselt suurem kui ROA väärtustel. Andmestike standardhälbed on vastavalt 33% ja 13%.

ROE andmestiku korrelatsioonimaatriks on toodud lisa 2 ja ROA andmestiku korrelatsioonimaatriks lisa 3. Kõige tugevamad korrelatsioonid on ettevõtte suurust iseloomustavate tunnuste käive, varad ja töötajate arv vahel. Tugev korrelatsioon on samuti tunnuste OKm ja CRi (korrelatsioonikordaja 56%), OKm ja VK vahel (-64%) ja OKm ja PVK (-54%) vahel. Teistel tunnuste paaridel ei ületanud korrelatsioonikordaja absoluutväärtus 40%.

2.2. Kasutatav metoodika

Empiirilise analüüsi viib autor läbi kasutades kahte meetodit, milleks on mitmene lineaarne regressioon ja masinõppel põhinev otsustusmets. Lineaarne regressioon viiakse läbi kasutades ökonomeetria tarkvara Gretl. Otsustusmetsa analüüsis kasutatakse R programmeerimiskeelt (R Core Team 2021), RStudio kasutajaliidest (RStudio Team 2020) ning „randomForest“ teeki (Liaw, Wiener 2002).

Varasemates empiirilistes uurimustes on kasutatud kapitali struktuuri ja kasumlikkuse seoste uurimiseks peamiselt lineaarset regressiooni (nt Rajan, Zingales 1995; Jong *et al.* 2008; Mokhova, Zinecker 2013; Margaritis, Psillaki 2010; Gill, Biger, Mathur 2011; Abdou *et al.* 2012), mistõttu on see valitud üheks meetodiks.

Teiseks meetodiks on valitud masinõppel põhinev otsustusmets. Finantsvõimenduse mudeldamisel lineaarseid mudeleid (mitmene lineaarne regressioon, LASSO, üldistatud aditiivne mudel) ja mitte-lineaarseid mudeleid (otsustusmets, gradientlaskumise regressioonipuud ja närvivõrgud) võrrelnud Amini *et al.* (2021) leidsid, et neist täpsem oli otsustusmets. Otsustusmetsa on tunnuste olulisuse ja seoste suundade tuvastamiseks kasutatud näiteks pangast laenu võtmise determinantide (Ozgur *et al.* 2021), FX-optsoonide investorite meeleolu determinantide (Washimi 2020) ja maksudeklaratsiooni valesti esitanud ettevõtte profiili (González-Martel *et al.* 2021) uurimustes.

Muudest meetoditest on kapitali struktuuri uurimustest kasutatud näiteks üldistatud momentide meetodit (Goddard *et al.* 2005; Pattitoni *et al.* 2014), kvantiilregressiooni (Margaritis, Psillaki 2010), Tobit meetodit (Rajan, Zingales 1995), DuPont meetodit (Aryantini, Jumono 2021) ning masinõppe meetoditest närvivõrke (Abdou *et al.* 2012; Pao 2008).

Erinevalt lineaarsest regressioonist on otsustusmets algoritmiline mitteparameetiline mudel, õppides andmete pealt (Breiman 2001a). Otsustusmets võimaldab tuvastada tunnuste vahelisi keerulisi mitte-lineaarseid suhteid (*Ibid*), mis on lineaarse regressiooniga võrreldes oluline eelis. Mitte-lineaarse seose kasumlikkuse ja võlataseme vahel on tuvastanud näiteks Euroopa Liidu ettevõtete uurimuses Pattitoni *et al.* (2014), Taiwani börsiettevõtete uurimuses Yang *et al.* (2010) ning Hiina tööstussektori börsiettevõtetes Dalci (2018).

Lisaks on otsustusmets ülesobitumisele robustne (Breiman 2001a) ning on mitmetes uurimustes saavutanud traditsiooniliste ökonomeetriliste mudelitega võrreldes suurema ennustustäpsuse (nt Li *et al.* 2014; Čeh *et al.* 2019; Amini *et al.* 2021). Otsustusmets suudab analüüsida koos kvalitatiivseid ja kvantitatiivseid selgitavaid tunnuseid (Genuer, Poggi 2020). Kuigi otsustusmetsa on mitteparameetrilise olemuse ja keerulisema algoritmi tõttu raskem tõlgendada kui lineaarset regressiooni, on selleks mitmeid meetodeid, mis võimaldavad hüpoteesi testida (McAlexander, Mentch 2020; Mentch, Hooker 2016).

Puudusena ei suuda otsustusmets ekstrapoleerida – kuna otsustusmetsa tulemus on otsustuspuude aritmeetiline keskmine, siis jääb tulemus alati valimi ekstreemumite vahele. Autori hinnangul ei saa aga magistritöö puhul seda oluliseks probleemiks pidada, sest valimisse jäänud sõltuvate muutujate ROE ja ROA miinimum- ja maksimumväärtused ei ole ettevõtetele tavaolukorras saavutatavad. Lisaks vajab otsustusmets täpseteks tulemusteks tavaliselt suurt andmehulka (Alaka *et al.* 2017).

Kahe erineva kompleksuse astmega mudelite ennustuvõime ja tuvastatud seoste kõrvutamine aitab probleemi paremini tõlgendada (Breiman 2001b). Mitmese lineaarse regressiooni ja otsustusmetsa tuvastatud seoseid kõrvutades saab uurimusküsimustele täpsemalt ja suurema usaldusväarsusega vastata.

Otsustusmetsa meetodika rakendamisel kasutati ristanndmeid ja samade ettevõtete eri aastate andmeid arvestati erinevate vaatlustena. Autor otsustas otsustusmetsa puhul paneelandmete asemel ristanndmeid kasutada tulenevalt kahest peamisest kaalutlusest. Esiteks, mitteparameetriline otsustusmets ei suuda ajaefektiga arvestada ning eeldab, et iga vaatlus on eraldiseisev. Paneelandmete kasutust võimaldavad otsustusmetsa semi-parameetrilised edasiarendused (Capitaine *et al.* 2019) on küll hiljuti loodud, kuid uuringutes vähe rakendatud. Samuti on need võrreldes klassikalise otsustusmetsaga vähemate tõlgendamise võimalustega, kuna neile pole välja arendatud vajalikke funktsioone (nt sõltuva tunnuse ja selgitava tunnuse osalise sõltuvuse funktsiooni). Teiseks, paneelandmetega teostatud regressioonanalüüsist ilmnis, et ajakomponent oli kõigi loodud mudelite puhul veenvalt statistiliselt ebaoluline. Seetõttu peab autor ebatõenäoliseks, et ajakomponent osutub oluliseks ka otsustusmetsa rakendamisel.

2.2.1. Mitmene lineaarne regressioon

Lineaarne regressioon võimaldab uurida sõltuva muutuja ning ühe või mitme selgitava muutuja seoseid. Selle üks kasulikumaid omadusi on võimalus tuvastada mitmete selgitavate muutujatega mudelis eraldi iga selgitava muutuja mõju sõltuvale muutujale. Lineaarset regressiooni kasutatakse kolmel eesmärgil: hindamine, hüpoteesi testimine ja ennustamine. (Greene 2018) Käesolevas töös keskendutakse neist kahele esimesele.

Mitmese lineaarse regressiooni mudeli matemaatiline üldkuju on järgnev (*Ibid*):

$$y = a + x_1\beta_1 + x_2\beta_2 + \dots + x_k\beta_k + \varepsilon \quad (2)$$

kus

a – vabaliige,

y – sõltuv muutuja,

x_1, \dots, x_k – selgitavad muutujad üks kuni k ,

β_1, \dots, β_k – selgitavate muutujate hinnatavad parameetrid,

ε – vealiige.

Magistritöö kontekstis on sõltuvateks muutujateks kasumlikkust iseloomustavad omakapitali tootlus (ROE) ja varade tootlus (ROA). Uurimisküsimustele vastamiseks viidi kummagi sõltuva tunnusega läbi kolm regressiooni, mis kirjeldasid nende ja võlakordaja, pikaajaliste võlakohustuste kordaja ja lühiajaliste võlakohustuste kordaja vahelist seost. Seega loodi kokku kuus regressioonimudelit.

Kasumlikkuse ja võlakordaja vahelist seost iseloomustava regressioonimudeli üldkuju on järgnev:

$$y = \alpha + \beta * VK + \beta * VANUS + \beta * OKm + \beta * \ln(KAIVE) + \beta * YoY + \beta * POV + \beta * CR + \varepsilon \quad (3)$$

kus

y – kasumlikkuse suhtarv (ROE või ROA),

α – vabaliige,

VK – võlakordaja,

$VANUS$ – ettevõtte vanus aastates,

OKm – omakapitali määr,

$\ln(KAIVE)$ – käibe naturaallogaritm,

YoY – käibe aastane kasvumäär,

POV – põhivarade osakaal varadest,

CR – lühiajaliste kohustuste kattekordaja,

ε – vealiige.

Kasumlikkuse ja pikaajaliste võlakohustuste kordaja vahelist seost iseloomustava regressioonimudeli üldkuju on järgnev:

$$y = \alpha + \beta * PVK + \beta * VANUS + \beta * OKm + \beta * \ln(KAIVE) + \beta * YoY + \beta * POV + \beta * CR + \varepsilon \quad (4)$$

kus

PVK – pikaajaliste võlakohustuste kordaja.

Kasumlikkuse ja lühiajaliste võlakohustuste kordaja vahelist seost iseloomustava regressioonimudeli üldkuju on järgnev:

$$y = \alpha + \beta * LVK + \beta * VANUS + \beta * OKm + \beta * \ln(KAIVE) + \beta * YoY + \beta * POV + \beta * CR + \varepsilon \quad (5)$$

kus

LVK – lühiajaliste võlakohustuste kordaja.

Regressioonis kasutatakse paneelandmeid, mistõttu on eelistatud fikseeritud või juhuslike efektidega regressioonimudelite kasutamine. Paneelandmeid kasutataval mudelitel on mitmeid eeliseid võrreldes ristanndmete ja aegridadega, sh efektiivsemad parameetrite hinnangud ja võimalus testida keerulisemaid hüpoteese (Hsiao 2007).

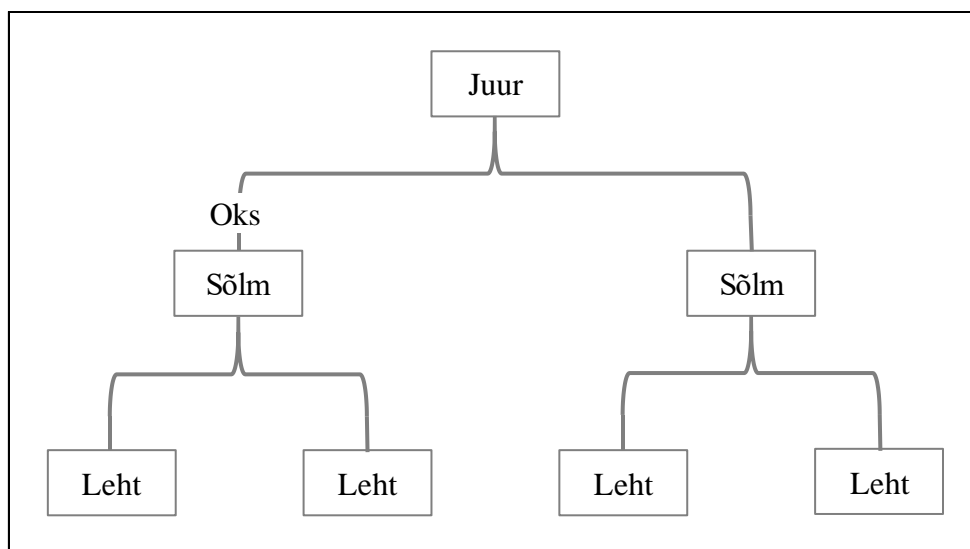
Esimesena hinnati fikseeritud efektidega mudeleid. Läbi viidud kitsenduste F-testi tulemusena tuli kõigi mudelite puhul vastu võtta nullhüpotees, mille kohaselt puudub objektide vaheline erinevus ja tuleks eelistada ühendatud (ingl *pooled OLS*) mudelit fikseeritud efektidega mudelile.

Teisena hinnati juhuslike efektidega mudeleid. Teostatud Hausmani testide tulemuste baasilt tuli kõigi mudelite puhul vastu võtta nullhüpotees, mille kohaselt hinnangute vahel erinevus puudub ja tuleks eelistada juhuslike efektidega mudelit fikseeritud efektidega mudelile. Samuti viitab tulemus selgitavate tunnuste sõltumatusele juhuslikust liikmest, mis on üks lineaarse regressiooni põhieeldustest. Breusch-Pagani testide tulemusena pidi kõigi mudelite puhul vastu võtma nullhüpoteesi, mille järgi on objektispetsiifilised veakomponendid nullid ja tuleks eelistada ühendatud mudelit.

Kokkuvõttes kasutati kõigi regressioonimudelitega ühendatud mudelit, tulenevalt kitsenduste F-testide ja Breusch-Pagani testide tulemust. Ühendatud mudelis käsitletakse sama ettevõtte erinevate aastate vaatlusi iseseisvate vaatlustena. Testide tulemuste koondvaade on toodud lisas 4.

2.2.2. Otsustusmetsa algoritm

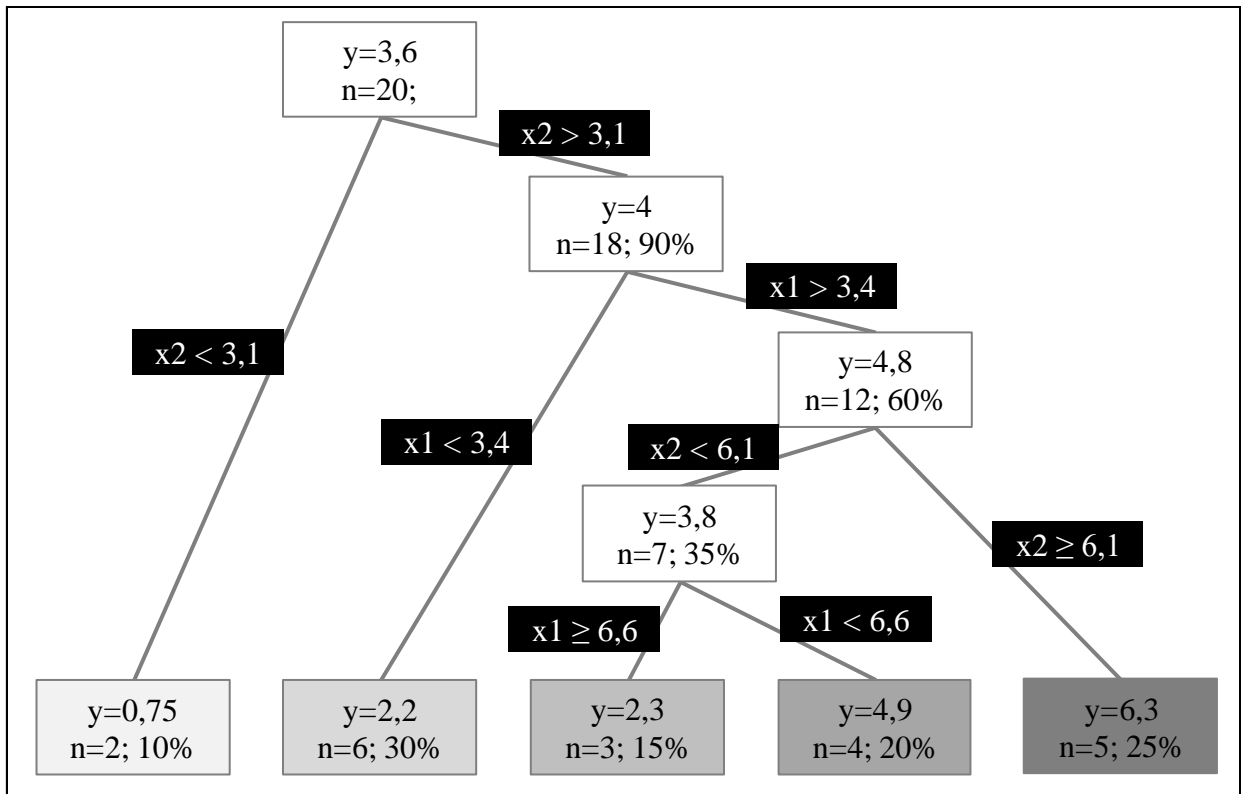
Otsustusmets on algoritmiline mudel, mis koosneb otsustuspuudest (Breiman 2001b; Marsland 2014). Joonisel 1 on toodud näide lihtsast otsustuspuust. Joonisel on tegemist kahendpuuga (ingl *binary tree*): juur ja iga sisendtipp jaguneb kaheks. Otsustuspuud on üldjuhul kahendpuud, kuigi on võimalik ka sõlmede hargnemine rohkem kui kaheks (Liu *et al.* 2005). Otsustuspuu algtipu, kust hakatakse probleemi lahendama, nimetatakse juureks. Juur jaguneb erinevateks harudeks, mille tippe nimetatakse sõlmedeks. Iga sõlm võib hargneda edasi järgnevateks sõlmedeks või olla ise viimane sõlm ehk leht, mis edasi ei hargne. (Marsland 2014; Duda *et al.* 2001)



Joonis 1. Lihtne otsustuspuu

Allikas: autori koostatud Iyer *et al.* 2015 alusel

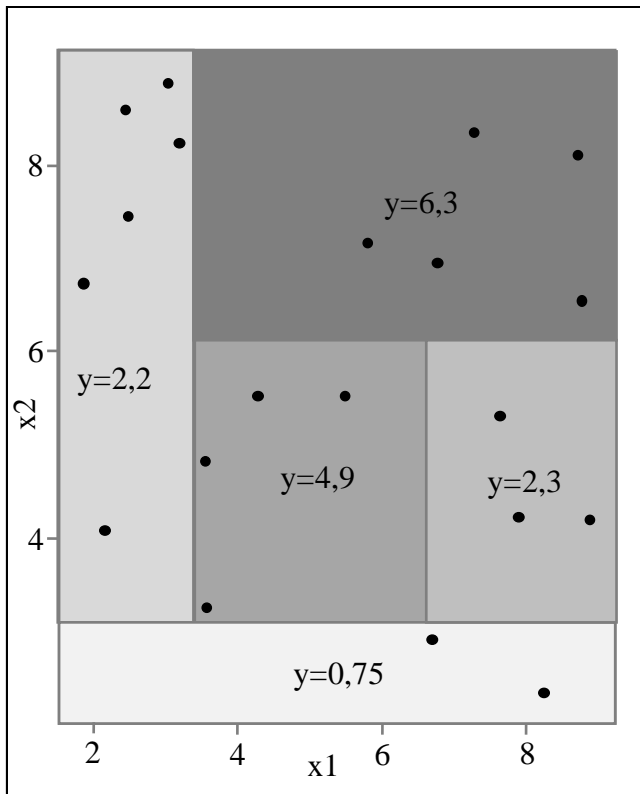
Iga otsustuspuu sõlme juures on test, mis määrab, millist haru mööda edasi minna, kuni jõutakse lõpptipuni ehk leheni, kus on ennustustulemus. (Marsland 2014; Duda *et al.* 2001) Tegemist on seega hierarhiliste testidega selgitavate muutujate väärtuste kohta, mille tulemus on sõltuva muutuja ennustatav väärtus. Joonisel 2 on toodud näide regressioonipuust. Otsustuspuu iga leht on arvuline ennustus y väärtuse kohta.



Joonis 2. Regressioonipuu.

Allikas: autori koostatud Sammut, Webb (2017, lk 1081) alusel

Otsustuspuu algoritmid jõuavad järjestikuse selgitavate tunnuste ruumi jaotamise ehk rekursiivse jaotamise (ingl *recursive splitting*, *recursive partitioning*) abil ennustuseni. (Strobl *et al.* 2009; Duda *et al.* 2001) Iga järgnev sõlm sõltub eelnevas sõlmes tehtud otsusest, jaotades eelneva sõlme otsuse tulemusena tekkinud alamhulka vaatlustest. Otsustuspuud võib seega kujutada ka selgitavate tunnuste ruumi jaotamisena, kus jaotuskoht on otsustuspuu sõlm. Sellisena kujutatakse otsustuspuud joonisel 3.



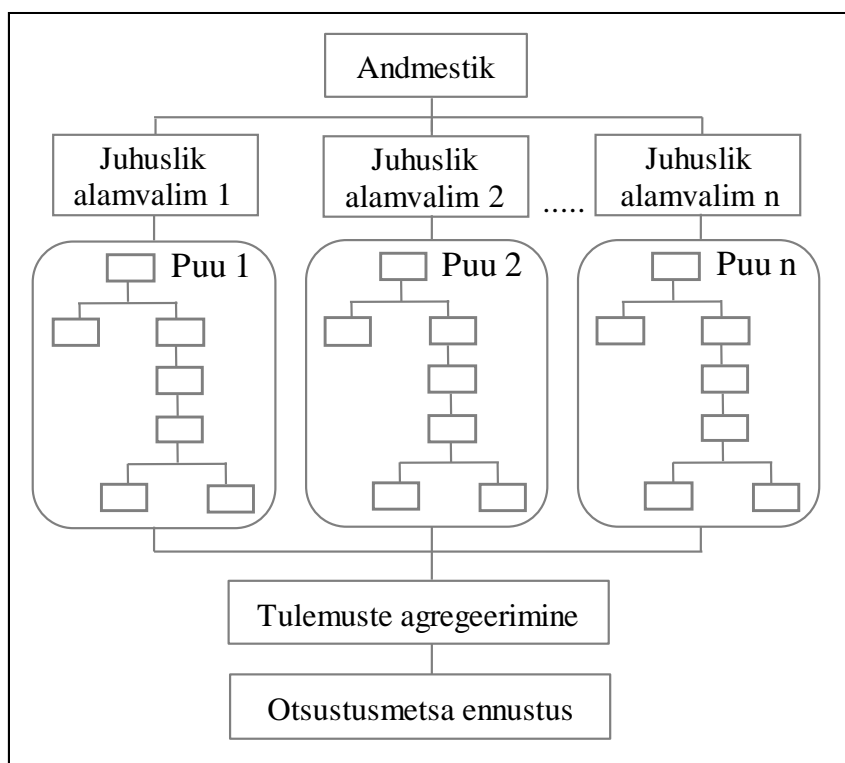
Joonis 3. Regressioonipuu alusel koostatud selgitavate tunnuste ruumi jaotus
Allikas: autori koostatud Sammut, Webb (2017, lk 1081) alusel

Paljudest otsustuspuudest koosnevat otsustusmetsa on võrreldes üksiku otsustuspuu algoritmiga keerulisem tõlgendada. Samas on otsustusmetsal oluline eelis – üksikutest mudelitest (puudest), millel igaühel eraldi on nõrgem ennustusvõime, moodustub kokku üks tugevam ennustaja (mets). Sellist mudelit nimetatakse ka ansambelmeetodiks. (*Ibid*)

Kõrgema täpsuse saavutamiseks on vajalik, et otsustusmetsas olevad puud oleksid üksteisest võimalikult erinevad (Hastie *et al.* 2017). Selleks kasutab otsustusmets kahte juhuslikkuse astet. Esiteks toimub iga otsustuspuu konstrueerimine erinevate andmete põhjal. Selleks võetakse iga puu jaoks treeningandmestikust juhuslikult sama arv vaatlusi nagu on treeningandmestikus. Sellega tekib iga puu jaoks nn kott vaatlusi (ingl *bag*), mis omavahel osaliselt erinevad. Sellist meetodit nimetatakse *bagging* meetodiks. Igast kotist jääb välja ca 1/3 vaatlusi, mille põhjal on võimalik arvutada puu ennustusviga (ingl *out-of-bag error*, *OOB error*). Otsustusmetsa ennustuse saamiseks puude ennustused agregeeritakse. Puude arv ja seega ka kottide arv on otsustusmetsa üks hüperparameetritest. (Breiman 2001a; Genuer, Poggi 2020; Sammut, Webb 2017)

Teine juhuslikkuse aste on igas sõlmes toimuv juhuslik tunnuste valik kogutunnuste seast (ingl *random subspace method*). Ainult valitud tunnused konkureerivad sõlme hargnemise otsustamisel. Tunnuste arv, mida iga otsustuspuu iga sõlme juures hargnemise valikus hõlmata, on otsustusmetsa kõige olulisem hüperparameeter. Kui see hüperparameeter on seadistatud selliselt, et juhuslikult valitavaid tunnuseid on sama arv kui kogutunnuseid, ei erine otsustusmets *bagging* meetodist. (*Ibid*) Parima hargnemise ehk tunnuse väärtuse alagruppideks poolitamise valik toimub vaikimisi sätetega regressioonimetsas selle põhjal, milline hargnemine minimaliseerib sõltuva tunnuse ruutvea summa (ingl *SSE, sum of squared errors*) (Boehmke, Greenwell 2020).

Joonisel 4 on kujutatud otsustusmetsa skeem, kust nähtub treeningandmestikust iga puu jaoks eraldi koti valimi ehk alamvalimi loomine ning nende baasilt otsustuspuude konstrueerimine, mille ennustuste aritmeetiline keskmine on regressioonimetsa ennustus.



Joonis 4. *Bagging* meetodit kasutava otsustusmetsa skeem
Allikas: autori koostatud Saeed et al. (2019) alusel

Peatükis kirjeldatu on klassikalise ehk Breimani otsustusmetsa algoritm selle vaikeväärtustega, mida on võimalik muuta – näiteks mitte kasutada andmete kotti valimisel asendamist (Liaw, Wiener 2002). Autor kasutab otsustusmetsa vaikeväärtustega, välja arvatud puude arvu ja

sõlmedes juhuslikult valitavate tunnuste arvu hüperparameetreid, mille osas on otsustusmets tundlik (*Ibid*).

2.2.3. Otsustusmetsa mudeldamise protsess

Masinõppe protsessi peamine eesmärk on leida mudel, mis oskaks võimalikult hästi ennustada (Kuhn, Johnson 2013). Protsess koosneb üldistatult probleemi mõistmisest, andmete hankimisest, analüüsimisest ja korrigeerimisest, tunnuste valikust, mudeli treenimisest ja ennustamisest, mudeli võimekuse hindamisest ning mudeli tulemuste tõlgendamisest. Seejuures võivad protsessi etapid korduda, et leida optimaalseim mudel. (Jablonka *et al.* 2020; Boehmke, Greenwell 2020) Seega toimub mudeli tõlgendamine, sh uurimisküsimustele vastamine, pärast kõige parema ennustusvõimega mudeli leidmist.

Erinevalt mitmesest lineaarsest regressioonist ei kasuta autor otsustusmetsa mudeldamisel kogu andmestikku, vaid jaotab selle treening- ja testandmeteks. Selleks kasutab autor teegi „Caret“ (Kuhn 2008) funktsiooni „createDataPartition“, mis teostab andmete kihilise juhusliku jagamise (ingl *stratified random split*), et treening- ja testandmestikus oleks võimalikult tasakaalustatud andmete jaotus. Andmete jagamise suhteks on 70%-30%. ROA andmestiku jagamisel tekib treeningandmestik 8894 vaatlusega ning testandmestik 3809 vaatlusega. ROE andmestiku jagamisel tekib treeningandmestik 8800 vaatlusega ja testandmestik 3768 vaatlusega. Treening- ja testandmestiku jaotused on sarnased koguandmestikule. ROE treening- ja testandmestiku kirjeldav statistika toodud lisas 5 ja ROA treening ja testandmestikku kirjeldav statistika lisas 6. Andmestike treening- ja testandmete jaotuse võrdlused on toodud lisades 7 ja 8.

Autor kasutab töös otsustusmetsa (Breiman 2001a) implementatsiooni „randomForest“ (Liaw, Wiener 2002), mis on tuntud ja laialt kasutatud otsustusmetsa teek R-is. Otsustusmetsa implementatsioone on R-is ka teisi - näiteks „ranger“ (Wright, Ziegler 2017) ja „randomForestSRC“ (Ishwaran, Kogalur 2021), mis on „randomForest“ teegist kiiremad. Samas ei ole nende teekidega võimalik kasutada mitmeid „randomForest“ teegiga ühtivaid teisi teeke, näiteks „permimp“ teegi (Debeer *et al.* 2021) funktsiooni „permimp“, mille abil saab leida tingimuslikku permutatsiooniolulisust, ning teegi „rfpermute“ (Archer 2020) funktsiooni tunnuste p-väärtuse leidmiseks.

Otsustusmets on tundlik kahele hüperparameetrile. Esiteks igas sõlmes juhuslikult valitavate tunnuste arvule (*mtry*), mille seast algoritm valib omakorda tunnuse, mille kaheks jaotamisel on

ruutvea summa minimaalne. Teiseks hüperparameetriks on puude arv (*ntrees*). (Liaw, Wiener 2002) Hüperparameetrit *mtry* optimeerib autor kasutades teegi „Caret“ (Kuhn 2008) funktsiooni „train“ ning k-korda ristvalideerimist (ingl *k-fold cross-validation*). Treeningandmestik jaotatakse juhuslikult sama suurusega kümneks osaks, mille iga osa kasutatakse ühe korra valideerimiseks ning üheksa korda mudeli treenimiseks. Ristvalideerimise protsessi korratakse kolm korda ning selle keskmine on mudeli valideerimise ennustusviga. K-korda ristvalideerimine on masinõppes laialt kasutusel valideerimismeetod, mida autor kasutab, kuna OOB viga võib olla liialt optimistlik (Han *et al.* 2016).

Teisi hüperparameetreid autor ei optimeeri – need jäävad vaikeväärtusesse. Regressioonimetsa vaikeväärtusteks on (Liaw, Wiener 2002):

- 1) iga puu jaoks kasutatakse treeningandmestikust vaatluste võtmisel asendamist, mistõttu kotti sattuvad osad vaatlused mitu korda;
- 2) iga puu jaoks treeningandmestikust võetavate vaatluste arv on andmestikus olevate vaatluste arv;
- 3) puu lehtede arvul pole piiri;
- 4) puu lehtedes olevate vaatluste minimaalne arv on 5.

Baggingu ja andmete asendamisega otsustusmetsa kasutamine võib tunnuste olulisuse hindamisel viia klassifikatsiooni puhul kallutatud tulemuseni, kuid seda probleemi ei esine juhul, kui kõik tunnused on arvulised (Strobl *et al.* 2007) nagu käesolevas töös.

Nii mudeli valideerimisel kui ka testimisel raporteerib autor nelja mõõdikut: ruutkeskmine viga RMSE, keskmine ruutviga MSE, keskmine absoluutne viga MAE ja determinatsioonikordaja R^2 . Esimeseks kolmeks kasutab autor teegi „Metrics“ (Hamner, Frasco 2018) vastavaid funktsioone, kuid determinatsioonikordaja funktsiooni loob autor ise selle puudumise tõttu vastavast teegist.

Determinatsioonikordaja valem on järgnev:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}, \quad (6)$$

kus

R^2 – determinatsioonikordaja,

SSR – mudeli kirjeldatud hajuvus (ingl *sum of squares due to regression*)

SST – koguhajuvus (ingl *total sum of squares*)

SSE – jääkhajuvus (ingl *sum of squares errors*)

Mudeli optimeerimisel lähtub autor ruutkeskmisest veast, mis on masinõppes levinud praktika ning on sarnaselt keskmise absoluutse veaga sõltuva muutujaga samades ühikutes. Ruutkeskmine viga on ruutjuur keskmisest ruutveast. Ruutkeskmist viga kasutavad kapitali struktuuri uurimuses otsustusmudeli täpsuse hindamiseks ka Amini *et al.* (2021). Keskmise absoluutne viga on tegelike ja ennustatud väärtuste keskmine absoluutne vahe. Kuna keskmine absoluutne viga ei võta viga ruutu, ei mõjuta selle tulemust suured vead niivõrd palju nagu keskmist ruutviga. (Boehmke, Greenwell 2020) Determinatsioonikordaja annab indikatsiooni mudeli kirjeldusvõime kohta, kuid olenevalt kasutatud andmetest võib see olla ebatäpne. Mitte-lineaarse mudeli puhul ei peaks sellest optimeerimisel ega mudeli headuse hindamisel lähtuma. Autor ei kasuta determinatsioonikordajat otsustusmetsa optimeerimiseks ning eelistab mudeli headuse hindamisel testandmetel arvatud ennustusviga Breimani (2001b) eeskujul.

Optimaalne *mtry* väärtus on selline, mille puhul on mudelil madalaim valideerimise RMSE. Läbi proovitakse kõik võimalikud *mtry* väärtused ehk 1-7. Optimaalse *mtry* väärtuse leidmisel on puude arv 500, mis on „randomForest“ teegi vaikeväärtus. Puude arvu piisavust kontrollib autor kasutades sama teegi „plot“ funktsiooni „randomForest“ objekti argumendiga, mis tagastab joonise *out-of-bag* keskmise ruutvea (OOB MSE) ja puude arvu sõltuvuse kohta. Joonistelt nähtub, et 300-400 puud on piisav vea stabiliseerimiseks. Samas on permutatsiooniolulisuse täpsemaks arvutamiseks vajalik, et puude arv oleks suurem (Debeer, Strobl 2020). Seetõttu valib autor „randomForest“ objekti loomisel pärast optimaalse *mtry* leidmist puude arvuks Debeeri ja Strobli (Debeer *et al.* 2021) eeskujul 1000 puud. Teegiga „Caret“ *mtry* optimeerimisel kasutab autor paralleelseks programmeerimiseks teeki „doParallel“ (Microsoft Corporation, Weston 2020).

Mtry väärtus mõjutab mudelis sõlmede hargnemisel valitavaid tunnuseid, mistõttu on sel efekt ka permutatsiooniolulisusele. Kui mudelis olevad tunnused on informatiivsed, võimaldab väike *mtry* väärtus sõlmes valituks saada ka nendel tunnustel, mis on vähem informatiivsed, kuid suurema *mtry* puhul kaotaksid kõige informatiivsemate tunnustega võistluse. Seetõttu võimaldab väike *mtry* väärtus tuvastada ka vähemolulisi, kuid siiski informatiivseid tunnuseid. Suurem *mtry* väärtus võimaldab aga tuvastada kõige informatiivsemaid tunnuseid, kuna need saavad vähem informatiivsete tunnustega võisteldes rohkem sõlmedes valituks. Kui *mtry* väärtus on 1, on sõlmes valitava tunnuse valik täielikult juhuslik, konkureerimist erinevate tunnuste vahel ei toimu. Permutatsiooniolulisuse leidmisel on soovitatav, et otsustusmetsa *mtry* väärtus oleks optimaalne, st selline, mis viib mudeli valideerimise ennustusvea miinimumi. (Debeer, Strobl 2020)

Optimeeritud *mtry* väärtusega ja 1000 puuga otsustusmetsa mudelite ülesobitumise testi ning mudelite p-väärtuse leiab autor teegi „rfUtilities“ (Murphy *et al.* 2010) funktsioonidega „rf.regression.fit“ ja „rf.significance“. Ülesobitumise testi kohaselt ei ole mudel ülesobitunud, kui ülesobitumise suhe (ingl *overfitting ratio*) on üle 10. Ülesobitumise suhte leiab algoritm sõltuva tunnuse väärtuste arvu ja tunnuste arvu jagatisena (nt ROE-VK mudelil 8800/7). Mudeli p-väärtuse leidmine toimub mudeli tunnuste permutatsioonil. Algoritm võrdleb permuteeritud väärtustega mudeli ja tegeliku mudeli erinevust. (*Ibid*)

Mudeli stabiilsuse kontrolliks autor treenib ja testib mudelit teise etteantud juhuslikkuse väärtusega (ingl *seed*) ning kõrvutab põhimudeli ja kontrollmudeli tulemused. Tulemuste varieeruvus peab jääma 0%-i lähedale.

Mudelite tõlgendamiseks ja uurimisküsimustele vastamiseks kasutab autor järgnevaid samme:

- 1) Tunnuste mõju tugevuse hindamine seades need olulisuse järjekorda:
 - a. tavalise permutatsiooniga, mille abil saab leida tunnuse marginaalse efekti;
 - b. tingimusliku permutatsiooniga, mis hindab tunnuse lisandväärtust võrreldes teiste tunnustega;
 - c. osalise sõltuvuse funktsiooniga, mis hindab tunnuse marginaalset efekti sõltuva tunnuse varieeruvuse põhjal.
- 2) Tunnuste p-väärtuse hindamine tavalise permutatsiooniga;
- 3) Sõltuva muutuja ja statistiliselt oluliste tunnuste vaheliste seoste tugevuse ja suuna kirjeldamine osalise sõltuvuse jooniste abil.

Optimeeritud *mtry* väärtusega ja 1000 puuga otsustusmetsa mudeli tunnused reastab autor olulisuse järjekorda ning eristab mudeli jaoks informatiivseid ja mitte-informatiivseid tunnuseid. Tunnuste olulisuse hindamise algoritmid jaotuvad mudeli ennustusvõimel põhinevateks (ingl *performance-based measure*) ning ennustatava tunnuse varieeruvusel põhinevateks (ingl *variance-based measure*) algoritmideks (Scholbeck *et al.* 2020). Autor kasutab tunnuse marginaalse efekti mõõtmiseks mõlemast kategooriast ühte meetodit, mis võimaldab tulemused kõrvutada. Esimeseks meetodiks on permutatsiooniolulisus ning teiseks meetodiks sõltuva tunnuse varieeruvuse alusel tunnuste olulisuse järjekorda reastamise meetod (ingl *feature importance ranking measure, FIRM*) (Greenwell *et al.* 2018).

Permutatsiooni võib defineerida kui olemasolevate objektide järjestuse muutmist (Merriam-Webster 2021). Tunnuse permutatsiooniolulisuse leidmiseks permuteeritakse ehk vahetatakse otsustusmetsas konkreetne tunnuse väärtused juhuslikult ning võrreldakse mudeli OOB ennustustäpsust enne tunnuse permuteerimist ja pärast. Mida suurem on erinevus kahel tulemusel, seda olulisem on tunnus. (Breiman 2001a) Tunnuse marginaalse permutatsiooniolulisuse leiab autor teegiga „rfpermute“ (Archer 2020). Marginaalne permutatsiooniolulisus võib kajastada tegelikust olulisemaks sõltuva tunnusega mitteseotud, kuid mõne tegelikult informatiivse tunnusega tugevalt korreleeruva tunnuse. Kuigi otsustusmetsa mudelis ei ole omavahel üle 70% korreleeruvusega tunnusepaare, võib ka väiksem korreleeruvus rolli mängida. Seetõttu leiab autor ka tunnuste osalise permutatsiooniolulisuse ehk tingimusliku permutatsiooniolulisuse teegiga „permimp“ (Debeer *et al.* 2021).

Tingimuslik permutatsiooniolulisus mõõdab tunnuse olulisust võttes arvesse temaga korreleeruvaid tunnuseid. Need seotud tunnused leitakse igas puus eraldi kasutades tunnuste sõlmede hargnevuste põhjal diskretiseerimist ja hii-ruut testi, mille p-väärtuseks on autor valinud vaikeväärtuse ehk 0,05. Tunnuse permutatsioon ei toimu otsustuspuu sõlmedes, vaid selleks luuakse iga puu kohta eraldi tingimusliku permutatsiooni võrk (ingl *conditional permutation grid*), kus iga kord jaotatakse tervet andmestikku, mitte eelnevatest sõlmedest sõltuvat alamandmestikku nagu otsustuspuu sõlmedes. Jaotamine toimub seotud tunnuste diskretiseeritud väärtustega, mis põhinevad otsustuspuu sõlmedel. Tunnuse tingimuslik permutatsiooniolulisus on kõikide võrkude keskmine esialgse ja pärast permutatsiooni saadud OOB vea vahe. (Debeer, Strobl 2020) Kui tunnuste vahel pole olulist korreleeruvust, siis peaks tavaline permutatsiooniolulisus ja tingimuslik permutatsiooniolulisus andma sarnase tunnuste olulisuse järjekorra.

Mudelile informatiivsed on tunnused, mille permutatsiooniolulisus on nullist suurem või mudelis, kus on negatiivse olulisusega tunnuseid, suurem kõige negatiivsema tunnuse absoluutväärtusest. Sellisel juhul saab öelda, et tunnus on sõltuva muutujaga seotud ning teeb mudelit paremaks. (Strobl *et al.* 2009) Kui leidub mitteinformatiivseid tunnuseid, eemaldab autor need mudelist ning kordab *mtry* optimeerimise ja tunnuste permutatsiooniolulisuse leidmise protsessi.

Tunnuste p-väärtuse leiab autor teegiga „rfpermute“ (Archer 2020), mis baseerub marginaalset efekti mõõtväl permutatsiooniolulisusel. Permutatsioonide arvuks valib autor 99, et algoritm ei oleks liialt arvutusressurssi nõudev. Nimetatud teeki on p-väärtuste leidmiseks kasutanud uurimustes ka Banerjee *et al* (2019) ja Levi (2021).

Sõltuva ja selgitava muutuja vaheliste seoste suundade kirjeldamiseks kasutab autor osalise sõltuvuse jooniseid, milleks vajalikud andmed arvutab teegi „pdp“ (Greenwell 2017) funktsioon „partial“. Osalise sõltuvuse funktsioon näitab, kuidas mõjutab mudelis selgitav muutuja sõltuva muutuja väärtust, võttes arve teiste tunnuste keskmist mõju (*Ibid*).

Eeldame, et x_1 on selgitav tunnus, mille väärtused on $\{x_{11}, x_{12}, \dots, x_{1k}\}$. Osalise sõltuvuse funktsiooni arvutamine ja seejärel joonisele kuvamine toimub järgnevalt (*Ibid*):

- 1) Andmestik, millel mudel on loodud, kopeeritakse ja asendatakse kõik x_1 väärtused konstandiga x_{1i} .
- 2) Eelneval sammul saadud andmestiku põhjal leitakse ennustatud sõltuva tunnuse väärtused. Igale kopeeritud andmestiku vaatlusele vastab seega üks ennustatud sõltuva tunnuse väärtus.
- 3) Unikaalsete x_{1i} väärtustega vaatluste ennustatud sõltuva tunnuse väärtused keskmistatakse.
- 4) Saadud paarid (unikaalne x_{1i} väärtus ja keskmine ennustatud sõltuva tunnuse väärtus) kuvatakse joonisele, kus y-teljel on ennustatud sõltuva tunnuse väärtus ja x-teljel x_{1i} väärtus.

Hea praktika, mida kasutab ka autor, on lisaks kuvada joonisele andmestikus olevate selgitava tunnuse (x_1) väärtuste jaotus (ingl *rug*) detiilide kaupa. See võimaldab joonise tõlgendamisel vältida järeldusi piirkondades, kus asetseb vähe selgitava tunnuse väärtusi.

Osalise sõltuvuse funktsiooni põhjal saab arvutada ka tunnuste variatsiooniolulisuse. Teegi „vip“ (Greenwell *et al.* 2018) funktsioon „vi“ mõõdab osalise sõltuvuse funktsiooni põhjal, milline on ennustatava väärtuse standardhälve. Mida suurem standardhälve, seda tugevama mõjuga on selgitav tunnus (Greenwell *et al.* 2018).

Otsustusmetsa mudelite taasloomiseks vajalik kood on lisas 9. Otsustusmetsa mudelite jääkliikmete histogrammid ja hajusgraafikud on toodud lisades 10-15.

3. TULEMUSED JA JÄRELDUSED

3.1. Lineaarse regressiooni mudelite tulemused

Lineaarse regressiooni mudelite tulemused, sh tunnuste koefitsiendid, statistiline olulisus ja standardhälbed, mudeli statistiline olulisus ja kirjeldusvõime ning täpsust iseloomustavad suhtarvud on toodud tabelites 4 ja 5.

Tabel 4. Lineaarse regressiooni mudelite I-III tulemused

	Sõltuv tunnus: ROE		
	I mudel (VK)	II mudel (PVK)	III mudel (LVK)
Konstant	0,0296 (0,0248)	0,0229 (0,0221)	0,0733*** (0,0213)
VK	0,0844** (0,0342)	-	-
PVK	-	0,166*** (0,0379)	-
LVK	-	-	-0,114** (0,0462)
ln (KAIVE)	0,0148*** (0,00186)	0,0152*** (0,00186)	0,0154*** (0,00185)
YoY	0,345*** (0,0119)	0,343*** (0,0119)	0,346*** (0,0120)
VANUS	-0,00716*** (4,25*10 ⁻⁴)	-0,00709*** (4,25*10 ⁻⁴)	-0,00719*** (4,24*10 ⁻⁴)
OKm	0,226*** (0,0308)	0,255*** (0,0285)	0,162*** (0,0217)
POV	-0,126*** (0,0163)	-0,147*** (0,0166)	-0,100*** (0,0136)
CR	-0,00322** (0,00143)	-0,00579*** (0,00155)	-0,00246* (0,00135)
R ²	0,131	0,133	0,131
Adj. R ²	0,131	0,133	0,131
F-statistik	236***	238***	240***
MSE	0,097	0,097	0,097
RMSE	0,312	0,311	0,312
MAE	0,194	0,194	0,194
Vaatluste arv	12 568	12 568	12 568

Allikas: autori arvutused

*** olulisuse nivoo 0,01, ** nivoo 0,05, *nivoo 0,10.

Tabel 5. Lineaarse regressiooni mudelite IV-VI tulemused

	Sõltuv tunnus: ROA		
	IV mudel (VK)	V mudel (PVK)	VI mudel (LVK)
Konstant	-0,0306*** (0,00681)	-0,033*** (0,00632)	-0,0133** (0,00633)
VK	0,0347*** (0,00866)	-	-
PVK	-	0,0694*** (0,00943)	-
LVK	-	-	-0,0509*** (0,0132)
ln (KAIVE)	0,00648*** (6,72*10 ⁻⁴)	0,00664*** (6,70*10 ⁻⁴)	0,00677*** (6,70*10 ⁻⁴)
YoY	0,155*** (0,00448)	0,155*** (0,00448)	0,156*** (0,00448)
VANUS	-0,00336*** (1,58*10 ⁻⁴)	-0,00333*** (1,58*10 ⁻⁴)	-0,00339*** (1,58*10 ⁻⁴)
OKm	0,215*** (0,00785)	0,227*** (0,00752)	0,190*** (0,00573)
POV	-0,0581*** (0,00553)	-0,670*** (0,00569)	-0,0471*** (0,00474)
CR	-7,09*10 ⁻⁴ (6,13*10 ⁻⁴)	-0,00178*** (6,49*10 ⁻⁴)	-4,73*10 ⁻⁴ (5,98*10 ⁻⁴)
R ²	0,254	0,256	0,255
Adj. R ²	0,254	0,256	0,254
F-statistik	494***	495***	437***
MSE	0,013	0,013	0,013
RMSE	0,116	0,116	0,116
MAE	0,084	0,084	0,084
Vaatluste arv	12 703	12 703	12 703

Allikas: autori arvutused

*** olulisuse nivoo 0,01, ** nivoo 0,05, *nivoo 0,1.

ROE-d sõltuva tunnusena kasutavate mudelite kirjeldusvõimet iseloomustav determinatsioonikordaja jäi vahemikku 13,1-13,3% ja ROA-d sõltuva tunnusena kasutatavatel mudelitel vahemikku 25,4-25,6%, mida võib pidada madalateks tasemeteks. Kõik regressioonmudelid osutusid tervikuna statistiliselt oluliseks nivool 0,01 ja kõigis mudelites olid

VK, PVK ja LVK statistiliselt olulised nivool 0,05 või 0,01. Kõige sagedamini ebaolulisemaks osutus tunnus CR, olles statistiliselt oluline vaid pooltes mudelites.

Kõik regressioonimudelid kinnitasid kasumlikkuse ja võlataseme vahelise seose statistilist olulisust, olenemata konkreetsete kasumlikkuse- ja võlamõõdikute kombinatsioonist. VK ja kasumlikkuse (nii ROE kui ROA) vahel on positiivne seos. Võlakordaja suurenemisel ühe protsendipunkti (pp) võrra suureneb ROE 0,08 pp võrra ja ROA 0,0347 pp võrra. Ka PVK ja kasumlikkuse vahel on positiivne seos – PVK suurenemisel ühe pp võrra suureneb ROE 0,166 pp võrra ja ROA 0,0694 pp võrra. Sealjuures on oluline märkida, et PVK ja kasumlikkuse vaheline seos on ligikaudu kaks korda tugevam kui VK ja kasumlikkuse vaheline seos. Vastupidiselt eelnevatele osutus LVK ja kasumlikkuse vahel negatiivne seos negatiivseks – LVK suurenemisel ühe pp võrra väheneb ROE 0,114 pp ja ROA 0,0509 pp.

Kõigi mudelite puhul on täidetud järgnevad lineaarse regressiooni eeldused:

1. Mudelid on lineaarsed parameetrite suhtes.
2. Vaatluste arv (minimaalselt 12 568) on suurem kui hinnatavate parameetrite (7) arv.
3. Selgitavate muutujate väärtused ei ole valimis ühesugused.
4. Ühte mudelisse lülitatud selgitavad muutujad ei ole lineaarselt sõltuvad. Magistritöö kontekstis on lineaarselt sõltuvad võlakordaja ja selle komponendid pikaajaliste võlakohustuste kordaja ja lühiajaliste võlakohustuste kordaja, kuid tunnused ei esine koos samas mudelis.
5. Juhuslike liikmete keskväärtnus on 0. Kuna kõikidesse mudelitesse on hõlmatud vabaliige, siis on see tingimus täidetud.
6. Selgitavad tunnused on sõltumatud juhuslikust liikmest. Kriteeriumit testiti Hausmani testiga, mille kõik mudelid läbisid.
7. Mudelite jääkliikmete vahel ei ole autokorrelatsiooni. Kriteeriumit testiti Wooldridge testiga, mille kõik mudelid läbisid.

Korrektsete standardhälvete jaoks peab olema täidetud homoskedastiivsuse tingimus. Kõigi mudelite puhul esines heteroskedastiivsus ehk tingimus ei olnud täidetud. Seetõttu kasutati kõigis regressioonimudelites Arellano (1987) väljapakutud heteroskedastiivsusega arvestavaid standardhälbeid.

Kõikide mudelite jääkliikmed on normaaljaotusega võrreldes veidi suurema püstakusega, kuid sümmeetrilised. Mitte ükski mudelitest ei läbi formaalset Jarque-Bera testi, mis võrdleb omavahel jääkliikmete ja normaaljaotuse kuju. Suurte valimite korral ei mõjuta aga jääkliikmete normaaljaotusest kõrvalekaldumine teststatistikute õigsust ja ei ole seetõttu oluliseks probleemiks. Regressioonimudelite jääkide histogrammid on toodud magistritöö lisades 16-21.

Mitte ühegi mudeli puhul ei esinenud tunnuste vahel multikollineaarsust. Mudelite varieeruvusindeksite (VIF) maksimaalne väärtus oli 3.0 (ROE-VK mudel), mis indikeerib, et kollineaarsust ei esine. Regressioonimudelid ei läbinud Ramsey RESET spetsifikatsioonitesti, mis viitab, et mudelites esineb mittelineaarseid seoseid. Mudelist mitte läbisaamine on magistritöö kontekstis aktsepteeritav, sest töö põhifookus on hüpoteesi testimisel ja parameetrite hindamisel, mitte ennustustäpsuse tagamisel.

3.2. Otsustusmetsa mudelite tulemused

Sarnaselt lineaarsele regressioonile koostas autor kuus otsustusmetsa mudelit. Kolm ROE ning kolm ROA mudelit, milles igaühes on selgitava muutujana vastavalt kas VK, PVK või LVK.

Optimaalne sõlmedes juhuslikult valitavate tunnuste arv *mtry* oli kõigil mudelitel 1, välja arvatud ROA-LVK mudelil, mille optimaalne *mtry* väärtus oli 2. See on üllatav tulemus, kuna enamasti toob suurem *mtry* väärtus väiksema ennustusvea, sest väiksema informatiivsusega tunnused jäävad kõrvale. Optimaalne *mtry* väärtus 1 viitab sellele, et kõik mudelid on informatiivsed – vastasel juhul oleks optimaalne *mtry* väärtus kõrgem. Kuigi ROA-LVK mudeli optimaalne *mtry* väärtus oli 2, oli sellel väga väike edumaa *mtry* väärtuse 1 ees. *Mtry* optimeerimise detailsed tulemused on lisas 22. Kasutatud hüperparameetrite väärtused on toodud tabelis 6.

Tabel 6. Otsustusmetsa mudelite I – VI hüperparameetrite väärtused (tk)

Hüperparameeter	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
<i>Mtry</i>	1	1	1	1	1	2
<i>Ntrees</i>	1000	1000	1000	1000	1000	1000

Allikas: autori koostatud

Otsustusmetsa ROA mudelid olid ROE mudelitest kõigi mõõdikute kohaselt täpsemad. Seejuures on ROE mudelite täpsus omavahel sarnane ning sama nähtub ka ROA mudelite võrdlusest. Sellest võib järeldada, et VK, PVK ja LVK ennustusvõime ei erine üksteisest olulisel määral. Mudelite täpsust iseloomustavad suhtarvud on toodud tabelist 7.

Tabel 7. Otsustusmetsa mudelite täpsust iseloomustavad mõõdikud

Mõõdik	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
MSE	0,091	0,091	0,090	0,013	0,013	0,012
RMSE	0,301	0,301	0,300	0,112	0,112	0,111
MAE	0,189	0,189	0,189	0,080	0,080	0,080
R ²	15,7%	15,8%	16,3%	29,3%	29,1%	29,8%

Allikas: autori arvutused

Ülesobitumise testide järgi polnud ükski mudel ülesobitunud. Ülesobitumise suhe oli ROE mudelitel 1257 ja ROA mudelitel 1271 (ülesobitunud mudel on väärtusega alla 10). Mudelite p-väärtuste testid näitasid kõigi mudelite olulisust nivool 0,01. Kontrollmudelite ja põhimudelite treening- ja testvea tulemuste kõrvutamisel nähtus, et ennustustäpsuse erinevus jääb nulli lähedale, seega on mudelid stabiilsed (vt lisa 23).

Tunnuste olulisuse hindamisel tavalise permutatsiooni, tingimusliku permutatsiooni ja osalise sõltuvuse funktsiooni abil nähtus, et kõikides mudelites olid kõik tunnused informatiivsed. Seetõttu ühtegi tunnust mudelitest ei eemaldatud. Meetodite tulemuste vahel on vähesed erisused ning on võimalik tuvastada trende.

Kõikides mudelites oli kõige olulisem tunnus käibe kasv. Olulise järjekorra tipus oli ka vanus. Omakapitali määr oli samuti mitmetes mudelites olulisuselt teisel või kolmandal kohal. Kõikides mudelites oli käive olulisuse järjekorra viimasel või eelviimasel kohal ehk võrreldes teistega kõige vähem sõltuva tunnusega seotud. Lühiajaliste kohustuste kattekordaja on kõigi kolme meetodi järgi enamasti keskmise olulisusega. Meetodite võrdluses eristus põhivara osakaal varadest - variatsiooniolulisus ja marginaalne permutatsioon seavad põhivara osakaalu varadest järjekorras viimaseks või eelviimaseks, kuid tingimuslik permutatsioon seab tunnuse ROA mudelites pigem keskele ja ROE mudelites olulisuselt teisele või kolmandale kohale.

Selgitavad muutujad VK, PVK ning LVK olid mudelites üldiselt keskmise või nõrgema olulisusega. Joonistub välja trend, et nii VK, PVK kui ka LVK on kasumlikkusega tugevamalt seotud kui näiteks käive, kuid üldiselt nõrgemalt seotud kui käibe kasv, vanus ja omakapitali määr. Samuti nähtus trend, et LVK on kasumlikkusega nõrgemalt seotud kui PVK ja VK. PVK ja VK on sarnase seose olulisuse tugevusega. Tunnuste olulisuse hindamise tulemused on lisades 24-29.

Uurimustöö fookuses olevad VK, PVK ja LVK olid kõigis mudelites statistiliselt olulised nivool 0,01. Kõigis mudelites olid statistiliselt olulised ka käibe kasv, käive ning vanus. Omakapitali määr oli statistiliselt oluline ROA mudelites ning ühes ROE mudelis. Põhivara osakaal varadest oli statistiliselt oluline vaid ROA mudelites. Lühiajaliste kohustuste kattekordaja polnud statistiliselt oluline üheski mudelis. Tunnuste statistilise olulisuse kokkuvõte on toodud tabelis 8.

Tabel 8. Tunnuste statistiline olulisus

Tunnus	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
VK	***	-	-	***	-	-
PVK	-	***	-	-	***	-
LVK	-	-	***	-	-	**
YoY	***	***	***	***	***	***
OKm	**	-	-	***	***	***
VANUS	***	***	***	***	***	***
CR	-	-	-	-	-	-
POV	-	-	-	***	**	***
KAIVE	***	***	**	***	***	***

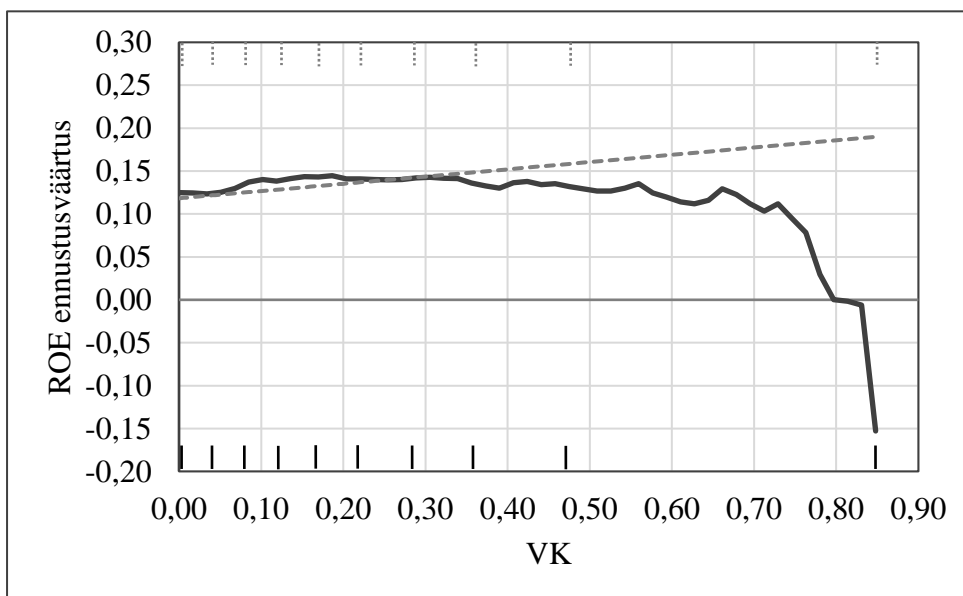
Allikas: autori koostatud

Märkus:

1. *** olulisuse nivoo 0,01, ** nivoo 0,05.
2. Statistilise olulisuse arutamiseks permuteeriti tunnuseid 99 korda.

3.2.1. Koguvõla ja kasumlikkuse seosed

Otsustusmetsa I mudeli kohaselt on ROE ja VK vahel nõrk positiivne seos VK vahemikus 0-19%, mille jooksul ROE tõuseb 12,5% pealt 14,4%-ni. Järgnevas VK vahemikus 19-34% on ROE stabiilne ja vahemikus 34-73% nõrgalt negatiivse, kuid ebaühtlase trendiga. Üle 73% VK väärtuste puhul ilmnenu järsk negatiivne seos võtab aluseks vaid 44 vaatlust, mistõttu ei ole seos suure usaldusväärusega. I mudeli osalise sõltuvuse graafik on toodud joonisel 5.



Joonis 5. ROE ennustusväärtuste osaline sõltuvus VK-st

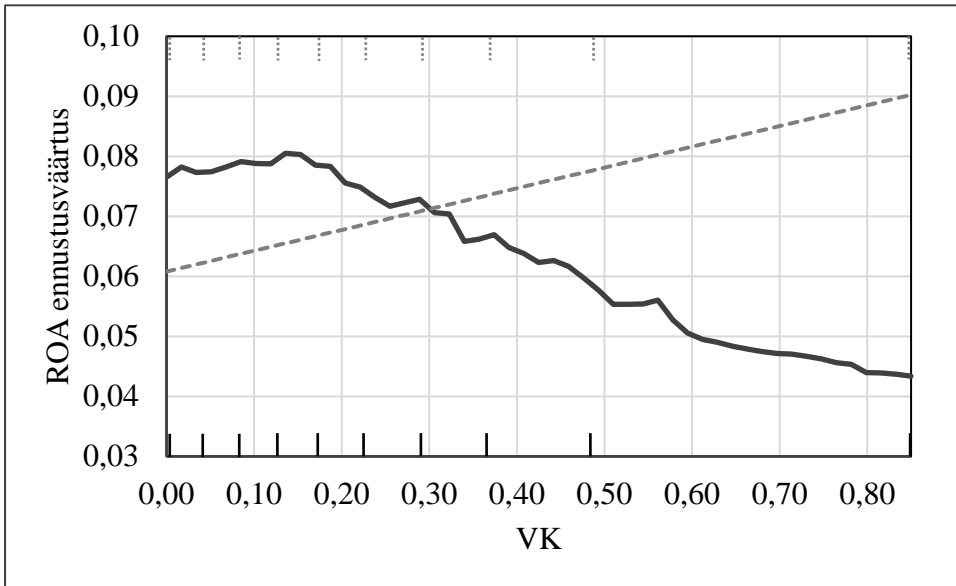
Allikas: autori koostatud

Märkused:

- 1) Pidevjoon näitab otsustusmetsa I mudeli ROE osalist sõltuvust VK-st.
- 2) Katkendjoon näitab lineaarse regressiooni I mudeli ROE osalist sõltuvust VK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detšiilid.
- 4) Jooned joonise ülasaosas on lineaarse regressiooni andmestiku detšiilid.

Seosed otsustusmetsa I mudeli statistiliselt oluliste kontrollnäitajate ja ROE vahel on toodud lisas 30. Omakapitali määra ja ROE vahel on tugev positiivne seos väikese omakapitaliga ettevõtete seas, so OKm vahemikus 0-30%. Kasumlikkus jätkab aeglast tõusu kuni OKm 50%-ni. Sellest tugevamate ettevõtete ROE on nõrgas langustrendis. Käibe ja ROE vahel on tugev positiivne seos kui käibe on kuni 1,7 miljonit eurot, mis moodustab käibe andmestikus esimesed 7 detšiili. Käibe aastase kasvu (tunnus YoY) ja ROE vahel on tugev positiivne seos. Kõige tugevam on seos YoY vahemikus -20% kuni 20% (>6 treeningandmete detšiili), mille jooksul ROE ennustusväärtus kasvab 2,3% pealt 20,8%-ni (*ceteris paribus*). Vanuse ja ROE vahel on tugev negatiivne seos. Seos on tugevam noorte, 0-7 aastaste ettevõtete puhul ning vanuse kasvades aeglustuv. ROE ennustusväärtus väheneb 0-7 aastaste ettevõtete vanusevahemikus 35% pealt 16%-ni ja edasi kuni 30-aastaste ettevõteten, mille ROE on 6,5%. Üle 30-aastaste ettevõtete seas on ROE stabiilne.

IV mudelist joonistub välja selge negatiivne seos VK ja ROA vahel. Kui väikeste VK väärtuste puhul ei ole mõju ROA-le oluline või on väga nõrgalt positiivne, siis VK 14% taset ületades hakkab ROA stabiilselt langema. IV mudeli osalise sõltuvuse graafik on toodud joonisel 6.



Joonis 6. ROA ennustusväärtuste osaline sõltuvus VK-st

Allikas: autori koostatud

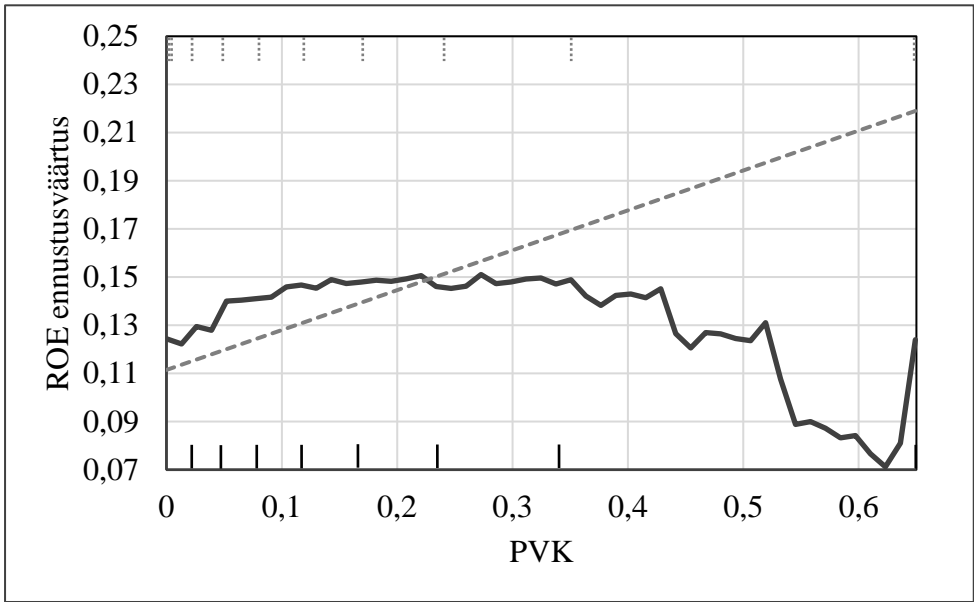
Märkused:

- 1) Pidevjoon näitab otsustusmetsa IV mudeli ROA osalist sõltuvust VK-st.
- 2) Katkendjoon näitab lineaarse regressiooni IV mudeli ROA osalist sõltuvust VK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detšiilid.
- 4) Jooned joonise ülaosas on lineaarse regressiooni andmestiku detšiilid.

Seosed otsustusmetsa IV mudeli statistiliselt oluliste kontrollnäitajate ja ROA vahel on toodud lisas 31. Põhivarade osakaalul koguvaradest (tunnus POV) ja ROA vahel on nõrk negatiivne seos kui POV ületab 45%. Ülejäänud kontrollnäitajate mõjud on analoogsed I mudeli kirjeldustele.

3.2.2. Pikaajalise võla ja kasumlikkuse seosed

PVK ja ROE vaheline seos on eriilmeline. PVK vahemikus 0-22% on ROE-l positiivne trend, kasvades vahemikus 12,4% pealt 15,1%-ni. Positiivne trend hõlmab ca 80% treeningandmestikust. PVK vahemikus 22-35% on ROE varieeruvus väikses vahemikus 14,5-15,1% ja trend on külgsuunaline. Viimases detšiilis (880 vaatlust) ennustab mudel negatiivset trendi, sh mõõdukalt PVK vahemikus 35%-52% ja tugevalt vahemikus 52-62%, millest viimane lõik hõlmab treeningandmetest 256 vaatlust ja ei tulene seega üksikutest ekstreemsetest väärtustest. Mudeli järgi hakkab ROE uuesti kiiresti kasvama, kui PVK ületab 62%, kuid ala hõlmab ainult 35 vaatlust ja trendi murdumine ei ole seega kõrge usaldusväärsusega. II mudeli osalise sõltuvuse graafik on toodud joonisel 7.



Joonis 7. ROE ennustusväärtuste osaline sõltuvus PVK-st

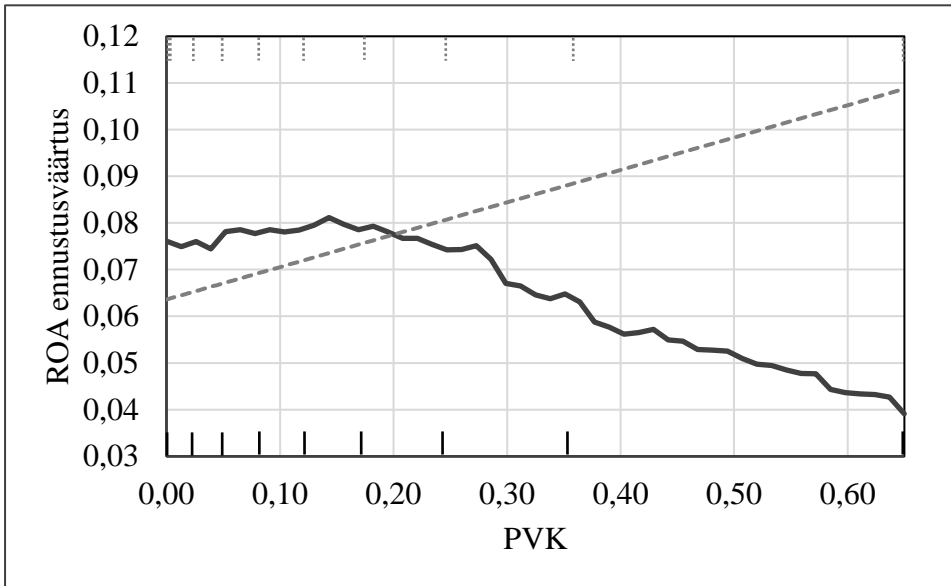
Allikas: autori koostatud

Märkused:

- 1) Pidevjoon näitab otsustusmetsa II mudeli ROE osalist sõltuvust PVK-st.
- 2) Katkendjoon näitab lineaarse regressiooni II mudeli ROE osalist sõltuvust PVK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detšiilid.
- 4) Jooned joonise ülaosas on lineaarse regressiooni andmestiku detšiilid.

Seosed otsustusmetsa II mudeli statistiliselt oluliste kontrollnäitajate ja ROE vahel on toodud lisa 32. Seosed ROE ja kontrollnäitajate VANUS, YoY ja KAIIVE vahel on analoogsed I mudeli juures kirjeldatule.

V mudeli tulemus on sarnane IV mudelile – PVK vahemikus 0-14% on ROA nõrgalt positiivne ja suuremate PVK väärtuste korral esineb selge langustrend. PVK vahemikus 14-65% langeb mudeli ROA ennustusväärtus 8,1% pealt 3,9%-ni. Nii tõusev kui langev trend katavad kumbki 50% treeningandmetest. V mudeli osalise sõltuvuse graafik on toodud joonisel 8.



Joonis 8. ROA ennustusväärtuste osaline sõltuvus PVK-st

Allikas: autori koostatud

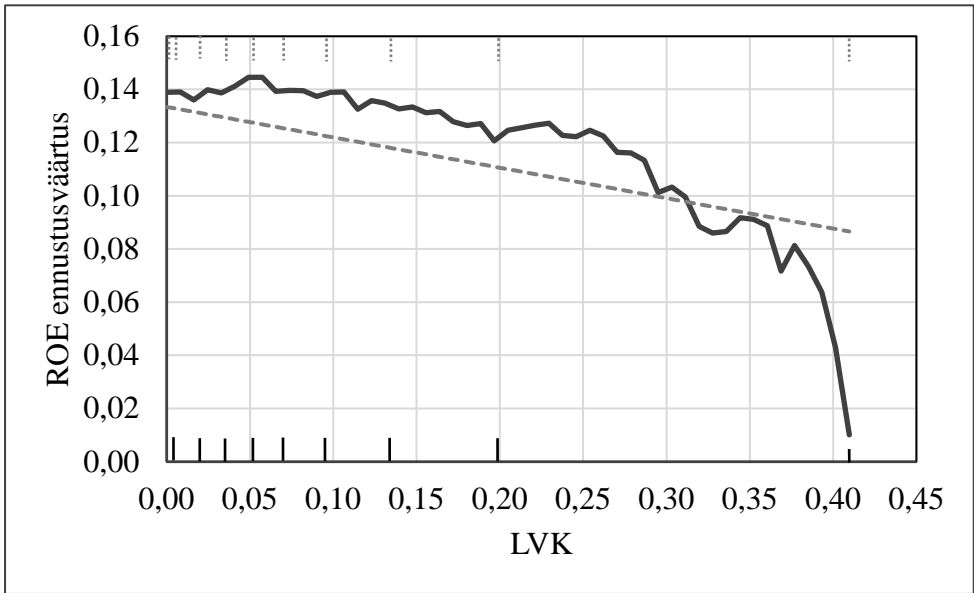
Märkused:

- 1) Pidevjoon näitab otsustusmetsa V mudeli ROA osalist sõltuvust PVK-st.
- 2) Katkendjoon näitab lineaarse regressiooni V mudeli ROA osalist sõltuvust PVK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detšiilid.
- 4) Jooned joonise ülaosas on lineaarse regressiooni andmestiku detšiilid.

Seosed otsustusmetsa V mudeli statistiliselt oluliste kontrollnäitajate ja ROE vahel on toodud lisas 33. Seosed ROE ja kontrollnäitajate VANUS, YoY, KAIVE, OKm ja POV vahel on analoogsed I ja IV mudeli juures kirjeldatule. V mudeli puhul tuleb ka käibe viimaste detšiilide puhul selgemalt välja stabiilne positiivne mõju ROA-le.

3.2.3. Lühiajalise võla ja kasumlikkuse seosed

III mudel kohaselt esineb ROE ja LVK vahel negatiivse ruutvõrdelise suhte laadne seos – madalate LVK väärtuste korral on ROE stabiilne või tõuseb aeglaselt, LVK kasvades kiireneb ROE ennustusväärtuste vähenemine. Mudeli kohaselt saavutab ROE maksimaalse väärtuse 14,5%, kui LVK on 6%, ning väheneb LVK vahemikus 6-20% 12,1%-ni. Viimases detšiilis negatiivne seos tugevneb ja LVK väärtusele 38% vastab ROE ennustusväärtus 8,1%. Vaatlusi, kus LVK oli väiksem kui 38%, esines treeningandmetes 49 tükki, mistõttu ei ole sealne tugev negatiivne seos usaldusväärne, kuna võib olla mõjutatud üksikute kõrge LVK ja madala ROE väärtustega ettevõtetest. III mudeli osalise sõltuvuse graafik on toodud joonisel 9.



Joonis 9. ROE ennustusväärtuste osaline sõltuvus LVK-st

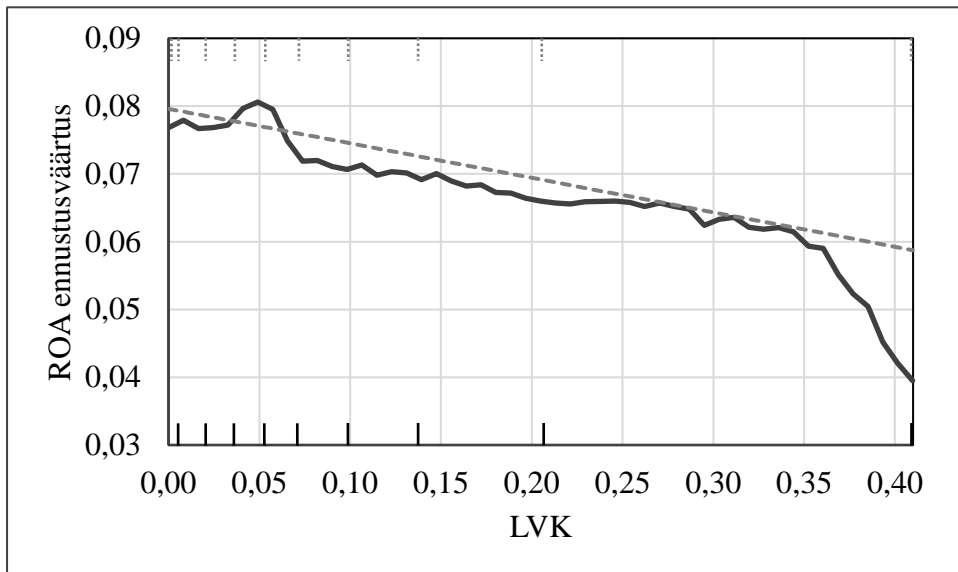
Allikas: autori koostatud

Märkused:

- 1) Pidevjoon näitab otsustusmetsa III mudeli ROE osalist sõltuvust LVK-st.
- 2) Katkendjoon näitab lineaarse regressiooni III mudeli ROE osalist sõltuvust LVK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detšiilid.
- 4) Jooned joonise ülasaosas on lineaarse regressiooni andmestiku detšiilid.

Seosed otsustusmetsa III mudeli statistiliselt oluliste kontrollnäitajate ja ROE vahel on toodud lisa 34. Seosed ROE ning kontrollnäitajate VANUS, YoY ja KAIVE vahel on analoogsed I mudeli juures kirjeldatule.

VI mudelist joonistub välja nõrk positiivne trend LVK ja ROA vahel LVK vahemikus 0-5% (esimesed viis treeningandmete detšiili) ja selge negatiivne trend suuremate LVK väärtuste puhul, mis katab ülejäänud viis treeningandmete detšiili. Sarnaselt III mudelile tugevneb tunnuste vaheline negatiivne seos, kui LVK ületab 35%, kuid VI mudelis on vastavate vaatluste arv suurem ja seos seega usaldusväärsem. Vaatlusi, milles LVK ületas 35%, oli 128. Vaatluste arv jaotus ühtlaselt kogu LVK 35-41% vahemikus. VI mudeli osalise sõltuvuse graafik on toodud joonisel 10.



Joonis 10. ROA ennustusväärtuste osaline sõltuvus LVK-st

Allikas: autori koostatud

Märkused:

- 1) Pidevjoon näitab otsustusmetsa VI mudeli ROA osalist sõltuvust LVK-st.
- 2) Katkendjoon näitab lineaarse regressiooni VI mudeli ROA osalist sõltuvust LVK-st.
- 3) Vertikaalsed jooned X-teljel on otsustusmetsa treeningandmestiku detiilid.
- 4) Jooned joonise ülaosas on lineaarse regressiooni andmestiku detiilid.

Seosed otsustusmetsa VI mudeli statistiliselt oluliste kontrollnäitajate ja ROA vahel on toodud lisas 35. Seosed ROA ja kontrollnäitajate VANUS, YoY, OKm, KAIVE ja POV vahel on analoogsed I ja IV mudeli juures kirjeldatule.

3.3. Järeldused ja ettepanekud

Otsustusmetsa ja lineaarse regressiooni tulemustest järeldus, et võlakohustuste taseme ja kasumlikkuse vahel on statistiliselt oluline seos. Järeldus sai kinnituse kõigi regressiooni mudelite ja otsustusmetsa mudelite puhul.

Lineaarse regressiooni kohaselt on ettevõtte kasumlikkuse ja koguvõla taseme ning pikaajaliste võlakohustuste taseme vahel positiivne seos. Sealjuures on märkimisväärne, et pikaajaliste võlakohustuste kordaja (PVK) ja kasumlikkuse vaheline seos on ligikaudu kaks korda tugevam kui koguvõla taset arvesse võtva võlakordaja (VK) ja kasumlikkuse vaheline seos. PVK suurenemisel ühe protsendipunkti (pp) võrra suureneb ROE 0,166 pp ja ROA 0,0694 pp, sellal kui VK suurenemisel ühe pp võrra suureneb ROE 0,0844 pp võrra ja ROA 0,0347 pp võrra.

Asjaolu aitab selgitada lühiajaliste võlakohustuste kordaja (LVK) ja kasumlikkuse vaheline negatiivseks osutunud seos – LVK suurenemisel ühe pp võrra väheneb ROE 0,114 pp võrra ja ROA 0,0509 pp võrra. Seega on LVK ja kasumlikkuse vaheline negatiivne mõju ca kolmandiku võrra nõrgem kui PVK ja kasumlikkuse vaheline positiivne mõju. Kuna VK on PVK ja LVK summa, sisaldab VK ja kasumlikkuse vaheline seos nii PVK positiivset mõju kui ka LVK negatiivset mõju. Kuna lisaks tugevamale seosele moodustavad pikaajalised võlakohustused ka suurema proportsiooni, so keskmiselt 63% ettevõtete kogu võlakohustustest, on VK ja kasumlikkuse seos tervikuna positiivne.

Otsustusmetsa mudelid viitavad võlakohustuste taseme ja kasumlikkuse vahel olevatele keerulisematele mitte-lineaarsetele seostele. Kuigi seoste kujud on mudelite lõikes veidi erinevad, saab üldistatult välja tuua järgneva kuju:

- 1) Väikeste võlatasemete korral avaldab täiendav võlataseme suurendamine kasumlikkusele nõrka positiivset või neutraalset mõju;
- 2) Teatud võlataseme juures on positiivse trendi murdepunkt;
- 3) Murdepunktist suuremate võlatasemete korral on võlataseme ja kasumlikkuse vahel negatiivne trend, mis on pikem ja tugevam kui esialgne positiivne trend.

ROE-d sõltuva tunnuse kasutavates otsustusmetsa mudelites oli positiivse trendi murdepunktiks VK tase 19%, PVK tase 22% ja LVK tase 6%. Kõige selgemini avaldus esialgne positiivne trend ROE-PVK mudelis ja kõige nõrgemini ROE-LVK mudelis. Murdepunktile järgnev negatiivne trend oli kõige tugevam ROE-LVK ja kõige nõrgem ROE-PVK mudelis.

ROA-d sõltuva tunnuse kasutavate otsustusmetsa mudelites oli positiivse trendi murdepunktiks VK tase 14%, PVK tase 14% ja LVK tase 5%. Erinevus ROE ja ROA mudelite trendi murdepunktides võib tuleneda raamatupidamislikust asjaolust, et erinevalt ROE-st väheneb ROA võlakohustuste lisandumisel ja ettevõtte bilansimahu suurenemisel. Kuna nimetatud vahetut mõju on tunnuse osalise sõltuvuse arvutuses keerukas arvestada, on järeldustes lähtutud konservatiivsematest murdepunktidest.

Kuigi otsustusmetsa mudelites on prominentsem võlataseme ja kasumlikkuse vaheline negatiivne seos, jaguneb otsustusmetsa mudelite treeningandmestik positiivse ja negatiivse trendi vahel kõigis mudelites ligikaudu pooleks. Kuna võlakordajate andmete jaotus on suure positiivse asümmeetriaga, siis ei ole väga kõrgete võlatasemete juures toimuvad järsud muutused tugevalt

põhjendatud. Sellised muutused esinesid ROE-PVK mudelis järsu tõusuna ja ROE-LVK mudelis järsu langusena suurte võlatasemetel korral.

Tabelis 9 on toodud kasumlikkuse mõõdikute ROE ja ROA väärtuste amplituud otsustusmetsa osalise sõltuvuse funktsiooni positiivse ja negatiivse trendi vahemikes. Mitmetes mudelites toimusid andmestike viimastes detšiilides sõltuva tunnuse järsud muutused, mis treeningandmete hajususest tulenevalt on madala usaldusväärsusega. Seetõttu on negatiivse trendi puhul välja toodud ka sõltuva tunnuse muutus maksimaalsest väärtusest treeningandmestiku 9. detšiili väärtuseni. Tabelist nähtub, et III ja VI mudeli positiivne trend on nõrgem kui II ja V mudeli trend. III mudeli negatiivne trend on tugevam kui II mudeli oma.

Tabel 9. ROE ja ROA väärtuste amplituud osalise sõltuvuse funktsiooni järgi

	Positiivse trendi amplituud	Negatiivse trendi koguamplituud	Negatiivse trendi amplituud 9. detšiilini
I mudel (ROE-VK)	2,0%	29,7%	1,3%
II mudel (ROE-PVK)	2,7%	8,0%	0,4%
III mudel (ROE-LVK)	0,6%	13,4%	2,4%
IV mudel (ROA-VK)	0,4%	3,7%	2,1%
V mudel (ROA-PVK)	0,5%	4,2%	1,6%
VI mudel (ROA-LVK)	0,3%	4,1%	1,5%

Allikas: autori arvutused.

Jättes väikse andmestikutihedusega treeningandmete viimase detšiili arvestamata, on ROE-PVK mudelis sõltuva tunnuse kasv positiivse trendi jooksul oluliselt suurem kui langus järgneva negatiivse trendi jooksul. ROA-PVK mudelis ületas negatiivse trendi jooksul toimunud sõltuva tunnuse vähenemine esialgset suurenemist.

Otsustusmetsa mudelid osutusid lineaarse regressiooni mudelitest vähesel määral täpsemaks nii keskmise ruutvea (RMSE) kui keskmise absoluutvea (MAE) arvestuses. Mudelite täpsust iseloomustav kokkuvõte on toodud tabelis 10.

Tabel 10. Lineaarse regressiooni ja otsustusmetsa mudelite täpsuse võrdlus.

Mudel	Lin. regressioon (LR)		Otsustusmets (OM)		Vahe (LR - OM)	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
I mudel (ROE-VK)	0,312	0,194	0,301	0,189	0,010	0,005
II mudel (ROE-PVK)	0,311	0,194	0,301	0,189	0,010	0,004
III mudel (ROE-LVK)	0,312	0,194	0,300	0,189	0,011	0,004
IV mudel (ROA-VK)	0,116	0,084	0,112	0,080	0,004	0,004
V mudel (ROA-PVK)	0,116	0,084	0,112	0,080	0,004	0,003
VI mudel (ROA-LVK)	0,116	0,084	0,111	0,080	0,004	0,004

Allikas: autori arvutused

Kahe meetodi tulemuste kokkuvõtteks võib järeldada, et Eesti töötleva tööstuse ettevõtete kasumlikkuse ja lühiajaliste võlakohustuste taseme vahel on peamiselt negatiivne seos. Otsustusmetsa kohaselt on seos negatiivne kui LVK ületab 6%, lineaarse regressiooni järgi on seos negatiivne olenemata LVK väärtusest. Seega ühtib mudelite tulemuste ühisosa otsustusmetsa negatiivse trendi vahemikuga. LVK vahemikus 0-6% on seos otsustusmetsa järgi nõrgalt positiivne ja lineaarse regressiooni järgi negatiivne. Kuigi otsustusmets osutus vähesel määral täpsemaks ja meetodi eeliseks on mitte-lineaarsete seoste tuvastamine, tuleks autori hinnangul LVK väärtuste 0-6% piirkonda täiendavalt uurida, et seose suuna ja tugevuse tulemuse osas suuremat usaldusväarsust saavutada.

LVK ja kasumlikkuse negatiivset trendi võib selgitada lühiajaliste võlakohustuste võtmisega seonduvad kõrgemad intressikulud, mis kasumlikkust vähendavad. Samuti ei saa välistada tunnuste vastupidist kausaalsust – nõrga kasumlikkuse korral võivad kreditorid nõuda lühemaid tagasimaksegraafikuid, mis suurendab ettevõtte lühiajaliste võlakohustuste osakaalu.

Kasumlikkuse ja pikaajaliste võlakohustuste taseme vahel on positiivne seos kui PVK tase on alla 14%, mida kinnitab nii otsustusmetsa kui ka lineaarse regressiooni tulemus. Suuremate PVK tasemete korral osutavad otsustusmetsa mudelid negatiivsele, aga lineaarne regressioon positiivsele seosele. Kuna tulemused nimetatud vahemikus on meetodite lõikes erinevad, siis ei saa magistritöö raames lõplikult järeldada, milline on kasumlikkuse ja PVK vaheline seos suuremate PVK väärtuste korral ja vajab täiendavat uurimist.

Koguvõla taseme ja kasumlikkuse seost uurinud mudelites olid kahe meetodi ühisosaks VK väärtused, mis olid väiksemad kui 14% – selles VK vahemikus on kasumlikkuse ja VK vahel

positiivne seos. Suuremate VK väärtuste puhul on otsustusmetsa tulemuste kohaselt kasumlikkuse ja koguvõla taseme vaheline seos negatiivne, lineaarse regressiooni puhul positiivne, mistõttu ei saa magistritöö raames lõpliku suuna ja tugevuse osas järeldust teha ning vajalik on piirkondade täiendav uurimine.




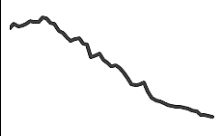


Autori hinnangul tuleks koguvõla taseme ja kasumlikkuse vahelist seost käsitleda kasumlikkuse ja pikaajaliste võlakohustuste taseme vahelise seose ning kasumlikkuse ja lühiajaliste võlakohustuste seose koondmõjuna. Vastasel juhul jääks koguvõlakohustuste komponentide eraldiseisvad ja magistritöö raames vastupidiseks osutunud seosed kasumlikkusega peitu. Töö raames osutus tunnuste vaheline seos väikeste VK tasemete korral positiivseks, sest kasumlikkuse ja pikaajaliste võlakohustuste vaheline positiivne seos oli tugevam kui kasumlikkuse ja lühiajaliste võlakohustuste vaheline seos, kuid muude andmestikega ei ole välistatud ka vastupidine olukord või olukord, kus VK komponentmõjud on võrdsed, mistõttu VK ja kasumlikkuse vahel seost agregeerituna ei esine.

Kokkuvõtlikult saab töös püstitatud uurimisküsimusele vastata järgmiselt:

- 1) Kogu võlakohustuste taseme ja kasumlikkuse vahel on positiivne seos VK vahemikus 0-14%. Tunnuste vaheline seos suuremate VK väärtuste korral ei ole magistritöö raames üheselt määratletud.
- 2) Pikaajaliste võlakohustuste taseme ja kasumlikkuse vahel on positiivne seos PVK vahemikus 0-14%. Tunnuste vaheline seos suuremate PVK väärtuste puhul ei ole magistritöö raames üheselt määratletud.
- 3) Lühiajaliste võlakohustuste taseme ja kasumlikkuse vahel on negatiivne seos, kui LVK on suurem kui 6%. Tunnuste vaheline seos LVK väärtuste vahemikus 0-6% ei ole magistritöö raames üheselt määratletud.
- 4) Kasumlikkuse (nii ROE kui ka ROA) ja kõigi kolme finantsvõimenduse näitaja vahel on statistiliselt oluline seos.

Tabelis 11 on toodud ROE ja ROA ning selgitavate tunnuste VK, PVK ja LVK vaheliste oodatud seoste ja regressioonanalüüsi ning otsustusmetsa tegelike tulemuste võrdlus.

Tabel 11. Võlakordajate ja kasumlikkuse oodatud ja tegelike seoste võrdlus.

Muutuja	VK	PVK	LVK
Oodatud seos kasumlikkusega	-	-	-
Lineaarse regressiooni tulemus	+	+	-
Otsustusmetsa tulemus (ROE andmestikuga)			
Otsustusmetsa tulemus (ROA andmestikuga)			

Allikas: autori koostatud

Tulemused vastavad ootusele osaliselt ja tingimuslikult. Kõige selgemalt sai kinnituse eeldus, et seos LKV ja kasumlikkuse vahel on negatiivne, aga seda vaid juhul, kui LVK on suurem kui 6%. Väiksemate LVK väärtuste korral ei ole tunnuste vaheline seos magistritöö raames üheselt määratud. Negatiivse seose lühiajaliste võlakohustuste taseme ja kasumlikkuse vahel on leidnud Eesti tööstusettevõtetes ka Mokhova ja Zineckeri (2013). Tulemus on kooskõlas finantshierarhia teooriaga, mille järgi eelistavad ettevõtted enda kohustusi finantseerida eelisjärjekorras omavahenditest, sest laenu võtmine on info asümmeetria tõttu kulukam. Kõrgema kasumlikkusega ettevõtetel on suurem võimekus lühiajalisi kohustusi omavahenditest finantseerida, mistõttu nende keskmine võlatase on madalam kui väiksema kasumlikkusega ettevõtetel.

Pikaajaliste võlakohustuste oodatud negatiivne seos langeb kokku otsustusmetsa ennustusega, mille järgi on PVK ja kasumlikkuse vaheline seos negatiivne kui PVK > 14%, kuid on vastupidine lineaarse regressiooni tulemusele. Mõlemat mudelit arvesse võttes ei ole magistritöö raames võimalik tunnuste vahelist seost > 14% PVK väärtuste puhul üheselt määrata. Kuna väiksemate PVK väärtuste korral on tunnuste vaheline seos positiivne, siis on kokkuvõtlikult PVK ja kasumlikkuse oodatud ja tegelikud suunad vastupidised. Sarnaselt PVK-le on vastupidised ka VK ja kasumlikkuse oodatud ja tegelikud suunad.



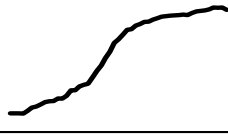
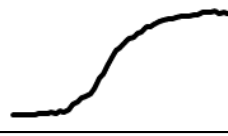
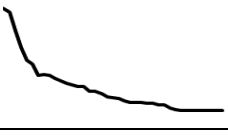
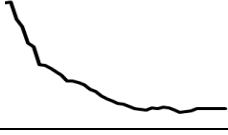
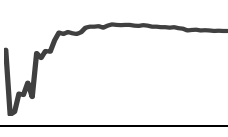
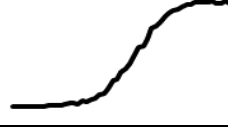
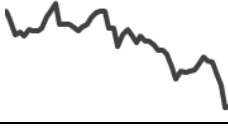
Positiivse seose PVK ja kasumlikkuse vahel on leidnud USA teenindus- ja tööstussektori börsiettevõtetes Gill *et al.* (2011) ning Itaalia keskmise suurusega tööstusettevõtetes de Luca (2014). Positiivse seose VK ja kasumlikkuse vahel on leidnud Prantsuse tekstiili- ja kemikaalitööstuses Margaritis ja Psillaki (2010). Samuti on leidnud VK ja kasumlikkuse vahel positiivse seose Itaalia keskmise suurusega tööstusettevõtetes de Luca (2014). Pattitoni *et al.* (2014) leidsid Euroopa tööstusettevõtetes mittelineaarse mudeliga, et madala VK korral on seos ROA-ga positiivne ja alates teatud tasemest negatiivne (nende uurimuses alates 32%).

PVK ja VK ning kasumlikkuse vahelised positiivsed seosed toetavad agendikulude teooriat, mille järgi teatud tasemini muudab võlg ettevõtte efektiivsemaks, kuna vähendab ettevõtte omanike ja ettevõtte siseste agentide vahelisi kulusid. Saavutades optimaalse punkti, hakkavad agendikulud uuesti suurenema, sel korral ettevõtte omanike ja kreditoride vahel.

VK ja PVK positiivset seost kasumlikkusega võib põhjendada ettevõtte suurenenud võimekusega ellu viia kasumlikke, positiivse nüüdisväärtusega investeerimisprojekte. Madalama võlakoormusega ettevõtete krediidirisk on väiksem, *ceteris paribus*, mistõttu nõuavad kreditorid väiksemat finantseeringuga kaasnevat riskipreemiat. Võlakoormuse tõustes kreditoride nõutav riskipremia suureneb, mistõttu teatud tasemest alates võlakapitali suurendamine kasumlikkust ei paranda. Samuti võib teatud võlataseme ületamisel väheneda ettevõtte kättesaadavus võõrfinantseerimisele.

Kontrollnäitajate oodatud ja tegelike tulemuste kokkuvõte on toodud tabelis 12. Kontrollnäitajate seosed kasumlikkusega osutusid regressioonanalüüsi ja otsustusmetsa rakendades samadeks. Varasematest uuringutest tulenenud oodatud seosed olid töö tulemustega vastuolus ainult tunnuse CR puhul.

Tabel 12. Kontrollnäitajate oodatud ja tegelikud tulemused.

Muutuja	Oodatav seos	Regressiooni tulemus	Otsustusmetsa tulemused (ROE andmestikuga)	Otsustusmetsa tulemused (ROA andmestikuga)
Käive (KAIVE)	+	+		
Käibe aastane kasvumäär (YoY)	+	+		
Ettevõtte vanus (VANUS)	-	-		
Omakapitali määr (OKm)	+	+		
Põhivarade osakaal koguvardadest (POV)	-	-	Stat. ebaoluline	
Lühiajaliste kohustuste kattekordaja (CR)	+	Negatiivne või stat. ebaoluline	Stat. ebaoluline	Stat. ebaoluline

Allikas: autori koostatud.

Märkused:

- 1) Tunnus CR oli statistiliselt ebaoluline lineaarse regressiooni III, IV ja VI mudelis. II ja V mudelites oli tunnus oluline nivool 0,01 ja I mudelis nivool 0,05.
- 2) Otsustusmetsa osalise sõltuvuse joonised on I (ROE-VK) ja IV (ROA-VK) mudelitest. Ülejäänud otsustusmetsa mudelite osalise sõltuvuse graafikud olid esitatutega analoogsed.
- 3) Lisaks tabelis märgitule osutus statistiliselt ebaoluliseks tunnus OKm otsustusmetsa II ja III mudelis.

Käibe ja kasumlikkuse seos osutus vastavalt oodatule positiivseks. Otsustusmetsa osalise sõltuvuse funktsioon viitab, et valdav osa positiivsest mõjust kasumlikkusele avaldub käibe kasvamisel kuni 1,7 miljoni euroni ja edasisel käibe kasvul on kasumlikkusele piiratud positiivne mõju. Positiivne trend on loogiline – tööstusettevõtte kasvades suureneb tööprotsesside efektiivsus ja uute seadmete soetamisega tootmisefektiivsus.

Seost käibe aastase kasvumäära (YoY) ja kasumlikkuse vahel iseloomustab otsustusmetsa mudelite järgi logistilise kõvera funktsiooni-laadne seos. Suurte käibe kahanemiste korral on kasumlikkuse ennustusväärtus minimaalne. Kasvumäära lähenemisel nullile kasumlikkuse kasv

kiireneb. Kui aastane käibekasvumäär ületab nulli, hakkab kasumlikkuse kasvu kiirus vähenema ning maksimaalsete kasvumäärade juures on ka kasumlikkus maksimaalne. Tulemus on ootuste kohane arvestades tööstussektori võrdlemisi suurt püsikulude osakaalu, mis tingib käibe kukkudes ebaproportsionaalselt suure kasumi languse ja käibe tõustes vastava kasumi kasvu.

Seost ettevõtte vanuse ja kasumlikkuse vahel iseloomustab otsustusmetsa mudelite järgi eksponentsiaalse kahanemise laadne funktsioon. Kasumlikkuse ennustusväärtus on maksimaalne äsja asutatud ettevõtetel ja langeb ettevõtte vananedes. Noorte ettevõtete puhul langeb kasumlikkuse ennustusväärtus kiiremini ja kasumlikkuse kahanemine aeglustub ettevõtte vananedes. Vanuse negatiivset mõju saab seostada väiksema innovatsiooni ja paindlikkusega ning uute konkurentsivõimelisemate ärimudelite lisandumisega.

Omakapitali määra (OKm) ja kasumlikkuse seost iseloomustab sarnaselt aastasele käibekasvumäärale logistilise S-kõveralaadne funktsioon. Tulemusi võib pidada loogiliseks – ettevõtte stabiilseks toimimiseks on vaja teatud määral finantspuhvleid, mis aitavad katta ootamatuid kahjusid ning on välistele osapooltele usaldusväärsuse märgiks. Väga tugeva kapitaliseerituse korral ei ole positiivne efekt nii suur ja ROE langeb, sest omakapital on ROE valemis jagajas.

Põhivarade osakaalu varadest (POV) ja kasumlikkuse vahel on negatiivne seos. Seos on küllaltki volatiilne, mis viitab, et tunnus on nõrga kirjeldusvõimega ning graafikul esinev volatiilsus on tingitud teiste tunnuste koosmõjust. Autori hinnangul võib negatiivset seost selgitada pidevate kasumlikkust vähendavate investeeringute ja käibevarade osakaalu vähenemisega. Käibevarade vähenemine ei ole tingitud põhivarade (osakaalu) kasvamisest, kuid võib iseloomustada vähemaktiivset (nt väiksemate nõuete ja varudega) ettevõtet.

Kokkuvõttes tuvastati statistiliselt oluliste kontrollnäitajate ja kasumlikkuse vaheliste seoste suunad järgnevalt:

- 1) Käibe ja kasumlikkuse vahel on positiivne mitte-lineaarne seos;
- 2) Käibe aastase kasvumäära ja kasumlikkuse vahel on positiivne mitte-lineaarne seos;
- 3) Ettevõtte vanuse ja kasumlikkuse vahel on negatiivne mitte-lineaarne seos;
- 4) Omakapitali määra ja kasumlikkuse vahel on positiivne mitte-lineaarne seos;
- 5) Põhivarade osakaalul varadest ja kasumlikkusel on negatiivne mitte-lineaarne seos.

Võttes arvesse magistritöös kasutatud andmeid, meetodikaid ja tulemusi ning nendest tulenevaid piiranguid, teeb autor ettepanekud edasisteks uurimusteks:

- 1) Keskenduda kasumlikkuse ja võlakohustuste taseme uurimisele nendes finantsvõimenduse piirkondades, milles näitasid magistritöös rakendatud meetodikad vastupidiseid tulemusi. Nendeks piirkondadeks on ettevõtted, mille VK ületab 14%, PVK ületab 14% või LVK on alla 6%.
- 2) Jaotada valim ettevõtte suuruse põhjal osadeks. Empiirilises kirjanduses on leitud (nt Luca 2014), et erineva suurusega ettevõtete seos kasumlikkuse ja võlataseme vahel on erinev.
- 3) Lisada selgitavaks tunnuseks ettevõtte alamsektor. Töötleva tööstuse alamsektorites (võimalik jaotus lisas 1) valitsevad erinevad ärimudelid ja turuolukorrad, mistõttu osa ettevõtte kasumlikkusest võib olla selgitatav tema alamsektoriga.
- 4) Eraldada lühiajaliste võlakohustuste hulgast pikaajaliste võlakohustuste lühiajaline osa. Vastavalt kehtivale raamatupidamiskorrale kajastatakse pikaajaliste võlakohustuste järgneva aasta jooksul tasutav osa lühiajalise võlakohustusena. Seetõttu sisaldab LVK ja kasumlikkuse vaheline seos osaliselt ka olemuse poolest PVK ja kasumlikkuse vahelist seost. Kuna pikaajalist ja lühiajalist finantseeringut kasutatakse erinevatel eesmärkidel, võivad PVK ja LVK seosed kasumlikkusega olla erinevad ning pikaajaliste võlakohustuste lühiajalise osa lühiajaliste võlakohustuste koosseisus esitamisel võib olla tulemustele moonutav mõju. Praktikas eeldaks see magistritöös kasutatud Orbis Europe andmebaasist muu allika kasutamist või suuremahulist tööd ettevõtete majandusaasta aruannetega.
- 5) Jagada võlakohustused konkreetsete võlatoodete vahel. Erinevatel võlatoodetel on erinevad kasutuse eesmärgid (näiteks investeerimislaenud on investeringute teostamiseks, käibekapitali laenud käibevahendite, nt varude ostmiseks jne), seonduvad kulud ja mõju ettevõtte tegevusele ja seetõttu ka potentsiaalselt erinev mõju ettevõtte kasumlikkusele. Ettepaneku elluviimine eeldaks tihedat koostööd ja andmete saamist krediitiasutustelt, mis võib praktikas andmekaitse kaalutlustel võimatuks osutada.
- 6) Kasutada otsustusmetsa paneelandmeid võimaldavaid edasiarendusi. Magistritöö kirjutamise seisuga olid sellised meetodid olemas, kuid ilma magistritöö eesmärgi saavutamiseks vajalike tugifunktsioonideta.
- 7) Rakendada kausaalsust tuvastavaid meetodeid. Üheks magistritöö võimalikuks edasiarenduseks on kasumlikkuse ja võlakohustuste vahelise kausaalsuse hindamine. Üheks potentsiaalseks meetodiks on magistritöös rakendatud otsustusmetsa edasiarendus, kausaalne otsustusmets.

KOKKUVÕTE

Kasumlikkuse kujunemise ja kapitali struktuuri valiku küsimused on ettevõtete jaoks olulise tähtsusega. Selles hoolimata ei ole tunnuste vahelise seose osas selget konsensust ei teoreetilisel tasandil ega empiirilistest uurimustest tulenevalt.

Magistritöö eesmärgiks oli hinnata Eesti töötleva tööstuse ettevõtete kasumlikkuse ja võlataseme vahelise seose olulisust, suunda ja tugevust. Eesmärgi saavutamiseks püstitati järgnevad uurimisküsimused:

- 1) Milline on kogu võlakohustuste taseme ja kasumlikkuse seos?
- 2) Milline on pikaajaliste võlakohustuste taseme ja kasumlikkuse seos?
- 3) Milline on lühiajaliste võlakohustuste taseme ja kasumlikkuse seos?

Töö empiirilises uurimuses oli sõltuvaks muutujaks kasumlikkus, mille mõõdikuna kasutati omakapitali tootlust (ROE) ja varade tootlust (ROA). Finantsvõimendust kirjeldavad selgitavad muutujad olid võlakohustuste ja varade jagatis ehk võlakordaja (VK) ja selle komponendid – pikaajaliste võlakohustuste kordaja (PVK) ja lühiajaliste võlakohustuste kordaja (LVK). Mudelites kasutati kontrollnäitajatena ettevõtte vanust, omakapitali määra, käivet, käibe aastast kasvumäära, põhivarade osakaalu koguvaradest ja lühiajaliste kohustuste kattekordajat.

Uurimisobjektiks olid Eesti töötleva tööstuse ettevõtted aastatel 2011-2019. Töö käigus moodustati kaks valimit, millest esimeses kasutati sõltuva tunnuseks ROE-d ja teises ROA-d. Valimite suuruseks kujunes vastavalt 12 568 ja 12 703 vaatlust 9-aastase perioodi jooksul. Unikaalseid ettevõtteid oli esimeses valimis 2064 ja teises valimis 2063.

Uurimisküsimustele vastamiseks rakendati kahte meetodit, millest üks on laialt kasutatud mitmene lineaarne regressioon ja teine valdkonnas vähe rakendatud otsustusmets. Kahe meetodi paralleelne rakendamine aitas tuvastada nii need uurimisprobleemi piirkonnad, mille kohta saab magistritöö raames usaldusväärseid järeldusi teha, kui ka need piirkonnad, millele üheselt vastata ei saanud.

Lineaarse regressiooni kohaselt esineb VK ja kasumlikkuse ning PVK ja kasumlikkuse vahel positiivne seos ning LVK ning kasumlikkuse vahel negatiivne seos. VK suurenemisel ühe protsendipunkti (pp) võrra suureneb ROE 0,0844 pp ja ROA 0,0347 pp. PVK suurenemisel ühe pp

võrra suureneb ROE 0,166 pp ja ROA 0,0694 pp. LVK suurenemisel ühe pp võrra väheneb ROE 0,114 pp ja ROA 0,0509 pp võrra.

Otsustusmetsa mudelid viitavad võlakohustuste taseme ja kasumlikkuse vahel keerulisematele mitte-lineaarsele seostele. Kuigi seosed on mudelite lõikes veidi erinevad, saab üldistatult välja tuua järgneva kuju:

- 1) Väikeste võlatasemete korral avaldab täiendav võlataseme suurendamine kasumlikkusele nõrka positiivset või neutraalset mõju;
- 2) Teatud võlataseme juures on positiivse trendi murdepunkt;
- 3) Murdepunktist suuremate võlatasemete korral on võlataseme ja kasumlikkuse vahel negatiivne trend, mis on pikem ja tugevam kui esialgne positiivne trend.

Nimetatud positiivse trendi murdepunktid olid ROE-d sõltuva tunnusega kasutatavates mudelites VK tase 19%, PVK tase 22% ja LVK tase 6% ja ROA-d sõltuva tunnusega kasutatavates mudelites VK tase 14%, PVK tase 14% ja LVK tase 5%.

Mõlema meetodi tulemusi arvesse võttes saab uurimisküsimusele vastata järgmiselt:

- 1) Kogu võlakohustuste taseme ja kasumlikkuse vahel on positiivne nõrk seos VK vahemikus 0-14%.
- 2) Pikaajaliste võlakohustuste taseme ja kasumlikkuse vahel on positiivne seos PVK vahemikus 0-14%. Seos on nõrgapoolne, kuid tugevam kui VK ja kasumlikkuse vahel.
- 3) Lühiajaliste võlakohustuste taseme ja kasumlikkuse vahel on negatiivne seos kui LVK on suurem kui 6%. Negatiivne seos on tugevam kui VK ja kasumlikkuse vaheline positiivne seos, kuid nõrgem kui PVK ja kasumlikkuse vaheline positiivne seos.
- 4) Mõlema kasumlikkuse mõõdiku (ROE ja ROA) ning kõigi kolme finantsvõimenduse näitaja vahel on statistiliselt olulised seosed.

Kasumlikkuse ja erinevate võlatasemete vahelised seosed võlakordajate piirkondades, kus $VK > 14\%$, $PVK > 14\%$ või $LVK < 6\%$, ei ole tunnuste vaheline seos üheselt määratletud, sest nendes piirkondades ennustavad lineaarne regressioon ja otsustusmets vastupidiseid tulemusi.

Magistritöö tulemused toetavad lühiajaliste võlakohustuste võtmise kontekstis finantshierarhia teooriat ja pikaajaliste võlakohustuste võtmise kontekstis agendikulude teooriat. Tulemustest järeldub, et ettevõtted eelistavad enda lühiajaliste kohustuste katteks kasutada esmajärjekorras

omavahendeid. Pikaajaliste võlakohustuste teatud tasemeni suurendamine (magistritöös kuni 14%-ni) vähendab agendikuluseid ja on seega kasumlikkusega positiivses seos.

Autori hinnangul võib koguvõla taseme ja pikaajalise võlataseme positiivset seost kasumlikkusega põhjendada ettevõtte suureneva võimekusega ellu viia kasumlikke, positiivse nüüdisväärtusega investeerimisprojekte. Madalama võlakoormusega ettevõtete krediidirisk on väiksem, *ceteris paribus*, mistõttu nõuavad kreditorid väiksemat finantseeringuga kaasnevat riskipremiat. Võlakoormuse tõustes kreditoride nõutav riskipremia suureneb, mistõttu teatud tasemest alates võlakapitali suurendamine kasumlikkust ei paranda. Samuti võib teatud võlataseme ületamisel väheneda ettevõtte kättesaadavus võõrfinantseerimisele.

Magistritöö järeldused on üldistatavad ainult Eesti töötleva tööstuse ettevõtetele. Esitatud tulemused ja järeldused on muu hulgas sõltuvad kasutatud andmestikest ja nende andmete kvaliteedist, valimi korrigeerimise meetodikast ja analüüsis hõlmatud ajaperioodist. Magistritöö ei hinnanud kasumlikkuse ja võlakohustuste taseme vahelist kausaalsust.

Edasistes uurimustes võiks autori hinnangul keskenduda kasumlikkuse ja võlakohustuste taseme seose põhjalikumale hindamisele nendes ettevõtete gruppides, mille kohta magistritöös vastukäivate tulemuste tõttu lõplikke järeldusi teha ei saanud. Need on ettevõtted, mille $VK > 14\%$, $PVK > 14\%$ või $LVK < 6\%$. Samuti võiks magistritöö edasiarendusena rakendada paneelandmeid kasutavaid ja kausaalsust tuvastavaid otsustusmetsa meetodeid. Neist viimane toetaks magistritöös tuvastatud seoste sisuliste põhjuste välja selgitamist.

SUMMARY

THE RELATIONSHIP BETWEEN CAPITAL STRUCTURE AND PROFITABILITY BASED ON ESTONIAN MANUFACTURING COMPANIES

Artur Kohv

Despite the apparent importance of profitability formation and capital structure choices, there is neither theoretical nor empirical consensus on how these factors are interrelated. The two most prominent theories in the field include the pecking-order theory and static trade-off theory. According to the former, companies prefer to use their own funds to cover costs as incurring debt is accompanied by costs. Thus, more profitable companies with more liquid funds use less debt compared to their less profitable counterparts. The theory suggests a negative relationship between leverage and profitability. According to the static trade-off theory, companies weigh the potential risks and benefits of incurring debt. As the risks for profitable companies are comparably lower than for companies with low profitability, profitable companies incur more debt. Overall, the static trade-off theory proposes a positive relationship between leverage and profitability, directly contradicting the pecking order theory. Prior academic studies have found support for both proposals, while more results point to a negative relationship between the variables.

In order to help clarify the currently mixed results in the local context, this thesis aims to determine the relationship between leverage and profitability in the case of Estonian manufacturing firms. The following research questions were asked:

- 1) What is the statistical importance, direction and strength of the relationship between total leverage and profitability?
- 2) What is the statistical importance, direction and strength of the relationship between long-term leverage and profitability?
- 3) What is the statistical importance, direction and strength of the relationship between short-term leverage and profitability?

Measuring firm profitability, ROE and ROA were selected as dependent variables. For leverage, three measures were tested, including debt to assets ratio (VK) and its components, long-term debt to assets ratio (PVK) and short-term debt to assets ratio (LVK). Control variables included

turnover, year-over-year turnover growth, firm age, equity ratio, non-current assets to total assets ratio and current ratio.

Data from the Orbis Europe database was used. The sample was restricted to Estonian manufacturing companies from 2011-2019, based on available data. Small firms with a turnover of less than 50 thousand euro and assets less than 30 thousand euro were excluded, along with companies with missing or extreme values. Two similar samples were constructed based on the used dependent variable (ROE or ROA). Final samples contained 12568 and 12703 observations over 9 years and included over 2000 unique companies.

In empirical analysis, two methodologies were employed – multiple linear regression and random forest. The latter has not been widely used in corporate finance topics for hypothesis testing, however, the methodology proved appropriate and insightful for answering the set research questions. In linear regression, pooled OLS was picked over fixed and random effects panel models based on the relevant F-tests and Breuch-Pagan tests. Tested time dummies proved reliably irrelevant to the regression model. By employing two models with different working logic, complexity, assumptions, and outputs enabled to assess the confidence level of different results and identify regions of data for which results are inconclusive.

Based on the linear regression results, the relationship between PVK and profitability was positive – one percentage point (pp) increase in PVK would increase ROE by 0.166 pp and ROA by 0.0694 pp. At the same time, the relationship between LVK and profitability was negative – one pp increase in LVK would result in 0.114 pp decrease in ROE and 0.0509 pp decrease in ROA. VK is the sum of PVK and LVK. As the positive effect of PVK became stronger than the negative effect of LVK, then the overall relationship of VK and profitability was also negative, whilst weaker than PVK. One pp increase in VK would result in 0.0844 pp increase in ROE and 0.0347 pp increase in ROA.

The results of random forest are more complex. Whilst models differ to some degree; the prevailing function could be described as follows:

- 1) At low leverage, the relationship between leverage and profitability is positive,
- 2) At certain leverage, the positive trend breaks,
- 3) A longer and stronger negative trend follows the breaking point.

The breaking point was at VK 19%, PVK 22% and LVK 6% for models using ROE as the dependent variable and VK 14%, PVK 14% and LVK 5% for models using ROA as the dependent variable. If one considers the results of both the linear regression and random forest, the following conclusions can be made about the research questions:

- 1) VK and profitability have a weak positive relationship when VK is between 0-14%,
- 2) PVK and profitability have a positive relationship when PVK is between 0-14%. The relationship between PVK and profitability is stronger than that of VK and profitability.
- 3) LVK and profitability have a weak negative relationship if LVK is larger than 6%. The relationship is weaker than for PVK and profitability but stronger than for VK and profitability.
- 4) The relationships between both profitability measures (ROE and ROA) and all three leverage measures (VK, PVK and LVK) were statistically important.

The results are inconclusive if $VK > 14\%$, $PVK > 14\%$ or $LVK < 6\%$ as the results of the linear regression and random forest are contradictory in these ranges.

The relationship between LVK and profitability supports the pecking order theory, suggesting that Estonian manufacturing companies prefer to cover their short-term financing needs from internal funds. On the other hand, the relationship between PVK and profitability rejects pecking order theory and seems to support the arguments made by the static trade-off theory, i.e. companies mostly consider the future gains and losses when deciding whether to incur debt. The positive relationship between profitability and low VK and PVK (up to 14%) supports agency theory that moderate long-term debt decreases agency costs.

The conclusions drawn in the thesis are only applicable to the Estonian manufacturing sector companies. The results and conclusions depend on the data used and its quality, the method of compiling a sample and the specific time period considered. No conclusions on the causality between leverage and profitability were drawn.

Future research might focus on the ranges of leverage where results obtained by the methods used in this thesis were insufficient to draw a reliable conclusion. New knowledge could also be obtained by employing newer random forest methods which can use panel data and test causality. The latter might help advance understanding of the reasons behind the relationships observed in this thesis.

KASUTATUD ALLIKATE LOETELU

- Abdou, H. A., Kuzmic, A., Pointon, J. (2012). Determinants of Capital Structure in the UK Retail Industry: A Comparison of Multiple Regression and Generalized Regression Neural Network. *Intelligent Systems in Accounting Finance & Management*, 19 (3), 151–169.
- Abor, J. (2007). Industry Classification and the Capital Structure of Ghanaian SMEs. *Studies in Economics and Finance*, 24 (3), 207–219.
- Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Oyedele, A. A., Akinade, O. O., Bilal, M., Ajayi, S. O. (2017). Critical Factors for Insolvency Prediction: Towards a Theoretical Model for the Construction Industry. *International Journal of Construction Management*, 17 (1), 25–49.
- Amini, S., Elmore, R., Öztekin, Ö., Strauss, J. (2021). Can Machines Learn Capital Structure Dynamics? *SSRN Working Paper*. The Netherlands, Amsterdam: Elsevier. Kättesaadav: <https://ssrn.com/abstract=3473322> (01.05.2021)
- Archer, E. (2020.). *rfPermute: Estimate Permutation p-Values for Random Forest Importance Metrics*. Kättesaadav: <https://CRAN.R-project.org/package=rfPermute>, 04.04.2021.
- Arellano, M. (1987). Computing Robust Standard Errors for Within-Groups Estimators. *Oxford Bulletin of Economics and Statistics*, 49 (4), 431–434.
- Aryantini, S., Jumono, S. (2021). Profitability and Value of Firm: An Evidence from Manufacturing Industry in Indonesia. *Accounting*, 7 (4), 735–746.
- Baker, M., Wurgler, J. (2002). Market Timing and Capital Structure. *The Journal of Finance*, 57 (1), 1–32.
- Balik, B. H., Gort, M. (1993). Decomposing learning by doing in new plants. *Journal of Political Economy*, 101 (4), 561-583.
- Bancel, F., Mittoo, U. R. (2004). Cross-Country Determinants of Capital Structure Choice: A Survey of European Firms. *Financial Management*, 33 (4), 103–132.
- Banerjee, S., Walder, F., Büchi, L., Meyer, M., Held, A. Y., Gattinger, A., van der Heijden, M. G. (2019). Agricultural Intensification Reduces Microbial Network Complexity and the Abundance of Keystone Taxa in Roots. *Multidisciplinary Journal of Microbial Ecology*, 13, 1722–1736.
- Baxter, N. D. (1967). Leverage, Risk of Ruin and the Cost of Capital. *The Journal of Finance*, 22 (3), 395–403.
- Boehmke, B., & Greenwell, B. (2020). *Hands-On Machine Learning with R*. USA, Boca Raton: CRC Press.

- Bolton, P., Scharfstein, D. S. (1990). A Theory of Predation Based on Agency Problems in Financial Contracting. *The American Economic Review*, 80 (1), 93–106.
- Brealey, R. A., Myers, S. C., Allen, F. (2019). *Principles of Corporate Finance* (13th ed). USA, New York: McGraw-Hill Education.
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45 (1), 5–32.
- Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Statistical Science*, 16 (3), 199–231.
- Brennan, M. J., Schwartz, E. (1978). Corporate Income Taxes, Valuation, and the Problem of Optimal Capital Structure. *The Journal of Business*, 51 (1), 103–114.
- Brounen, D., Jong, A. d., Koedijk, K. (2006). Capital Structure Policies in Europe: Survey Evidence. *Journal of Banking & Finance*, 30 (5), 1409–1442.
- Capitaine, L., Genuer, R., Thiébaud, R. (2019). Random forests for high-dimensional longitudinal data. *Eprint arXiv:1901.11279v1 Working Paper*. USA, New York: Cornell Tech.
- Čeh, M., Kilibarda, M., Lisec, A., Bajat, B. (2019). Estimating the Performance of Random Forest versus Multiple Regression for Predicting Prices of the Apartments. *ISPRS International Journal of Geo-Information*, 7 (5), 168.
- Dalci, I. (2018). Impact of Financial Leverage on Profitability of Listed Manufacturing Firms in China. *Pacific Accounting Review*, 30 (4), 410–432.
- DeAngelo, H., Masulis, R. W. (1980). Optimal Capital Structure Under Corporate and Personal Taxation. *Journal of Financial Economics*, 8 (1), 3–29.
- Debeer, D., Strobl, C. (2020). Conditional Permutation Importance Revisited. *BMC Bioinformatics*, 21 (1), 1–30.
- Debeer, D., Hothorn, T., & Strobl, C. (2021). *permimp: Conditional Permutation Importance*. Kättesaadav: <https://CRAN.R-project.org/package=permimp>, 04.04.2021.
- Doğan, M. (2013). Does Firm Size Affect the Firm Profitability: Evidence from Turkey Data of 200 Companies Active in the Istanbul Stock Exchange between 2008-2011. *Research Journal of Finance and Accounting*, 4 (4), 53-59.
- Duda, R. O., Hart, P. E., Stork, D. G. (2001). *Pattern Classification* (2nd ed). USA, New York: John Wiley & Sons, Inc.
- Eesti Statistikaamet. (2021). EM001: Ettevõtete majandusnäitajad tegevusala ja tööga hõivatud isikute arvu järgi [E-andmebaas]. Kättesaadav: <https://andmed.stat.ee/>, 04.04.2021.
- Eesti Statistikaamet. (2021). EM001A: Ettevõtete majandusnäitajad tegevusala ja tööga hõivatud isikute arvu järgi [E-andmebaas]. Kättesaadav: <https://andmed.stat.ee/>, 04.04.2021.

- Eesti Statistikaamet. (2021). EM009: Ettevõtete vara, kohustused ja omakapital tegevusala (EMTAK 2008) ja tööga hõivatud isikute arvu järgi [E-andmebaas]. Kättesaadav: <https://andmed.stat.ee/>, 04.04.2021.
- Eesti Statistikaamet. (2021). RAA042: Lisandväärtus tegevusala järgi (ESA 2010) [E-andmebaas]. Kättesaadav: <https://andmed.stat.ee/>, 04.04.2021.
- Eesti Statistikaamet. (2021). VK02: Kaupade eksport ja import majandusüksuse registreerimise koha, tegevusala (EMTAK 2008) ja KN kaubajaotise järgi [E-andmebaas]. Kättesaadav: <https://andmed.stat.ee/>, 04.04.2021.
- Farrar, D. E., Selwyn, L. L. (1967). Taxes, Corporate Financial Policy and Return to Investors. *National Tax Journal*, 20 (4), 444–454.
- Fisher, E. O., Heinkel, R., Zechner, J. (1989). Dynamic Capital Structure Choice: Theory and Tests. *The Journal of Finance*, 44 (1), 19–40.
- Frank, M., Goyal, V. K. (2009). Capital Structure Decisions: Which Factors are Reliably Important? *Financial Management*, 38 (1), 1–37.
- Genuer, R., Poggi, J.-M. (2020). Random Forests with R. Springer International Publishing.
- Gill, A., Biger, N., & Mathur, N. (2011). The effect of Capital Structure on Profitability: Evidence from the United States. *International Journal of Management*, 28 (4), 3–15.
- Goddard, J., Tavakoli, M., Wilson, J. O. (2005). Determinants of Profitability in European Manufacturing and Services: Evidence from a Dynamic Panel Model. *Applied Financial Economics*, 15 (18), 1269–1282.
- González, V. M., González, F. (2011). Firm Size and Capital Structure: Evidence Using Dynamic Panel Data. *Applied Economics*, 44 (36), 4745–4754.
- González-Martel, C., Hernández, J. M., Manrique, C. (2021). Identifying Business Misreporting in VAT Using Network Analysis. *Decision Support Systems*, 141:113464.
- Graham, J. R., Harvey, C. R. (2001). The Theory and Practice of Corporate Finance: Evidence from the Field. *Journal of Financial Economics*, 60 (2-3), 187–243.
- Greene, W. H. (2018). *Econometric Analysis* (8th ed). UK, London: Pearson.
- Greenwell, B. (2017). Pdp: An R Package for Constructing Partial Dependence Plots. *The R Journal*, 9 (1), 421–436.
- Greenwell, B., Boehmke, B., McCarthy, A. (2018). A Simple and Effective Model-Based Variable Importance Measure. *Eprint arXiv:1805.04755 Working Paper*. USA, New York: Cornell Tech.
- Haas, R., Peeters, M. (2006). The Dynamic Adjustment Towards Target Capital Structures of Firms in Transition Economies. *The Economics of Transition*, 14 (1), 133–169.

- Hall, G., Hutchinson, P., Michaelas, N. (2000). Industry Effects on the Determinants of Unquoted SMEs' Capital Structure. *International Journal of the Economics of Business*, 7 (3), 297–312.
- Hamner, B., Frasco, M. (2018). *Metrics: Evaluation Metrics for Machine Learning*. Kättesaadav: <https://CRAN.R-project.org/package=Metrics>, 04.04.2021.
- Han, H., Guo, X., Yu, H. (2016). Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest. *2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, China, Beijing, 26-18.08.2016 (219-224). USA, New York: IEEE.
- Hannan, M. T., Freeman, J. (1984). Structural Inertia and Organizational Change. *American Sociological Review*, 49 (2), 149-164.
- Hanousek, J., Shamshur, A. (2011). A Stubborn Persistence: Is the Stability of Leverage Ratios Determined by the Stability of the Economy? *Journal of Corporate Finance*, 17 (5), 1360–1376.
- Hapfelmeier, A., Hothorn, T., Ulm, K., Strobl, C. (2014). A New Variable Importance Measure for Random Forests with Missing Data. *Statistics and Computing*, 24, 21–34.
- Harris, M., & Raviv, A. (1985). A Sequential Signaling Model of Convertible Debt Call Policy. *The Journal of Finance*, 40 (5), 1263–1281.
- Hastie, T., Tibshirani, R., Friedman, J. (2017). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). USA, New York: Springer.
- Horne, J. V., Wachowicz, J. M. (2008). *Fundamentals of Financial Management* (13th ed). USA, New York: Pearson.
- Hsiao, C. (2007). Panel Data Analysis – Advantages and Challenges. *Test*, 16 (1), 1–22.
- Ishwaran, H. K., Kogalur, U. B. (2021). *Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC)*. Kättesaadav: <https://cran.r-project.org/web/packages/randomForestSRC/index.html>, 04.04.2021.
- Iyer, A., Jeyalatha, S., Sumbaly, R. (2015). Diagnosis of Diabetes Using Classification Mining Techniques. *International Journal of Data Mining & Knowledge Mangament Process*, 5 (1), 1–14.
- Jablonka, K. M., Ongari, D., Moosavi, S. M., Smit, B. (2020). Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chemical Reviews*, 120 (16), 8066–8129.
- Jensen, M. C. (1986). Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review*, 76 (2), 323–329.
- Jensen, M. C., Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3 (4), 305–360.

- Jong, A., Kabir, R., Nguyen, T. T. (2008). Capital Structure Around the World: The Roles of Firm- and Country-Specific Determinants. *Journal of Banking & Finance*, 32 (9), 1954–1969.
- Jõeveer, K. (2018). Firm Capital Structure in Europe: A Comparative Analysis of CEE Firms vs. Western Firms in the Changing Financial Environment. Suder, G., Riviere, M., Lindeque, J. (Eds.). *The Routledge Companion to European Business*, (267-281). UK, Oxfordshire: Routledge.
- Kraus, A., Litzenberger, R. H. (1973). A State-Preference Model of Optimal Financial Leverage. *The Journal of Finance*, 28 (4), 911–922.
- Kroajczyk, R. A., Lucas, D. J., McDonald, R. L. (1992). Equity Issues with Time-Varying Asymmetric Information. *The Journal of Financial and Quantitative Analysis*, 27 (3), 397–417.
- Kuhn, M. (2008). Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 28 (5).
- Kuhn, M., Johnson, K. (2013). *Applied Predictive Modeling*. USA, New York: Springer.
- Lazaridis, I., Tryfonidis, D. (2006). Relationship Between Working Capital Management and Profitability of Listed Companies in the Athens Stock Exchange. *Journal of Financial Management & Analysis*, 19 (1), 26-35.
- Leland, H. E. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *Journal of Finance*, 49 (4), 1213–1252.
- Leland, H. E., Pyle, D. H. (1977). Informational Asymmetries, Financial Structure, and Financial Intermediation. *The Journal of Finance*, 32 (2), 371–387.
- Levi, S. (2021). Country-Level Conditions Like Prosperity, Democracy, and Regulatory Culture Predict Individual Climate Change Belief. *Communications Earth & Environment*, 2 (51).
- Li, H., Leung, K.-S., Wong, M.-H., Ballester, P. J. (2014). Substituting Random Forest for Multiple Linear Regression Improves Binding Affinity Prediction of Scoring Functions: Cyscore as a Case Study. *BMC Bioinformatics*, 15 (1), 291.
- Liaw, A., Wiener, M. (2002). Classification and Regression by randomForest. *R news*, 2/3, 18–22.
- Liu, B., Xia, Y., Yu, P. (2005). Clustering Via Decision Tree Construction. Chu, W., Lin, T. Y. (Eds.). *Foundations and Advances in Data Mining* (97-124). USA, New York: Springer.
- Luca, P. D. (2014). Capital Structure and Economic Performance of the Firm: Evidence from Italy. *International Journal of Management*, 5 (3), 1–20.

- Lucas, D. J., McDonald, R. L. (1990). Equity Issues and Stock Price Dynamics. *The Journal of Finance*, 45 (4), 1019–1043.
- Maksimovic, V., Titman, S. (1991). Financial Policy and Reputation for Product Quality. *The Review of Financial Studies*, 4 (1), 175–200.
- Margaritis, D., Psillaki, M. (2010). Capital Structure, Equity Ownership and Firm Performance. *Journal of Banking & Finance*, 34 (3), 621–632.
- Marsland, S. (2014). *Machine Learning: An Algorithmic Perspective* (2nd ed). USA, Boca Raton: CRC Press.
- Mazur, K. (2007). The Determinants of Capital Structure Choice: Evidence from Polish Companies. *International Advances in Economic Research*, 13 (4), 495–514.
- McAlexander, R. J., Mentch, L. (2020). Predictive Inference with Random Forests: A New Perspective on Classical Analyses. *Research & Politics*, 7 (1).
- Mentch, L., Hooker, G. (2016). Quantifying Uncertainty in Random Forests via Confidence Intervals and Hypothesis Tests. *Journal of Machine Learning Research*, 17 (26), 1–41.
- Michaelas, N., Chittenden, F., Poutziouris, P. (1999). Financial Policy and Capital Structure Choice in UK SMEs: Empirical Evidence from Company Panel Data. *Small Business Economics*, 12 (2), 113–130.
- Microsoft Corporation ja Steve Weston. (2020). *doParallel: Foreach Parallel Adaptor for the 'parallel'*. Kätesaadav: <https://CRAN.R-project.org/package=doParallel>, 04.04.2021.
- Miglo, A. (2010). The Pecking Order, Trade-Off, Signaling, and Market-Timing Theories of Capital Structure: A Review. K. H. Baker, G. S. Martin (Eds.). *Capital Structure and Corporate Financing Decisions: Theory, Evidence, and Practice* (lk 171–191). USA, New York: John Wiley & Sons.
- Miller, M. H. (1977). Debt and Taxes. *The Journal of Finance*, 32 (2), 261–275.
- Modigliani, F., Miller, M. H. (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, 48 (3), 261–297.
- Modigliani, F., Miller, M. H. (1963). Corporate Income Taxes and the Cost of Capital: A Correction. *The American Economic Review*, 53 (3), 433–433.
- Mokhova, N., Zinecker, M. (2013). The Determinants of Capital Structure: The Evidence from the European Union. *Acta Univ. Agric. Silvic. Mendelianae Brun*, 61 (7), 2533–2546.
- Murphy, M. A., Evans, J. S., Storfer, A. (2010). Quantifying *Bufo boreas* Connectivity in Yellowstone National Park with Landscape Genetics. *the Ecological Society of America*, 91 (1), 252–261.
- Myers, S. C. (1984). Capital Structure Puzzle. *The Journal of Finance*, 39 (3), 575–592.

- Myers, S. C., Majluf, N. S. (1984). Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have. *Journal of Financial Economics*, 13 (2), 187–221.
- Myers, S. C. (2015). Finance, Theoretical and Applied. *Annual Review of Financial Economics*, 7, 1-34.
- Nivorozhkin, E. (2005). Financing Choices of Firms in EU Accession Countries. *Emerging Markets Review*, 6 (2), 138–169.
- Norvaišiene, R. (2012). The Impact of Capital Structure on the Performance Efficiency of Baltic Listed Companies. *Inžinerine Ekonomika*, 23 (5), 505–516.
- Ozgur, O., Karagol, E. T., Ozbugday, F. C. 2021. Machine Learning Approach to Drivers of Bank Lending: Evidence from An Emerging Economy. *Financial Innovation*, 7:20.
- Pao, H. T. (2008). A Comparison of Neural Network and Multiple Regression Analysis. *Expert Systems with Applications*, 35 (3), 720–727.
- Papadogonas, E. M. (2007). Financial Performance of Large and Small Firms: Evidence from Greece. *International Journal of Financial Service Management*, 2 (1/2), 14-20.
- Pattitoni, P., Petracci, B., Spisni, M. (2014). Determinants of Profitability in the EU-15 Area. *Applied Financial Economics*, 24 (11), 763–775.
- Permutation. (2021). *Merriam-Webster's online dictionary*. Kättesaadav: <https://www.merriam-webster.com/dictionary/permutation>, 04.04.2021.
- R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. Viin: R Foundation for Statistical Computing. Kättesaadav: <https://www.R-project.org/>, 04.04.2021.
- Rahman, A., Sarker, S. I., Uddin, J. (2019). The Impact of Capital Structure on the Profitability of Publicly Traded Manufacturing Firms in Bangladesh. *Applied Economics and Finance*, 6 (2), 1–5.
- Rajan, R. G., Zingales, L. (1995). What Do We Know about Capital Structure? Some Evidence from International Data. *Journal of Finance*, 50 (5), 1421–1460.
- Robichek, A. A., Myers, S. C. (1966). Problems in the Theory of Optimal Capital Structure. *The Journal of Financial and Quantitative Analysis*, 1 (2), 1–35.
- Roden, D. M., Lewellen, W. G. (1995). Corporate Capital Structure Decisions: Evidence from Leveraged Buyouts. *Financial Management*, 24 (2), 76–87.
- Ross, S. A. (1977). The Determination of Financial Structure: The Incentive-Signaling Approach. *The Bell Journal of Economics*, 8 (1), 23–40.
- RPS RT I, 15.03.2019, 12, § 3.

- RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. USA, Boston: RStudio, PBC. Kättesaadav: <http://www.rstudio.com/>, 04.04.2021.
- Saeed, M. S., Mustafa, M. W., Sheikh, U. U., Jumani, T. A., Mirjat, N. H. (2019). Ensemble Bagged Tree Based Classification for Reducing Non-Technical Losses in Multan Electric Power Company of Pakistan. *Electronics*, 8, 860.
- Sammut, C., & Webb, G. I. (Eds.). (2017). *Encyclopedia of Machine Learning and Data Mining*. USA, Boston: Springer.
- Sander, P. (2005). Laenukapitali maksueelis Eestis – müüt või tegelikkus? *Eesti ettevõtlike perspektiivid Euroopa Liidus* (170-178). Tallinn: Mattimar OÜ.
- Sander, P. (2014). *Konkurentsiameti poolt väljatöötatud kaalutud keskmise kapitali hinna (WACC) arvutamise meetodika analüüs*. Kättesaadav: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUEwjly5HquTvAhXBFXcKHRauDSYQFjAAegQIBRAD&url=https%3A%2F%2Fwww.konkurentsiamet.ee%2Fsites%2Fdefault%2Ffiles%2F3_wacc_analuus_2014_konkurent ameti_tellimus_21.07.14.pdf&usg=AOvVaw1oWd9jLaKVGQoHOj0s-RVU, 04.04.2021.
- Scholbeck, C. A., Molnar, C., Heumann, C., Bischl, B., Casalicchio, G. (2020). Sampling, Intervention, Prediction, Aggregation: A Generalized Framework for Model-Agnostic Interpretations. *Eprint arXiv:1904.03959v4 Working Paper*. USA, New York: Cornell Tech.
- Scott, J. H. (1976). A Theory of Optimal Capital Structure. *The Bell Journal of Economics*, 7 (1), 33–54.
- Seppa, R. (2008). Capital Structure Decisions: Research in Estonian Non-Financial Companies. *Baltic Journal of Management*, 16 (2), 55–70.
- Solomon, E. (1963). Leverage and the Cost of Capital. *The Journal of Finance*, 18 (2), 273–279.
- Stiglitz, J. E. (1972). Some Aspects of the Pure Theory of Corporate Finance: Bankruptcies and Take-overs. *The Bell Journal of Economics and Management Science*, 3 (2), 458–482.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., Hothorn, T. (2007). Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. *BMC Bioinformatics*, 8 (25).
- Strobl, C., Malley, J., Tutz, G. (2009). An Introduction to Recursive Partitioning: Rationale, Application and Characteristics of Classification and Regression Trees, Bagging and Random Forests. *Psychological Methods*, 14 (4), 323–348.
- Škuláňová, N. (2019). Influence of Selected Determinants on the Financial Structure in the Civil Engineering Companies in the Selected Countries. *Socialiniai tyrimai / Social Research*, 42 (2), 5–16.

- Talberg, M., Winge, C., Frydenberg, S., Westgaard, S. (2008). Capital Structure Across Industries. *International Journal of the Economics of Business*, 15 (2), 181–200.
- Taub, A. J. (1975). Determinants of the Firm's Capital Structure. *The Review of Economics and Statistics*, 57 (4), 410–416.
- Titman, S. (1984). The Effect of Capital Structure on a Firm's Liquidation Decision. *Journal of Financial Economics*, 13 (1), 137–151.
- Titman, S., Wessels, R. (1988). The Determinants of Capital Structure Choice. *The Journal of Finance*, 43 (1), 1–19.
- Washimi, K. (2020). Revisiting Determinants of Investor Sentiment in the FX Option Market by Machine Learning Approaches. *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, Australia, Canberra 1-4.12.2020 (22-27). USA, New York: IEEE.
- Wright, M. N., Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*, 77 (1), 1–17.
- Yang, J-A., Chou, S-R., Cheng, H-C., Lee, C-H. (2010). The Effects of Capital Structure on Firm Performance in the Taiwan 50 and Taiwan Mid-Cap 100. *Journal of Statistics and Management Systems*, 13 (5), 1069–1078.

LISAD

Lisa 1. Valimisse lülitatud NACE Rev. 2 koodid ja nende tähendus

- 10 – toiduainetööstus,
- 11 – joogitööstus,
- 12 – tubakatööstus,
- 13 – tekstiilitööstus,
- 14 – rõivatööstus,
- 15 – nahktoodete tootmine,
- 16 – puidutööstus,
- 17 – paberitööstus,
- 18 – trükitööstus,
- 19 – koksi ja rafineeritud petrooliumtoodete tootmine,
- 20 – keemiatööstus,
- 21 – farmaatsiatööstus,
- 22 – kummi ja plastikute tootmine,
- 23 – mitte-metalseste mineraaltoodete tootmine,
- 24 – baasmetallide tootmine,
- 25 – metalltoodete tootmine,
- 26 – arvutite ja optikaseadmete tootmine,
- 27 – elektroonikatööstus,
- 28 – masinatööstus,
- 29 – mootorsõidukite tootmine,
- 30 – muu transpordivahendite tootmine,
- 31 – mööblitööstus.

Allikas: Orbis Europe andmebaas

Lisa 2. ROE andmestiku korrelatsioonimaatriks

Tunnus	ROE	VANUS	KAIIVE	YoY	VARA	TOOT	CR	OKm	POV	VK	PVK	LVK
ROE		-17%	2%	31%	-1%	-2%	6%	6%	-10%	-4%	-2%	-6%
VANUS	-17%		15%	-18%	19%	24%	2%	20%	11%	-13%	-13%	-4%
KAIIVE	2%	15%		0%	91%	76%	-5%	-1%	-6%	1%	-3%	6%
YoY	31%	-18%	0%		-3%	-5%	-3%	-8%	-2%	7%	7%	3%
VARA	-1%	19%	91%	-3%		74%	-2%	3%	4%	3%	0%	6%
TOOT	-2%	24%	76%	-5%	74%		-9%	-3%	-2%	2%	-2%	6%
CR	6%	2%	-5%	-3%	-2%	-9%		57%	-36%	-34%	-19%	-40%
OKm	6%	20%	-1%	-8%	3%	-3%	57%		-3%	-66%	-55%	-43%
POV	-10%	11%	-6%	-2%	4%	-2%	-36%	-3%		32%	32%	12%
VK	-4%	-13%	1%	7%	3%	2%	-34%	-66%	32%		88%	58%
PVK	-2%	-13%	-3%	7%	0%	-2%	-19%	-55%	32%	88%		13%
LVK	-6%	-4%	6%	3%	6%	6%	-40%	-43%	12%	58%	13%	

Allikas: autori arvutused

Lisa 3. ROA andmestiku korrelatsioonimaatriks

Tunnus	ROA	VANUS	KAIVE	YoY	VARA	TOOT	CR	OKm	POV	VK	PVK	LVK
ROA		-15%	2%	33%	0%	-3%	21%	30%	-12%	-19%	-14%	-16%
VANUS	-15%		15%	-18%	19%	24%	3%	21%	11%	-14%	-14%	-5%
KAIVE	2%	15%		0%	90%	76%	-4%	0%	-6%	0%	-3%	6%
YoY	33%	-18%	0%		-3%	-4%	-3%	-6%	-2%	7%	7%	3%
VARA	0%	19%	90%	-3%		74%	-1%	4%	4%	3%	0%	5%
TOOT	-3%	24%	76%	-4%	74%		-8%	-2%	-2%	1%	-2%	7%
CR	21%	3%	-4%	-3%	-1%	-8%		56%	-35%	-34%	-19%	-39%
OKm	30%	21%	0%	-6%	4%	-2%	56%		0%	-64%	-54%	-43%
POV	-12%	11%	-6%	-2%	4%	-2%	-35%	0%		31%	31%	11%
VK	-19%	-14%	0%	7%	3%	1%	-34%	-64%	31%		88%	58%
PVK	-14%	-14%	-3%	7%	0%	-2%	-19%	-54%	31%	88%		13%
LVK	-16%	-5%	6%	3%	5%	7%	-39%	-43%	11%	58%	13%	

Allikas: autori arvutused

Lisa 4. Lineaarse regressiooni testide tulemused

	Sõltuv tunnus: ROE		
	I mudel (VK)	II mudel (PVK)	III mudel (LVK)
Kitsenduste F-test, p-väärtus	0,824	0,905	0,879
Hausmani test, p-väärtus	0,342	0,740	0,479
Breusch-Pagani test, p-väärtus	0,257	0,145	0,192
Multikollineaarsus, maksimaalne VIF väärtus	2,989	2,761	1,881
Heteroskedastiivsus, p-väärtus	0,000	0,000	0,000
Ramsey RESET test, p-väärtus	0,000	0,000	0,000
Jääkliikmete vastavus normaaljaotusele, p-väärtus	0,000	0,000	0,000
Jääkliikmete autokorrelatsioon, p-väärtus	0,728	0,613	0,684
Ajaefekti olulisus (Gretl, <i>add variable</i>), p-väärtus	0,670	0,660	0,659
Pesarani CD test, p-väärtus	0,961	0,937	0,868

	Sõltuv tunnus: ROA		
	IV mudel (VK)	V mudel (PVK)	VI mudel (LVK)
Kitsenduste F-test, p-väärtus	0,077	0,182	0,142
Hausmani test, p-väärtus	0,182	0,464	0,153
Breusch-Pagani test, p-väärtus	0,135	0,277	0,194
Multikollineaarsus, maksimaalne VIF väärtus	2,864	2,662	1,868
Heteroskedastiivsus, p-väärtus	0,000	0,000	0,000
Ramsey RESET test, p-väärtus	0,000	0,000	0,013
Jääkliikmete vastavus normaaljaotusele, p-väärtus	0,000	0,000	0,000
Jääkliikmete autokorrelatsioon, p-väärtus	0,711	0,825	0,727
Ajaefekti olulisus (Gretl, <i>add variable</i>), p-väärtus	0,307	0,280	0,298
Pesarani CD test, p-väärtus	0,178	0,162	0,102

Allikas: autori arvutused

Lisa 5. Otsustusmetsa ROE andmestiku treening- ja testandmestiku kirjeldav statistika

Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROE (%)	14%	12%	34%	-205%	223%
VANUS (aastat)	13,6	13,0	7,1	0,0	39,0
KAIVE (tuh eur)	2 764	720	6 134	50	82 013
YoY (%)	10%	6%	28%	-92%	119%
CR (ühikuta)	2,6	1,8	2,5	0,0	16,6
OKm (%)	55%	57%	23%	0%	99%
POV (%)	47%	48%	24%	0%	100%
VK (%)	21%	17%	18%	0%	85%
PVK (%)	13%	8%	15%	0%	65%
LVK (%)	8%	5%	8%	0%	41%

Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROE (%)	14%	12%	33%	-202%	224%
VANUS (aastat)	13,7	13,0	7,1	1,0	37,0
KAIVE (tuh eur)	2 677	780	5 629	50	68 563
YoY (%)	10%	6%	27%	-96%	120%
CR (ühikuta)	2,5	1,7	2,3	0,0	16,3
OKm (%)	54%	56%	23%	1%	100%
POV (%)	48%	48%	24%	0%	99%
VK (%)	22%	18%	18%	0%	85%
PVK (%)	14%	9%	15%	0%	65%
LVK (%)	8%	5%	8%	0%	41%

Allikas: autori koostatud

Märkus: Treeningandmestiku statistika on üleval ja testandmestikul all

Lisa 6. Otsustusmetsa ROA andmestiku treening- ja testandmestiku kirjeldav statistika

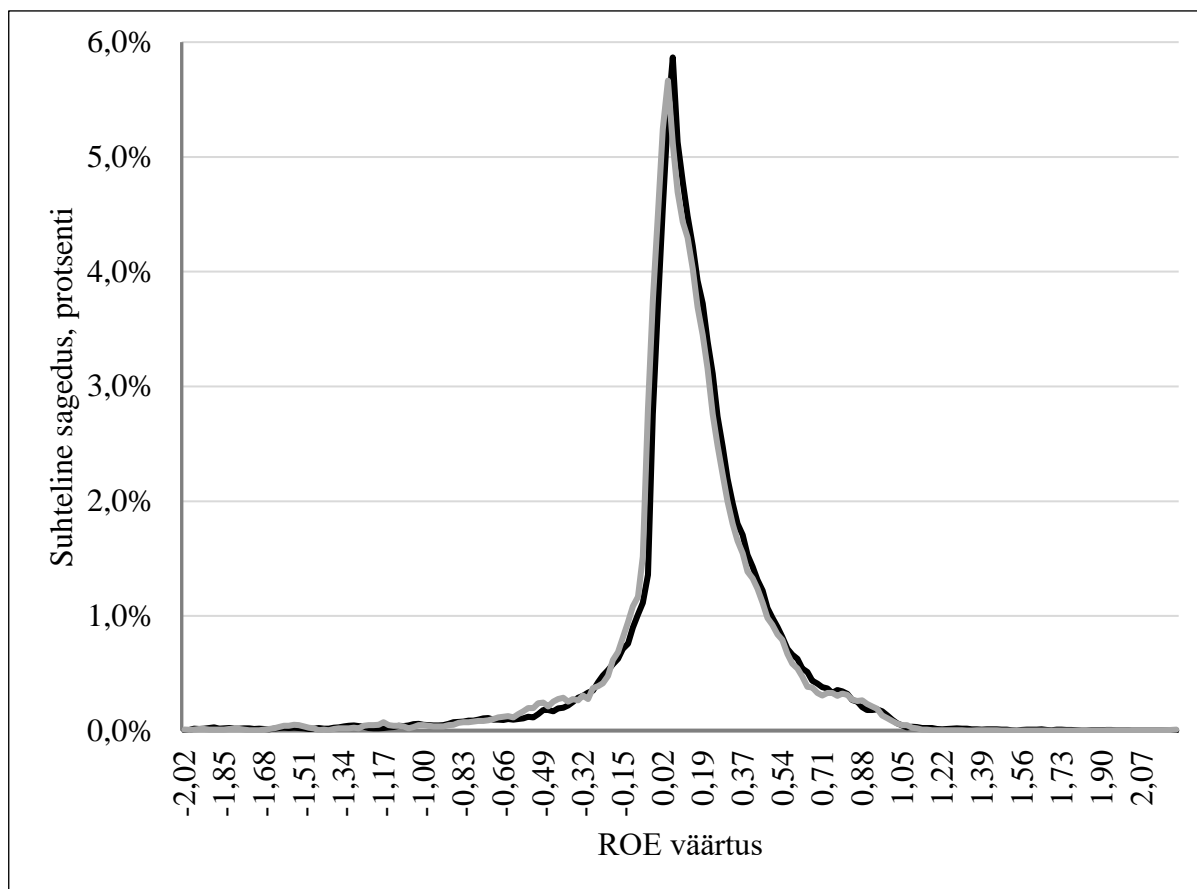
Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROA (%)	8%	6%	13%	-44%	55%
VANUS (aastat)	13,6	13,0	7,1	0,0	39,0
KAIIVE (tuh eur)	2 713	744	5 948	50	82 013
YoY (%)	10%	6%	27%	-92%	118%
CR (ühikuta)	2,5	1,7	2,4	0,0	16,6
OKm (%)	53%	56%	25%	-74%	100%
POV (%)	48%	48%	24%	0%	100%
VK (%)	21%	17%	18%	0%	85%
PVK (%)	13%	8%	15%	0%	65%
LVK (%)	8%	5%	9%	0%	41%

Tunnus	Keskmine	Mediaan	St. hälve	Min	Max
ROA (%)	7%	6%	13%	-44%	54%
VANUS (aastat)	13,6	13,0	7,1	0,0	33,0
KAIIVE (tuh eur)	2 833	764	6 037	50	79 282
YoY (%)	9%	6%	27%	-96%	120%
CR (ühikuta)	2,5	1,7	2,4	0,0	16,3
OKm (%)	53%	54%	25%	-74%	99%
POV (%)	48%	48%	24%	0%	100%
VK (%)	22%	18%	18%	0%	84%
PVK (%)	14%	8%	15%	0%	65%
LVK (%)	8%	6%	8%	0%	41%

Allikas: autori arvutused

Märkus: Treeningandmestiku statistika on üleval ja testandmestikul all

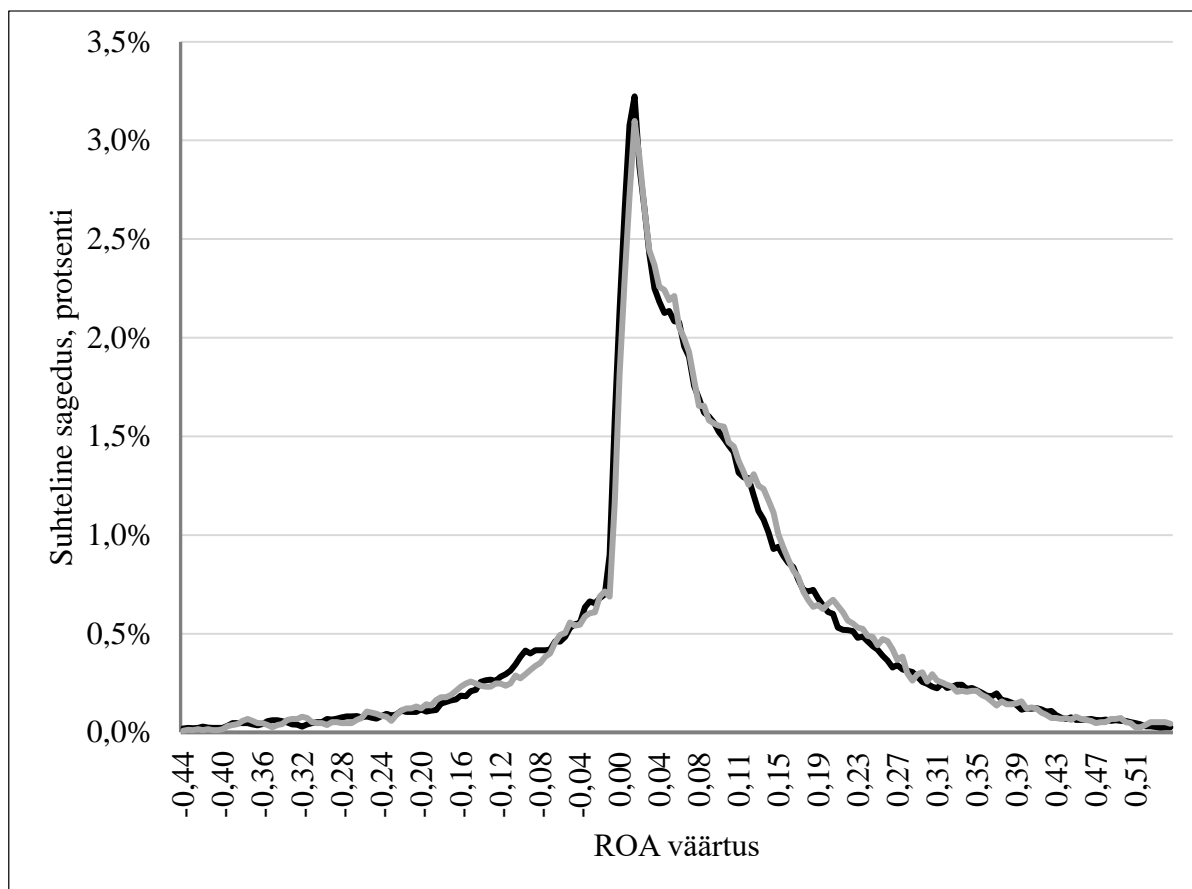
Lisa 7. Otsustusmetsa ROE treening- ja testandmete ROE suhtelise sageduse graafik



Allikas: autori koostatud

Märkus: Must funktsioon on treeningandmete ja hall funktsioon testandmete jaotus

Lisa 8. Otsustusmetsa ROA treening- ja testandmete ROA suhtelise sageduse graafik



Allikas: autori koostatud

Märkus: Must funktsioon on treeningandmete ja hall funktsioon testandmete jaotus

Lisa 9. R skript otsustusmetsa I mudeli näitel

```
#####  
# Vajalike teekide laadimine  
#####  
library(randomForest) # Otsustusmetsa teek  
library(caret) # Teek andmestiku jaotamiseks ja mtry optimeerimiseks  
ristvalideerimisega  
library(Metrics) # MSE, RMSE ja MAE arvutamise teek  
library(rfUtilities) # Mudeli ülesobitumise testi ja p-väärtuse testi teek  
library(rfPermute) # Marginaalse permutatsiooniolulisuse ja p-väärtuste teek  
library(permp) # Tingimusliku permutatsiooniolulisuse teek  
library(vi) # Teek, mis võimaldab arvutada variatsiooniolulisuse  
library(pdp) # Osalise sõltuvuse funktsiooni arvutamise teek  
library(xlsx) # Teek, mis võimaldab salvestada andmed Excelisse  
  
library(doParallel) # Caret train funktsiooniga paralleelse  
programmeerimise teek  
cl <- makePSOCKcluster(4)  
registerDoParallel(cl)  
# Lõpetades saab kasutada käsklust stopCluster(cl)  
  
#####  
# ROE andmestiku jaotamine treening- ja testandmestikuks suhtes 70:30 Caret teegiga. Eelnevalt  
# on ROE andmestik RStudio'sse laetud kasutades valikut "Import Dataset".  
#####  
  
set.seed(123) # siin ja järgnevalt tulemuste reprodutseerimiseks  
index <- createDataPartition(ROE_data$ROE, p = 0.7,  
                             list = FALSE)  
trainROE <- ROE_data[index, ]  
testROE <- ROE_data[-index, ]  
  
#####  
# Mtry väärtuste analüüsimine Caret funktsiooniga  
# train kasutades ristvalideerimist  
#####  
  
control <- trainControl(method='repeatedcv',  
                        number=10,  
                        repeats=3)  
set.seed(123)  
caret_cf_ROE_VK <- train(ROE ~ VANUS+KAIIVE+YoY+CR+OKm+POV+VK,  
                        data=trainROE,  
                        method='rf',  
                        metric='RMSE',  
                        tuneGrid = expand.grid(.mtry = c(1:7)),  
                        trControl=control)  
  
#####
```



```

# Optimiseeritud mtry väärtusega otsustusmetsa mudeli loomine randomForest teegiga.
# Otsustusmetsa puude arv 1000, kuna suurem puude arv
# võimaldab leida stabiilsemat permutatsiooniolulisust. "importance = TRUE,
# keep.forest = TRUE ja keep.inbag = TRUE" on vajalikud permimp teegiga
# tingimusliku permutatsiooniolulisuse arvutamiseks
#####

set.seed(123)
rf_ROE_VK <- randomForest(ROE ~ VANUS+KAIVE+YoY+CR+OKm+POV+VK, data =
trainROE, mtry = 1, ntrees = 1000,
                        importance = TRUE, keep.forest = TRUE, keep.inbag = TRUE)

# Puude arvu piisavuse kontroll. Kuvab MSE vastavalt puude arvule.
plot(rf_ROE_VK, main = "Puude arvu piisavuse kontroll")

#####
# Mudeli OOB treeningtäpsuse ja testandmetel ennustustäpsuse leidmine. MSE, RMSE ja MAE
# arvutamine Metrics teegiga.
#####

# Funktsioon determinatsioonikordaja leidmiseks
RSquared <- function(actual, predicted) {
  rss <- sum((predicted - actual) ^ 2)
  tss <- sum((actual - mean(actual)) ^ 2)
  rsq <- 1 - rss/tss
  return(rsq)
}

#OOB vead
mse(trainROE$ROE, rf_ROE_VK$predicted)
rmse(trainROE$ROE, rf_ROE_VK$predicted)
mae(trainROE$ROE, rf_ROE_VK$predicted)
RSquared(trainROE$ROE, rf_ROE_VK$predicted)

# Testandmetel ennustamine
predicted_ROE_VK <- predict(rf_ROE_VK, testROE)

# Test vead
mse(testROE$ROE, predicted_ROE_VK)
rmse(testROE$ROE, predicted_ROE_VK)
mae(testROE$ROE, predicted_ROE_VK)
RSquared(testROE$ROE, predicted_ROE_VK)

#####
# Mudeli stabiilsuse kontroll
#####

set.seed(876) # Teine etteantud juhuslikkus
stability_ROE_VK <- randomForest(ROE ~ VANUS+KAIVE+YoY+CR+OKm+POV+VK,

```

```

data = trainROE, mtry = 1, ntrees = 1000)

# Kontrollmudeli OOB vead
mse(trainROE$ROE, stability_ROE_VK$predicted)
rmse(trainROE$ROE, stability_ROE_VK$predicted)
mae(trainROE$ROE, stability_ROE_VK$predicted)
RSquared(trainROE$ROE, stability_ROE_VK$predicted)

# Kontrollmudeli ennustus testandmetel
stability_predicted_ROE_VK <- predict(stability_ROE_VK, testROE)

# Kontrollmudeli test vead
mse(testROE$ROE, stability_predicted_ROE_VK)
rmse(testROE$ROE, stability_predicted_ROE_VK)
mae(testROE$ROE, stability_predicted_ROE_VK)
RSquared(testROE$ROE, stability_predicted_ROE_VK)

#####
# Mudeli ülesobitumise test rfUtilities teegiga
#####

rf.regression.fit(rf_ROE_VK)

#####
# Mudeli p-väärtuse arvutamine rfUtilities teegiga
#####

set.seed(123)
ROEVK_sign <- rf.significance(rf_ROE_VK, trainROE, p = 0.05, nperm = 99)

#####
## Tunnuste marginaalse permutatsiooniolulisuse ja p-väärtuse arvutamine
#####

set.seed(123)
pvalues_ROE_VK <- rfPermute(ROE ~ VANUS+KAIVE+YoY+CR+OKm+POV+VK,
trainROE, mtry = 1, ntrees = 1000, nrep = 100, num.cores = 4)
rp.importance(pvalues_ROE_VK)

#####
# Tingimusliku permutatsiooniolulisuse leidmine
#####

set.seed(123)
condperm_ROE_VK <- permimp(rf_ROE_VK, data = trainROE, nperm =2, conditional =
TRUE, asParty = TRUE)

#####
# Tunnuste variatsiooniolulisuse leidmine osalise sõltuvuse
# funktsiooni põhjal. Kuna funktsiooni tulemus ei sõltu juhuslikkusest,

```

```

# pole vajalik kasutada etteantud juhuslikkust.
#####

vipd_ROAVK <- vi(rf_ROA_VK, method = "firm")

#####
# Osalise sõltuvuse funktsioonid pdp teegiga
#####

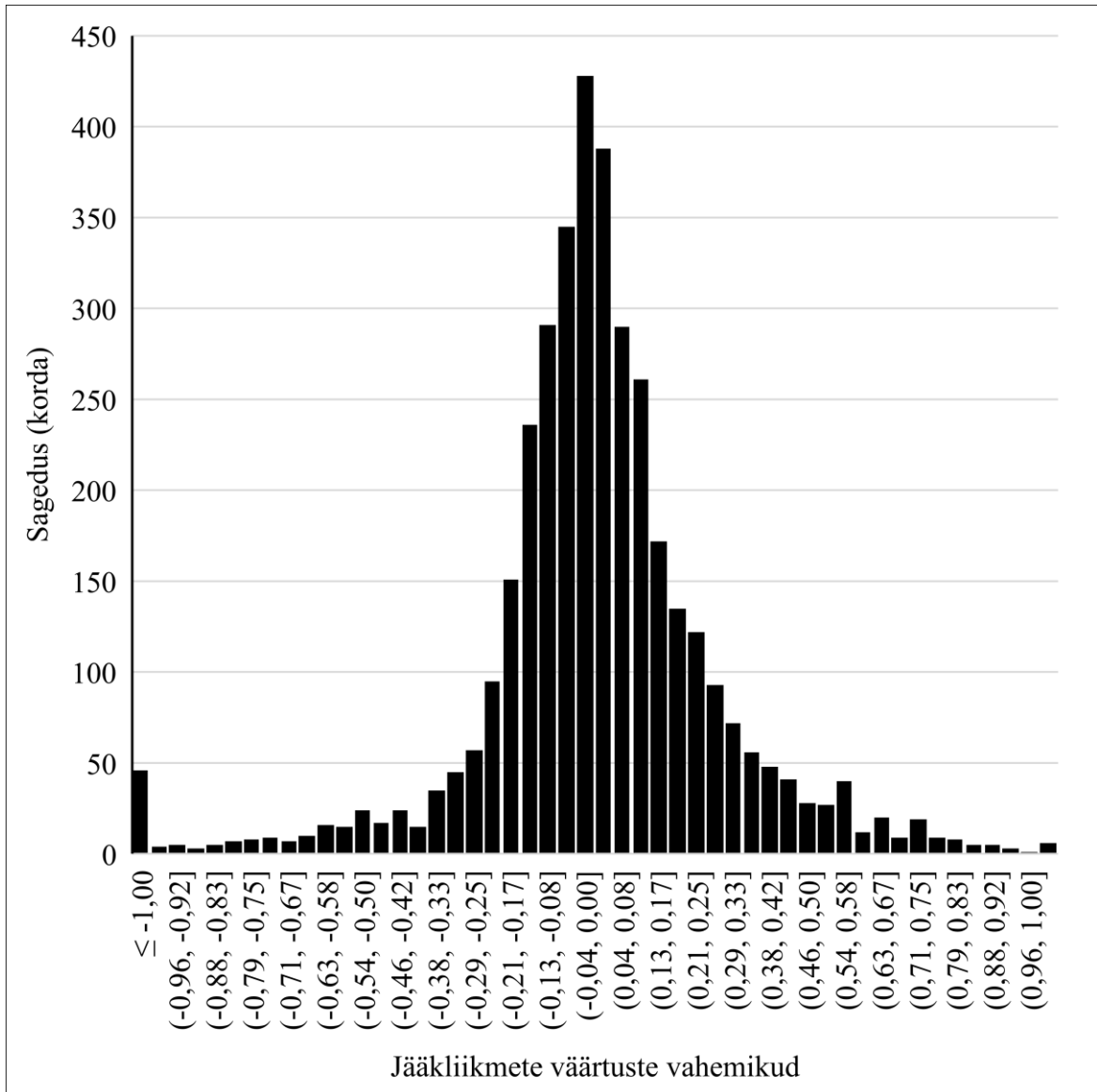
pd_ROEVK_VANUS <- partial(rf_ROE_VK, "VANUS")
pd_ROEVK_KAIVE <- partial(rf_ROE_VK, "KAIVE")
pd_ROEVK_YoY <- partial(rf_ROE_VK, "YoY")
pd_ROEVK_CR <- partial(rf_ROE_VK, "CR")
pd_ROEVK_OKm <- partial(rf_ROE_VK, "OKm")
pd_ROEVK_POV <- partial(rf_ROE_VK, "POV")
pd_ROEVK_VK <- partial(rf_ROE_VK, "VK")

# Kvantiilide arvutamine
df_VK <- data.frame(quantile(trainROE$VK, seq(0.1, 0.9, by = 0.1)))

# Autor kasutab läbivalt andmete Excelisse salvestamiseks xlsx teeki. Näide:
write.xlsx(df_VK, file = " df_VK.xlsx", col.names = TRUE, row.names = TRUE)

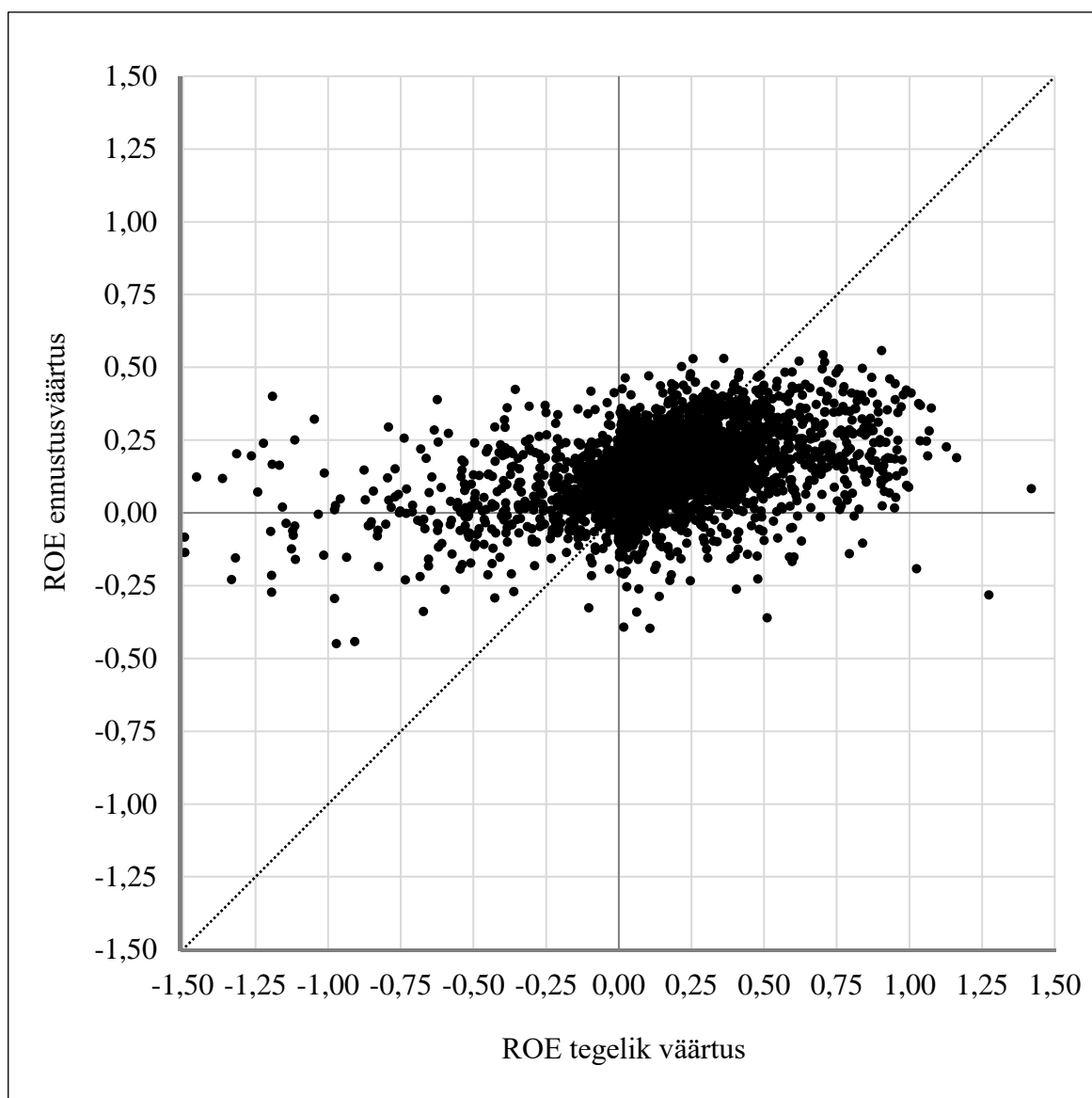
```

Lisa 10. Otsustusmetsa I mudeli jääkliikmete histogramm ja hajususgraafik



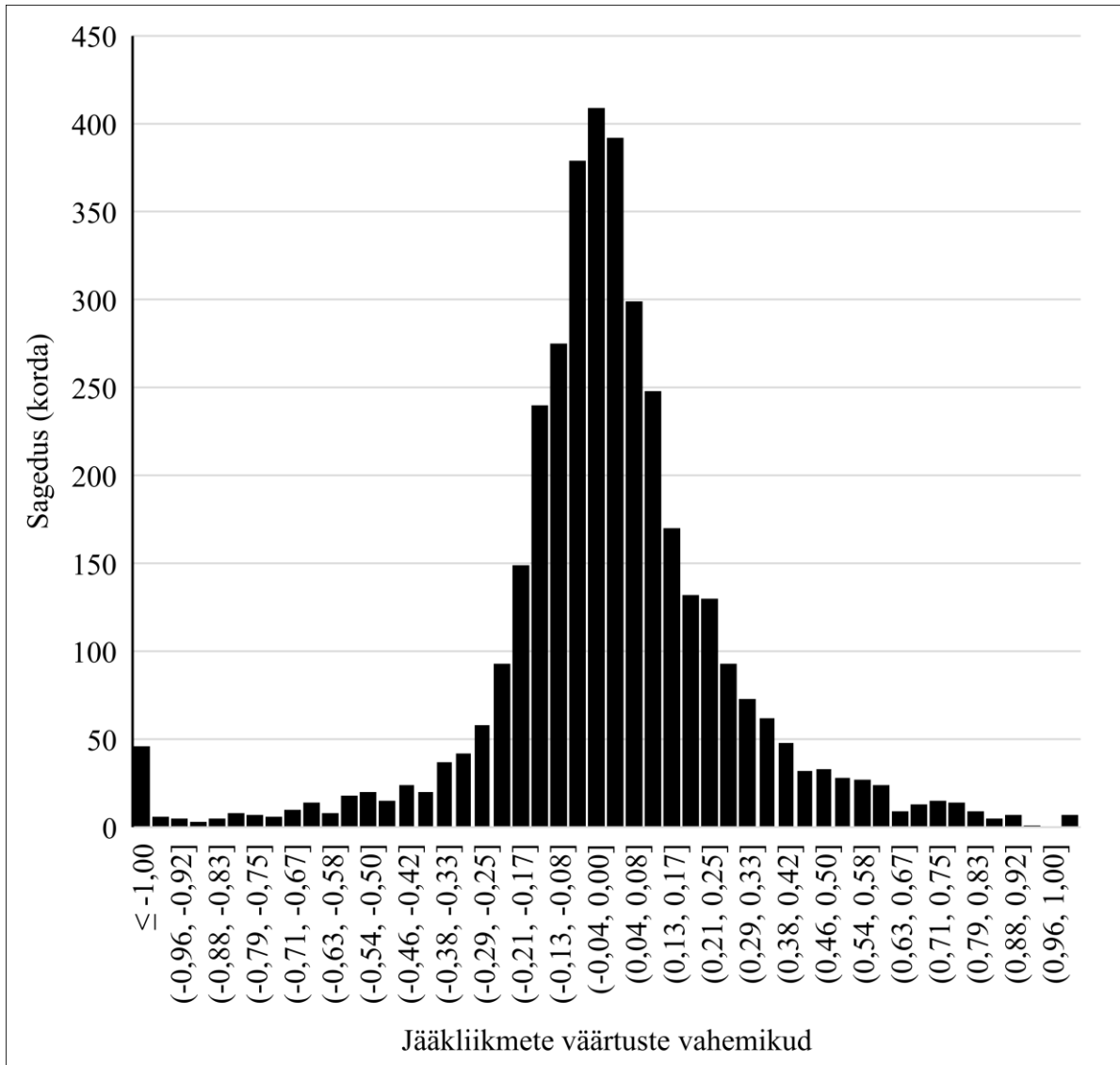
Allikas: autori koostatud

Lisa 10 järg



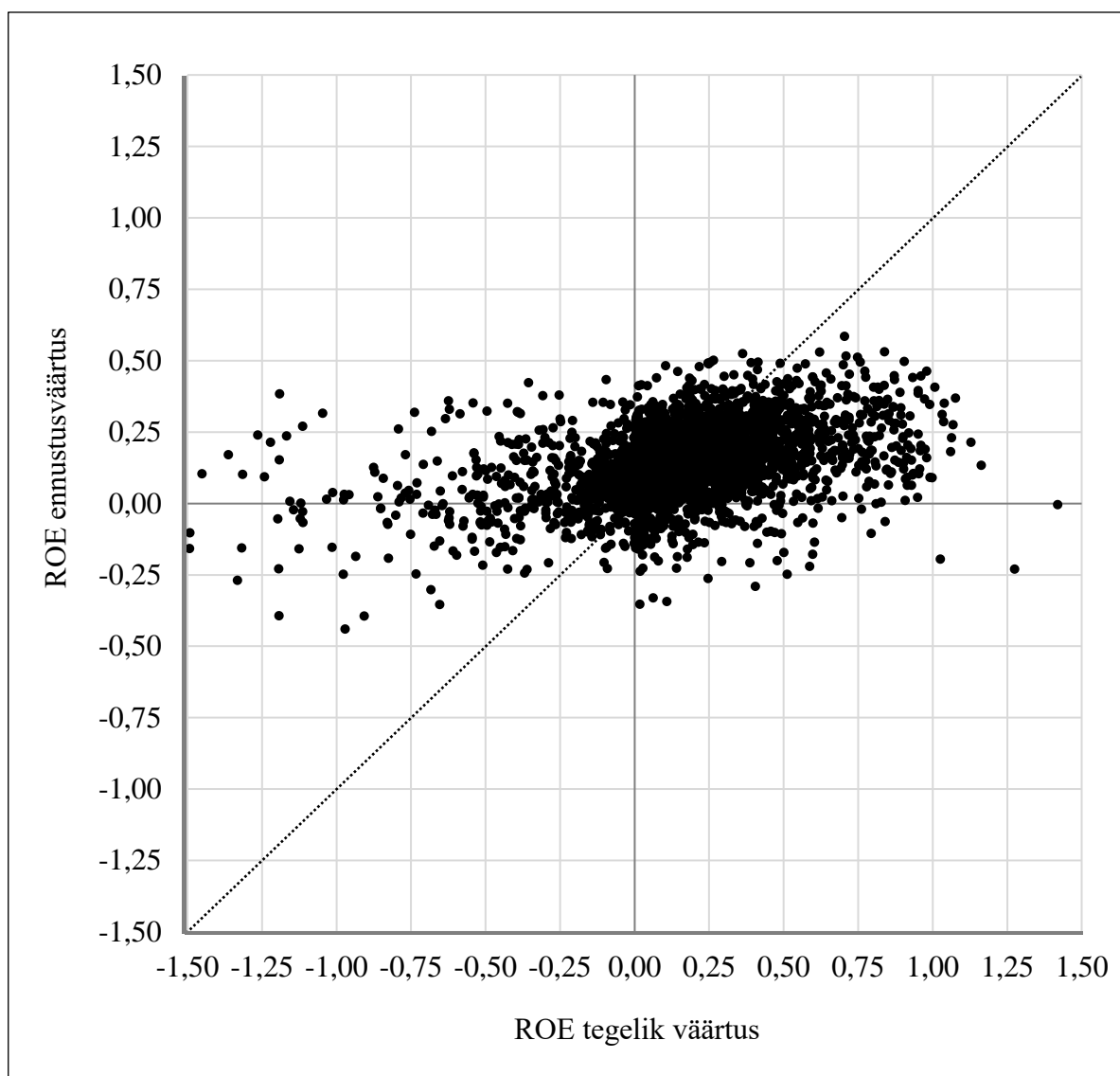
Allikas: autori koostatud

Lisa 11. Otsustusmetsa II mudeli jääkliikmete histogramm ja hajususgraafik



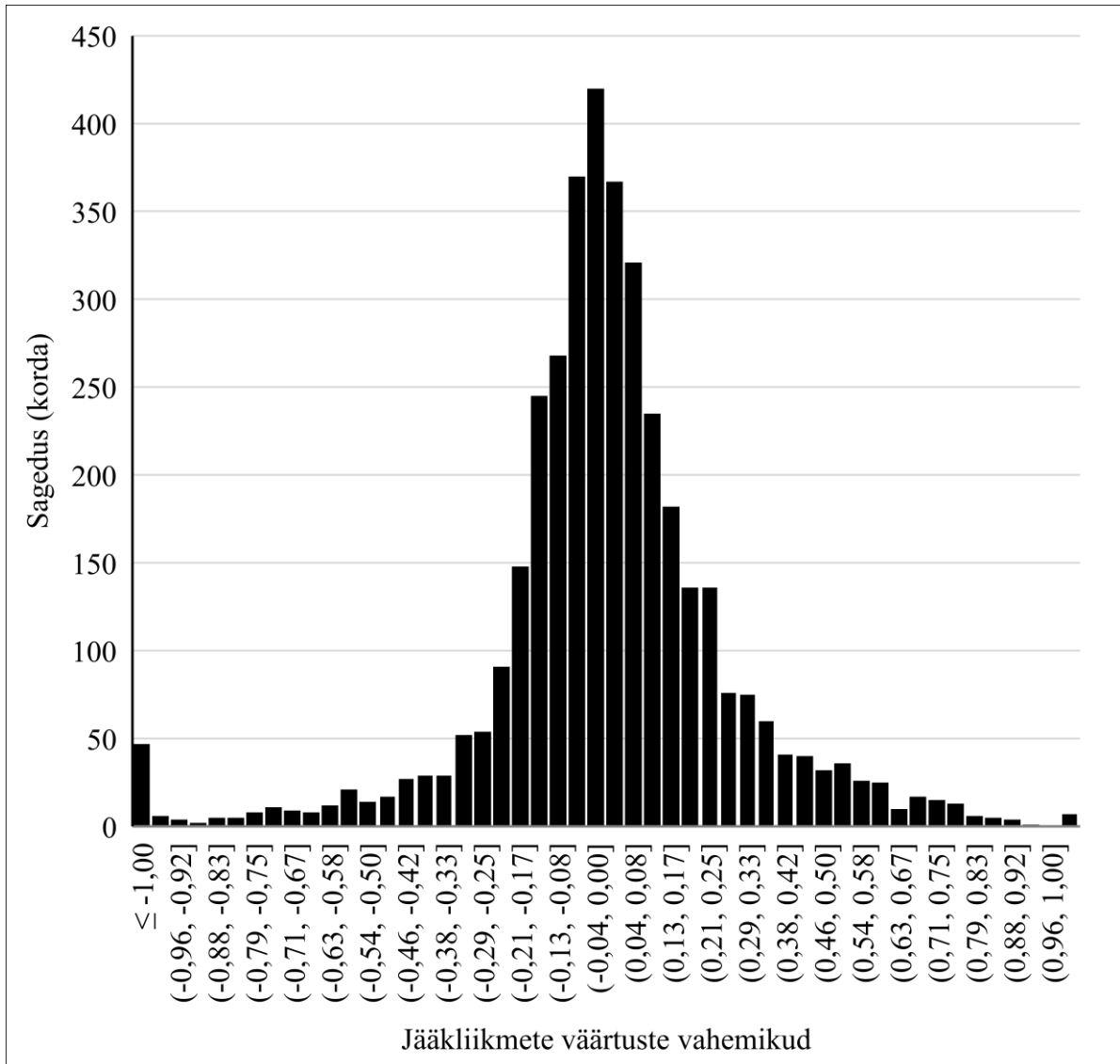
Allikas: autori koostatud

Lisa 11 järg



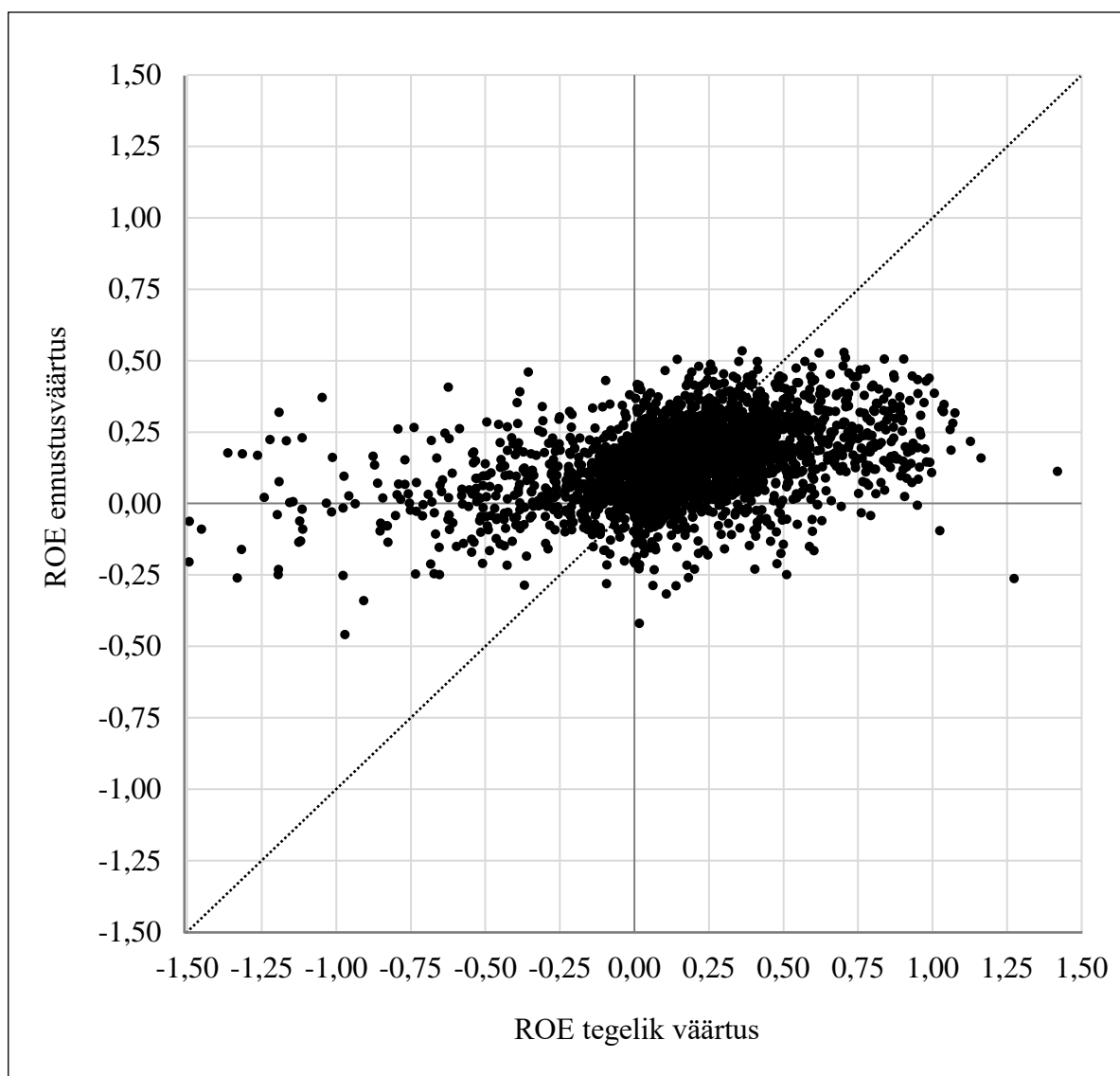
Allikas: autori koostatud

Lisa 12. Otsustusmetsa III mudeli jääkliikmete histogramm ja hajususgraafik



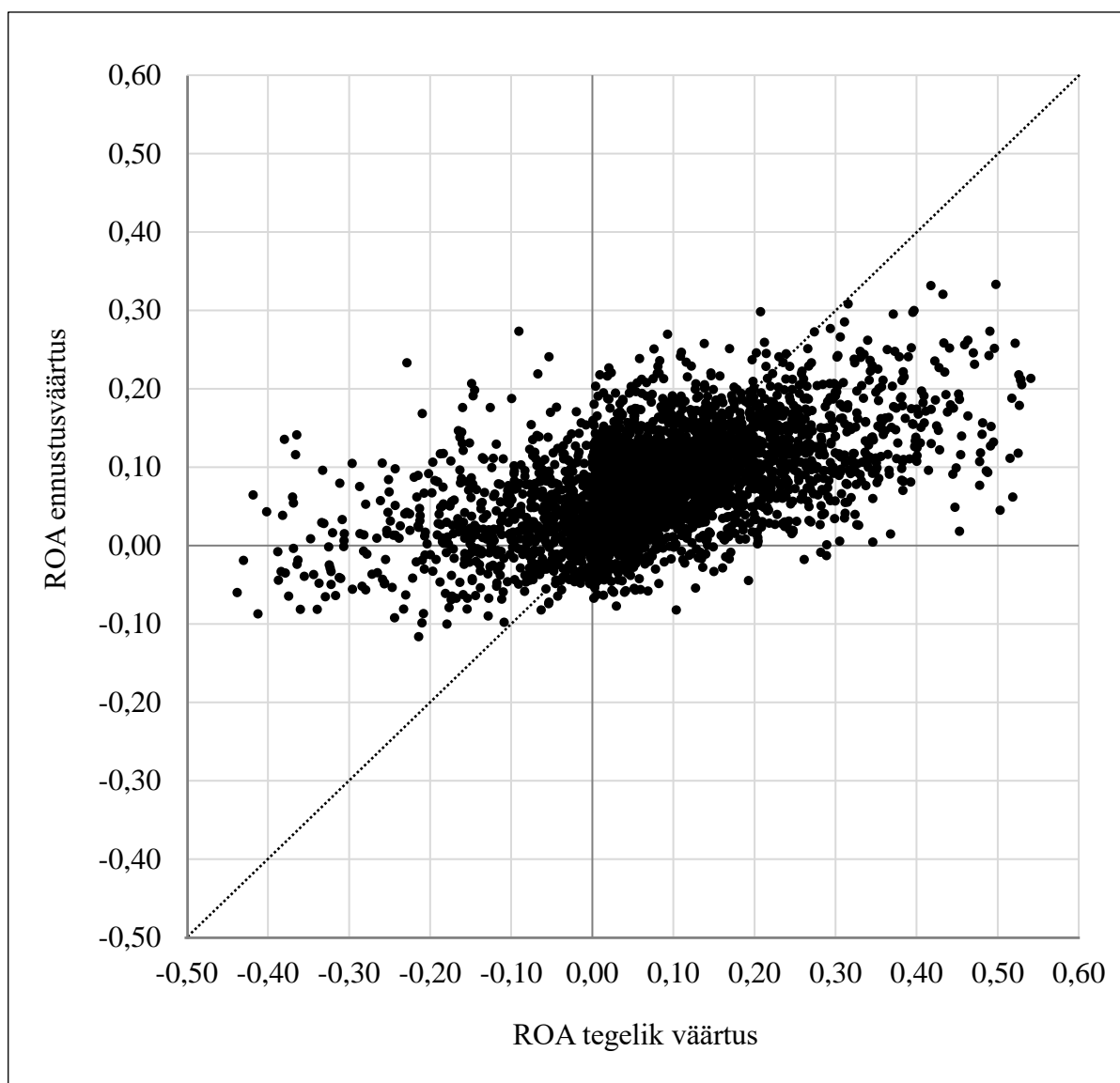
Allikas: autori koostatud

Lisa 12 järg



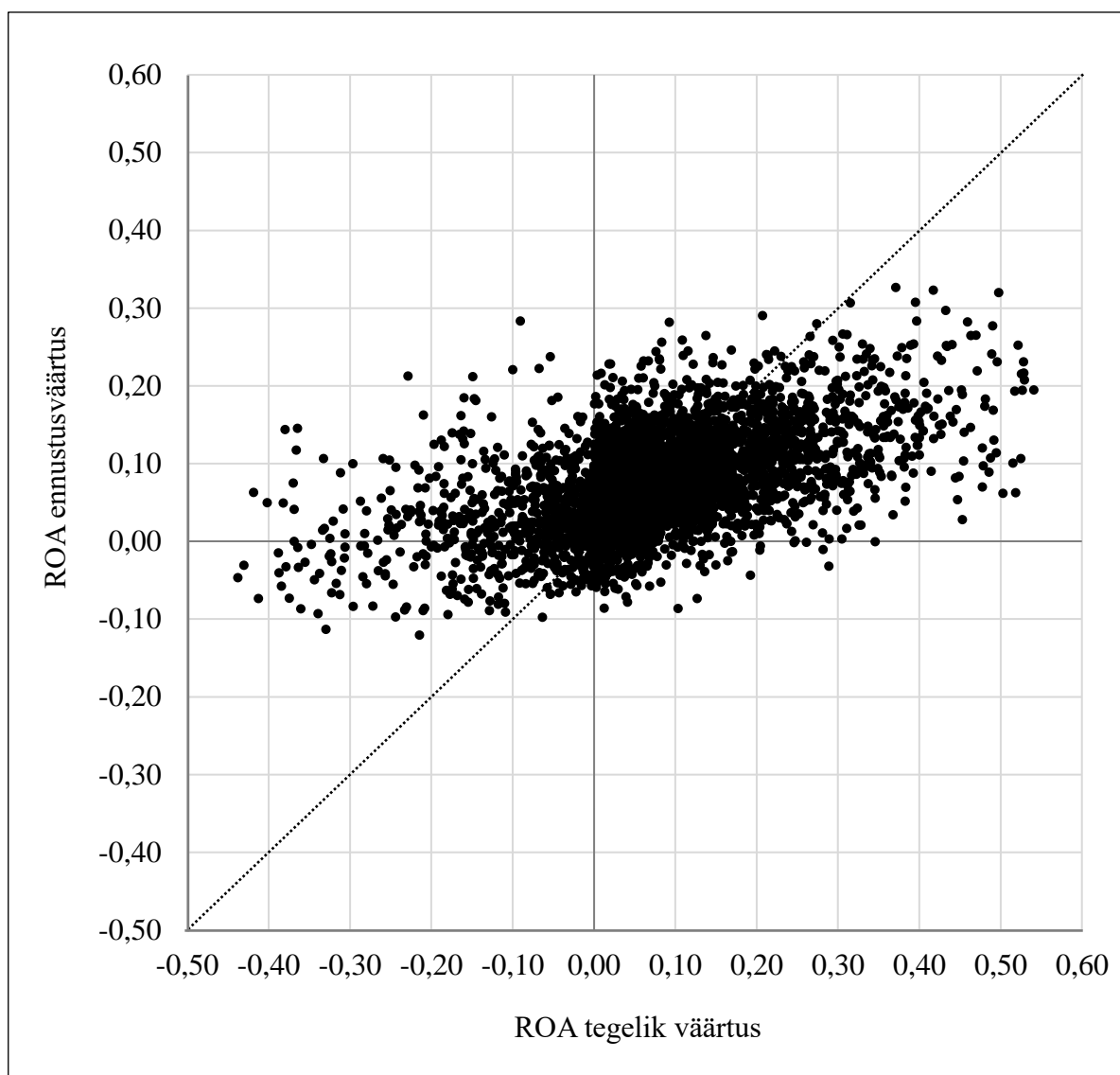
Allikas: autori koostatud

Lisa 13 järg



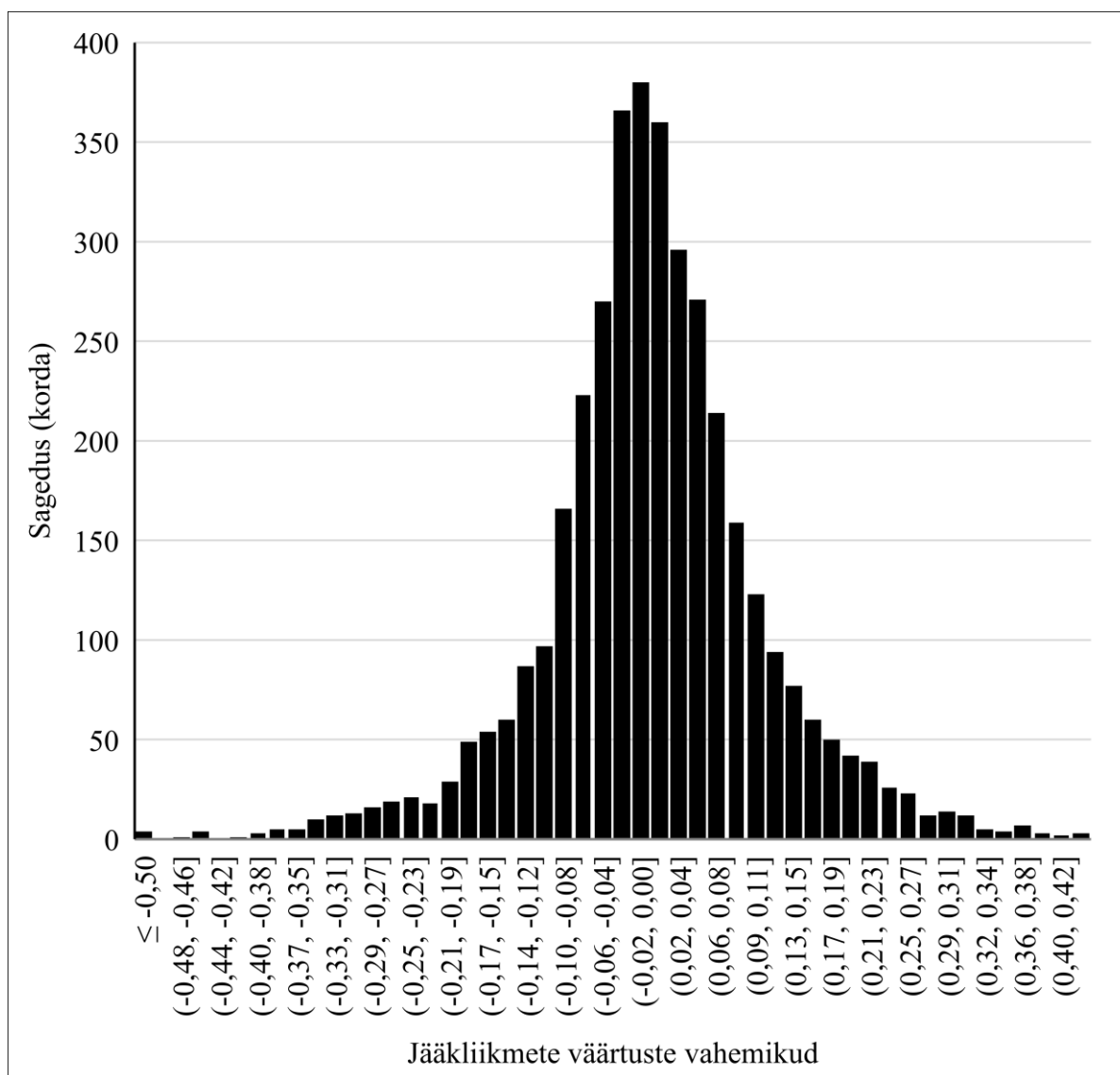
Allikas: autori koostatud

Lisa 14 järg



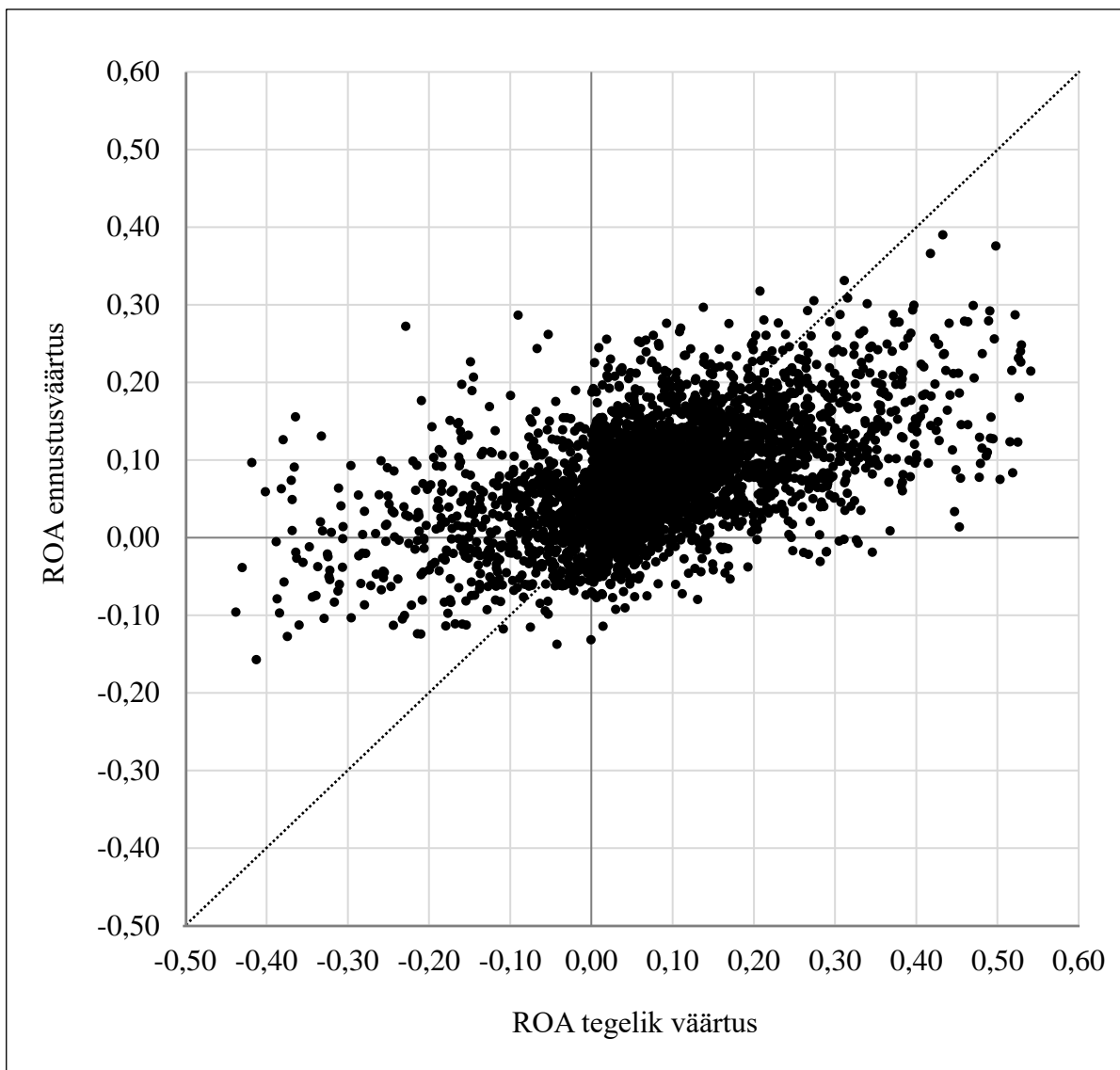
Allikas: autori koostatud

Lisa 15. Otsustusmetsa VI mudeli jääkliikmete histogramm ja hajususgraafik



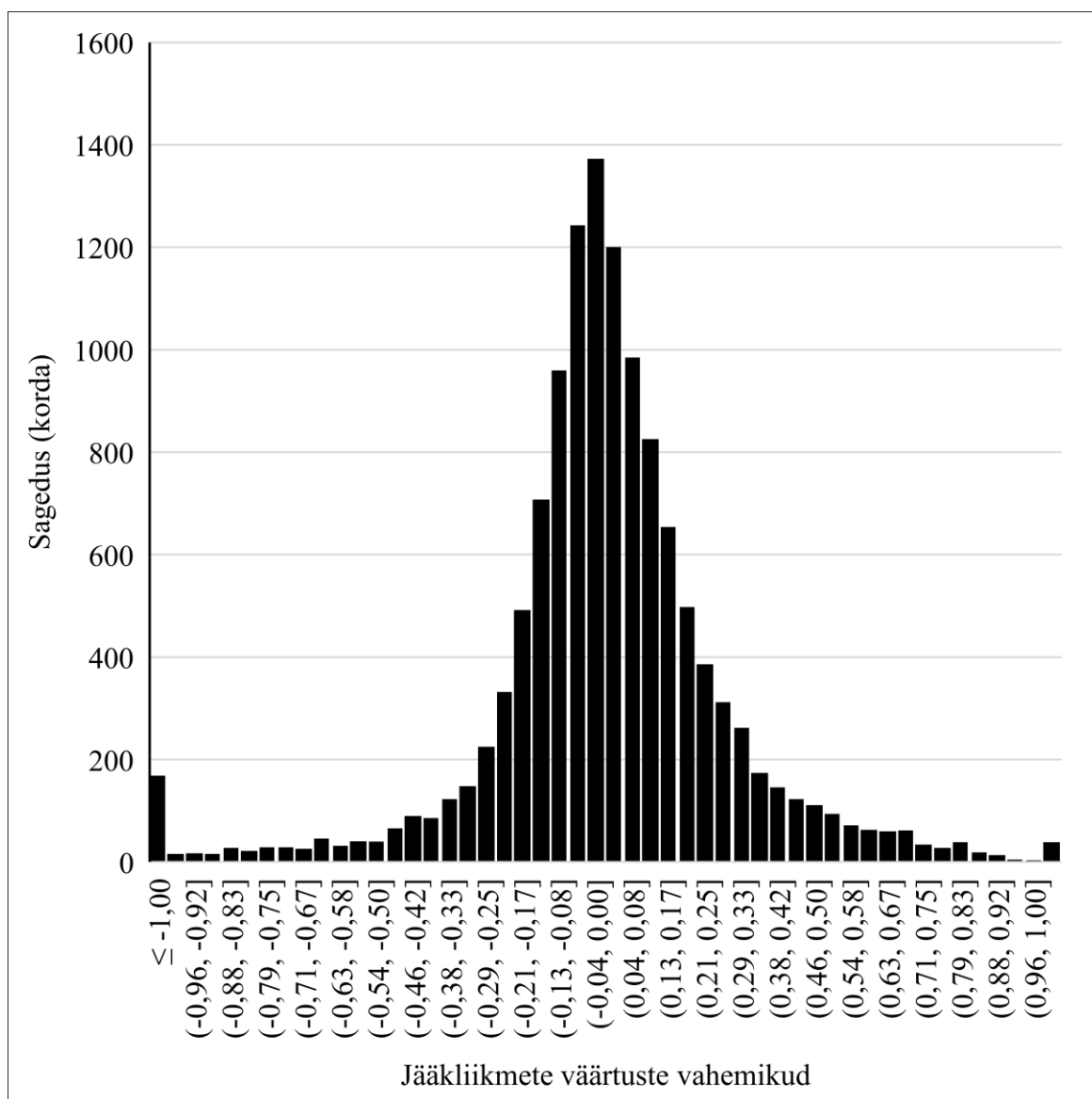
Allikas: autori koostatud

Lisa 15 järg



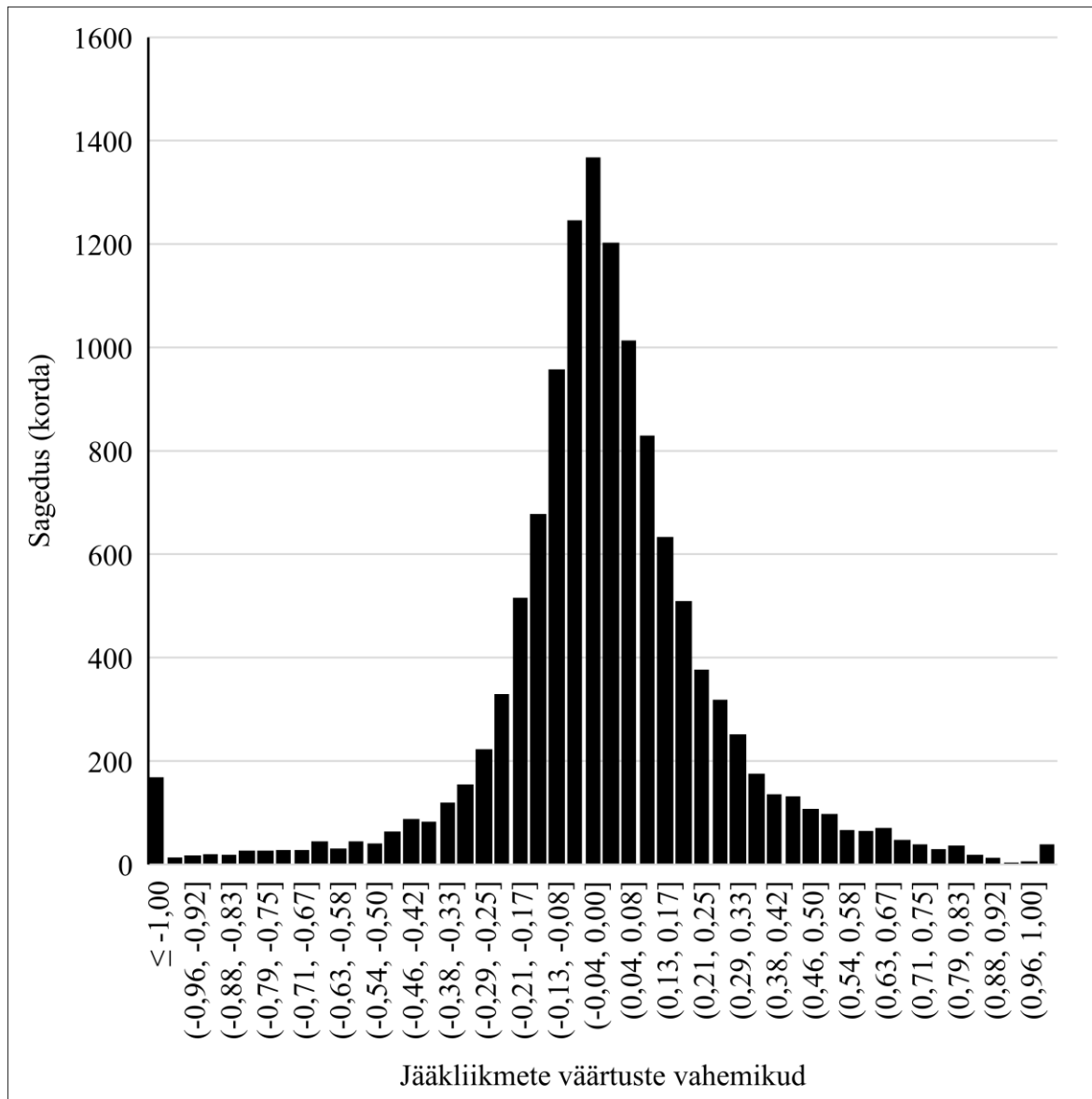
Allikas: autori koostatud

Lisa 16. I regressioonimudeli jääkliikmete histogramm



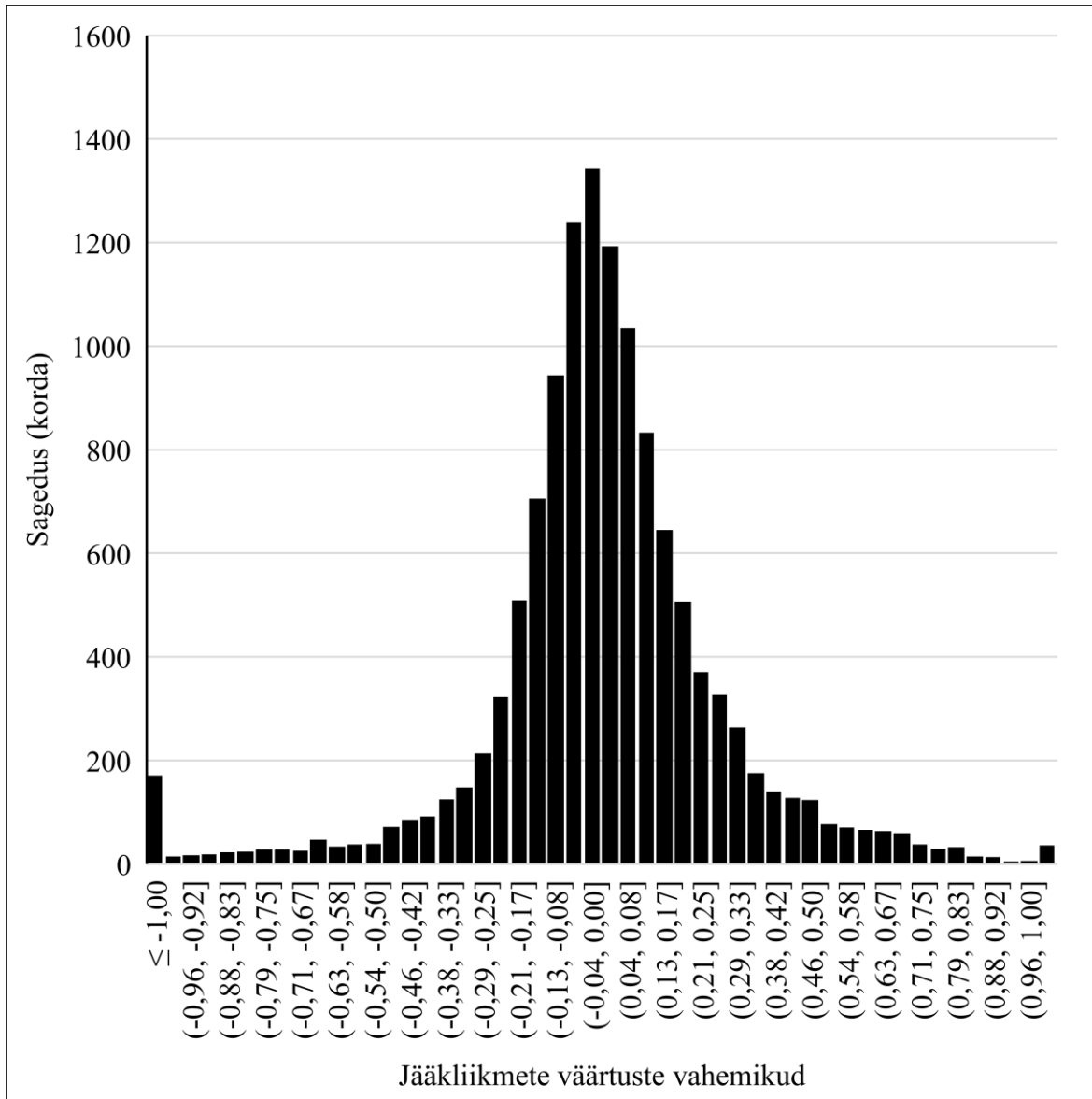
Allikas: Gretl, autori koostatud

Lisa 17. II regressioonimudeli jääkliikmete histogramm



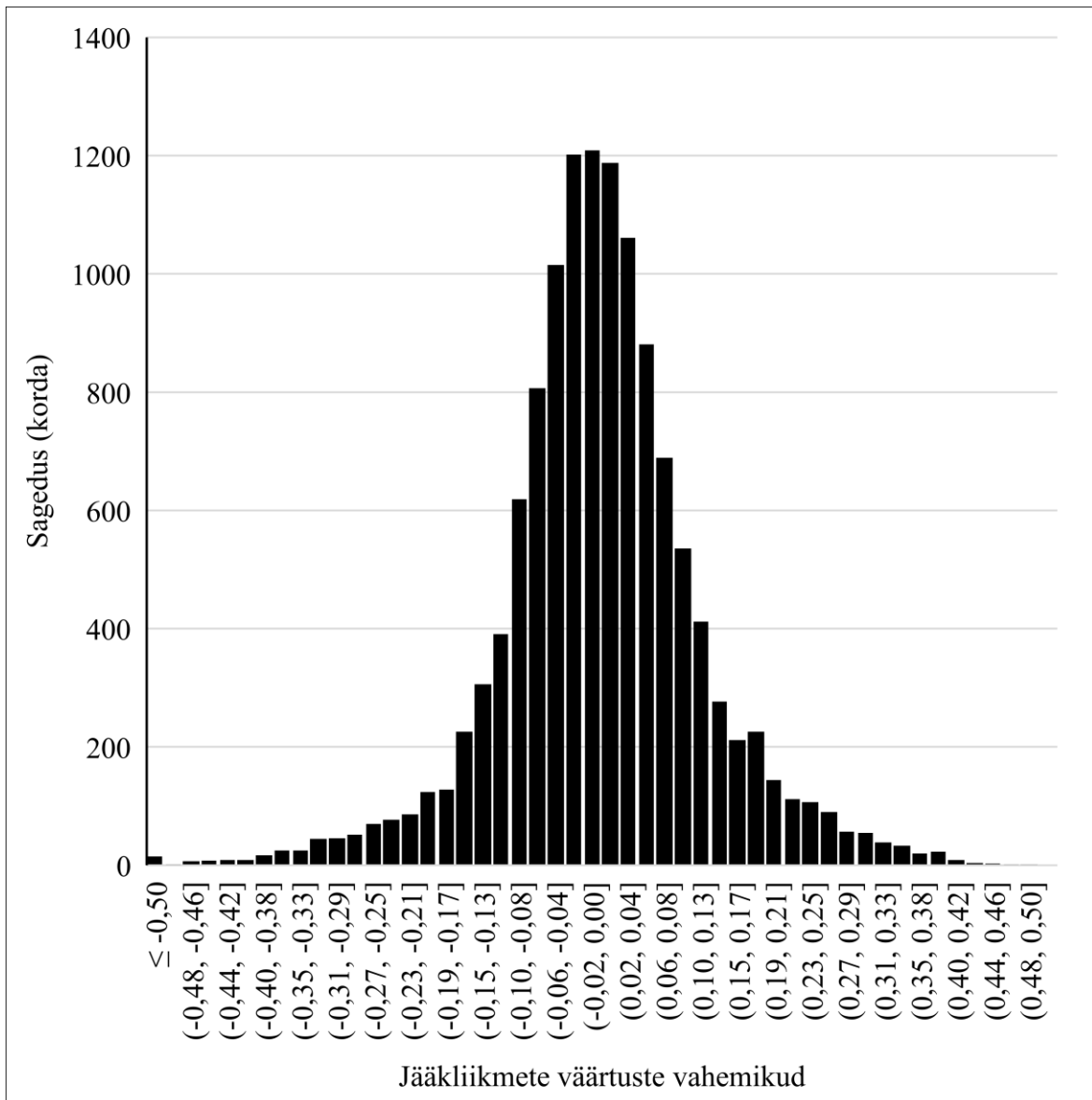
Allikas: Gretl, autori koostatud

Lisa 18. III regressioonimudeli jääkliikmete histogramm



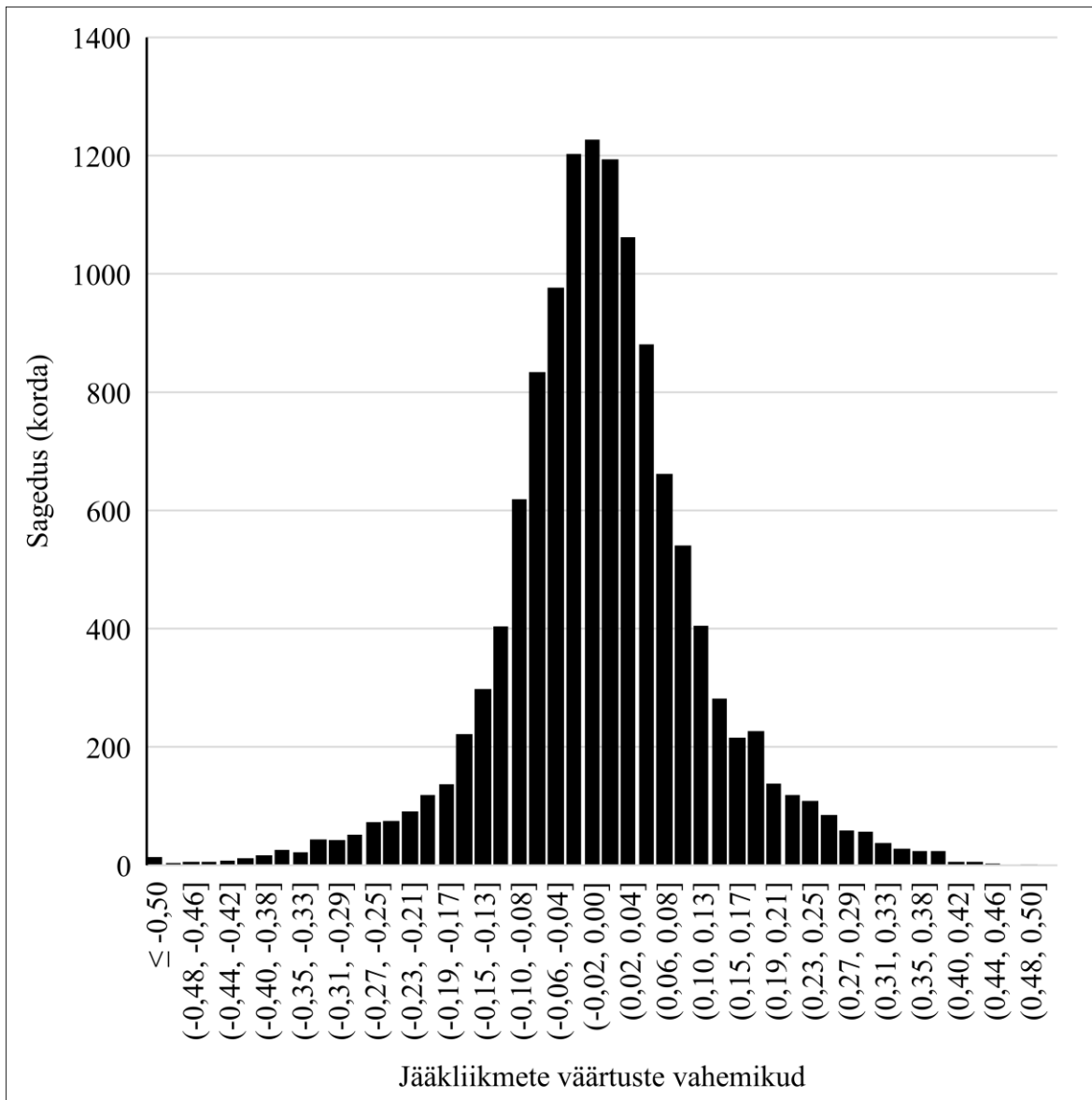
Allikas: Gretl, autori koostatud

Lisa 19. IV regressioonimudeli jääkliikmete histogramm



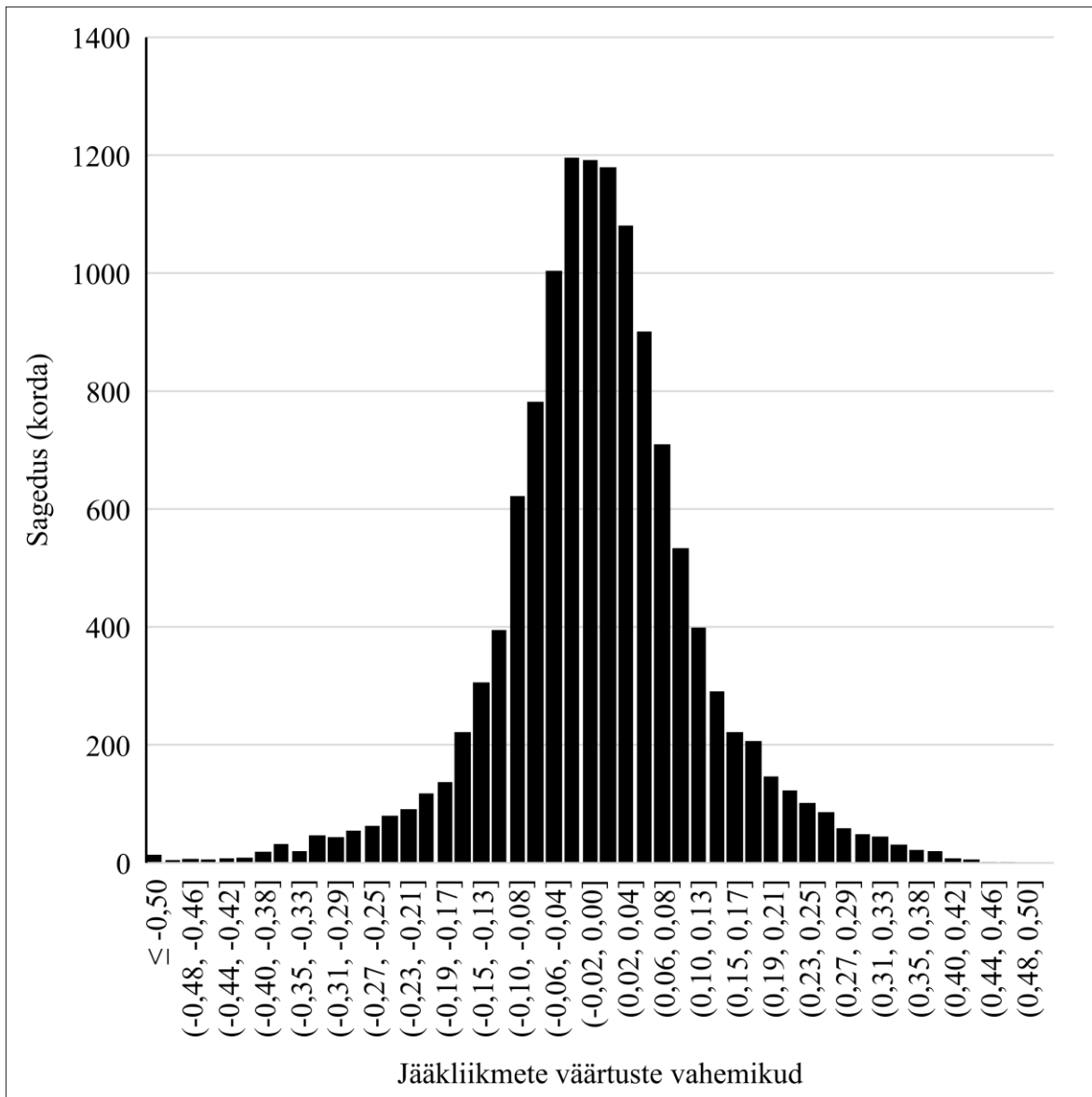
Allikas: Gretl, autori koostatud

Lisa 20. V regressioonimudeli jääkliikmete histogramm



Allikas: Gretl, autori koostatud

Lisa 21. VI regressioonimudeli jääkliikmete histogramm



Allikas: Gretl, autori koostatud

Lisa 22. Hüperparameeter mtry ristvalideerimisega optimeerimise tulemused

RMSE väärtused mtry optimeerimisel 10-kordse ristvalideerimise abil (3 kordust)

Mtry	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
1	0,3132	0,3131	0,3123	0,1146	0,1146	0,1144
2	0,3148	0,3148	0,3138	0,1147	0,1148	0,1143
3	0,3162	0,3164	0,3149	0,1152	0,1152	0,1147
4	0,3171	0,3175	0,3160	0,1154	0,1156	0,1151
5	0,3181	0,3184	0,3169	0,1157	0,1159	0,1154
6	0,3188	0,3192	0,3177	0,1159	0,1161	0,1155
7	0,3193	0,3199	0,3182	0,1161	0,1162	0,1157

Allikas: autori koostatud

Märkus: optimaalsed mtry väärtused on toodud rasvases kirjas

MAE väärtused mtry optimeerimisel 10-kordse ristvalideerimise abil (3 kordust)

Mtry	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
1	0,1934	0,1932	0,1928	0,0819	0,0820	0,0818
2	0,1956	0,1954	0,1951	0,0823	0,0824	0,0821
3	0,1968	0,1971	0,1964	0,0828	0,0829	0,0825
4	0,1978	0,1979	0,1973	0,0831	0,0832	0,0828
5	0,1983	0,1987	0,1980	0,0834	0,0835	0,0831
6	0,1989	0,1993	0,1985	0,0835	0,0836	0,0832
7	0,1992	0,1996	0,1990	0,0837	0,0838	0,0834

Allikas: autori koostatud

Märkus: optimaalsed mtry väärtused on toodud rasvases kirjas

R² väärtused mtry optimeerimisel 10-kordse ristvalideerimise abil (3 kordust)

Mtry	Sõltuv tunnus: ROE			Sõltuv tunnus: ROA		
	I (VK)	II (PVK)	III (LVK)	IV (VK)	V (PVK)	VI (LVK)
1	13,5%	13,6%	14,0%	27,7%	27,6%	27,9%
2	12,9%	12,9%	13,4%	27,3%	27,2%	27,7%
3	12,5%	12,5%	13,1%	26,8%	26,8%	27,3%
4	12,3%	12,1%	12,7%	26,7%	26,4%	27,0%
5	11,9%	11,8%	12,5%	26,3%	26,2%	26,7%
6	11,7%	11,6%	12,2%	26,2%	26,0%	26,6%
7	11,6%	11,4%	12,0%	26,0%	25,8%	26,4%

Allikas: autori koostatud

Märkus: optimaalsed mtry väärtused on toodud rasvases kirjas

Lisa 23. Otsustusmetsa kontroll- ja põhimudelite treening- ja testvea täpsus

I mudeli (ROE-VK) ennustustäpsus				II mudeli (ROE-PVK) ennustustäpsus			
Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus	Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus
OOB MSE	0,098	0,098	0,014%	OOB MSE	0,099	0,099	0,030%
OOB RMSE	0,313	0,314	0,023%	OOB RMSE	0,314	0,314	0,048%
OOB MAE	0,193	0,194	0,024%	OOB MAE	0,194	0,193	0,030%
OOB R ²	13,40%	13,27%	0,126%	OOB R ²	12,89%	13,16%	0,265%
TEST MSE	0,091	0,091	0,021%	TEST MSE	0,091	0,091	0,014%
TEST RMSE	0,301	0,301	0,034%	TEST RMSE	0,301	0,301	0,023%
TEST MAE	0,189	0,189	0,004%	TEST MAE	0,189	0,190	0,030%
TEST R ²	15,69%	15,88%	0,191%	TEST R ²	15,77%	15,90%	0,131%
III mudeli (ROE-LVK) ennustustäpsus				IV mudeli (ROA-VK) ennustustäpsus			
Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus	Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus
OOB MSE	0,098	0,098	0,053%	OOB MSE	0,013	0,013	0,003%
OOB RMSE	0,313	0,313	0,085%	OOB RMSE	0,115	0,115	0,013%
OOB MAE	0,193	0,193	0,033%	OOB MAE	0,082	0,082	0,005%
OOB R ²	13,44%	13,91%	0,467%	OOB R ²	26,94%	27,10%	0,168%
TEST MSE	0,090	0,090	0,026%	TEST MSE	0,013	0,013	0,001%
TEST RMSE	0,300	0,301	0,043%	TEST RMSE	0,112	0,112	0,006%
TEST MAE	0,189	0,189	0,004%	TEST MAE	0,080	0,080	0,004%
TEST R ²	16,30%	16,06%	0,240%	TEST R ²	29,32%	29,24%	0,077%
V mudeli (ROA-PVK) ennustustäpsus				VI mudeli (ROA-LVK) ennustustäpsus			
Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus	Mõõdik	Põhi-mudel	Kontroll-mudel	Erinevus
OOB MSE	0,013	0,013	0,001%	OOB MSE	0,013	0,013	0,006%
OOB RMSE	0,115	0,115	0,005%	OOB RMSE	0,115	0,114	0,024%
OOB MAE	0,082	0,082	0,004%	OOB MAE	0,082	0,082	0,058%
OOB R ²	27,07%	27,14%	0,068%	OOB R ²	27,31%	27,62%	0,308%
TEST MSE	0,013	0,013	0,002%	TEST MSE	0,012	0,013	0,014%
TEST RMSE	0,112	0,112	0,007%	TEST RMSE	0,111	0,112	0,061%
TEST MAE	0,080	0,080	0,003%	TEST MAE	0,080	0,080	0,012%
TEST R ²	29,05%	28,96%	0,093%	TEST R ²	29,83%	29,07%	0,764%

Allikas: autori koostatud

Märkus: etteantud juhuslikkus on põhimudelil 123, kontrollmudelil 876

Lisa 24. Otsustusmetsa I mudeli (ROE-VK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	50,5%
2.	OKm	32,7%
3.	VK	32,0%
4.	VANUS	28,8%
5.	CR	27,3%
6.	POV	25,2%
7.	KAIVE	18,4%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	3,50%
2.	VANUS	3,33%
3.	POV	3,26%
4.	CR	1,76%
5.	OKm	1,41%
6.	VK	1,40%
7.	KAIVE	1,10%

Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

Jrk	Tunnus	Ennustatud ROE standardhälve
1.	YoY	13,48%
2.	VANUS	7,41%
3.	VK	5,26%
4.	OKm	3,96%
5.	CR	2,48%
6.	POV	1,57%
7.	KAIVE	0,65%

Allikas: autori koostatud

Lisa 25. Otsustusmetsa II mudeli (ROE-PVK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	51,2%
2.	OKm	31,2%
3.	VANUS	30,7%
4.	PVK	28,7%
5.	POV	26,4%
6.	CR	26,0%
7.	KAIVE	15,4%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	4,14%
2.	POV	3,47%
3.	VANUS	2,60%
4.	CR	2,48%
5.	PVK	2,04%
6.	OKm	1,15%
7.	KAIVE	0,90%

Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

Jrk	Tunnus	Ennustatud ROE standardhälve
1.	YoY	5,92%
2.	OKm	4,92%
3.	VANUS	2,80%
4.	CR	1,70%
5.	PVK	1,34%
6.	KAIVE	0,53%
7.	POV	0,41%

Allikas: autori koostatud

Lisa 26. Otsustusmetsa III mudeli (ROE-LVK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	51,5%
2.	VANUS	30,4%
3.	CR	26,0%
4.	OKm	25,9%
5.	LVK	25,2%
6.	POV	23,5%
7.	KAIVE	15,5%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	4,08%
2.	VANUS	2,99%
3.	POV	2,34%
4.	CR	2,11%
5.	OKm	1,55%
6.	KAIVE	1,31%
7.	LVK	1,12%

Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

Jrk	Tunnus	Ennustatud ROE standardhälve
1.	YoY	13,39%
2.	VANUS	7,26%
3.	OKm	4,73%
4.	CR	3,05%
5.	LVK	2,84%
6.	POV	1,39%
7.	KAIVE	0,61%

Allikas: autori koostatud

Lisa 27. Otsustusmetsa IV mudeli (ROA-VK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	82,3%
2.	OKm	59,4%
3.	VANUS	56,4%
4.	CR	34,4%
5.	VK	33,3%
6.	POV	32,7%
7.	KAIVE	23,6%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	6,47%
2.	VANUS	5,59%
3.	OKm	3,78%
4.	POV	1,98%
5.	CR	1,96%
6.	VK	1,67%
7.	KAIVE	1,19%

Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

Jrk	Tunnus	Ennustatud ROA standardhälve
1.	YoY	5,93%
2.	OKm	4,74%
3.	VANUS	2,83%
4.	CR	1,57%
5.	VK	1,30%
6.	KAIVE	0,53%
7.	POV	0,44%

Allikas: autori koostatud

Lisa 28. Otsustusmetsa V mudeli (ROA-PVK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	84,5%
2.	OKm	63,4%
3.	VANUS	56,1%
4.	CR	36,7%
5.	PVK	36,3%
6.	POV	34,6%
7.	KAIVE	25,1%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	6,48%
2.	VANUS	5,99%
3.	OKm	4,17%
4.	CR	2,43%
5.	POV	2,34%
6.	PVK	1,90%
7.	KAIVE	1,27%

Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

Jrk	Tunnus	Ennustatud ROA standardhälve
1.	YoY	5,92%
2.	OKm	4,92%
3.	VANUS	2,80%
4.	CR	1,70%
5.	PVK	1,34%
6.	KAIVE	0,53%
7.	POV	0,41%

Allikas: autori koostatud

Lisa 29. Otsustusmetsa VI mudeli (ROA-LVK) tunnuste olulisus

Marginaalne permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	103,1%
2.	OKm	79,1%
3.	VANUS	63,9%
4.	POV	34,6%
5.	CR	34,5%
6.	KAIVE	29,2%
7.	LVK	28,5%

Allikas: autori koostatud

Tingimuslik permutatsiooniolulisus

Jrk	Tunnus	MSE suurenemine
1.	YoY	8,25%
2.	VANUS	6,16%
3.	OKm	4,88%
4.	CR	2,29%
5.	POV	1,85%
6.	LVK	1,51%
7.	KAIVE	1,38%

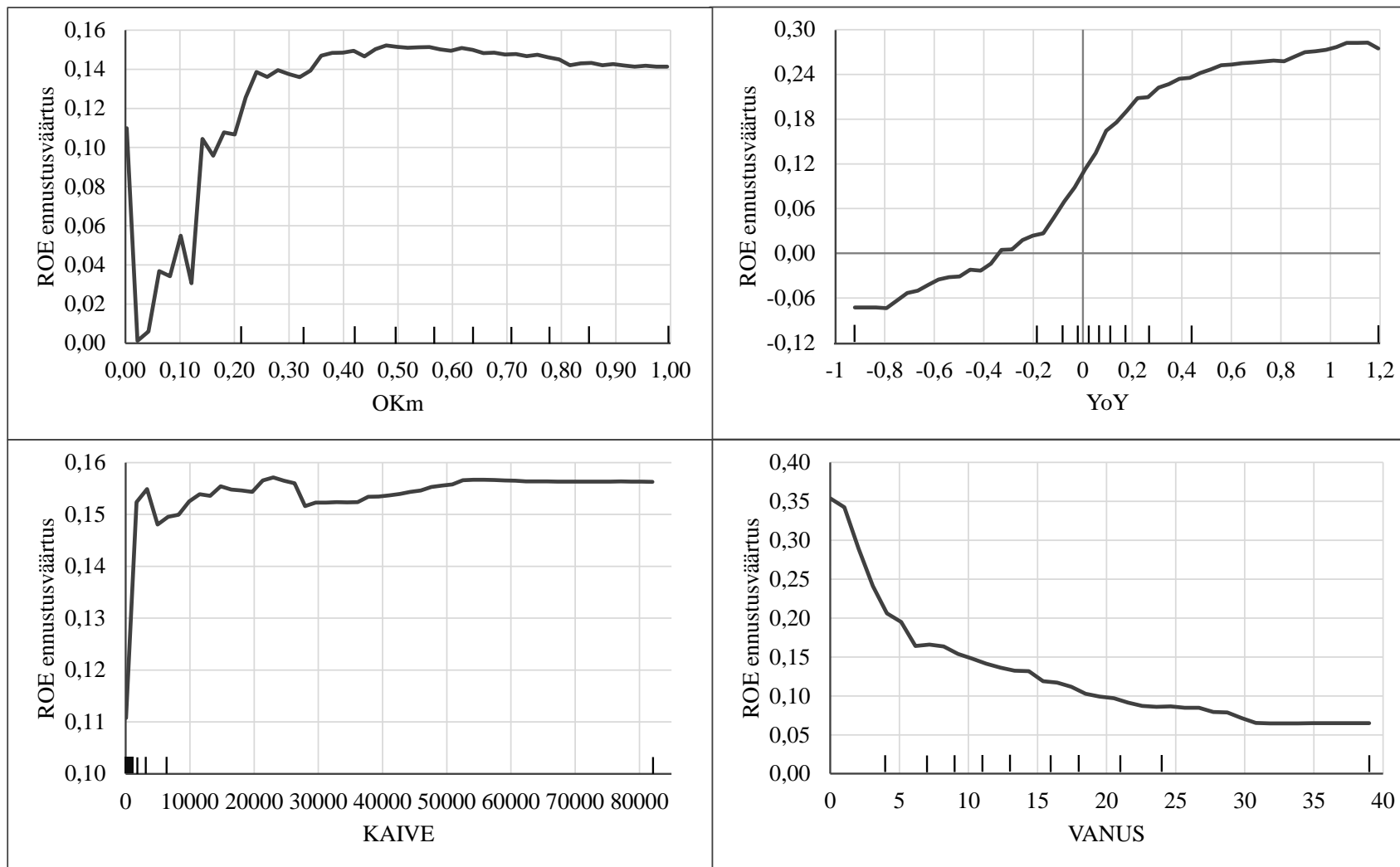
Allikas: autori koostatud

Variatsiooniolulisus osalise sõltuvuse funktsiooni põhjal

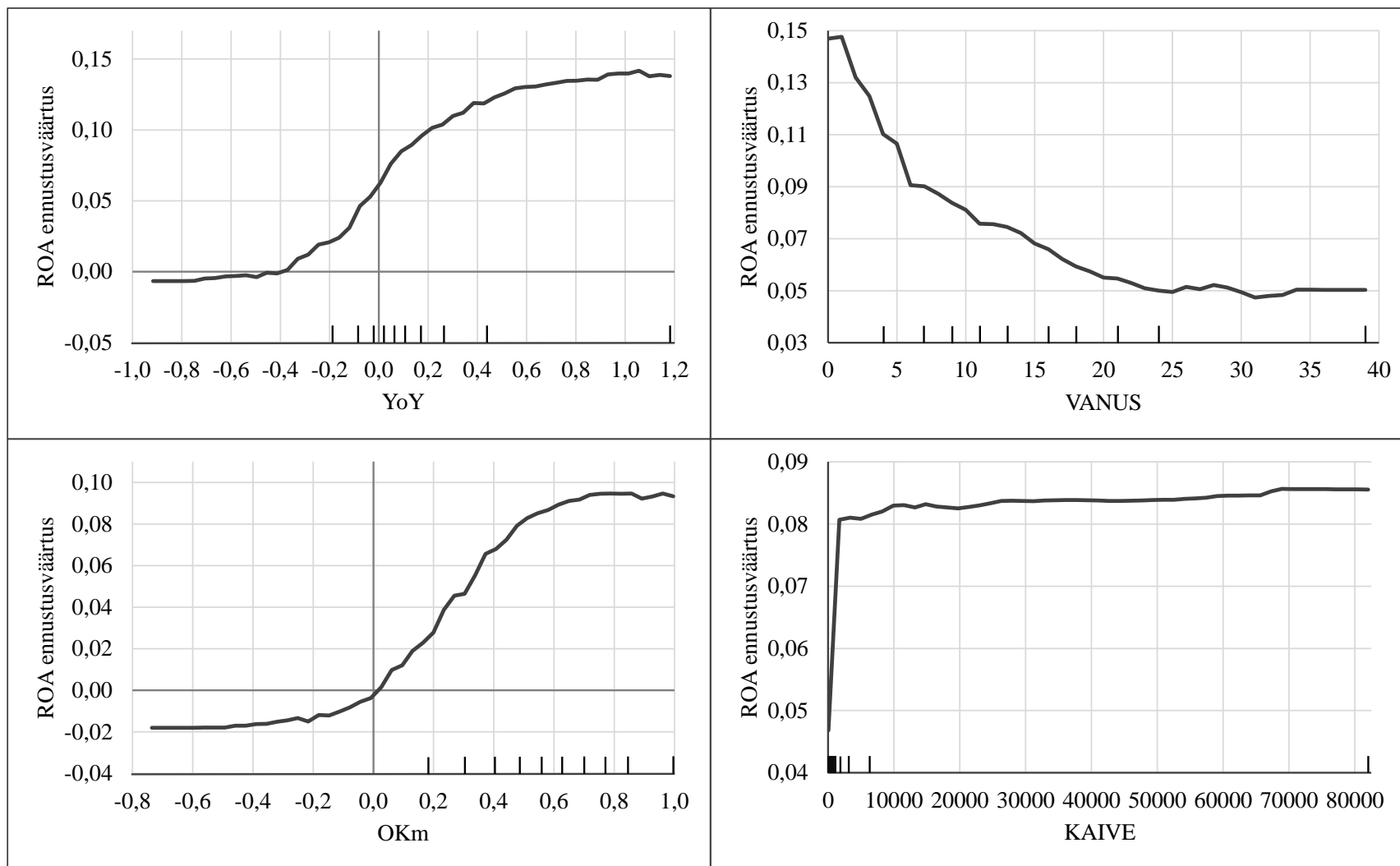
Jrk	Tunnus	Ennustatud ROA standardhälve
1.	YoY	7,16%
2.	OKm	6,63%
3.	VANUS	3,12%
4.	CR	1,44%
5.	LVK	0,89%
6.	POV	0,57%
7.	KAIVE	0,56%

Allikas: autori koostatud

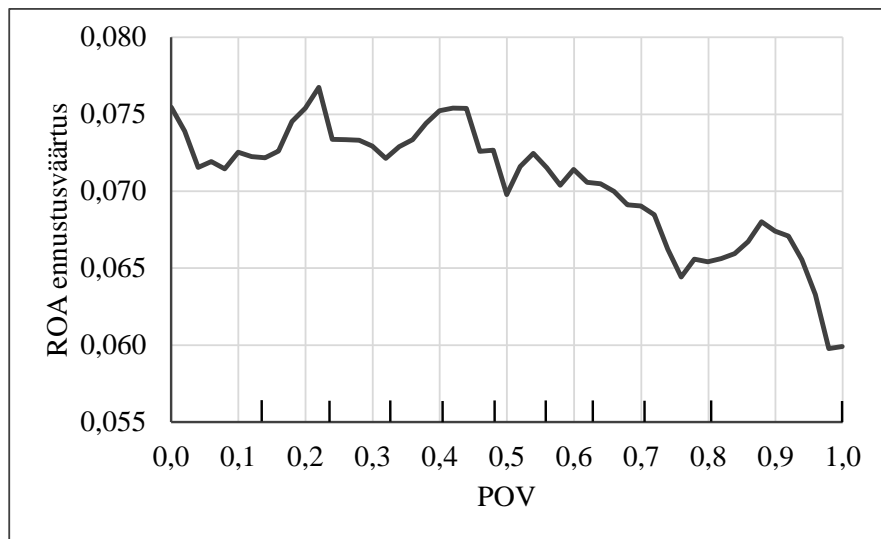
Lisa 30. Otsustusmetsa I mudeli kontrollnäitajate osalise sõltuvuse joonised



Lisa 31. Otsustusmetsa IV mudeli kontrollnäitajate osalise sõltuvuse joonised

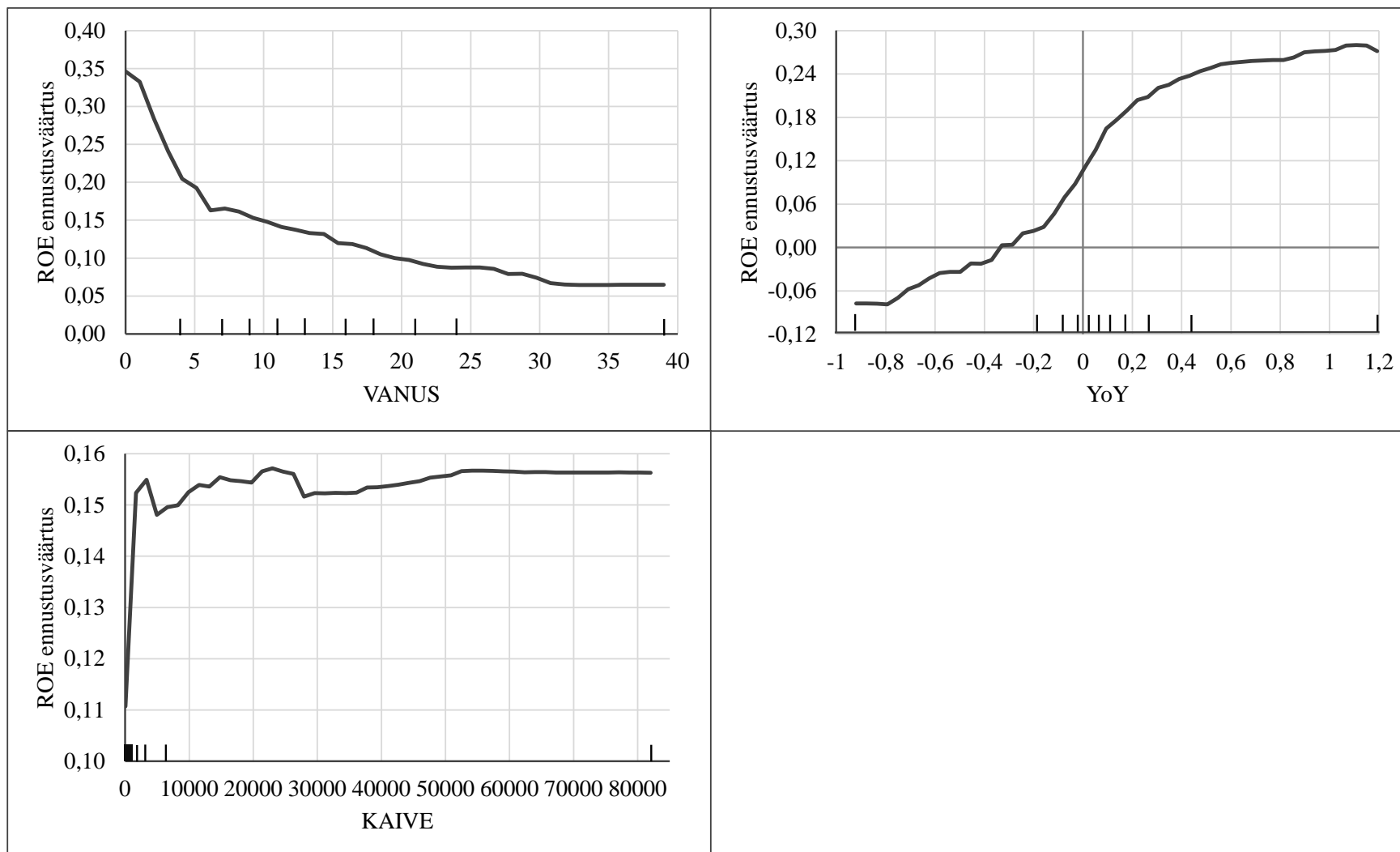


Lisa 31 järg

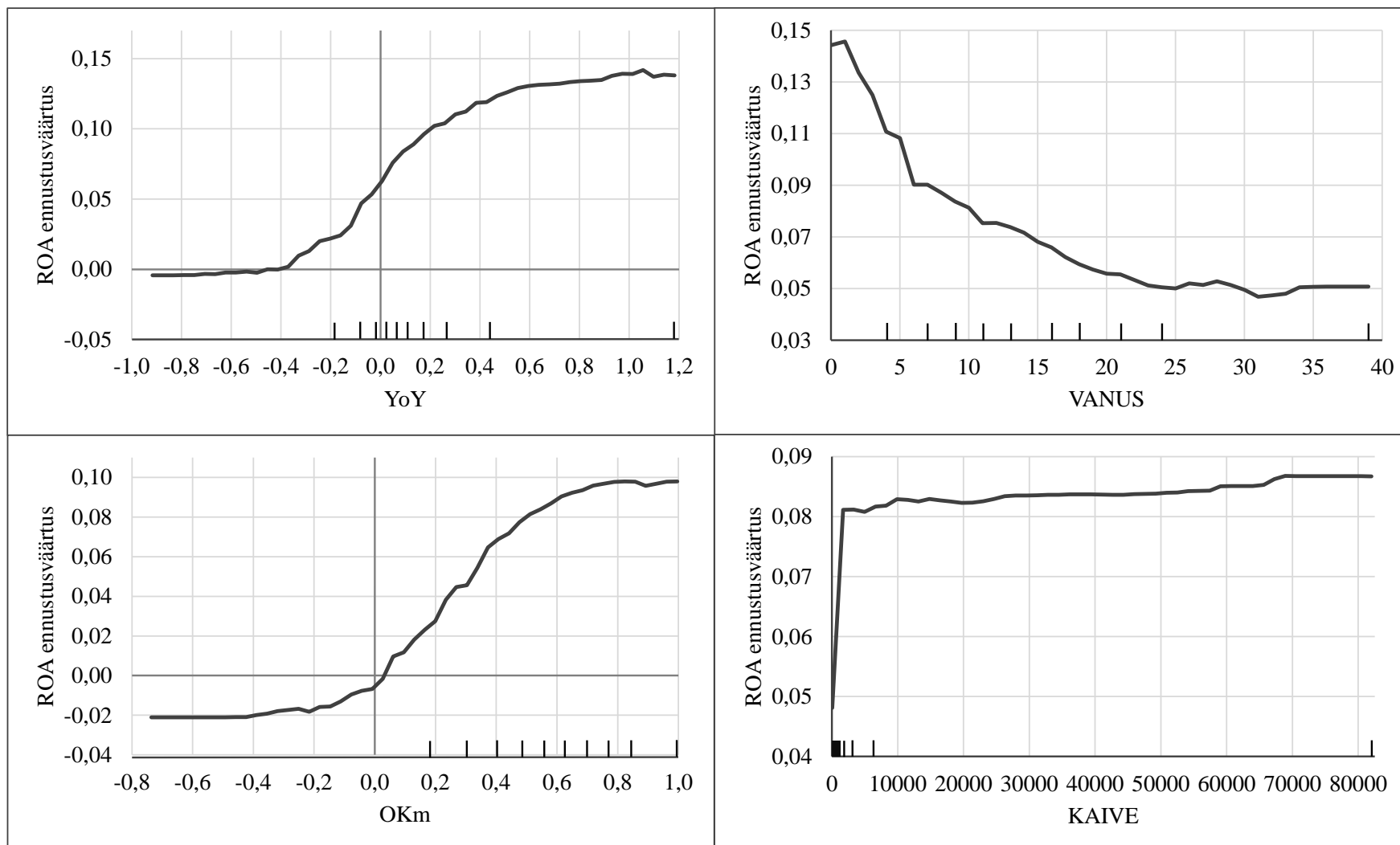


Allikas: autori koostatud

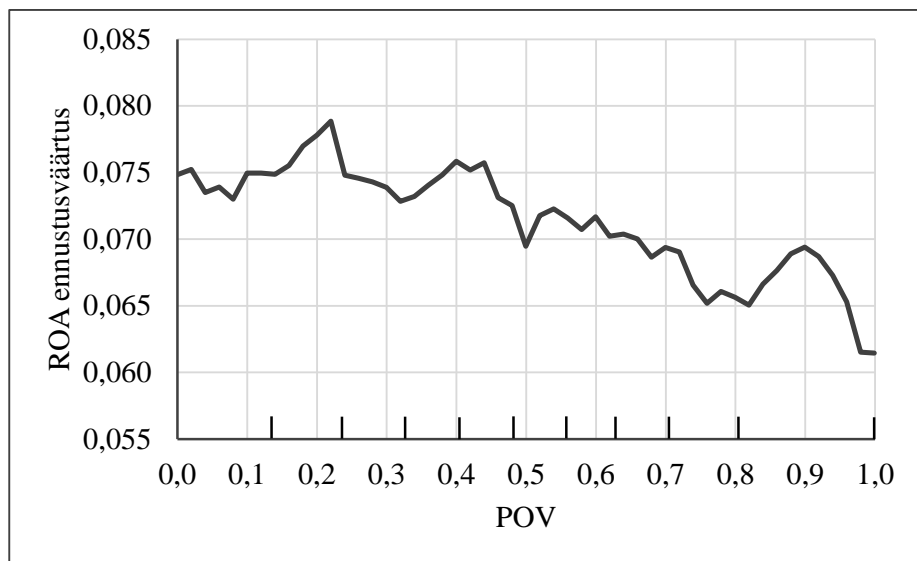
Lisa 32. Otsustusmetsa II mudeli kontrollnäitajate osalise sõltuvuse joonised



Lisa 33. Otsustusmetsa V mudeli kontrollnäitajate osalise sõltuvuse joonised

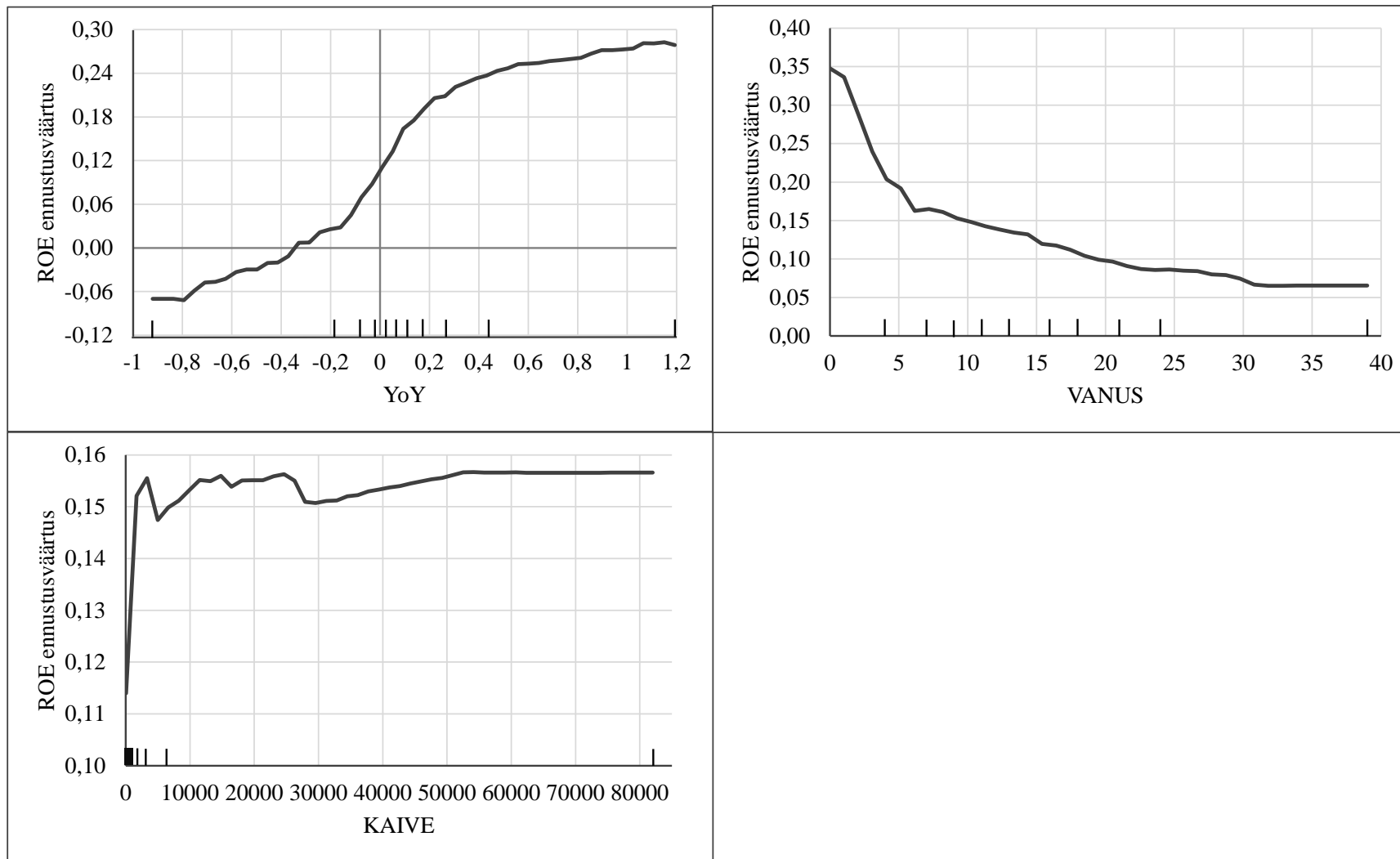


Lisa 33 järg

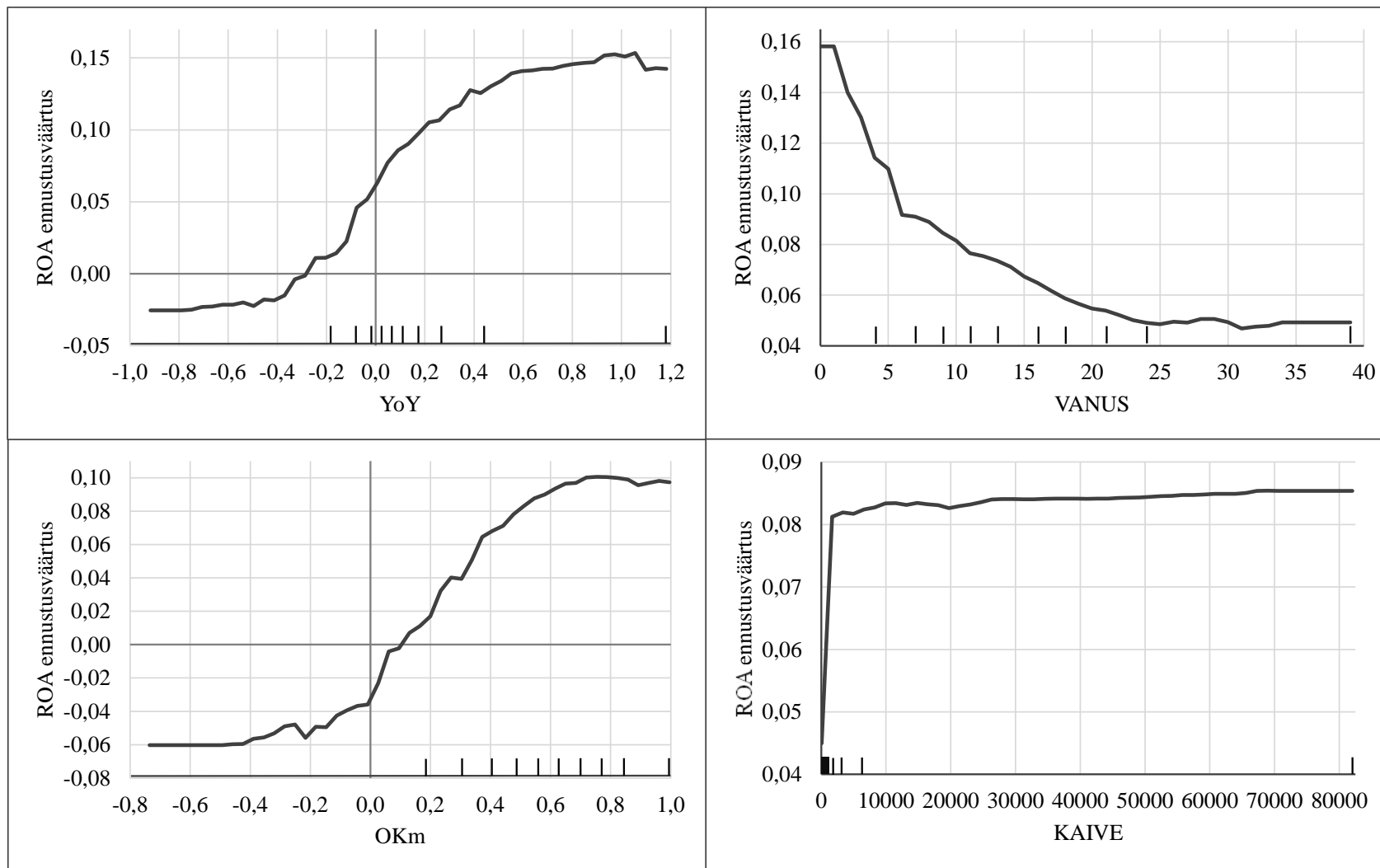


Allikas: autori koostatud

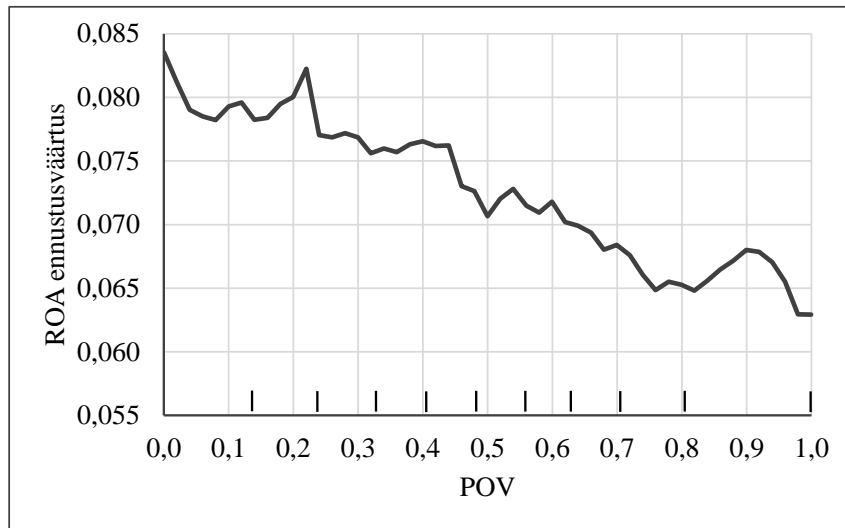
Lisa 34. Otsustusmetsa III mudeli kontrollnäitajate osalise sõltuvuse joonised



Lisa 35. Otsustusmetsa VI mudeli kontrollnäitajate osalise sõltuvuse joonised



Lisa 35 järg



Allikas: autori koostatud

Lisa 36. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Artur Kohv,

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose

Eesti tööstusettevõtete kapitali struktuuri seosed kasumlikkusega,

mille juhendaja on Kaido Kepp,

1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

08.05.2021