

TALLINNA TEHNIKAÜLIKOOL
Majandusteaduskond
Majandusanalüüsi ja rahanduse instituut

Marko Mitt
SOTSIAALSETEL TUNNUSTEL PÕHINEV DISKRIMINEERIMINE BONDORA
LAENUPLATVORMIL
Magistritöö
Õppekava TARM02/02, peaeriala ärirahandus

Juhendaja: Mari-Liis Kukk, 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 11 999 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

.....04.01.2021

(allkiri, kuupäev)

Üliõpilase kood: 106591TARM

Üliõpilase e-posti aadress: marko.mitt@zoho.com

Juhendaja: Mari-Liis Kukk, MA

Töö vastab kehtivatele nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(nimi, allkiri, kuupäev)

SISUKORD`

LÜHIKOKKUVÕTE	4
SISSEJUHATUS	5
1. LAENUPLATVORMID JA DISKRIMINEERIMINE.....	8
1.1. Laenuplatvormid	8
1.1.1. Pangad vs. laenuplatvormid	9
1.1.2. Risk ja reitingud	11
1.2. Diskrimineerimine	13
1.2.1. Algoritmiline diskrimineerimine	14
1.2.2. Sugu	17
1.2.3. Päritolu	19
1.2.4. Muud sotsiaalsed tunnused.....	21
1.2.5. Finantsnäitajad.....	22
2. ANDMED JA METOODIKA	24
2.1. Andmed.....	24
2.1.1. Valim	24
2.1.2. Muutujad ja kirjeldav statistika	25
2.2. Metoodika.....	31
3. MUDEL JA JÄRELDUSED	32
3.1. Mudel.....	32
3.1.1. Mudeli testimine	33
3.1.2. Lõplik mudel	35
3.2. Järeldused.....	36
KOKKUVÕTE	44
SUMMARY	47
KASUTATUD ALLIKATE LOETELU	52
LISAD	60
Lisa 1. Korrelatsioonimaatriks (Gretl).....	60
Lisa 2. Regressioonanalüüsi mudel I (Gretl).....	61
Lisa 3. Regressioonanalüüsi mudel II (Gretl).....	62
Lisa 4. VIF (Gretl).....	63
Lisa 5. Lihtlitsents	64

LÜHIKOKKUVÕTE

Töö eesmärk on välja selgitada, kas ja millises ulatuses tuleb ette diskrimineerimist Bondora laenuplatvormil sotsiaalsete tunnuste põhjal. Uurimisküsimused on eelkõige erinevate vähemusgruppide võimalikku ebavõrdset kohtlemist analüüsivad. Täpsemalt, kas esineb diskrimineerimist soo, päritolu, vanuse või muude sotsiaalsete tunnuste järgi.

Peale andmete kohandamist on valimi suuruseks 69 902 heaks kiidetud laenuaotlust. Analüüsimiseks kasutatakse vähimruutude meetodil põhinevat regressioonanalüüsi, kus sõltuvaks muutujaks on laenuintress. Kõik analüüsi kaasatud muutujate tulemused on statistiliselt olulised ja nende märgid loogilised. Samuti on lõplik mudel statistilised oluline.

Tulemuste põhjal saab asuda seisukohale, et naissoost, mitte eestlased, noored ja ilma kõrghariduseta laenajad tasuvad kõrgemat intressi. Selle alusel on võimalik asuda seisukohale, et neid vähemusgruppe diskrimineeritakse. Samas tuleb rõhutada, et ei saa välistada, et Bondora laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleeruvad töös kasutatud tunnustega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. Seega pole võimalik põhjanevalt väita, et Bondora laenuplatvormil on tegemist sotsiaalsete tunnuste järgi diskrimineerimisega.

Võtmesõnad: Bondora, laenuplatvorm, sotsiaalsed tunnused, regressioonanalüüs, diskrimineerimine

SISSEJUHATUS

Võrdõiguslikkus leiab aasta-aastalt aina rohkem kõlapinda. Sellest tulenevalt on meie keskel uurima ja tuvastama hakatud erinevaid diskrimineerimisvorme. Peamiselt diskrimineeritakse vähemusgruppe nende sotsiaalsete tunnuste põhjal. Näiteks soo ja rahvuse, kuid ka vanuse, puude, religiooni, seksuaalse orientatsiooni jms erinevuste järgi enamusgrupist. Tüüpilised näited on tööhõive, kus eelistatakse nt meest naisele, noort vanale, või palgalõhe, kus meestel on naistest kõrgem sissetulek. Viimase osas on paslik näide Eesti. Nimelt Euroopa Liidu riikide võrdluses selgus, et vastav näitaja on Eestis kahjuks suurim (Eurostat 2019).

Lisaks on võimalik indiviide või gruppe ebavõrdselt kohelda teenuste pakkumisel. Nt müüa teenust kallimalt, näha ette koormavamaid tingimusi või selle pakkumisest üldse keelduda. Üheks selliseks objektiks on finantsteenused. Täpsemalt on siinkohal küsimuseks, kas krediidi pakkumisel koheldakse erinevaid gruppe olenemata nende sotsiaalsetest tunnustest võrdselt. Olukorras, kus näiteks nn halli passi omanikud saaks Eestis laenu harvem või kõrgema intressiga, kuid samas nende kohustuste suhe sissetulekusse on enamusgrupiga võrrelduna vähemalt sama, oleks tegemist diskrimineerimisega.

Sotsiaalseid tunnuseid väljendavate andmete kättesaadavus on paraku kommertspankade puhul keeruline. Samas on turule tulnud veebiplatvormid, mis keskenduvad laenude vahendamisele. Eestis paremini tuntud kui ühisrahastuse platvormid. Üldjuhul on need laenuplatvormid teinud kättesaadavaks oma klientide kohta põhjaliku andmestiku, mis hõlmavad ka erinevaid sotsiaalseid näitajaid. Üheks selliseks platvormiks Eestis on Bondora, mis vahendab laene eraisikutele. Selliste andmete analüüsimine annab hea võimaluse uurida vastavat probleemi Eesti kontekstis. Tulenevalt eeltoodud viitest Eesti kohta, kui suurima palgalõhega riik, on seda enam vastava temaatika uurimine oluline.

Töö olulisuse osas saab lisaks välja tuua sarnaste uurimuste puudumise, kus Eestis oleks uuritud diskrimineerimist finantsteenuste pakkumisel (siinkirjutajale teadaolevalt). See annab võrdlusmomendi muu maailmaga ja aitab paremini mõista sellist tähtsat ühiskonna toimimise

küsimust, nagu seda on võrdõiguslikkus. Võttes arvesse Bondora andmete head kättesaadavust ja suurt turuosa Eesti mõttes, saab seda pidada heaks uurimise subjektiks diskrimineerimise välja selgitamisel.

Töö eesmärk on välja selgitada, kas ja millises ulatuses tuleb ette diskrimineerimist Bondora laenuplatvormil sotsiaalsete tunnuste järgi. Sellest johtuvalt on uurimisprobleemiks laenuplatvormil Bondora finantsteenuste pakkumisel võimalik diskrimineerimine sotsiaalsete tunnuste järgi ja kõrvutada saadud tulemusi vastava erialakirjandusega. Uurimisküsimused on eelkõige erinevate vähemusgruppide võimalikku ebavõrdset kohtlemist analüüsivad. Kas esineb diskrimineerimist:

- soo;
- päritolu;
- vanuse;
- muude sotsiaalsete tunnuste järgi?

Uurimisülesandeks on erialakirjanduse põhjal luua teoreetiline alus diskrimineerimise kohta finantsteenuste pakkumisel ja samas selgitada ka platvormide toimimisprotsessi. Nimelt laenuplatvormid määravad laenudele intressid läbi enda poolt loodud krediidireitingute. Erinev intress, *ceteris paribus*, on üheks võimalikuks teguriks, mille järgi saab hinnata diskrimineerimise olemasolu. Peale selle valida ja kohandada muutujad Bondora andmestikust arvutuste läbiviimiseks ning valida selleks sobiv meetodika. Järgnevalt katsetada mudelit, esitada tulemused ja teha otsused vastavalt null- või sisuka hüpoteesi vastuvõtmise teel. Seejärel hinnata, kas töös püstitatud hüpoteesid leidsid kinnitust või mitte. Kokkuvõtteks saab tulemuste põhjal esitada seisukohad uurimisküsimuste kohta, kas Eesti andmestikul põhineval laenuplatvormil esineb viiteid võimalikule diskrimineerimisele sotsiaalsete tunnuste järgi. Nimelt, Bondora tegutseb küll kolmes erinevas riigis, kuid valdav enamus tegevusest on Eestis ning välisriikide, Soome ja Hispaania, kohta on hulganisti vigaseid andmeid. Sellest tulenevalt keskendub töö analüüs vaid Eesti andmetele.

Käesolev töö on jaotatud kolme peatüki alla. Neist esimene on teoreetiline käsitlus erialakirjanduse põhjal, mis keskendub diskrimineerimise lahkamisele finantsteenuste pakkumisel. Seda nii laiemalt panganduses kui ka otseselt sarnastel laenuplatvormidel, nagu seda on Bondora. Mõistmaks paremini uurimisobjekti, annab teoreetiline käsitlus lisaks ülevaate laenuplatvormide

toimimispõhimõtetest üldiselt, tuues välja ka erinevusi traditsioonilisest pangandusest. Olenevalt asjaolust, et riski tajumine ja selle hinnang on keskne tegur intressimäära juures, leiab käsitlust ka laenuplatvormidega seotud riskid ning nende poolt loodavad reitingud. Töö teises osas selgitatakse, mis on valimiks, kust see pärineb, kuidas on andmed kohandatud ning esitatud ka kirjeldav statistika. Lisaks leiab aset arutelu erinevate meetodikate vahel vastavalt erialakirjandusele ning selgitus valiku osas. Töö kolmandas osas testitakse mudelit, esitatakse tulemused ning tehakse teoreetilisest käsitlusest juhitud järeldused.

Uurimuse aluseks olev andmestik pärineb Bondora avalikust andmebaasist, kuhu on koondatud kõik tema poolt väljastatud laenud kogu tegevusajaloo käigus. Peale andmete kohandamist ja perioodi kitsendamist on valimi suuruseks 69 902 heaks kiidetud laenuaotlust Eesti piires. Andmete analüüsimiseks kasutatakse käesolevas lõputöös vähimruutude meetodil põhinevat lineaarset regressiooni, kus sõltuvaks muutujaks on määratud laenuintress.

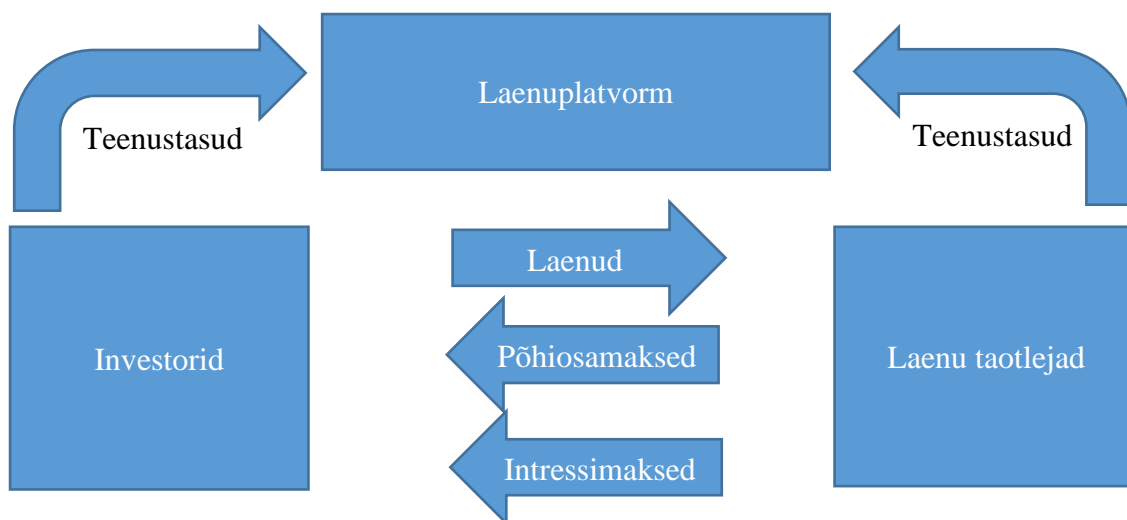
Autor jääb lootma, et see töö tekitab huvi samalaadsete uurimuste läbiviimiseks, mis annaks Eestis parema üldpildi võrdõiguslikkuse küsimuses. Kokkuvõtteks tänab töö autor juhendajat põhjaliku tagasiside eest. Sellest nähtus selgelt, et juhendaja võttis esitatu süvitsi ette. Sellest tulenevalt aitas tagasiside märgavalt tajuda töö kirjutamise suunda. Head!

1. LAENUPLATVORMID JA DISKRIMINEERIMINE

Tulenevalt asjaolust, et töö keskendub *peer-to-peer* (P2P) laenamisele selleks ette nähtud veebiplatvormil, siis esmalt terminoloogiline selgitus. Eesti keeles on *peer-to-peer*'i tõlkeks võrdõigusvõrk¹ ning erialakirjanduses on kasutusel ka termin partnervõrk (Laht 2016). Haakuvaks terminiks saab pidada ka ühisrahastust, kuid erialakirjanduses käsitletakse neid üldjuhul eraldi (Tao *et al.* 2017). Viimane selgitab finantseerimisviisi, kus eelkõige rahastatakse mõnda projekti, mitte tarbimislääne, nagu seda on üldjuhul P2P laenamine (Belleflamme *et al.* 2021; Cumming, Johan 2016). Kuna käesoleva töö fookus on suunatud P2P laenuplatvormidele siis edaspidi on parema loetavuse nimel kasutatud läbivalt terminit laenuplatvorm või platvorm.

1.1. Laenuplatvormid

Alljärgnev joonis (1) illustreerib laenuplatvormi osapoolte vahelisi maksete liikumisi.



Joonis 1. Laenuplatvormi tööpõhimõte
Allikas: Davis ja Murphy (2016), autori koostatud.

¹ IT terministandardi sõnastik <http://www.eki.ee/dict/its/>

Laenuplatvormi teenusel on lihtsustatult väljendades kolm osapoolt: platvormi haldaja, laenu taotlejad ja laenuandjad ehk investorid. Haldaja teenuseks on teiste osapoolte kokkuviimine ning laenu ja selle tagasimaksete vahendamine. Seejuures viib ta läbi krediidi hinnangu ja paneb paika intressid. Haldaja tegutseb seega vahendaja rollis, võtmata otsest finantsriski. St haldaja ise raha välja ei laena. Tema tulu tuleb mõlemale osapooltele määratud teenustasudest. (Davis, Murphy 2016)

1.1.1. PANGAD vs. LAENUPLATVORMID

Pankade roll on pakkuda erinevaid finantsteenuseid. Tänapäeval on panga poolt pakutavate teenuste hulk väga suur ja mitmetahuline. Klassikalises mõttes maksab pank hoiustajatele intressi ja laenab raha kõrgema intressi eest välja. Seejuures võtab ta laenaja ees finantsriski ning vastutab hoiuste eest – ta on selle suhte vahelülis, kus pooled usaldavad teda kui spetsialisti informatsiooni asümmeetria vähendamisel. Pank kasutab laenu andmiseks hoiustajate raha, omakapitali, oma rahavooge, finantsturgudelt laenatud raha jms. Ta viib läbi taustakontrolli, väljastab laenu ning kogub laenumakseid. Pankade finantsvahenduse kulu on 130 aasta jooksul varieerunud 1.5–2.5 protsendi vahel, millest viimased 40 aastat on olnud trend üles (Philippon 2015).

Finantsinnovatsiooniliste lahenduste eesmärgiks on seda kulu vähendada (Thakor 2019). Nn digitaalse revolutsiooni käigus 90ndatel kasvas turuosa enim just neil ettevõtetel, kes suutsid juurutada digitaalsetel lahendustel põhinevat andmeanalüüsi (Stein 2002). See muutis ka kogu laenu turgu. Kui varasemalt oli võimaliku kliendi finants- ja sotsiaalandmete analüüsija laenuhaldur, siis info kiire ning odava digitaliseerimise tulemusena on nüüd otsustajaks laiahaardelised algoritmid (Wang 2018). See on turule toonud laenuplatvormid, kes kasutavad neid erinevaid infotehnoloogilisi ja finantsinnovaatilisi lahendusi tõhusalt ära. Tänu sellele on nad suutnud lühikese ajaga tekitada märkimisväärse turuosa samaväärses sektoris traditsiooniliste pankade kõrvale.

Nimelt esimene laenuplatvorm loodi Suurbritanniasse 2005. aastal ning aasta hiljem USAsse (Pokorna, Sponer 2016). Aastal 2010 vahendasid platvormid USAs ühe protsendi kogu tagatiseta tarbimislaenudest. Seitse aastat hiljem oli see juba kolmandik (Wang 2018). Alternatiivrahastus üldiselt (hõlmates ühisrahastus- ja P2P laenuplatvorme) moodustas 2018. aastal üle 180 riigis

käibeks 304,5 miljardit dollarit (Ziegler *et al.* 2020).² Seejuures kaks kolmandikku globaalsest turust moodustab Hiina, kus tegutseb üle nelja tuhande laenuplatvormi (Tharok 2019). Samuti ei ole enam tegemist pelgalt idufirmade ja erainvestorite pärusmaaga. Nt sellele turule on sisenenud ka finantsasutus Goldman Sachs (läbi ettevõtte Marcus), kes väljastas esimese kaheksa kuuga üle miljardi dollari laene (Wang 2018). Lisaks eraisikutele on platvormidel investoriteks suurte mahtudega nii era- kui riiklikud institutsioonid (Wang 2018; Pokorna, Sponer 2016).

Saksamaal läbi viidud uurimuses selgus, kui võeti arvesse laenuplatvormide riskimäärasid, et intressid on pankadega tegelikult samaväärsed (de Roure *et al.* 2016). Põhjused, miks pangad seda turuosa käest ära annavad, ei pruugi seega peituda üksnes asjaolus, et laenuplatvormid pakuvad klientidele odavamaid laene. Platvormide kiirele arengule on aidanud kaasa samaaegne pangandusregulatsioonide ja kapitalinõuete rangemaks muutmine peale 2008. aasta globaalset finantskriisi (Traci *et al.* 2014; Moenninghoff, Wieandt 2013). Kriisi tõttu vähendasid pangad oma laenuportfelli ja klientuuri (Wang 2018). Samas on laenuplatvormil turule sisenemiseks vaja vaid veebirakendust koos haldusega ning turundust. Erinevalt pangast pole vaja teenindussaale ja olenevalt riiklikust regulatsioonist ei pruugi vaja olla ka litsentse.

Lisaks on pankade arvutisüsteemid suured ja kulukad, kuna üldjuhul on tegemist eelmiste kümnendite päranditega, kuhu peale on pidevalt ehitatud uuendusi.³ Laenuplatvormid on aga äsja programmeeritud kasutades tänapäevaseid efektiivsemaid mooduleid, mille halduskoormus on madal (*Ibid.*). Kui platvorm on programmeeritud nii, et toimimis- ja otsustusprotsess on läbinisti automaatne, puudub vajadus sellise hulga töötajaskonna järele nagu on samaväärsel pangandusüksusel (*Ibid.*). Täiendavalt on positiivseks argumendiks läbipaistvus. Platvormid keskenduvad suures osas regulaarse statistika väljatoomisele erinevate laenutoodete, reitingute, tootlikkuse jms kohta, et investorid saaksid iseseisvalt hinnata ümbritsevat keskkonda. Suurt rõhku on pandud selgitamaks intresside ja reitingute määramisprotsessi, mis pole nt pankade puhul tavapraktika. (*Ibid.*)

Laenuplatvormide klientideks on peamiselt eraisikud, kes kasutavad võetavat laenu olemasoleva kallima laenu tasumiseks ja muudeks suuremateks kuludeks nagu pulmad, reisid, haridus, äri, auto, ravikulud jne (Wang 2018). Mõned platvormid keskenduvad vaid äri-laenudele väikeettevõtluks,

² Aastal 2017 oli käibeks 419 miljardit dollarit. Langus tuli peamiselt Hiinast. Kui Hiina välja arvata siis globaalne turu tõus 2018. aastal oli 48%. (Ziegler *et al.* 2020)

³ Kaks kolmandikku pankade IT eelarvest kulub vana süsteemi hoolduse peale (Celent 2013).

kinnisvarale või öppelaenude refinantseerimisele (Gaigalienė, Česnys 2018). Platvormid on suurendanud märgatavalt madalama sissetulekuga inimeste võimalusi mõistliku intressiga laenu saamiseks. Varasemalt oli neil valikuks vaid kallid laenuhood (kiiralaenu, krediitkaardid) ja nn laenuhaide teenused (Santoso *et al.* 2020).

Laenuplatvormide investoriks meelitab eelkõige selle suurem võimalik tootlikkus võrreldes nt hoiuste ja võlakirjadega (Malekipirbazari, Aksakalli 2015), kuna ajalooliselt on rahaturgude intressid laenuplatvormide kiire arengu perioodil olnud väga madalal tasemel (Haewon *et al.* 2012). Lisaks läheb võlakirja investeeringu tarvis vaja märksa suuremat algkapitali, et tulusus oleks rahuldav. Laenuplatvormi investoriks on võimalik saada aga miinimum kapitali piiranguta ja teenida sellelt kahekohalist määra. Oluliseks argumendiks on ka kasutusmugavus, kus investeringuid tehakse internetipangast või deebet-/krediitkaardiga. Lisaks on võimalik seadistada pangakontolt igakuine passiivne automaatmaks, mille korral platvorm investeerib raha vastavalt kasutaja poolt ette antud oodatava tulususe ja riskimäära järgi. (Wang 2018)

Investorite vaatenurgast on kognitiivsete põhjustena välja toodud veel ka platvormide mängulisust, kus investor üritab ära tabada vähem riskantsemaid taotlejaid. Lisaks enda esiletoomist nt sõprade seltsis kui asjatundjat ja ülemäärast enesekindlust, et saab riskihindamisega pankadest paremini hakkama. (*Ibid.*)

1.1.2. Risk ja reitingud

Viimati mainitu, riskihinnang, on laenuhoode alustalaks. Sellise pädevuse evimine on olnud läbivalt pankade põhirolliks, mis mõnel juhul võib rajaneda sadade aastate pikkusele kogemusele (Diamond 1984). Nüüd on aga turul laenuplatvormid, kelle praktika on väga lühike, nad ei kohtu laenuhoodejatega ja laenu väljastatakse üldjuhul ilma tagatiseta. Samas eelistavad ja otsivad uusi laenuvõimalusi eelkõige riskantsemad laenajad (Cornaggia *et al.* 2018) või need, kes pangast enam laenu ei saa (Pokorna, Sponer 2016). Nt platvormide maksejõuetusele keskenduvast uurimuses samastasid Serrano-Cinca *et al.* (2015) selle kaudu investeerimise riskitaset derivatiividega, kuna neile pole kapitali nõudeid ja nad ei võta tagatist. Kui laenaja tagasimakseid ei tee, siis suure tõenäosusega peab investor leppima raha kaotusega (Pokorna, Sponer 2016).

Laenuplatvormide roll lisaks osapoolte kokkuvõimisele ja haldusele on hinnata riski iga taotleja kohta. Tegemist on informatsiooni asümmeetria probleemiga (Stiglitz, Weiss 1981), kus laenuhoode puhul puudub investoril täielik ülevaade laenuhoodejate kohta, mis võib viia mitte

tõepärase riskihinnanguni. Informatsiooni asümmeetria ja krediidiriski vahel on tugev positiivne korrelatsioon (Cai *et al.* 2016). Seega platvormi tegevuse aluseks on oskus koguda infot ja selle põhjal tuvastada riskitase⁴, et investor saaks teha adekvaatsema otsuse. Kui platvorm eksib tihti, kaotab ta usaldusväarsuse, mis omakorda peletab tulevase investoreid eemale. Sellest tulenevalt on riski hindamine platvormi jaoks üks olulisemaid pädevusi (Pokorna, Sponer 2016).

Selle tarvis on nad üldjuhul välja töötanud riskitasemed, mis koosnevad reitingutest. Nt USA turuliider Lending Club⁵ kasutab astmesüsteemi A1 (nii risk kui intress on madalaim) kuni G5 (risk ja intress kõrgeim).⁶ Laenuplatvormide reitingute loomisprotsess põhineb mitmetahulistel algoritmidel (masinõpe, tehisintellekt jms) (Wang 2018). Algusaastatel kasutasid platvormid reitingute määramisel peamiselt vaid maksehäirete andmebaase (*Ibid.*) (nagu Krediidiinfo Eestis).

Nüüdseks on vastavaid algoritme arendatud sedavõrd, et kasutatakse lisaks finantsvõimekuse näitajatele (pangaväljavõte, töökoht, -kogemus, varad, kohustused ja maksehäire raport (*Ibid.*) ka sotsiaalseid näitajaid (perekonnanais, hobid ja sotsiaalmeedia aktiivsus) (Guo *et al.* 2016; Ge *et al.* 2017). Nimelt on leitud, et ülevaade n-ö pehmetest andmetest (*soft data*) ja sotsiaalsetest võrgustikest aitab vähendada informatsiooni asümmeetriat (Lyer *et al.* 2009; Bellefamme *et al.* 2015; Liu *et al.* 2015; Dorfleitner *et al.* 2016).⁷ Kokkuvõtvalt on erinevad uurimused näidanud, et need reitingud aitavad ennustada maksejõuetuse riski adekvaatselt (Herzenstein *et al.* 2008; Lyer *et al.* 2009; Emekter *et al.* 2015; Pokorna, Sponer 2016).⁸

Ülaltoodud riskihinnangut saab pidada operatsiooniliseks riskiks, kus vähesest oskusest või kogemata määratakse vale reiting (Lei 2016). Täheldatud on, et info haldamise kvaliteet pole platvormide lõikes alati samaväärne (Vallee, Zang 2019). Investori seisukohast on veel riskideks laenaja maksevõime (Davis, Murphy 2016) ja vähene hajutus (Gaigalienė, Česnys 2018). Osa

⁴ Prantsusmaa notarid olid sajandeid laenuvahendajad nagu seda on praegusel juhul P2P laenuplatvormid. Nimelt notarid olid erinevate võla- ja asjaõiguslepingute tõestajateks ja tunnistajateks. See andis neile pädeva ülevaate inimeste varade üle. Selle info põhjal nad teadsid, kellel on raha laenamiseks, kes tahab laenu saada ja kas viimane on võimeline ka tagasi maksma. Sel viisil vahendatud P2P laenude osakaal oli 19. sajandil riigi suurim (Hoffmann *et al.* 2019).

⁵ Lending Club Loan Grades. <https://www.lendingclub.com/foiofn/rateDetail.action> (11.11.2021).

⁶ Suurt turuosa omavad platvormid järgivad sarnast reitingusüsteemi: Prosper (USA), Renrendai (Hiina), Symecredit (Tšehhi), Alpha (Indonesia), Smava (Saksamaa) (Pokorna, Sponer 2016; Chen *et al.* 2019; Santoso *et al.* 2020; Barasinska 2011).

⁷ Saab argumenteerida, et sellise info omamine on turueeliseks finantsteenuste pakkumisel. Nt Apple on hakanud makseid vahendada (Apple Pay), Amazon on hakanud laene väljastama (Amazon Lending) ja Facebook üritab luua oma krüptovaluutat (Wang 2018).

⁸ Samuti krediitkaartide riskihindamisel on algoritmide kasutamine aidanud vähendada kahjumeid neljandiku võrra (Khandani *et al.* 2010).

platvorme on näinud ette, et järelturul on võimalik oma investeeringut edasi müüa. See annab investorile võimaluse halvast laenust väljuda. Iseküsimus, kas ja mis hinnaga seda on võimalik teha (*Ibid.*). Hajutamise osas on enamikes platvormides põhimõte, et investori raha ei lähe kogu ulatuses ühele taotlejale (Hollas 2013). Lisaks on ka platvorme, kus võimalik ise valida nt erinevaid piirkondi, et sellist süstemaatilist riski omakorda vähendada (Alyakoob *et al.* 2020).

Investori riskiks saab pidada ka teda ennast. Enamikel laenuplatvormide investoritel puudub piisav finantskirjaoskus (Klafft 2008b) ja seda ka ei nõuta (Gibilaro, Mattarocci 2018). Nad ei teadvusta võetavat riski adekvaatselt (Freedman, Jin 2017; Haewon *et al.* 2012) ning nad hindavad seda erinevalt (Cai *et al.* 2016). Probleemkohaks on info limiteeritus, kuna pole otseselt teada, kellele raha läheb (Pokorna, Sponer 2016). Otsuste tegemisel on peamiselt vaid valik reitingu, riski ja tootlikkuse suhte vahel. Kui laenaja kohta on olemas täiendavad andmed, siis investorid tuginevad rohkem nn pehmetele andmetele (nt pilt), olgugi et sellel pole toetavat mõju investeeringu õnnestumisele (Pope, Sydnor 2012; Barasinska, Schäfer 2014; Gonzales, Loureiro 2014).

Platvormidel on täheldatud ka palju karjakäitumist (*herding behaviour*) – indiviidi otsused on oluliselt mõjutatud teiste valikutest (Herzenstein *et al.* 2011). Nimelt platvormide pakkumiskeskond sarnaneb veebioksjonitele, kus antud juhul on kaubaks laenu (Davis, Murphy 2016). Sellelt nähtub, kui palju pakkumisi on erinevatele laenudele tehtud. Investorid eelistavad neid, kus on rohkem osavõtjaid (Lee, Lee 2012; Eunkyoungh *et al.* 2012).

Peale selle on riskiks ka platvorm (Milne, Parboteeah 2016). Kui äri ei ole jätkusuutlik võib platvorm oma tegevuse ootamatult lõpetada. Laenuperioodid on aga aastateks ja enamik platvorme on kindlustamata (Thakor 2020) ning neil puudub riiklik tagatis, mis on panganduslitsentside puhul tavapärane (Gaigalienė, Česnys 2018). Samas on mõned platvormid lisatasu eest ette näinud garantiifondid (Zhongming *et al.* 2019). Täiendavateks riskideks on veel ka õigusloome (nt seadusandja võib hakata nõudma platvormilt nt tagatisfondi, mis võib kujuneda üle jõu käivaks) (Milne, Parboteeah 2016) ja küberrünnak (Wei 2015).

1.2. Diskrimineerimine

Erialakirjanduses jaotub diskrimineerimine peamiselt kahte gruppi: (1) eelistustel põhinev (*taste-based*) ja (2) statistiline. Neist esimese liigituse autoriks oli Gary Becker 1957. aastal (Chen *et al.*

2019, viidatud Becker 1957). Tema teooria seisneb selles, kus vähemusgrupi töötajad peavad kompenseerima enamusgruppi, kas siis rohkem töötades või saades vähem palka, *ceteris paribus*. Teooria aluseks on näide, kus tööandja eelistab tööle võtta tema poolt soositud grupi liikmeid isegi siis, kui peab neile teistega võrreldes rohkem maksma. Selleks, et vähemusgrupp tema juurde üldse tööle saaks, peavad nad kas oluliselt vähem palka saama või oluliselt rohkem selle tarvis tööd tegema. Vastasel juhul see tööandja ei eelista viimasega kokku puutuda.

Panganduse kohta saab tuua vastava näite, kus Tšiili pankades ütlevad laenuhaldurid naiste poolt esitatud laenuaotlustele oluliselt rohkem ära, kui samadel tingimustel olevatele meeste taotlustele. Seejuures riiklik statistika on toonud välja, et naistel on laenude teenindamise osas vähem maksehäireid. Sellest tulenevalt teenivad pangad 10 protsenti vähem aastast kogutulu (Montoya *et al.* 2020). Kokkuvõttes on selle tagajärjeks otsene majanduslik kahju diskrimineerijale, mis on siis tingitud teatud grupi eelistamisest. St tehingu lõpule viimiseks peab ta tegema lisakulutusi eelisgrupi tarvis. (Chen *et al.* 2019, viidatud Becker 1957)

Statistiline diskrimineerimine lähtub vastupidiselt aga kuluefektiivsuse vaatenurgast. Selle liigituse autoriteks on Edmund Phelps (1972) ja Kenneth Arrows (1973) (Chen *et al.* 2019, viidatud Phelps 1972 ja Arrows 1973). Turul, kus on tegemist informatsiooni asümmeetriaga on odavam teha järeldusi mingi grupi kohta tervikuna, kui üksikult selle grupi liikmete kohta. Nt tööpakkuja varasemast kogemusest on jäänud mulje, et perekondlikel põhjustel puuduvad töölt rohkem naised. Järeldub, et efektiivsem oleks palgata meesterahvas.

Sellisel juhul on tegemist statistilise diskrimineerimisega, kus otsus tehakse mitte indiviidi põhised, vaid üldistatakse vastavalt grupikuuluvusele. Samal põhimõttel toimub see ka finantsinstitutsioonides. Nt eelarvamuse põhjal või sisestatistika on näidanud, et immigrantidel on rohkem maksehäireid. Sellest tulenevalt pank otsustab, kas laenuandmisest keelduda või teha seda kallimalt, kuna see rahvusgrupp on tema jaoks riskantsem. Kokkuvõttes on tegemist panga vaatevinklist kulupõhise ratsionaalse otsusega. (*Ibid.*)

1.2.1. Algoritmiline diskrimineerimine

Eelnevalt sai avatud, milliste teooriate alusel selgitatakse isikute vahelist diskrimineerimist. Kuna suur osa või kogu otsustusprotsess on laenuplatvormidel algoritmidel põhinev (vt peatükk 1.1.2.) on kohane selgitada diskrimineerimist ka sellest vaatenurgast lähtuvalt.

Inimeste hindamine nende ajalooliste saavutuste järgi on tavapärase. Finantsinnovatsiooni areng on muutnud väga lihtsaks andmete kogumise ja analüüsimise ning nende põhjal otsuste tegemise indiviidide üle krediitdireitingu määramiseks, tööpakkumiseks või nt reklaami esitamiseks. Seejuures saab ka algoritm teha diskrimineerivaid otsuseid. Selle põhjuseks võib olla programmeerija eelistused või viimasele antud kallutatud lähteülesanne. Neil juhtudel on tegemist algoritmi tööpõhimõtte suunamisega vastavalt kellegi otsusele, mille aluseks on eelistustel põhinev või statistiline diskrimineerimine.

Peale selle on aga võimalus, et algoritm ilma otsese käsuta võib jõuda iseseisvalt diskrimineerimiseni seetõttu, et ta eesmärgiks on seatud kuluefektiivsus või kasumi maksimeerimine (Bartlett *et al.* 2019; Fuster *et al.* 2021).⁹ Nimelt analüüsi aluseks olevad suurandmed (*big data*) võivad olla ajalooliselt juba eelneva diskrimineerimise tõttu kallutatud (Williams *et al.* 2018). Nt võib see olla olukord, kus teatud vähemusgrupp on läbivalt maksnud laenude eest kõrgemat intressi, saanud vähem laenu, periood on olnud lühem jms — tegemist on olnud kehvamate tingimustega. Algoritm tuvastab selle ja teeb samaväärsed otsused. Selle koostoimel tekibki (statistiline) algoritmiline diskrimineerimine.

Andmeanalüüside programmid nagu masinõpe ning tehisintellekt õpivad mustrite järgi ja teevad tulevikku vaatavaid otsuseid korrelatsioonide põhjal. St kui on piisavalt andmeid, mida keegi on teinud ja mida teised on teinud, saab ennustada, mida see keegi tõenäoliselt võib edaspidi teha. Kui indiviidi kohta kogu infot pole, on võimalik vaadata üldkogumit, et täita puuduolevad andmelüngad ja teha sellest järeldused. Sotsiaalse taustaga info nagu päritolu, sugu jms on olemas suurtes andmehulkades ja isegi juhul, kui seda ei ole spetsiaalselt kogutud (Williams *et al.* 2020). Seda võivad reeta nt nimed, veebiajalugu, tekstianalüüs (Caliskan *et al.* 2017), elukoht (Bertrand, Weill (2021) ja näotuvastus (Winter 2015). Vastavat seisukohta ilmestab alljärgnev uurimus.

Bertrand ja Weill (2021) analüüsisid USA laenuplatvormi Lending Club andmeid aastatel 2016 kuni 2017. Neis andmetes pole otsest infot etnilise päritolu kohta. Samas on teatud nimesid, mis eelkõige on just mustanahalistel¹⁰. Peale selle on linnaosaid, kus enamik elanikke on

⁹ Nimelt Lambrecht ja Tucker (2018) uurisid algoritmi käitumist inseneride töökuulutuse näitamisel. Olgugi, et see reklaam oli mõeldud ja disainitud kui sooneutraalne, näidati seda rohkem meestele. Põhjuseks oli asjaolu, et algoritm õppis selle reklaami varajases faasis selgeks, et selle näitamine on kallim naistele, kuna nad avavad seda harvem. Tulemuseks oli naiste diskrimineerimine sellele töökohale.

¹⁰ Töös kasutusel olevad uurimused on nimetanud päritolu valdavalt inglise keeles *black, white, hispanic, asian ja arabic*. Siinkirjutaja on võtnud nõuks lähtuda sarnasest (tõlke-)jaotusest eesti keeles, ilma laskumata süviti

mustanahalised. Neid linnaosasid on võimalik posti sihtnumbrite järgi eristada. Need olid mh andmed, millele oli algoritmil ligipääs. Autorid võtsid need muutujad oma uurimuse aluseks ja leidsid, et mustanahaliste laenuaotlustele öeldakse tihedamini ära, nad maksavad kõrgemat intressi ja nende laenudele antakse lühem tagasimakse periood. Nad leidsid, et tegemist võib olla just algoritmilise diskrimineerimisega.

Sarnasele tulemusele jõudsid ka Bartlett *et al.* (2019), kes uurisid laenuaotlusi, mis olid läbinud otsustusprotsessi (vaid) algoritmiliselt vs. vahetult (n-ö näost näkku kohtumise järgselt pangatöötaja poolt). Vaatluse all olid USA eralaenu perioodil 2009–2015 ja taotlejate päritolu. Selgus, et algoritmid ei diskrimineerinud laenuaotluste aktsepteerimisel (eeldatav põhjus kasumi maksimeerimisele suunatus). Kuid diskrimineerimine on jätkuv laenu intresside osas, olles küll kolmandiku võrra madalam, kui seda on vahetu kohtumise järgsete laenulepingute puhul. Andmetest selgus veel, et mõlemas vaatlusgrupis (algoritm vs. vahetu) diskrimineerimine langes ajas. Selle eeldatav põhjus on just erinevate laenuplatvormide turuletulek, mis on suurendanud konkurentsi.

Ülaltoodu põhjal saab asuda seisukohale, et isegi juhul, kui kogu otsustusprotsess on algoritmil põhinev, võib viimane teha diskrimineerivaid otsuseid.¹¹ Olenemata, kas otseselt on kogutud andmeid nt etnilise päritolu kohta jms. Tehisintellekt ja masinõpe on võimelised selle info välja selekteerima ajaloolise diskrimineerimise tulemusel. See viib omakorda statistilise algoritmilise diskrimineerimiseni. Juhindudes nii sellest, kui ka eelnevast peatükist (1.2.1.), saab asuda seisukohale, et diskrimineerimine võib leida aset nii isikute kui ka algoritmide poolt.

Pankadel on võimaliku maksejõuetuse tuvastamisel pikaajaline kogemus ja oskus (Diamond 1984). Selleks, et platvormid oleks antud küsimuses konkurentsivõimelised, on nad rohkem pannud rõhku sotsiaalsete andmete kasutamisele (Guo *et al.* 2016; Lyer *et al.* 2016; Barasinska, Schafer 2014). Peatükis 1.1.2. sai välja toodud, et see aitab vähendada informatsiooni asümmeetriat. Samuti oli käesolevas peatükis avatud diskrimineerimisel põhinevates uurimustes keskne koht sotsiaalsetel andmetel. Sotsiaalne taust on eelistustel põhineva ja statistilise

sotsiaalteaduslikele liigitustele. Võib argumenteerida, kas nt negriid, europiid ja mongoliid oleks sobivamad. Kuid siinkirjutaja arvamusel need ei kõla sihipäraselt eesti keeles ega ka ühti siinsete ingliskeelsete tõlgetega. Niisiis on kasutusel mustanahalised, valged, asiaadid, latiinod (*hispanic*) ning araablasted. Toodud tõlked on samuti olemas Eesti õigekeelsussõnaraamatus ja mingit halvustavat märkust pole nende kohta välja toodud.

¹¹ Eesmärgil, et algoritmid oleks vähem mõjutatud diskrimineerimiseni viivatest teguritest on loodud erinevaid ülemaailmse haardega ettevõtmisi nagu *Fairness, Accuracy, and Transparency in Machine Learning* (www.fatml.org) ja *Association for Computing Machinery – Code of Ethics* (<https://www.acm.org/code-of-ethics>).

diskrimineerimise lähtepunktiks. Alljärgnevalt esitatakse täpsem käsitlus neist andmegruppidest, mille puhul on tuvastatud diskrimineerimist finantssektoris.

1.2.2. Sugu

Mitmes Araabia ja Aafrika riigis on naistel keelatud erinevaid tegevusi teha ilma meessoost pereliikme nõusolekuta, sh pangakontot avada (World Bank 2018). Siinpool tunduvad need keelud kohatud. Samas ka n-ö lääneühiskonnas on traditsioonilises panganduses naisi koheldud ebavõrdselt. Nimelt USAs pangalaenude väljastamisel on naistelt nõutud kirjalikku kinnitust, et nad kasutaksid rasestumisvastaseid vahendeid, teeksid hüsterektomia (emaka eemaldamise operatsiooni) või katkestaks raseduse (Cyr 2011). Seda siis eesmärgil, et vältida tööturult puudumist ja sissetulekute vähenemist, mis omakorda raskendaks laenu teenindamist. See oli pigem levinud 70ndatel, kuid sellelaadset diskrimineerimist tuleb ette ka praegusel ajal (*Ibid.*).¹²

Üldiselt on leitud, et naistel on halvem positsioon finantsturgudele pääsemisel traditsioonilise panganduse kaudu (Baydas *et al.* 1994). Mh on välja toodud, et naised saavad laenu harvem (Asiedu *et al.* 2013), vähem (Agier, Szafarz 2013), see võib oleneda rohkem perekondlikust seisust ja töökogemusest (Chen *et al.* 2019), nad peavad pakkuma suuremat tagatist (Belucci *et al.* 2010), neilt nõutakse tihedamalt käendajat (Brock, de Haas 2020) ning neil on kõrgem intress (Muravyev *et al.* 2009; Alesina *et al.* 2013).

Pank määrab laenule intressi vastavalt riskile. Mida suurem risk seda kallim laen. Finantsmaailmas naiste ja meeste riskikäitumist uurinud tööd on toonud välja, et naised on riski osas ettevaatlikumad (Eckel, Fullbrunn 2015; Sunden, Surette 1998; Barber, Odean, 2001; Bernasek, Shwiff 2001), neil on kõrgem moraal (Bernardi, Arnold 1997; Croson, Gneezy 2009) ja nad panevad toime vähem pettusi (Cumming *et al.* 2015). Argumendiks on toodud, et naised plaanivad finantstehinguid rohkem tulevikku vaatavalt, kuna nende otsused võivad mõjutada enam nende perekonda, lapsi (Wiswall, Zafar 2018). St valdavalt naine laste juurest ära ei lähe n-ö uut elu alustama. Peale selle on leitud, et naistel on ka parem krediidi ajalugu (Alesina *et al.* 2013).

Niisiis, naistel on laenu saada keerukam ja kulukam. Samas on panga jaoks nende puhul tegemist vähemriskantsete klientidega. Sellest tulenevalt saab asuda seisukohale, et nn traditsioonilises

¹² Selle põhjuseks saab tuua asjaolu, et USAs puudub üleriigiline vanemahüvitis. Eestis on see pikim periood maailmas (Bryant 2020).

panganduses esineb soolist diskrimineerimist. Tõenäosus suureneb, et mehi ja naisi koheldakse võrdsemalt, kui piirkonna finantsturg on hästi välja arenenud (Asiedu *et al.* 2012). Nt on leitud, et finantseeringu saamisel ei kohelda Euroopa Liidu ettevõtete naisjuhte kehvemini (Moro *et al.* 2017).¹³ Siinkohal saab ka laenuplatvorme võtta kui finantsturu arengu osa, mis on andnud võimaluse selles osaleda oluliselt lihtsamalt ja suurendanud konkurentsi (Bartlett *et al.* 2019), kui seda on olnud panganduses. See on omakorda parendanud naiste võimalusi laenuteenuste kasutamisel.

Nimelt Chen *et al.* (2017) uurisid Hiina ühte suurematest laenuplatvormidest (*Ppdai.com*), kus on üle 100 miljoni registreeritud kasutaja. Nad leidsid, et naiste laenutaotlusi kiideti meeste omadest isegi rohkem heaks. Peale selle, uuringud Chen *et al.* (2019) poolt, mis keskendus viimase konkurendile Hiinas (*Renrendai.com* – registreeritud kasutajaid hinnanguliselt 50 miljonit), ja Duarte *et al.* (2015) poolt, mis põhines USA ühel turuliidril (*Prosper.com*), meeste ja naiste vahel taotluste heakskiitmise osas erinevust ei täheldanud. Ka Saksa turul jõudsid oma uurimusega samale tulemusele Barasinska, Schäfer (2014). Viimasele lisaks leidsid Weizsäcker ja Zankiewicz (2017), et erinevalt meestest, kiideti samaväärselt heaks nii kehvasid kui ka häid laenutaotlusi, mis olid esitatud naiste poolt. Seega saab asuda seisukohale, et selles osas diskrimineerimist laenuplatvormidel aset ei leia.

Samas ülaltoodud Chen *et al.* (2017) võrdlesid ka intressimäärasid ja leidsid, et naistele on see määratud kõrgem. Lisaks sellele leidsid Dietrich ja Wernli (2020) Šveitsi laenuplatvormi (*Cashare.ch*) andmete analüüsimise tulemusena, et naisterahvas, kes ei ole abielus ja tal on vähemalt üks laps alla 16 aasta, maksab 2,3 protsendipunkti kõrgemat intressi kui mehed. Seejuures ilma lasteta vallalisel oli intress 0,25 protsendipunkti meestest kõrgem. Siinkohal on ilmne, et tegemist on neil juhtudel soolise diskrimineerimisega, kuna *ceteris paribus* on laen naiste jaoks kallim.

Lisaks eeltoodule täheldas nii Chen *et al.* (2017) kui ka Chen *et al.* (2019), et naistel oli parem krediidikäitumine. St naistele antavate laenude puhul oli maksejõuetuse tõenäosus madalam —

¹³ Siinseks täiendavaks teguriks saab pidada ka võrdse kohtlemise printsiipi, mida juurutatakse järjepidevalt direktiivide tasandil. Paslik näide on kindlustusturgu reguleeriv direktiiv, mis keelab soo arvesse võtmise riskipreemia arvutamisel. Samas on leitud, et naised teevad autodega vähem õnnetusi ja neil on vähem tõsiseid terviseprobleeme. Seega mõneti irooniliselt, naised maksavad kõrgemat hinda suhtes oma riskitasemega. Olgugi, et printsiibi eesmärk oli diskrimineerimise vältimine, on see selle teises vormis hoopis tekitanud. (Petkantchin 2010)

suurem võimalik kasum ja väiksem võimalik kahju vastaspoolele. Samale järeldusele jõudsid oma uurimustöodes ka Lin *et al.* (2017), kus vaatluse all oli järgmine Hiina laenuplatvorm (*Yooli.com* – registreeritud kasutajaid 33 miljonit) ning Gaigalienė, Česnys (2018), kes uurisid Leedu laenuplatvormi (*Finbee.lt*). Ka siinkohal saab asuda seisukohale, et tegemist on soolise diskrimineerimisega. Nimelt Chen *et al.* (2019) argumenteerib, et naised peavad kompenseerima parema kediidikäitumisega laenuturul osalemist.

Ülaltoodu põhjal saab tuua välja, et laenuplatvormide turuletulek on eemaldanud naiste laenuaotluste heakskiitmisel esinenud diskrimineerimise, mis esines traditsioonilise panganduse puhul. Samas ebavõrdne kohtlemine võib olla jätkuvalt laenu intresside osas. Peale selle teenindavad naised paremini võetud laene, kuid see neile eelist meeste ees ei anna. Seega saab väita, et laenuplatvormide (algoritmiline) otsustusprotsess on vaatluse all olnud uurimustes jätkuvalt sooliselt diskrimineeriv. Soo kõrval on teiseks suureks diskrimineerimise uurimisobjektiks päritolu. Järgnevalt leiab aset selle küsimuse põhjalikum analüüs.

1.2.3. Päritolu

Siia alla kuulub erinevus prevalveerivast grupist rahvuse, etnilise kuuluvuse, rassi või nahavärvi alusel vaatluse all olevas paigas, kus vähemusgruppi koheldakse ebavõrdselt võrreldes enamusgrupiga. Eesti mõttes saaks rääkida nt setode, venekeelse elanikkonna või välismaalaste diskrimineerimisest.

Traditsioonilises panganduses on diskrimineerimist uuritud üksjagu. Ühendkuningriigis läbi viidud uurimus (Deku *et al.* 2015) vaatlus perioodi 2001–2009. Leiti, et võrreldes valgetega oli vähem võimalusi asiaatidel laenu saada ning mustanahalistel krediitkaarte. Kusjuures see lõhe suurenes majandusbuumi ajal, aastatel 2004–2007, kus laenude andmistingimusi küll tehti leebemaks, kuid eelkõige said sellest kasu just valged. Euroopas viidi läbi ka järgmine uurimus, kus Stefan *et al.* (2018) saatsid seitsme Euroopa riigi 1218 panka päringud (mille sisuks oli laenu taotlemise osas kontaktandmete küsimine). Kõigist kirjadest vastati valgete moodi kõlavatele nimedele 55.2% ja araabipäraselt kõlavatele nimedele 31.6%.

Samas võtmes saab USA turu kohta tuua järgneva. Quillian *et al.* (2020) uurimuse sisuks oli viimase 40 aasta (1976–2016) kohta läbi viidud 35 erineva teadustöö meta-analüüs. Need teadustööd uurisid kinnisvaralaene diskrimineerimise kontekstis. Autorid järeldasid, et diskrimineerimine on aastate jooksul olnud langustrendis, kuid siiski märkimisväärne.

Vähemusgruppidel on raskem laenu saada (öeldakse tihedamini ära) ja tingimused on kehvemad (kõrgem intress, väiksem laen, suurem tagatis, lühem periood) võrreldes *ceteris paribus* valgetega. Diskrimineerimine on üldjuhul kõrgeim mustanahaliste vastu, veidi madalam latiinode ning madalaim asiaatide vastu. Neist kahe uurimuse valimis olid ka araablased. Selgus, et ka neid diskrimineeritakse kinnisvara laenuturul palju.

Ülaltoodu põhjal saab asuda seisukohale, et panganduses on päritolu järgne diskrimineerimine kinnitust leidnud. Paraku on ka laenuplatvormide kohta tehtud uurimustes jõutud samale seisukohale. Peatükis 1.2.2. on toodud tulemused kahe uurimise (Bertrand, Weill 2021 ja Bartlett 2019) kohta, mis käsitlesid algoritmilist diskrimineerimist. Selgus, et ajalooliste andmete kallutatuse tõttu teeb ka algoritm inimeste päritolu järgi diskrimineerivaid otsuseid. Lisaks on Šveitsi laenuplatvormil (*Cashare.ch*) leitud (Dietrich, Wernli 2020), et need kellel on kodakondsus (omavad Šveitsi passi), maksavad oma laenudelt madalamat intressi. Samas nende kohustuste suhe sissetulekusse (*debt-to-income ratio*) on keskmiselt kõrgem, kui seda on seal elavate välismaalaste puhul.

Pope ja Sydnor (2011) viisid läbi uurimuse USA laenuplatvormil *Prosper.com*. Erinevalt ülalmainitud teadustöödest, kus diskrimineeriva otsuse tegi algoritm, on siinkohal tegemist indiviidide poolt tehtud valikutega. Nimelt vaatlusperioodil oli sellel platvormil üleval ka laenuaotleja pilt koos tutvustava tekstiga. Seega võimalikud laenuandjad said mh taotleja välimuse järgi otsustada, kas ja kellele anda laenu. Selgus, et mustanahalistel oli 25–35 protsenti madalam tõenäosus laenu saada ning positiivse vastuse korral oli neil 0,6–0,8 protsendipunkti kõrgem intress, võrreldes samaväärsete valgetega. Lisaks leiti, et mustanahaliste seas tuleb rohkem ette maksejõuetust.

Pope ja Sydnor argumenteerivad, et sellest tulenevalt võib tegemist olla eelistustel põhineva diskrimineerimisega või süstemaatilise eelarvamusega, kus võetakse arvesse mustanahaliste seas esinevat kõrget maksejõuetuse määra. Samale tulemusele jõudis *Prosper.com* uurimisel ka Ravina (2012): mustanahalistel on raskem laenu saada, neil on kõrgem intress ja rohkem tuleb ette maksejõuetust.

Kokkuvõtvalt on järelduseks, et nii panganduses üldiselt kui ka laenuplatvormidel esineb päritolu põhjal diskrimineerimist. Järgmisena on vaatluse all muud sotsiaalsed tunnused.

1.2.4. Muud sotsiaalsed tunnused

Diskrimineerimine võib aset leida veel ka puude, usutunnistuse, seksuaalse sättumuse, vanuse ja välimuse tõttu. Panganduses on selles kontekstis ette tulnud järgnevad juhtumid. Bank of America keeldus aastatel 2010–2016 laenude andmisest puudega inimestele, kellele oli määratud hooldaja (Department of Justice 2020). Sun ja Gao (2019) analüüsisid USA viimase kolme kümnendi kinnisvaralaenude andmeid. Nad jõudsid järeldusele, et mitte-heteroseksuaalsetele paaridel on 73 protsenti suurem tõenäosus saada laenutaotlusele eitav vastus. Kui nad saavad laenu siis nad maksavad 0,2 protsendipunkti kõrgemat intressi. Seejuures ei esine nende seas rohkem maksejõuetust, kui heteroseksuaalsete paaride seas.

Samuti on panganduses täheldatud diskrimineerimist usutunnistuse (Beck ja Brown, 2011) ja vanuse tõttu (Financial Ombudsman, 2015). Coen *et al.* (2021) analüüs tõi välja, et pangad diskrimineerivad kallimate laenudega just noori, tulenevalt nende vähesest finantskirjaoskusest.

Platvormide osas saab välja tuua järgnevad teadustööd. Laenuplatvormil *Prosper.com* oli investoritele nähtav ka laenutaotlejate fotod. Seda platvormi on uurinud selles võtmes järgnevad teadustööd. Selgus, et laenu saada oli raskem ülekaalulistel ja kes nägid välja õnnetumad (Pope, Sydnor 2011). Ravina (2012) leidis, et ilusamate inimeste laenutaotlusi kiideti rohkem heaks ja nad tasusid 0,6 protsendipunkti madalamat intressi. Duarte *et al.* (2012) tõi välja, et ka usaldusväärsema välimusega inimestel on suurem tõenäosus laenu saamiseks. Peale selle leidsid Duarte *et al.* (2015), et diskrimineeritakse lisaks ülekaalulistele ka noores eas inimesi.

Täiendavalt on platvorme uurinud töedes välja toodud haridus. Hiinas on leidnud nii Xu *et al.* (2018), Tao *et al.* (2017) kui ka Chen *et al.* (2020) *Renrendai.com* andmete põhjal, et kõrgharitud saavad kergemini laenu ja nad maksavad madalamat intressi. Seejuures, mida parema reputatsiooniga ülikool, seda väiksem maksejõuetuse tõenäosus (Li, Hu 2019).

Leidlikuma lähenemisega tööd on uurinud ka taotlejate tutvustustekste. Dorfleitner *et al.* (2016) analüüsis Euroopa kahe juhtiva laenuplatvormi taotlejate tekste. Autorid leidsid, et kirjavead, teksti pikkus ning sotsiaalsete ja emotsionaalsete sõnade kasutamine mõjutas negatiivselt laenu saamise tõenäosust, kuid ei mõjutanud maksejõuetuse tõenäosust. Lyer *et al.* (2016) tõi *Prosper.com* platvormi kohta välja, et tekstidele pööratakse rohkem tähelepanu, kui tegemist on madalama reitingu laenajaga.

Maailma suurimal laenuplatvormide turul, Hiinas, puudub maksehäire register ning puuduvad efektiivsed võimalused tsiviilasjades võlgade välja nõudmiseks (Ge *et al.* 2017). Selle tühimiku täitmiseks on seda enam kasutusel sotsiaalse taustaga andmed. Mh keskendutakse ka sotsiaalmeedia kontodele (Tao *et al.* 2017). Seda infot kontrollitakse ja kogutakse avalikult. Nimelt, eesmärgiks on laenajale teadvustada, et võla mitte tasumise korral võib ees oodata võlgniku maksejõuetuse staatuse avalikustamine. Selle tulemusel võib laenaja sõpruskond sellest teada saada. Seega rõhutakse sotsiaalsele häbistamisele (Ge *et al.* 2017).

Oluliseks teguriks on lisaks sotsiaalmeedia aktiivsusele ka sotsiaalne kapital. Siia alla kuuluvad sõprussuhted (Lin *et al.* 2013; Li, Hu 2019) või teatud grupi liikmelisus (Chen *et al.* 2009). Neis uurimistöodes on leitud, et mida suurem sotsiaalne kapital, seda suurem tõenäosus laenu saamiseks ja seda madalam intress ning neil tuleb ette vähem maksehäireid.

Sotsiaalse taustaga andmete kasutamine võib tekitada olukorra, mille puhul mingite parameetrite järgi üldiselt üksikule järelduste tegemine paneb teatud inividid halvemasse seisu. Nt ülaltoodud Hiina puhul saab asuda seisukohale, et sotsiaalmeedia järgi otsuseid tehes kannatavad need, kes seda kasutavad vähe või üldse mitte (Wang 2018). Sotsiaalmeedia andmete kasutamise kaitseks saab argumenteerida, et see on alternatiiv näost näkku kohtumisele (*Ibid.*).

Samas avalik konto koos piltide, postituste, sõpruskonna jms-ga avab selle isiku kohta oluliselt rohkem isiklikke detaile, kui seda oleks pangatöötaja juures intervjuul. See info võib tekitada aga eelarvamuslikke otsuseid, mille tagajärjeks on diskrimineerimine sotsiaalsete suhete iseärasuste tõttu. Edasi on vaatluse laenuaotleja majanduslik pool.

1.2.5. Finantsnäitajad

Siinkohal on vaieldav, kas laenuaotleja varaline seis saab olla diskrimineerimise aluseks. Ühelt poolt see näitab otseselt laenuaotleja võimet võetavate kohustuste kandmise osas. Madal sissetulek, palju kohustusi, vähe varasid või on need juba tagamas teisi laene, on pankade tavapärased riskihindamise alused.

Samas saab argumenteerida, et see on diskrimineeriv, kui võetavate kohustuste osas on taotleja sissetulek piisav. Seda enam laenuplatvormide osas, kus tegemist on üldjuhul väiksemate tarbimislauenudega. Peatükis 1.2.4. sai välja toodud, et immigrandid maksavad Šveitsis kõrgemat

intressi, kuid nende kohustuste suhe sissetulekusse on madalam, kui kohalikel (Dietrich, Wernli 2020).

Platvormi *Prosper.com* puhul olid paremad väljavaated laenu saamisel kinnitatud pangaväljavõtte esitajatel (Klafft 2008a). Hiina laenuplatvormi (*Renrendai.com*) andmete analüüsimisel leidsid Chen (2019) ja Ning (2013) mh, et suurem tõenäosus laenu saamiseks oli neil taotlejatel, kes omasid elamukinnisvara või kes elasid parema mainega piirkonnas (Hasan *et al.* 2017). Samuti oli *Ppdai.com* platvormi korduvlaenajatel lihtsam laenu saada ja kohaldus madalam intress, kui seda oli esmakordsetel (Cai *et al.* 2016). Šveitsis (*Cashare.ch*) oli majaomanikel madalam intress (Dietrich, Wernli 2020). Lisaks analüüsisid Guo *et al.* (2016) USA platvormide *Lendingclub.com* ja *Prosper.com* andmeid. Nad jõudsid järeldusele, et laenajatel, kes elavad üüripinnal, tuleb rohkem ette maksejõuetust.

Sellega on võimalikud diskrimineerimist puudutavad näitajad avatud. Järgmine peatükk käsitleb uurimismetoodikat ja kasutatavaid andmeid.

2. ANDMED JA METOODIKA

2.1. Andmed

2.1.1. Valim

Uurimus viiakse läbi kasutades laenuplatvormi Bondora andmeid. Bondora kodulehe (www.bondora.ee) põhjal selgub, et tegemist on Eestis asutatud ettevõttega, mis tegutseb alates 2008. aastast. Lisaks Eestile pakub ta sama teenust veel Soomes ja Hispaanias ning on seda teinud ka Slovakkias. Samas Eesti turul on olnud kauem ja see on ka suurim, moodustades 62% laenudest. Laene on vahendatud 2021. aasta septembrikuu lõpuks 482 miljoni euro ulatuses.

Bondora veebilehelt on 7. juulil alla laetud Exceli fail nimega *LoanDataset.csv*, millel asub iga väljastatud laenu kohta detailne informatsioon 112 teguriga (samal peab märkima, et paljudel juhtudel ei ole kõik väljad täidetud). Mh on välja toodud sellised andmed nagu sugu, vanus, rahvus, perekonnaseis, haridus, tööstaaž ja kinnisvara olemasolu – teoreetilises osas selgus, et need tegurid võivad olla diskrimineerimise objektiks finantsteenuste pakkumisel. Laenuaotlusi on selleks kuupäevaks tabelisse märgitud 177 275.

Andmete vaatlusel selgus, et välisriikide puhul esineb märksa rohkem puudujääke. Nt Hispaanias pole poolte andmete puhul sugu avaldatud, mis on aga käesoleva uurimuse üheks oluliseks muutujaks võimaliku diskrimineerimise tuvastamisel. Seega tulenevalt välisriikide laenuaotlejate väiksemast hulgast ja puudulikest andmetest on autor otsustanud keskenduda vaid Eestis väljastatud laenu analüüsile. Võttes arvesse Eesti rahvaarvu ja laenu suurte hulka, saab seda põhjapanevamad uurimustulemused järelduste tegemiseks.

Lisaks selgus, et Bondora on oma tegevusajaloo jooksul täiendanud reitingumääramise mudelit/algoritmi. Tegemist on olnud seitsme erineva versiooniga, millest algseid on kasutatud mõnel juhul isegi alla aasta ja läbisegi teistega. Siinkirjutaja on arvamisel, et Bondora on otsinud ja katsetanud erinevaid versioone praktikas ning leidnud, et eelnevad ei ole eesmärki täitnud. Neis mudelites võisid olla sees erinevad vead ning nende tulemuste omavaheline võrdlemine käesolevas

töös ei pruugi anda adekvaatset ja ühest seisukohta. Seevastu viimane versioon on kasutusel 2018. aastast. Sellest võib järeldada, et Bondora on leidnud parima algoritmilise mudeli reitingute määramiseks. Seega alates viimasest versioonist on reiting määratud samadel alustel. Järeldub, et need tulemused on omavahel võrreldavad. Sellest tulenevalt on autor otsustanud võtta aluseks just seda perioodi hõlmavad andmed.

Peale välisriikide ning muude puudulike andmeridade eemaldamist (nt oli mõni oluline lahter tühjaks jäetud või selle väärtus oli -1 , mille kohta puudus Bondoral selgitav kriteerium) ja keskendudes vaid viimasele reitingumääramise versioonile on valimiks 69 902 heaks kiidetud laenuaotlust.

Alljärgnevas alapeatükis (2.1.2.) on esitatud uuringu jaoks oluliste muutujate kohta arutelu ja selgitused nende kohandamise kohta regressioonanalüüsi läbiviimiseks. Kõigi analüüsis kasutatavate muutujate kohta on kirjeldav statistika esitatud kokkuvõtva tabelina (vt Tabel 1.) järgneva alapeatüki (2.1.2.) lõppu.

2.1.2. Muutujad ja kirjeldav statistika

Sõltuv muutuja on **intress**. Eeltoodud teadustöödele tuginedes (Hasan *et al.* 2017; Freedman, Jin 2017; Xu *et al.* 2018; Dietrich, Wiernli 2020; Duarte *et al.* 2012; Tao *et al.* 2017) saab asuda seisukohale, et diskrimineerimise tuvastamiseks finantsteenuste pakkumisel on see tegur Bondora andmeid arvesse võttes parim valik. Seeläbi saab hinnata, kas vähemusgrupile on määratud kõrgem intress.

Üksjagu töid hindas seda ka ära öeldud laenuaotluste põhjal (Chen *et al.* 2019; Gaigalienė, Česnys 2018; Tao *et al.* 2017; Freedman, Jin 2017; Cai *et al.* 2016). Paraku sellised andmed Bondora puhul puuduvad. Peale selle on uuritud ka maksejõuetuse tõenäosuse järgi (Barasinska, Schafer 2014; Hasan *et al.* 2017; Tao *et al.* 2017; Duarte *et al.* 2012; Freedman, Jin 2017; Lin *et al.* 2013; Weizsäcker, Zankiewicz 2017), kuid lõputöö autor on arvamusel, et intressi järgi on võimalik teha diskrimineerimise kohta sihipärasemaid järeldusi. Vastupidisel juhul, kus aluseks võtta maksejõuetus, saab seda tuletada kaudsemalt. Nimelt eeltoodud tööd leidsid, et naistel tuleb ette vähem maksejõuetust. St et tegemist on vähemriskantsete klientidega. Ainuüksi see nüanss ei anna teavet diskrimineerimise kohta. Seda on vaja kõrvutada ära öeldud laenuotsuste või intressimääraga. Seega diskrimineerimisega pole tegemist, kui võrreldes meestega nende taotlusi kiidetakse samas hulgas heaks ja/või neil on samaväärne intress. Paraku esimese kohta andmed

puuduvad, millest tulenevalt taandub küsimus siinkohal vaid intressile. Sel põhjusel peab lõputöö autor otseselt viimase uurimist paremaks valikuks, kuna järelduse alus (kas intressimäär on gruppide vahel erinevusi) on sama.

Sõltumatud muutujad. Diskrimineerimise uurimisel on kesksed tegurid sugu, vanus ja rahvus. Eespool välja toodud uurimustes selgus, et finantsteenuste pakkumisel, seda nii traditsioonilises panganduses kui ka laenuplatvormidel, on leidnud aset diskrimineerimist nende kolme teguri puhul. Alljärgnevalt on välja toodud arutelu nende kolme muutuja kohandamise kohta analüüsi läbiviimiseks ja püstitatud vastavad hüpoteesid teoreetilise raamistiku baasilt. Samuti on esitatud muud asjakohased tegurid diskrimineerimise kontekstis ning kontrollmuutujad.

Sugu. Naised moodustavad 40% valimist. Kuna tegemist on kategoorilise muutujaga, siis on need näitajad väljendatud indikaatoritena, kus mehe esitatud laenuaotlust tähistab null ja naise esitatud laenuaotlust tähistab üks. Hüpotees: naistele on määratud kõrgem intress (põhinedes alapeatüki 1.2.2. arutelul).

Vähemusrahvus.¹⁴ Otseselt pole rahvust Bondora andmetes välja toodud. Samas on olemas taotluse esitaja keele valik, mida saab teha peale eesti ka vene, inglise, soome, hispaania, slovakkia ja saksa keeles. Seega vähemuseks on arvestatud muukeelne laenuaotlus. Kokku on Eesti andmete osas 22% muukeelset laenuaotlust. Andmed on väljendatud indikaatoritena, kus eestikeelne laenuaotlus on tähistatud väärtusega null ja muukeelne laenuaotlus väärtusega üks. Hüpotees: vähemusrahvusele on määratud kõrgem intress (vastavalt alapeatüki 1.2.3. arutelule).

Vanus. Laenuvõtjate vanusevahemik on 18–70 aastat. Võttes arvesse käsitletud uuringuid (Duarte *et al.* 2015; Coen *et al.* 2021), kus noori diskrimineeriti kallimate laenudega, on siinkohal uurimusfookus sama. Noored (*youth; young adult*) on defineeritud vastavalt ÜRO määratlusele, s.o 18–24 aastat (United Nations 2007). Noori on valimis kokku 13%. Nn vanema grupi mediaanvanuse on 39 aastat. Andmed on indekseeritud vastavalt noored üks ja ülejäänud null. Hüpotees: nooremad tasuvad kõrgemat intressi.

Haridus. Teoreetilises osas juhiti tähelepanu, et kõrgharidusega laenuaotlejatele määrati madalam intress. Laenuandja seisukohalt võib olla see üldiseks kriteeriumiks, et sellisel inimesel on suurema

¹⁴ Teoreetilises osas nimetatud ka kui päritolu järgi inimeste/gruppide liigitamine.

sissetulekuga töökoht või on tal töö otsimisel konkurentsieelis ehk lihtsam tööd leida. Samuti võib anda see aimu sellise inimese suuremast kohusetundlikkusest, oskusest kriitilisemalt mõelda, asju pikemalt ette planeerida ja suuremast finantskirjaoskusest. Samas, olenevalt laenuaotleja asukoha tööhõive võimalustest, võib teatud oskustöö olla tasuvam kui n-ö kontoritöö ning kohusetundlikkus jms võib olla väga subjektiivne. Kokkuvõtteks peab märkima, et seda tegurit on keeruline diskrimineerimise võtmes hinnata. Samas vähemalt mudeli parendamise mõttes ja kontrollmuutujana on seda põhjendatud kasutada.

Bondora andmetes on haridus jaotatud viite eri rühma, algharidusest kõrghariduseni. Eesmärgiga kontrollida sama hüpoteesi, nagu eelnevas lõigus selgitatud, on moodustatud kõrgharidusest eraldi rühm ning ülejäänud on kokku pandud. Selle tulemusena on valimis 18% kõrgharidusega laenajat. Näitajad on indekseeritud vastavalt kõrgharidusega null ja vastaspool üks. Tulemuse eeldatav märk: pluss (kõrgharitul madalam intress).

Kinnisvara. Mitmed uurimused töid välja, et selle olemasolu on tähendanud madalamat intressi. Seda ei saa pidada otseselt diskrimineerivaks teguriks. Eelkõige annab see aimu laenuaotleja finantsolukorrast ja võimest täita võetud kohustusi, mis on laenuandmise juures riski maandamise puhul olulised asjaolud. Teisalt, *ceteris paribus*, kus ilma kinnisvarata inimese Bondorast võetava laenu teenindamise võime on hoopis parem, nt muude pikaajaliste laenude puudumise tõttu (nagu kodulaen), saab argumenteerida, et viimast selles osas diskrimineeritakse. Samas tuleb tõdeda, et teadmata kohustuste suhet sissetulekusse (kohustuste kohta selgitus allpool), pole võimalik seda tegurit diskrimineerimise võtmes hinnata. Mudeli parendamise mõttes ja kontrollmuutujana on see analüüsi lisatud.

Bondora andmed on jaotatud 11-nesse eri rühma, alates kodutust kuni koduomanikuni välja. Siinkohal on eesmärk võrrelda kinnisvara omamist selle puudumisega. Sellest tulenevalt on ühte rühma pandud omanik, kaasomanik ja hüpoteegiga omanik (st et võetud on kodulaen) ning kõik muu (nagu üürimine, vanematekodu jms) on pandud teise rühma. Koduomanikke on 51%. Andmed on indekseeritud vastavalt koduomanik null ja vastaspool üks. Tulemuse eeldatav märk: pluss (koduomanikul madalam intress).

Tööstaaž. Sarnaselt hariduse ja kinnisvara kohta toodud argumentatsioonile saab põhjendada ka tööstaaži kaasamist uuringusse. Nimelt kauem töötanud inimesed evivad endas eeldatavalt suuremat distsiplineeritust, kindlamat ja suuremat sissetulekut, mis on laenu teenindamise puhul

olulised tegurid laenuandjale. Selle alusel otsuste tegemist diskrimineerivaks pidada on muidugi keerukas, kuna sissetuleku stabiilsus on igasuguse krediidilepingu sõlmimise aluseks. Kuna vastav tegur on aga Bondoras välja toodud, kasutab autor seda võimalikuks mudeli parendamiseks ning kui kontrollmuutujat.

Bondora andmetes on välja toodud ka töötud, katseaeg jms kuni tööstaažini rohkem kui viis aastat. Siinkirjutaja otsustas grupeerida näitajad nii, et oleks võrreldavad kindlama ja eeldatavalt kõrgema sissetulekuga inimesi töötutest ja äsja tööle asunutest (töötanud kuni aasta viimasel töökohal). Nn ebakindlama tööstaažiga laenutaotlejaid on 28% valimist. Väärtused on indekseeritud vastavalt pikem tööstaaž null ja vastaspool üks. Tulemuse eeldatav märk: pluss (vähese tööstaaži puhul kõrgem intress).

Ülejäänud uuringusse kaasatavate tegurite puhul on tegemist otseselt konkreetse laenu või selle taotleja finantsnäitajatega. Need täidavad otseselt kontrollmuutujate funktsiooni ning võivad olla olulised tegurid mudeli seletusvõime parendamiseks.

Laenu pikkus on välja antud laenudel vahemikus 1 kuni 60 kuud, millest mediaan 60 kuud. Väärtused on logaritmitud. Tulemuse eeldatav märk: miinus (mida pikem laenuperiood, seda madalam intress). **Laenu suurus** on vahemikus 100 kuni 10 632 eurot, kus mediaan on 1594 eurot. Väärtused on logaritmitud. Tulemuse eeldatav märk: miinus (mida suurem laen, seda madalam intress). **Esmakordne laenaja** – laenuandja seisukohalt on laenaja, kellel on eelnev positiivne laenu teenindamise ajalugu, madalama riskiga. Seega annab isiku Bondora laenuajalugu hea indikatsiooni võimaliku riskitaseme kohta. Esmakordseid laenajaid on andmetes kolmandik. Korduvlaenaja on indekseeritud väärtusega null ja esmakordne taotleja väärtusega üks. Tulemuse eeldatav märk: pluss (esmakordsel laenajal on kõrgem intress). **Sissetulek** – eeldades, et kõrgema sissetulekuga laenajal on madalam intress tulenevalt tema paremast finantsvõimekusest. Väärtused on logaritmitud. Tulemuse eeldatav märk: pluss.

Laenureiting. See on Bondora poolt väljastatud reiting vahemikus AA kuni HR, kus viimasega kaasneb eeldatavalt suurim risk (HR tähendab *high risk*). Kuidas Bondora täpselt oma reitingut välja arvutab, pole paraku teada. Kodulehel on öeldud, et selle tarvis on kaasatud üle saja erineva teguri. Eeldatavalt tegemist sarnase lähenemisega, mis on kasutusel paljude laenuplatvormide puhul (vt alapeatükki 1.1.2.). Välja on töötatud algoritm, mis suudab kogu andmebaasi alusel teha otsuseid indiviidi kohta. Võib arvata, et nende üle saja kaasatava teguri seas on ka sugu, vanus ja

rahvus. Samas, kui reitingu osakaal saab olema regressioonianalüüsi mudeli seletusvõimel kaalukas ning selgub, et toimub soo, vanuse ja rahvuse järgi diskrimineerimine (st neile määratud intress on kõrgem), saab asuda seisukohale, et ka Bondora reitingu algoritmiline määramisprotsess võib olla diskrimineeriv.

Reitingud on Bondoral madalamast suurema riskini vastavalt AA, A, B, C, D, E, F ja HR. Mediaan näitajaks on reiting C. Reitingute väärtused on indekseeritud kõrgemast madalamani vastavalt üks kuni kaheksa. Tulemuse eeldatav märk: pluss (mida madalam reiting, seda kõrgem intress).

Siinkohal tuleb rõhutada, et pole teada palju tegelikult mõjutab intressi nt reiting AA võrreldes ülejäänud reitingutega. Seega see võrdväärne suhe ühest kuni kaheksani on pigem pseudosuhe, mitte otseselt täpne. Samas on tegemist vaid kontrollmuutujaga, mille eesmärk ei ole välja selgitada palju mõjutab intressi iga erinev reiting. Sarnast indekseerimist on kasutanud enamik käeoleva uurimuse aluseks olnud teadustööd (nt Chen *et al.* 2019; Cai *et al.* 2016).

Alljärgnevalt muutujatest, mis oleks olnud asjakohased, kuid puudulike andmete tõttu pole võimalik arvutustesse kaasata. Esiteks perekondlik seis ja ülalpeetavate hulk. Võttes arvesse teoreetilise osas välja toodut, et naisi võidakse nende perekondlikust seisust ja laste hulgast tulenevalt diskrimineerida (Dietrich, Wernli 2020), on siinkirjutaja arvamusel, et ka see tegur oleks põhjendatud uurimusse kaasata. Paraku on korrektsed väärtused märgitud vaid kuuendikul valimireal. Ülejäänutel on väärtuseks märgitud miinus üks (-1), mille tähenduse kohta puudub Bondora lehel selgitus. Sellest tulenevalt on otsustatud need muutujad analüüsist välja jätta ja arvutused läbi viia valimit kitsendamata vaid kuuendikule kogu kõne all olevast mahust.

Teiseks laenu suhe sissetulekusse. See näitab laenuaotleja võimet tasuda võetavat kohustust. Samas andmete uurimisel selgus, et sarnaselt perekondlikule seisule ja ülalpeetavate hulga on sarnane hulk andmeid puudu ka selle muutuja kohta. Täpsemalt on väärtuseks neil juhtudel märgitud null. See on siinkirjutajas kahtlust tekitav. Nimelt teoreetilises osas sai välja toodud, et eelkõige on selle teenuse kliendid inimesed, kes mingil põhjusel tavapangast laenu ei saa. Eeldatavalt ei saa laenu need, kellel on panga poolt ette nähtud sissetulekute suhe kohustustesse ületatud. Kuidas aga enamik laenuplatvormi klientidest on ilma igasuguste kohustusteta, on arusaamatust tekitav. Pigem võib olla tegu selliste andmete kogumata või märkimata jätmisega. Sellest tulenevalt ei ole selle muutuja kaasamine arvutustesse põhjendatud, kuna eeldatavalt on tegemist vigaste andmetega.

Alljärgnevalt on esitatud tabel (Tabel 1.) kõigi analüüsis kasutatavate tegurite kohta.

Tabel 1. Kirjeldav statistika

Muutuja		Valimist osa (%)	Kesk- väärtus	Kesk- mine	Sandard- hälve	Miini- mum	Maksi- mum
Intress (protsenti)		100%	24,4	26,1	10,9	7,27	48,8
Sugu	mees	60%	–	–	–	–	–
	naine	40%	–	–	–	–	–
Päritolu	eestlane	78%	–	–	–	–	–
	vähemusrahvus	22%	–	–	–	–	–
Vanus (aastat)	noor	13%	22	22	2	18	24
	vana	87%	39	41	11	25	70
Haridus	kõrgharidus	18%	–	–	–	–	–
	muu	82%	–	–	–	–	–
Kinnisvara	koduomanik	51%	–	–	–	–	–
	muu	49%	–	–	–	–	–
Tööstaaž	üle aasta	72%	–	–	–	–	–
	muu	28%	–	–	–	–	–
Laenu pikkus (kuudes)		100%	60	48	15	3	60
Laenu suurus (eurot)		100%	1594	2462	2292	112	10632
Esma- või korduvlaenaja	esmakordne	33%	–	–	–	–	–
	korduv	67%	–	–	–	–	–
Sissetulek (eurot kuus)		100%	1000	1168	601,8	0,26	4950
Laenureiting	AA	6%	–	–	–	–	–
	A	6%	–	–	–	–	–
	B	20%	–	–	–	–	–
	C	25%	–	–	–	–	–
	D	25%	–	–	–	–	–
	F	18%	–	–	–	–	–
	E	0,1%	–	–	–	–	–
	HR	0,01%	–	–	–	–	–

Allikas: Bondora andmed; autori arvutused

2.2. Metoodika

Teoreetilises osas sai käsitletud suur hulk teadustöid, mis analüüsisid diskrimineerimist finantsteenuste pakkumisel. Otseselt laenuplatvorme uurinud tööde regressioonanalüüside mudelid saab jagada kolme rühma. Neist esimene keskendub riskantsemate laenajate kindlaks tegemisele või nende ette ennustamisele. Analüüside läbiviimiseks on kasutatud peamiselt **Probit** (Tao *et al.* 2017; Freedman, Jin 2017) või **Logit** (Chen *et al.* 2019; Gaigalienė, Česnys 2018; Cai *et al.* 2016;) regressioonanalüüsi, kus sõltuvaks muutujaks on määratud maksejõuetus (*default*) ning hinnatud selle tõenäosust.

Järgmiseks rühmaks saab pidada töid, mis on hinnanud, millistel gruppidel on väiksem tõenäosus laenu saada. Sõltuvaks muutujaks on laenu taotluse otsus. Analüüside läbiviimiseks on kasutatud samuti **Probit** (Barasinska, Schafer 2014; Hasan *et al.* 2017; Tao *et al.* 2017; Duarte *et al.* 2012; Freedman, Jin 2017; Lin *et al.* 2013) või **Logit** (Weizsäcker, Zankiewicz 2017) regressioone. Bondora andmetes on välja toodud vaid väljastatud laenude andmed. Seega see lähenemine langeb ära.

Viimaseks rühmaks on intressimäära alusel diskrimineerimist uurinud tööd. Analüüsiks on kasutatud peamiselt vähimruutude (**OLS** – *ordinary least squares*) (Hasan *et al.* 2017; Freedman, Jin 2017; Xu *et al.* 2018) või **Tobit** (Duarte *et al.* 2012; Tao *et al.* 2017) regressioone.

Kuna Bondora andmetes puuduvad ära öeldud laenu taotlused ning maksejõuetuse puhul taandub diskrimineerimise tuvastamine intressile, on käesoleva uurimuse sõltuvaks muutujaks määratud intress (vt vastavat arutelu peatüki alguses). Lähtudes ülaltoodud teadustöödest on regressioonmudelite valik OLS ja Tobit vahel. Viimase puhul on tegemist mudeliga, mille eesmärk on kõrvaldada probleeme tulenevalt puudulikest andmetest või kui sõltuva muutuja väärtuseks on kohati null (Gujarati 1995). Autor on andmed kohandanud juba vastavalt, kus puudulikud read või arusaamatu koodiga näitajad (nagu miinus üks) on kõrvaldatud. Kuna tegemist on intressiga, mis saab laenu väljastamise puhul olla vaid positiivne ja nullist alati suurem, siis pole ka sel juhul Tobit mudeli kasutamine põhjendatud. Sellele tuginedes on autor otsustanud uurimuse läbi viia OLS mudelit kasutades.¹⁵

¹⁵ Analüüsi läbiviimiseks on kasutatud Gretl ökonomeetria programmi, mis on vabavarana allalaetav Gretli veebilehelt (www.gretl.sourceforge.net).

3. MUDEL JA JÄRELDUSED

3.1. Mudel

Alltoodud tabelis (Tabel 2.) on selguse huvides esitatud veelkord sõltumatud muutujad. Juhul kui regressioonanalüüsis (täpsemalt Gretli programmis) kasutatav nimetus erineb tegelikust siis on see märgitud selle järgi sulgudes. Tabeli parempoolses tulbas on märgitud regressioonanalüüsi tulemuse eeldatav märk vastavalt teoreetilises käsitluses toodule.

Tabel 2. Sõltumatud muutujad

Muutuja	Eeldatav märk
Sugu	Pluss
Vähemusrahvus (Vahemusrahvus)	Pluss
Vanus	Pluss
Haridus	Pluss
Kinnisvara	Pluss
Tööstaaž (Toostaaz)	Pluss
Laenu pikkus (l_LaenuPikkus)	Miinus
Laenu suurus (l_LaenuSuurus)	Miinus
Laenureiting	Pluss
Esmakordne laenaja (EsmakLaenja)	Pluss
Sissetulek (l_Sissetulek)	Miinus

3.1.1. Mudeli testimine

OLS mudeli puhul on tegemist mitmese lineaarse regressiooniga, kus sobivate kordajate leidmiseks kasutatakse vähimruutude meetodit.¹⁶ Regressioonanalüüsi aluseks on ristanometel põhinevad arvutused. Mudel on esitatud kujul:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 X_i + \dots + \beta_n X_i + u_i,$$

kus Y_i on laenu i intress (sõltuv muutuja), α on vabaliige, β on vastava muutuja parameeter (regressioonikordaja), X_i laenu i seletav tunnus (sõltumatu muutuja) ja u_i on laenu i jääkliige. Kordaja β näitab seega, kui X suureneb ühe ühiku võrra siis kõikidel juhtudel suureneb Y väärtus keskmiselt β võrra. Kõikidel juhtudel tähendab, et analüüs võtab arvesse muutust üle i , tähendades, et hindamine toimub gruppide vaheliselt ja gruppide siseselt.

Esmalt on läbi viidud korrelatsioon-maatriksi analüüs. Siinkohal on vaatluse all seletatavate muutujate omavaheline korrelatsioon, kus muutujate paaride kohta on arvutatud paariskorrelatsiooni kordajad. Vastav näitaja ei tohiks olla üle 0,85 (Vainu 2006). Vastasel juhul võib tegemist olla multikollineaarsusega.

Kõrgeimaks korrelatsioonikoefitsiendiks osutus 0,29 (vt Lisa 1.). Saab asuda seisukohale, et multikollineaarsust ei esine. Seega võib kõik seletatavad tunnused regressioonanalüüsi kaasata. (Allpool leiab mudeli testimise käigus aset ka dispersiooni inflatsiooniteguri (VIF) järgi multikollineaarsuse kontrollimine.)

Esialgmesse mudelisse on kaasatud kõik sõltumatud muutujad. Esmase regressioonanalüüsi läbiviimisel selgus, et kõik kaasatud muutujad on statistiliselt olulised nivool 0,01. Seega saab kõigi tunnuste parameetrite osas vastu võtta sisuka hüpoteesi: statistiliselt olulised nivool 0,01. Seejuures mudeli determinatsioonikordaja on 0,815. Samuti saab vastu võtta sisuka hüpoteesi, et mudel on statistiliselt oluline: P väärtus allub olulisuse nivoole 0,01. Lisaks saab analüüsi lülitatud kontrollmuutujate põhjal asuda seisukohale, et ka märgid tunduvad loogilised (vt Tabel 2.; Lisa 2.). Samas neist üks, sissetulek, andis vastuseks vastupidise märgi (vastav arutelu peatükis 3.2.).

¹⁶ Vähimruutude meetodi puhul tuleb leida selline sirge võrrand, kus empiiriliste punktide ja sirge punktide vaheliste kauguste ruutude summa on minimaalne (Sauga 2005).

Heteroskedastiivsuse kõrvaldamiseks on kasutatud Gretlis toodud arvutust *Heteroskedasticity-corrected*. See põhineb White'i (1980) poolt välja toodud meetodikal. Nimelt heteroskedastiivsuse puhul on koefitsendid korrektsed. Probleem on aga standardvigadega, mis sel juhul muutuvad valeks. Teisisõnu, see ei mõjuta mitte parameetrite hinnangut, vaid parameetrite usalduspiire. Heteroskedastiivsuse korrastamise arvutuse teel korrigeeritakse standardvead, mille tulemusena on hüpoteeside testimine võimalik ja usaldusnivoo paikapidav. (Vt Lisa 2.)

Multikollineaarsust on lisaks korrelatsioonanalüüsile kontrollitud dispersiooni inflatsiooniteguriga (VIF). Juhul, kui VIF analüüsi tulemusena on mõne tunnuse näitaja kümnest suurem, võib tegemist olla multikollineaarsusega. Selgus, et multikollineaarsust ei esine. (Vt Lisa 4.)

Jääkide kontrollimine andis vastuseks, et need ei allu normaaljaotusele. Samas, Lumley *et al.* (2002) on välja toonud, et piisavalt suure valimi juures kehtib tsentraal piirteoreem, mille korral vähimruutude regressiooni koefitsendid on normaalselt jaotatud ja t -testi p väärtused tulevad adekvaatsed. Selles uurimuses oli piisavalt suureks valimiks ligemale viissada. Lisaks sedastavad Knief ja Forstmeier (2021), kui lineaarse regressiooni muud eeldused on täidetud siis p väärtuste hindamist võib pidada (peaaegu alati) usaldusväärseks ka andmete puhul, mis ei allu normaaljaotusele. Nad leiavad, et andmete kohandamine või erinevate muude meetodite kasutamine normaaljaotuse saavutamiseks võib olla pigem kahjulikum. Tuginedes siinsele arutelule ja käesoleva uurimuse valimi suurusele (ligemale 70 tuhat), saab asuda seisukohale, et andmete allumine normaaljaotusele ei ole siinkohal olulise kaaluga.

Võttes arvesse, et muutuja laenureiting on Bondora intressimääramisel peamine tegur, mis sisuliselt peaks hõlmama kõiki käesolevas analüüsis kasutatavaid muutujaid, on paslik testida mudelit ka ilma selleta. Võib oletada, et reitingul on sedavõrd suur kaal regressioonanalüüsis, mille tulemusel see mõjutab oluliselt kõigi muutujate statistilist olulisust. Sellisel juhul, kus suur osa muutujaid ei oleks enam statistiliselt olulised, seaks see kahtluse alla tulemuste adekvaatsuse mõlema mudeli puhul.

Peale kitsendatud mudeli (vt Lisa 3.) testimist ilma muutujata laenureiting selgus, et kõik muutujad on jätkuvalt statistiliselt olulised nivool 0,01. Samuti on mudel statistiliselt oluline ja vastu saab võtta sisuka hüpoteesi. Märkide osas on erinevus vaid kontrollmuutujas laenu periood, kus tulemus on positiivne negatiivse asemel. Samas on mõne muutuja puhul muutunud suurusjärgud. Nt

muutujate vanus ja haridus puhul on siinkohal tulemuseks üle kahe prostendipunkti, kuid algse mudeli puhul oli see alla ühe.

Eeltoodud tulemusi kitsendatud mudeli osas saab pidada hindamiseks adekvaatseteks, kuid vastaka seisukoha aluseks on mudeli determinatsioonikordaja: R^2 on vaid 0,12. Seevastu esialgse mudeli korral oli 0,81. Seega mudeli seletusvõime halvenes oluliselt kitsendatud mudeli korral. St et mudelist on puudu oluline muutuja laenureiting. Seega pole mudeli kitsendamine põhjendatud.

3.1.2. Lõplik mudel

Alljärgnevalt on esitatud lõpliku mudeli tulemused võimaliku diskrimineerimise tuvastamiseks Bondora andmete põhjal (Tabel 3.). Valimi suuruseks on 69 902 ning mudeli seletusvõime R^2 on 0,815 (vt Lisa 2.). Siinkohal kokkuvõtvalt tulemused: naistel, vähemusrahvuse esindajatel, nooremaealistel, kõrghariduseta laenajatel, vähe töötanutel, esmakordsetel laenajate, madalama sissetuleku ja halvema reitingu korral on kõrgem intress ning pikemal laenuperioodil ja suuremal laenusummal on madalam intress. Järgnevas peatükis (3.2.) leiab aset nende tulemuste tõlgendamine ning vastav arutelu.

Tabel 3. Tulemused lõpliku mudeli põhjal

Muutuja	Tulemus
Konstant	-6,210 *** (0,275)
Sugu	1,170 *** (0,032)
Vähemusrahvus	0,334 *** (0,037)
Vanus	0,400 *** (0,052)
Haridus	0,389 *** (0,039)
Kinnisvara	0,329 *** (0,032)
Tööstaaž	0,161 *** (0,036)
Laenu pikkus	-0,545 *** (0,036)
Laenu suurus	-0,147 *** (0,018)
Laenureiting	6,736 *** (0,013)
Esmakordne laenaja	0,943 *** (0,032)
Sissetulek	0,812 *** (0,033)

Märkused: tulemustes on esitatud esmalt koefitsiendi väärtus, järgneb tärnidega statistilise olulisuse tase (* kui $p < 0,1$; ** kui $p < 0,05$; *** kui $p < 0,01$) ning sulgudes standardhälve.

3.2. Järeldused

Sissejuhatavalt saab märkida, et tulemuste märgid on loogilised (v.a kontrollmuutuja sissetulek) ja vastavad enamikele töös käsitletud uurimustele. Alljärgnevalt on esitatud vastavad järeldused iga uuringus kasutatud muutuja kohta.

Soo analüüsimisel selgus, et naistele väljastatud laenudele on määratud kõrgem intress (1,2 protsendipunkti võrra). See tulemus langeb kokku teiste eelnevalt käsitletud uurimustega (Muravyev *et al.* 2009; Alesina *et al.* 2013; Chen *et al.* 2017). Nt Šveitsis läbi viidud uurimuses (Dietrich, Wernli 2019) oli intressimääradel perekonnaseisust tulenevalt oluline vahe. Nimelt vallalisel ja lapsega naisterahval oli intress 2,3 protsendipunkti meestest kõrgem. Samas, kui lapsi polnud, oli vahe vaid 0,25 protsendipunkti.

Paraku Bondora andmete põhjal ei olnud võimalik perekonnaseisu (hõlmates siinkohal abielu, kooselu, vallalisust jms) arvesse võtta, kuna vastavad andmed olid puudulikud. Seega ei saa tulemust samaväärselt Šveitsi uurimusega kõrvutada. Käesoleva juhul on tulemuseks 1,2 protsendipunkti kõigi naiste osas. Šveitsi puhul see varieerus olenevalt määratud grupist (vallalised, abielus, lastega jms) vahemikus 0,25 kuni 2,3 protsendipunkti. Siinkohal ei ole meelevaldne eeldada, et ka Eestis võidakse üksikuid lastega naisterahvaid (eelarvamuslikult) pidada riskantsemateks laenuklentideks. Seega on võimalik, et grupeerides naisi samadel alustel, nagu tegid seda Dietrich ja Wernli, võib samuti ilmnedä sarnane tendents.

Eelnevalt sai välja toodud üksjagu uurimusi, kus olid esitatud erinevad positiivsed naisi iseloomustavad karakteristikud. Täpsemalt oli tegemist uurimistulemustega, mille põhjal ilmnnes, et finantsmaailmas naised evivad meestest madalamat riski (Eckel, Fullbrunn 2015; Sunden, Surette 1998; Barber, Odean 2001; Bernasek, Shwiff 2001). Erinevate laenuplatvormide uurimustes selgus, et naistel on parem krediidikäitumine ja nende puhul tuleb vähem ette maksejõuetust (Chen *et al.* 2017; Chen *et al.* 2019; Lin *et al.* 2017; Gaigalienė, Česnys 2018). Sellest saab üheselt järeldada, et naisi diskrimineeritakse.

Siinkohal on paslik rõhutada, et diskrimineerimise aluseks on mitte üks, vaid kaks aspekti. Nimelt lisaks sellele, et naistel on *ceteris paribus* (1) kõrgem intress, on nende puhul varasemate uurimuste kohaselt tegemist ka (2) vähemriskantsete laenuklentidega. N-ö ideaaljuhul peaks madalama riski juures (naistele) kohalduma võrreldavast poolest (meestest) madalam intress.

Seega saaks argumenteerida võimaliku diskrimineerimise üle isegi juhul, kui naiste ja meeste intress oleks võrdne. Sellises võtmes on arutlenud Chen *et al.* (2019). Ta täheldas, et meestel on naistega võrdne positiivse otsuse saanud laenuaotluste hulk, kuid samas naistele väljastatud laenud on laenu andjale kasumlikumad madalama riski ja parema laenu teenindamise tõttu. Seega naised peavad kompenseerima laenuturul osalemist parema krediitkäitumisega.

Sellest johtuvalt tekib põhjendatud küsimus, millest selline märgatavalt erinev tulemus (või suhtumine) alguse saab või tekib. Teise peatüki alguses on sissejuhatavalt välja toodud kaks peamist diskrimineerimisteooriat, s.o (1) eelistustel põhinev ja (2) statistiline diskrimineerimine. Neist esimene põhineb teesil, kus otsustav osapool on nõus kandma kasvõi suuremaid kulutusi selle nimel, et teatud gruppi vältida, kuna sõna otseses mõttes ta ei eelista nendega tegemist teha. Juhul, kui selle grupi vältimine pole võimalik, võib ta üritada neid eemale hoida määrates kõrgema hinna. Viimane võib olla otsustavale poolele ka teguriks, millest tulenevalt on ta nõus oma eelistusi muutma ja seda gruppi aktsepteerima.

Bondora platvormil määratakse intress Bondora poolt. Seega naissoost laenuaotlejatele rakendatava kõrgema intressimäära küsimus taandub Bondora kui laene vahendava platvormi valikutele või otsustele, mitte laenuandjatele. Hindamaks eelistustel põhineva diskrimineerimise kohaldumise võimalikkust, peab võtma seisukoha, kus Bondora mingil põhjusel ei soovi naissoost klientidega pigem tegemist teha. Selleks on nad läbivalt ette näinud kõrgema intressi, mis siis peaks neid eemale peletama või kompenseerib selle stressi, et neid teenindada. Peab märkima, et see põhjendus ei tundu pädev. Seega tuleb asuda seisukohale, et tegemist ei ole siinkohal eelistustel põhineva diskrimineerimisega Bondora platvormil.

Statistiline diskrimineerimine põhineb kulupõhisel lähenemisel, kus eelarvamus teatud grupi võimekuse osas kohaldatakse kõigile grupi liikmetele. Grupikuuluvuse tõttu on sisse arvatud teatud lisakoefitsent ja otsus tehakse pühendamata liialt aega individuaalsetele näitajatele. Toetudes teoreetilise osa raamistikule on saadud tulemuste põhjal võimalik argumenteerida, et naistele intressi määramisel võib olla selline lisakoefitsent Bondora poolt arvesse võetud.

Bondora on oma veebilehel selgitanud, et laenureitingu koostamisel kliendi kohta võtavad nad arvesse üle saja erineva teguri. Mida selle all täpselt mõeldakse, on välja toomata. Eeldatavalt, sarnaselt teistele suurtematele laenuplatvormidele, tehakse reitinguarvutused läbi vastava matemaatilise programmi. Seega on tegemist algoritmilise otsustusprotsessiga. Sellisel juhul on

võimalik, et Bondora otseselt ei ole sellist käsklust ette näinud, et naiste puhul tuleks sisse arvutada teatud koefitsient. Olenevalt kasutatava algoritmi võimekusest, võib see masinõppe või tehisintellektina sellise diskrimineeriva otsuse ise teha. Hindamise aluseks olevates ajaloolistes nn suureandmetes (*big data*) võib diskrimineeriv hoiak sees olla. See omakorda kandub edasi programmi otsustusprotsessi, mille tulemuseks on algoritmiline diskrimineerimine (Bartlett *et al.* 2019; Fuster *et al.* 2021; Williams *et al.* 2018).

Peale selle on võimalikuks vaatenurgaks otsene finantsvõimekus. Madala sissetuleku juures on laenu teenindamise võime väiksem. Sellisel juhul võib riskihinnang anda tulemuseks, et määrata tuleb kõrgem intress. Eesti on saanud haakuvas küsimuses Euroopa Liidu kontekstis üksjagu kõlapiinda. Nimelt sooline palgalõhe on liikmesriikide võrdluses suurim (Eurostat 2018). Bondora andmed on kohustuste osas puudulikud, kuid olemas on info sissetulekute kohta. Vastavate arvutuste läbiviimisel selgus, et naissoost laenuaotlejate mediaanpalk on 950 eurot ja meessoost 1100 eurot. See võib osaliselt selgitada, miks naistele on määratud kõrgem intress. Samas tuleb märkida, et sissetulek on regressioonanalüüsi juba lisatud ning tulemus oli statistiliselt oluline. Seega nõrgendab see seisukohta, et finantsvõimekusel võib olla intressi määramise juures oluline kaal meeste ja naiste vahel.

Ülaltoodu põhjal taandub järeldus võimalikule soolisele diskrimineerimisele Bondora platvormil. Seejuures ei saa välistada, et laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleerub naissoost laenuaotlejatega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. Võimalik, et see seaks kahtluse alla seisukoha diskrimineerimise osas Bondora laenuplatvormil.

Vähemusrahvuse (päritolu) puhul oli tulemuseks 0,33 protsendipunkti kõrgem intress. Tegemist on vähemusgrupiga, kelle laenuaotlus oli esitatud muus keeles, kui eesti. Sarnaselt soolisele diskrimineerimisele, leidsid Dietrich ja Wernli (2020) laenuplatvormil *Cashare.ch*, et Šveitsi passi omavatel inimestel on 0,54 protsendipunkti madalam intressimäär. Pope ja Sydnor (2011) poolt läbi viidud uurimuse tulemuseks USA laenuplatvormil *Prosper.com* oli mustanahalistel 0,6–0,8 protsendipunkti kõrgem intress. (See uurimus erineb selle poolest, et laenuandjad said otsustada taotlejate piltide järgi, kellele laenu anda. Üldjuhul teistes töös käsitletud platvormidel, k.a Bondoras, see teave puudub.) Selle põhjalt saab asuda seisukohale, et saadud tulemus langeb kokku siinkohal välja toodud Šveitsi ja USA uurimustega. Sellele lisaks langeb see kokku ka teiste

teadustöödega, mis selles kontekstis on esimeses peatükis lahatud (Deku *et al.* 2015; Stefan *et al.* 2018; Quillian *et al.* 2020; Ravina 2012).

Argumentatsioon selle üle, kas tegemist võib olla eelistustel põhineva või statistilise diskrimineerimisega, ei erine eelkirjeldatud soolise diskrimineerimise vastavatest seisukohtadest. Tegemist võib olla Bondora poolt teadliku otsusega, et käsitleda muus keeles esitatud taotlusi suurema ettevaatlikkusega ning näha ette riskikoeffitsient. Põhjendades seda nt sellega, et eesti keelt kõnelevatel on paremad võimalused tööturul ning on Eestiga tihedamalt seotud. Sisulises mõttes on sel juhul tegemist statistilise diskrimineerimisega.

Bertrand, Weill (2021) ja Bartlett (2019) uurisid päritolu järgset diskrimineerimist ning leidsid, et tõenäoliseks põhjuseks oli algoritmiline diskrimineerimine just ajalooliste andmete kallutatuse tõttu. Samuti võib Bondora poolt kasutatavates andmetes olla sees üks või mitu tegurit, mis evivad taolist tooni. Nt Eesti Vabariigi aastakümnete vältel võib olla olnud nn halli passi omanikel keerulisem tööd leida. See madalam finantsvõimekus on olemas suurandmetes. Algoritmi ülesandeks on üldjuhul seatud kuluefektiivsus või kasumi maksimeerimine. Selle tulemusel leiab ta need erisused gruppide vahel üles ja kasutab neid vastavalt oma lähteülesandele. Tulemuseks algoritmiline diskrimineerimine, millest Bondoral endal ei pruugi aimu olla.

Sarnaselt soole on siinkohal paslik peatuda ka finantsvõimekusel. Vähemusrahvuse mediaanpalgaks on 1000 eurot ning eestlaste (lähtudes siin laenuaotleja poolt kasutatud keelest) vastav näitaja on 1050 eurot. Suurem sissetulek aitab paremini kohustusi täita, mis omakorda tähendab laenajale väiksemat riski. Viimane väljendub madalama intressimäärana. Kahtlemata tuleks siinkohal arvesse võtta osapoolte kohustusi. Paraku, nagu eelnevalt selgitatud, vastavad andmed puuduvad. Samas Dietrich ja Wernli (2020) töid Šveitsi puhul välja, et mittekodanikel oli kõrgem intress, kuid kohustuste suhe sissetulekusse oli madalam.

Sarnane tendents võib olla ka Bondora laenuaotlejate puhul. Põhjendades viimast sellega, et eestlastele võidakse väljastada kergemini kodulaene jms (sellel samal põhjusel, et tegemist panga jaoks madala riski kliendiga). Seega eestlastel võib olla küll suurem sissetulek, kuid neil on ka rohkem kohustusi. Sellisel juhul saab asuda seisukohale, et tegemist on pigem päritolu järgi diskrimineerimisega.

Vanuse kohta selgus, et noored tasuvad kõrgemat intressi (0,4 protsendipunkti võrra). See langeb kokku hüpoteesi aluseks olnud uurimustega (Duarte *et al.* 2015; Coen *et al.* 2021). Seega vastav hüpotees leidis tõestust.

Coen *et al.* (2021) uurimusest selgus, et ebavõrdse kohtlemise aluseks ei olnud mitte suurem risk, vaid võimalus noortele kallimat laenu müüa. Sel juhul on tegemist statistilise diskrimineerimise vormiga, kus ajendiks on kasumi maksimeerimine. Suurandmete põhjal võib välja joonduda, et noored aktsepteerivad kõrgemat intressimäära kergekäelisemalt. Algoritm tuvastab erisuse ja teeb vastava otsuse selle grupi kahjuks.

Samas noorte mediaanpalgaks on 850 eurot ja ülejäänutel 1100 eurot. Vahe on suur. Seejuures tuleb aga märkida, et valdavalt pole noortel suuri laenukohustusi. Peale selle ei pruugi neist paljudel olla ka kulutusi elamispinnale. Nimelt tegemist on alates 18-aastastega, kellest nii mõnigi võib elada veel vanematega. Sellest tulenevalt pole pelgalt sissetuleku võrdlemine pädev. Lisaks kohustuste osas poleks üllatav, et paljude noorte jaoks võib Bondora laen olla elu esimesi. N-ö vanematel inimestel on see tõenäoliselt aga mitmes. Samuti on sissetulek juba analüüsi kaasatud iseseisva muutujana. Eeltoodut arvesse võttes ei ole meelevaldne järeldada, et tegemist võib olla diskrimineerimisega Bondora platvormil.

Siinkohal tuleb samuti rõhutada, sarnaselt soole, et ei saa välistada, et laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleerub võõrkeelsete või noorte laenuaotlejatega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. See võib oluliselt mõjutada argumenti diskrimineerimise kohta Bondora laenuplatvormil.

Hariduse kohta ilmnes, et kõrghariduseta laenajatele kohaldub kõrgem intressimäär 0,4 protsendipunkti võrra. Ei oleks imeks pandav, kui laenuandja hindaks riski madalamaks neil, kes on kõrgharidusega. Sellisel juhul on Bondora oma algoritmi sisse pannud käsu, kus madalama haridusega kliendile on ette nähtud kallim laen, et kompenseerida suuremat riski. Samas sissetulekute võrdluses ilmneb, et kõrgharitudel on mediaan 1300 eurot ning vastaspoolel 1000 eurot. Seega finantsvõimekus on tegelikult viimasel oluliselt kehvem. Selle põhjal võib tegemist olla riskihinnanguga, kus diskrimineerimine pole otsustamise aluseks olnud. Seevastu tuleb märkida, et sissetulek on analüüsi juba kaasatud. Seega muude tingimuste samaks jäädes on

madalama haridusega laenajatele kohaldatud kõrgem intress. Selle põhjal saab asuda seisukohale, et tegemist on diskrimineerimisega.

Sarnaselt eelnevatele punktidele tuleb ka siin välja tuua, et ei saa välistada, et laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleerub madalama haridusega laenajatega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. See omakorda võib ümber lükata argumendi diskrimineerimise kohta Bondora laenuplatvormil.

Muutuja **kinnisvara** puhul selgus, et isiklikku elamispinda omavatel inimestel on 0,33 protsendipunkti madalam intress. Antud muutujat on mõnevõrra keeruline diskrimineerimise kontekstis hinnata. Laenuandja seisukohast on õigustatud riski juures arvesse võtta laenuvõtja varalist seisu. Samas kinnisvara omaval inimesele on eeldatavalt koos Bondora laenuga üldjuhul kaks kohustust, mida igakuiselt täita. Võib argumenteerida, et kodu omaval inimesel on madalam igakuine elamiskulu. Samas, kui võrrelda seda üürikorteriga, siis viimasel puuduvad amortisatsiooni, kindlustus, sisustus, intressi jms kulu. Seega olenevalt turuolukorrast ja asukohast, võib jooksev üürikulu olla ka madalam elamispinna omamise igakuisest kulust (Mitt 2019).

Diskrimineerimisega oleks käesoleval juhul tegemist, kui elamispinda omaval poolel on tegelik majanduslik olukord kehvem, kui elamispinda üürival laenajal. Võimalik, et kinnisvara on Bondora algoritmi pandud sellise käsuga sisse, kus tema puhul on ette nähtud, *ceteris paribus*, madalam risk. Sellisel juhul on tegemist statistilise diskrimineerimisega. Mediaansissetulekute võrdluses on koduomanikel 1100 eurot ja vastaspoolel 1000 eurot. Selgub, et vahe on. Samas jääb õhku küsimus, milliste kohustuste suhte juures.

Pikem **tööstaaž** andis tulemuseks madalama intressi (0,16 protsendipunkti). Selle teguri puhul on eelkõige tegemist regressiooni kontrollmuutujaga. Tööstaaž väljendab sissetulekute stabiilsust ja ajalugu. Mida kauem on inimene töötanud, seda paremini on ta oma kulutusi suutnud katta. Kui aga töö on saadud käia vähe või kaootiliselt, seda tõenäolisemalt on igakuised kohustused olnud sissetulekutest suuremad. Sellisel inimesel on seega raskem uut laenu teenindada samaväärselt, kui stabiilse töökogemusega laenajal. Kokkuvõtteks saab märkida, et antud kontrollmuutuja tulemus vastab eeldusele.

Kontrollmuutujate **laenu pikkus** ja **laenu suurus** eeldused olid: mida pikem laenu periood ja mida suurem laen, seda madalam intress. Need leidsid ka tõestust. Tulemuseks oli $-0,54$ protsendipunkti kuu ja $-0,15$ protsendipunkti euro kohta. Lühiajalist laenu on võimalik taotleda Bondoral alates üheks kuuks. Samas pikaajaline laen on kuni viis aastat. Laenu suurused on vahemikus 100 kuni 16 632. Sellest ka oluline intressi erinevus.

Samuti oli muutuja **esmakordne laenaja** regressiooni lisatud tulemuste kontrolliks. Eelduseks oli, et korduvlaenaja, tänu oma krediidiajaloole, on Bondora jaoks väiksema riskiga klient. Sellest tulenevalt peaks kohalduma madalam intressimäär. Vastupidiselt esmakordse laenaja puhul on tegemist suurema teadmatusena ja on põhjendatud vastavat laenu kallimalt hinnastada. Intress on esmakordsel laenajal $0,94$ protsendipunkti kõrgem.

Kontrollmuutuja **sissetulek** tulemus tuli ainsa muutujana teoreetilisele eeldusele vastupidise märgiga. Selle põhjal on väiksema sissetuleku juures madalam intress. Pelgalt selle põhjalt järelduse tegemine oleks vastuolus tavapärasele finantsriski hindamise loogikale. Seega tuleb asuda seisukohale, et Bondora on tõenäoliselt arvesse võtnud ka mingil moel kohustusi. Niisiis oleks argumendiks, et olenemata suuremast sissetulekust on laenu teenindamise võimekus neil laenajatel kõrgemate muude kohustuste tõttu riskantsem. Sellest ka suurem intress.

See omakorda toetab kõiki eelnevaid tulemusi järgnevalt. Nimelt, kui suurema sissetuleku puhul on Bondora ette näinud kõrgema intressi ja eeldades sellest suuremat kohustuste suhet sissetulekutesse, saab asuda seisukohale, et nii naiste, vähemusrahvuse, noorte ja madalamalt haritute puhul on olenemata väiksemast sissetulekust tegemist diskrimineerimisega.

Kontrollmuutuja **laenureiting** oli regressioonanalüüsi tulemustest suurima väärtusega. Iga reitingu taseme langusega on intress $6,7$ protsendipunkti võrra kõrgem. Vastav hüpotees leidis seega kinnitust. Kahtlemata on tegemist intressi enim mõjutava teguriga, mis on ka igati loogiline. Bondora intressi määramisprotsess lähtub just reitingust. See on indikatsiooniks investoritele, millised laenud on riskantsemad ja samas suurema võimaliku tulususega. Vastupidiselt AA reitingu puhul on laenuandjale teada, et laenukoote risk on madalaim. See väljendub omakorda väiksemas intressimääras.

Välja jäid puudulike andmete tõttu muutujad perekondlik seis ja ülalpeetavate hulk. Need näitajad on eelkõige olulised soolise diskrimineerimise kontekstis. Šveitsis oli olulise kaaluga naiste

diskrimineerimine, kui ta polnud abielus ja vähemalt üks ülalpeetav oli alla 16 aasta (Dietrich, Wernli 2020). Selle põhjalt võib eeldada, et ka Eesti puhul on võimalik, kus need tegurid võiks tulemusi samaväärselt väljendada. Peale selle jäi samal põhjusel välja muutuja laenu suhe sissetulekusse. See oleks aidanud paremini mõista laenajate finantsvõimekust.

Selle probleemi ületamiseks tõi autor välja erinevate muutujate juures gruppide mediaanpalgad. Seda ei saa võtta samaväärse infoallikana, kuid annab arusaama, kas võrreldavate poolte vahel võib olla märkimisväärne sissetulekute erinevus. Nimelt diskrimineerimise asemel võib tegemist olla ka majanduslikel kaalutlustel tehtud otsustega, kus madalam finantsvõimekus evib suuremat tehinguriski. Samas selgus, et suurema sissetuleku puhul oli kõrgem intress. See võib viidata suurtematele kohustustele ja viimane argument langeks ära.

Kokkuvõtteks saab asuda seisukohale, et tulemused vastavad erialakirjanduses käsitletule. Kontrollmuutujate kaasamine näitas, et märgid tulid loogilised ja ebakõlasid ei tekkinud. Kõigi muutujate puhul saab asuda seisukohale, et lisaks loogilistele märkidele olid ka suurusjärgud mõistlikud ja vastavad oma kategooriale. Seejuures muutuja sissetulek vastupidine märk võib olla samuti põhjendatud, nagu sai eelnevalt argumenteeritud. Kõik püstitatud hüpoteesid leidsid kinnitust.

KOKKUVÕTE

Töö eesmärk oli välja selgitada, kas ja millises ulatuses tuleb ette diskrimineerimist Bondora laenuplatvormil sotsiaalsete tunnuste järgi. Seatud uurimisülesannetest lähtuvalt sai loodud laiapõhjaline teoreetiline raamistik. Sellele tuginedes omakorda valiti analüüsi läbiviimiseks sobiv meetodika, kohandati andmed ning viidi läbi vähimruutude meetodil põhinev regressioonanalüüs. Sõltuvaks muutujaks oli intress.

Valimi suurus on 69 902 Bondora klienti. Periood sai valitud vastavalt Bondora poolt kasutatava laenureitingute loomise viimasele mudeliversioonile. Seeläbi on tulemused omavahel samaväärselt võrreldavad. Periood on 2018. aasta algusest kuni 2021. aasta juulini. Kõik uurimusse kaasatud muutujate regressioonanalüüsi tulemused on statistiliselt olulised ja nende märgid loogilised. Samuti on mudel statistilised oluline ning puudub multikollineaarsus ja heteroskedastiivsus. Mudeli seletusvõimet saab pidada kõrgeks, s.o $R^2 = 0,81$.

Töö eesmärgist lähtuvalt on siinkohal esitatud oluliste tunnuste kohta käivad kokkuvõtavad järeldused. **Sugu** – selgus, et naised tasuvad kõrgemat intressi. Selle alusel saab võtta seisukoha, et naised diskrimineeritakse laenuplatvormil Bondora. Kuna intressi määramisel on Bondoras kasutusel algoritm, mis võtab otsuse tegemise aluseks palju erinevaid tegureid¹⁷, on võimalik, et selline diskrimineeriv suund tuleb sisse kasutatavate suurandmete (*big data*) põhjal. Tegemist on nn algoritmilise diskrimineerimisega. Vastuargumentiks saaks justkui esitada naiste madalama sissetuleku. Samas on sissetulek analüüsi kaasatud ja tulemus statistiliselt oluline. Seega langeb see vastuargument ära.

Päritolu – selgus, et vähemusrahvuse esindajad tasuvad kõrgemat intressi. Siinkohal oli grupeerimise aluseks võõrkeelne laenuaotlus. **Vanus** – samuti tasuvad kõrgemat intressi noored. Sellesse vähemusgruppi kuulusid laenajad vanusevahemikus 18 kuni 24 aastat. Nii päritolu kui vanuse puhul on argumentatsioon ülaltoodud sooga sarnane. Tulemuste põhjal saab asuda

¹⁷ Bondora lehel on selgituseks öeldud, et nad kasutavad laenureitingu määramisel üle saja erineva teguri. Mis need täpselt on, pole välja toodud.

seisukohale, et Bondora platvormil toimub nende sotsiaalsete tunnuste järgi diskrimineerimine. Võib oletada, et selle põhjuseks on kasutatava algoritmi mõjutatus ajaloolistest suurandmetest.

Samas tuleb rõhutada, et ei saa välistada, et Bondora laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleerub naissoost, võõrkeelsete või noorte laenuaotlejatega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. See võib oluliselt mõjutada ja ümber lükata eelnevalt toodud väite diskrimineerimise kohta Bondora laenuplatvormil.

Muude sotsiaalsete andmete puhul olid veel vaatluse all haridus, kinnisvara olemasolu ja tööstaaž. Tulemuseks oli, et kõrgemat intressi tasuvad ilma kõrghariduseta laenajad, kellel puudub kinnisvara või kes on töötanud viimasel töökohal vähem kui aasta. Neist esimese puhul on võimalik argumenteerida, et tegemist võib olla diskrimineerimisega. Muude tingimuste samaks jäädes ei ole siinkirjutaja arvates põhjendatud pelgalt kõrgema hariduse puudumise tõttu eeldada, et tegemist on riskantsema kliendiga ja näha ette kõrgem intress. Seevastu tööstaaž ja kinnisvara olemasolu on otseselt seotud jätkusuutlikuma finantsvõimekuse või võimega tasuda kohustusi muude varade arvelt – tegemist on tavapärase krediidiriski hindamise aluseks olevate teguritega.

Kontrollmuutujate tulemused nagu laenu pikkus, suurus, reiting ja esmakordne vs. korduvlaenamine tulid loogiliste märkidega ja vastasid teoreetilistele eeldustele. Seevastu kontrollmuutuja sissetulek tuli vastupidise märgiga. St mida suurem sissetulek, seda kõrgem intress. Tegemist on ärirahanduse vaatepunktist ebaloogilise tulemusega.

Samas juhib see tähelepanu ühele olulisele puudusele käesolevas töös. Nimelt pole Bondora andmetes järjepidevat ülevaadet laenuaotlejate kohustuste kohta. Vastav andmerida oli enamjaolt kas tühjaks jäetud või väärtuseks oli null. Viimane pole aga loogiline, kuna eeldatavalt on Bondora kliendid inimesed, kellel on sedavõrd palju kohustusi n-ö traditsiooniliste pankade ees, et nad sealt enam laenu ei saa. Sellest tulenevalt võib pidada tõenäoliseks, et Bondoral on klientide muude kohustuste kohta parem ülevaade ja see on reitingumääramise protsessis arvesse võetud. Millest johtuvalt on tegelik kohustuste suhe sissetulekutesse suurem just kõrgema sissetulekuga inimestel – tagajärjeks kõrgem intress.

Eelnevalt sai rõhutatud, et ei saa välistada, et Bondora laenuplatvormi otsustusprotsessis esineb muid olulisi tunnuseid, sh finantsilisi tunnuseid, mis osaliselt korreleeruvad töös kasutatud

tunnustega, kuid mida autoril ei olnud võimalik käesolevas lõputöös esitatud mudelisse kaasata. Selles võtmes heaks näiteks ongi laenajate muud kohustused. Seega pole võimalik üheselt väita, et käesolevate uurimustulemuste põhjal saab võtta seisukoha, et Bondora platvormil toimub sotsiaalsete tunnuste järgi diskrimineerimine.

Autor loodab, et saadud tulemused aitavad paremini mõista diskrimineerimist finantsteenuste pakkumisel ning seda nii kitsas valdkonnas nagu laenude väljastamisel Eestis. Kahtlemata parendaks taolise uurimuse adekvaatsust põhjalik ülevaade kõigist teguritest, mida Bondora kasutab laenureitingu määramisel. Sellisel juhul saaks esitada põhjaneva seisukoha diskrimineerimise kohta.

SUMMARY

DISCRIMINATION ON THE BASIS OF SOCIAL CHARACTERISTICS ON THE BONDORA LOAN PLATFORM

Equality is gaining ground every year. As a result, various forms of discrimination have begun to be investigated and identified among us. Minority groups are mainly discriminated on the basis of their social characteristics. For example, their differences in gender and ethnicity, but also in age, disability, religion, sexual orientation, etc. from the majority group. Typical examples are employment, where, for example, men are preferred to women, young to old, or the pay gap, where men earn more than women. Estonia is a good example of the latter. Namely, the comparison of the European Union countries revealed that the corresponding indicator is unfortunately the highest in Estonia (Eurostat 2019).

In addition, individuals or groups may be treated unequally in the provision of services. For example, to sell a service at a higher price, to impose more onerous conditions or to refuse to offer it at all. One such object is financial services. More specifically, the question here is whether different groups are treated equally in the provision of credit, regardless of their social characteristics. In a situation where, for example, holders of a so-called Gray Passport in Estonia (Alien / Stateless Passport) would receive a loan in Estonia less often or with a higher interest rate, but at the same time their debt to income ratio is at least the same compared to the majority group, it would be discrimination.

Unfortunately, to obtain data about clients social characteristics is difficult from commercial banks. At the same time, online platforms focusing on loan intermediation have entered to the market. (Better known in Estonia as crowdfunding platforms and in English academic literature as peer-to-peer lending platforms). In general, these loan platforms have made available comprehensive data about their customers, which also includes various social indicators. One such platform in Estonia is Bondora, which mediates loans from individuals to individuals. The analysis of such data provides a good opportunity to study the respective problem about Estonian.

Regarding to the importance of the thesis, it is possible to point out the lack of similar studies, where discrimination in the provision of financial services would have been studied in Estonia (as far as the author is aware). It provides a point of comparison with the rest of the world and helps to understand better such an important issue in the society as equality. Given the good availability of Bondora's data and its large market share in Estonia, it can be considered a good object to investigate discrimination.

The aim of the study is to find out whether and to what extent there is discrimination on the basis of social characteristics on the Bondora loan platform. Consequently, the research problem in the provision of financial services on the Bondora financial platform is the possible discrimination on the basis of social characteristics and the comparison of the obtained results with the relevant literature. The research questions primarily analyze the possible unequal treatment of different minority groups. Is there discrimination by:

- gender;
- national minority;
- age;
- other social characteristics?

The task of the research is to create a theoretical basis for discrimination in the provision of financial services on the basis of the relevant scientific literature. In addition to explain the operation of the loan platforms. Namely, loan platforms set interest rates on loans through the credit ratings they create. The different interest, *ceteris paribus*, is one of the possible factors for to evaluate the existence of discrimination. In addition, select and adjust variables from the Bondora dataset for calculations and choose the appropriate methodology. After test the model, present the results, and make decisions by accepting a null or alternative hypothesis. Then assess whether the hypotheses made beforehand were confirmed or not.

In conclusion, based on the results, it is possible to present opinions on the research questions as to whether there are any indications of possible discrimination on the basis of social characteristics in the loan platform based on Estonian data. Namely, Bondora operates in three different countries, but the vast majority of its operations are in Estonia, and there are many erroneous data on foreign countries, i.e. Finland and Spain. Consequently, the analysis of the work focuses only on Estonian data.

This thesis is divided into three chapters. The first is a theoretical approach based on the literature focusing on dissection of discrimination in the provision of financial services. This is the case in banking more broadly, as well as on similar loan platforms as Bondora. In order to better understand the object of research, the theoretical approach also provides an overview of the principles of operation of loan platforms in general, also pointing out the differences from traditional banking. Depending on the fact that the perception of risk and its assessment is a key factor in the interest rate, the risks associated with credit platforms and the ratings they create are also considered.

The second part of the paper explains what the sample is, where it comes from, how the data have been adjusted. Plus the descriptive statistics have been presented. In addition, there is a discussion between different methodologies according to the literature and an explanation of the choice. The third part of the work tests the model, presents the results and draws conclusions based on the theoretical approach.

The data is derived from Bondora's public database, which contains all its loans throughout its history. After adjusting the data and narrowing the period, the sample size is 69,902 approved loan applications within Estonia. In this thesis, a linear regression based on the ordinary least squares method is used, where the loan interest rate is set as a dependent variable.

The period was selected according to the latest version of the credit rating model used by Bondora. As a result, the results are equally comparable. The period is from the beginning of 2018 to July 2021. All results of the regression analysis of the variables included in the study are statistically significant and their signs are logical. The model is also statistically significant and lacks multicollinearity and heteroskedasticity. The explanatory power of the model can be considered high, i.e. $R^2 = 0.81$.

Based on the goal of the thesis, summary conclusions about the important features are presented here. **Gender** - It turned out that women pay higher interest. On this basis, it can be argued that women are discriminated against on the Bondora loan platform. As Bondora uses an algorithm to determine interest rate based on many different factors, it is possible that this discriminatory trend is influenced by the big data. This is so-called algorithmic discrimination. The lower income for women could be argued as a counter-argument. At the same time, income is included in the analysis and the result is statistically significant. This counter-argument is therefore rejected.

National minority - It turned out that the representatives of the national minority pay higher interest rate. Here, the grouping was based on a loan application in a foreign language. **Age** - young adults also pay higher interest. This minority group included borrowers between the ages of 18 and 24. In terms of both origin and age, the argument is similar to the above mentioned gender. Based on the results, it can be concluded that the Bondora platform discriminates on the basis of these social characteristics. It can be assumed that this is due to the influence of the algorithm that have been using historical big data.

However, it cannot be ruled out that there are other important variables in the decision-making process of the Bondora loan platform, including financial variables that correlate in part with female, national minority or young adult applicants, but which the author was unable to include in this analysis. This could significantly affect and refute the above allegation of discrimination on the Bondora loan platform.

For other social data, as education, real estate and work experience were also analysed. As a result, borrowers without higher education who have no real estate or who have worked in their last job for less than a year pay higher interest rates. In the first case, it can be argued that there may be case about discrimination. Otherwise, if the other conditions remain the same, it is not justified to assume that this is a riskier customer and to provide for a higher interest rate simply due to the lack of higher education. In contrast, seniority and the ownership of real estate are directly related to a more sustainable financial capacity or the ability to meet liabilities at the expense of other assets - these are the usual factors taken into account in credit risk assessment.

Results for control variables such as loan period, size, rating, and first-time borrower *vs.* multiple time came with logical signs and met theoretical assumptions. In contrast, the the control variable named income had the opposite sign. That is, the higher the income, the higher the interest rate. This is an illogical result from the point of view of business finance.

However, it draws attention to one important shortcoming in this analysis. Namely, Bondora's data do not provide a consistent overview of the loan applicants' liabilities. The corresponding data series were mostly left blank or had a value of zero. However, the latter is not logical, as Bondora's customers are usually expected to be people who have already multiple liabilities that so-called traditional banks will no longer provide loans for them. Consequently, it can be considered likely that Bondora has a better overview of their customers other liabilities and this has been taken into

account in the rating making process. As a result, the actual ratio of liabilities to income is higher for people with higher incomes - resulting in higher interest rates.

It was previously emphasized that it cannot be ruled out that there may have be other important features in the decision-making process of the Bondora loan platform, including financial variables that correlate in part with the variables used in this analysis but could not be included in the used model. Other liabilities of borrowers are a good example of this. It is therefore not possible to state unequivocally that the results of the present study suggest that there is discrimination on the basis of social characteristics on the Bondora platform.

A comprehensive overview of all the factors used by Bondora in determining a credit rating would undoubtedly improve the adequacy of such a study. In that case would be possible to state unconditionally about possible discrimination on the Bondora's peer-to-peer loan platform.

KASUTATUD ALLIKATE LOETELU

- Agier, I., Szafarz, A. (2013). Microfinance and gender: Is there a glass ceiling on loan size? *World Development* 42, 165–181. Alesina, A. F., Lotti, F., & Mistrulli, P. E. (2013). Do women pay more for credit? evidence from Italy. *Journal of the European Economic Association*, 11(s1), 45–66.
- Alyakoob, M., Rahman, M. S. & Wei, Z. (2020) Where You Live Matters: The Impact of Local Financial Market Competition on Banks' Response to Peer-to-Peer Lending. *Information Systems Research*.
- Asiedu, E., Freeman, J.A., Nti-Addae, A. (2012). Access to credit by small businesses: How relevant are race, ethnicity, and gender? *American Economic Review*, 102, 532–537.
- Asiedu, E., Kalonda-Kanyama, I., Ndikumana, L., Nti-Addae, A. (2013). Access to credit by firms in sub-Saharan Africa: How relevant is gender? *American Economic Review*, 103, 293–297.
- Barasinska, N. (2011). Does Gender Affect Investors' Appetite for Risk?: Evidence from Peer-to-Peer Lending. *Discussion Papers of DIW Berlin*, No 1125. DIW Berlin, German Institute for Economic Research.
- Barasinska, N. & Schäfer, D. (2014). Is Crowdfunding Different? Evidence on the Relation between Gender and Funding Success from a German Peer-toPeer Lending Platform. *German Economic Review*, 15(4), 436–452.
- Barber, B.M. & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116, 261–292.
- Bartlett, R., Morse, A., Stanton, R & Wallace, N. (2019). Consumer-Lending Discrimination in the FinTech Era. *NBER Working Papers*, No 25943. National Bureau of Economic Research, Inc.
- Baydas, M. M., Meyer, R. L. & Aguilera-Alfred, N. (1994). Discrimination against women in formal credit markets: reality or rhetoric? *World Development*, 22(7), 1073–1082.
- Beck, T. & Brown, M. (2011). Which Households Use Banks? Evidence From the Transition Economies. *ECB Working Paper*, No. 1295. Germany: ECB Publishing.
- Belleflamme, P., Omrani, N. & Peitz, M. (2021). The Economics of Crowdfunding Platforms. *Information Economics and Policy*, 33(C), 11-28.
- Bellucci, A., Borisov, A. & Zazzaro, A. (2010). Does gender matter in bank–firm relationships? Evidence from small business lending. *Journal of Banking & Finance*, 34, 2968–2984.

- Bernardi, R.A. & Arnold, D.F. (1997). An examination of moral development within public accounting by gender, staff level, and firm. *Contemporary Accounting Research*, 14, 653–668.
- Bernasek, A. & S. Shwiff (2001). Gender, risk, and retirement. *Journal of Economic Issues*, 35(2), 345–356.
- Bertrand, J. & Weill, L. (2021). Do Algorithms Discriminate Against African Americans in Lending? *Economic Modelling*, 104.
- Brock, J. M. & De Haas, R. (2020). Discriminatory Lending: Evidence from Bankers in the Lab. *Discussion Paper*, No. DP14340. CEPR Publishing.
- Bryant, M. (27, jaanuar 2020). Maternity leave: US policy is worst on list of the world’s richest countries. *The Gaurdian*.
- Cai, S., Lin, X., Xu, D. & Fu, X. (2016). Judging Online Peer-to-Peer Lending Behavior: A Comparison of First-Time and Repeated Borrowing Requests. *Information & Management*, 53(7), 857–867.
- Caliskan, A., Bryson, J. & Narayanan. A. (2017). Semantics Derived Automatically from Language Corpora Contain Human-Like Biases. *Science*, 356, 183–86.
- Celent. (2012). *IT Spending in Banking: A Global Perspective*. (2012). New York: Celent.
- Chen, D., Li, X. & Lai, F. (2017). Gender discrimination in online peer-to-peer credit lending: evidence from a lending platform in China. *Electronic Commerce Research, Springer*, 17(4), 553-583.
- Chen, J. & Xin, N. (2013). Empirical research on the influence of personal information on P2P lending’s success rate. *Accounting and Finance*, 6, 13-17.
- Chen, S., Gu, Y., Liu, Q. & Tse, Y. (2020). How do lenders evaluate borrowers in peer-to-peer lending in China? *International Review of Economics & Finance, Elsevier*, 69(C), 651–662.
- Chen, X., Huang, B. & Ye, D. (2019). The Gender Gap in Peer-to-Peer Lending: Evidence from the People’s Republic of China. *ADB Working Paper*, No 977. Asian Development Bank Institute Publishing.
- Chen, Z., Li, B., Keung, G., Yin, H., Lin, C., & Wang, Y. (2009). *How scalable could P2P live media streaming system be with the stringent time constraint?* In IEEE International Conference on Communications (1–5).
- Coen, J., Kashyap, A. & Rostom, M. (2021). Price discrimination and mortgage choice. *Working Paper*, No. 926. Bank of England Publishing.

- Cornaggia, J., Wolfe, B. & Yoo, W. (2018). Crowding Out Banks: Credit Substitution by Peer-To-Peer Lending. Kättesaadav: <https://projects.iq.harvard.edu/fintechlaw/publications/crowding-out-banks-credit-substitution-peer-peer-lending> 11. november 2021.
- Cumming, D., Leung, T.Y. & Rui, O. (2015). Gender diversity and securities fraud. *Academy of Management Journal*, 58, 1572–1593.
- Cumming, D.J. & Johan, S. J. (2016). Crowdfunding and Entrepreneurial Internationalization. *Entrepreneurial Finance: Managerial and Policy Implications. The World Scientific Publishers, chapter 5.*
- Cyr, M. (2011). Gender, Maternity Leave, and Home Financing: A Critical Analysis of Mortgage Lending Discrimination against Pregnant Women. *University of Pennsylvania Journal of Law and Social Change*, 15, 109-141.
- Davis, K. & Murphy, J. (2016). Peer to Peer Lending: Structures, Risks and Regulation. *The Finsia Journal of Applied Finance*, 2016(3), 37-44.
- Deku, S., Kara, A. & Molyneux, P. (2015). Access to consumer credit in the UK. *The European Journal of Finance*, 22, 1-24.
- Department of Justice. (2020). Justice Department Reaches Settlement with Bank of America to Resolve Claims of Disability Discrimination and Compensate Victims. *Office of Public Affairs*. Kättesaadav: <https://www.justice.gov/opa/pr/justice-department-reaches-settlement-bank-america-resolve-claims-disability-discrimination> 11. november 2021.
- Diamond, D. W. (1984). Financial intermediation and delegated monitor. *Review of Economic Studies*, 51(3), 393-414.
- Dietrich, A. & Wernli, R. (2020). Determinants of Interest Rates in the P2P Consumer Lending Market: How Rational are Investors? *Institute of Financial Services IFZ, Lucerne University of Applied Sciences.*
- Dorfleitner, G., Priberny, C., Schuster, S., Stoiber, J., Weber, M., de Castro, I. & Kammler, J. (2016). Description-text related soft information in peer-to-peer lending—evidence from two leading European platforms. *Journal of Banking & Finance*, 64(C), 169–187.
- Duarte, J, S. Siegel and L. Young (2012). Trust and credit: The Role of Appearance in Peer-to-Peer Lending, *Review of Financial Studies*, 25, 2455-2483.
- Duarte, J., Siegel, S. & Young, L. (2015). To Lend or Not to Lend: Revealed Attitudes towards Gender, Ethnicity, Weight, and Age in the U.S. Kättesaadav: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2473240 11. november 2021.
- Eckel, C.C. & Fullbrunn, S.C. (2015). Thar SHE blows? Gender, competition, and bubbles in experimental asset markets. *American Economic Review*, 105, 906–920.
- Emekter, R., Tu, Y., Jirasakuldech, B. & Lu, M. (2015). Evaluating Credit Risk and Loan Performance in Online Peer-to-Peer (P2P) Lending. *Applied Economics*, 47(1), 54–70.

- Eunkyoung, L. & Byungtae, L. (2012). Herding behavior in online P2P lending: An empirical investigation. *Electronic Commerce Research and Applications*, 11, 495–503.
- Financial Ombudsman. (2015). Just a number? Age, Complaints and the Ombudsman. *Financial Ombudsman United Kingdom Service Insight Briefing*. Kättesaadav: <https://www.financial-ombudsman.org.uk/files/1762/age-insight-briefing-2015.pdf> 11. november 2021
- Freedman, S. & Jin, G.Z. (2017). The information value of online social networks: Lessons from peer-to-peer lending. *International Journal of Industrial Organization, Elsevier B.V.*, 51, 185–222.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T. & Walther, A. (2021). Predictably Unequal? The Effects of Machine Learning on Credit Markets. *Journal of Finance (accepted)*.
- Gaigalienė, A. & Česnys, D. (2018). Determinants of Default in Lithuanian Peer-To-Peer Platforms. *Management of Organizations: Systematic Research, Sciendo*, 80(1), 19-36.
- Ge, R., Feng, J., Gu, B. & Zhang, P. (2017) Predicting and Deterring Default with Social Media Information in Peer-to-Peer Lending. *Journal of Management Information Systems*, 34(2), 401-424.
- Gibilaro, L. & Mattarocci, G. (2018). Peer-to-peer lending and real estate mortgages: evidence from United Kingdom. *Journal of European Real Estate Research*, 11(3).
- Gonzalez, L. & Loureiro, Y. (2014). When can a Photo Increase Credit? The Impact of Lender and Borrower Profiles on Online Peer-to-Peer Loans. *Journal of Behavioral and Experimental Finance*, 2(1), 44–58.
- Gujarati, D. N. (1995). *Basic Econometrics*. International Third Edition, McGraw-Hill, Inc. New York.
- Guo, Y., Zhou, W., Luo, C., Liu, C. & Xiong, H. (2016). Instance-based credit risk assessment for investment decisions in P2P lending. *European Journal of Operational Research, Elsevier*, 249(2), 417-426.
- Haewon, Y., Byungtae, L. & Myungsin, C. (2012). From the wisdom of crowds to my own judgment in microfinance through online peer-to-peer lending platforms. *Electronic Commerce Research and Applications*, 11(5), 469–483.
- Hasan, I., He, Q. & Lu, H. (2017). Stereotypes in Person-to-Person Lending: Evidence from Debt Crowdfunding.
Kättesaadav:https://business.depaul.edu/about/centers-institutes/financial-services/Documents/1_3_Hasan.pdf 11. november 2021.
- Herzenstein, M., Dholakia, U.M. & Andrews, R.L. (2011). Strategic Herding Behavior in Peer-to-Peer Loan Auctions. *Journal of Interactive Marketing*, 25(1), 27–36.

- Hoffman, P., T., Postel-Vinay, G. & Rosenthal, J., L. (2019). *Dark Matter Credit: The Development of Peer-to-Peer Lending and Banking in France*. Princeton University Press.
- Hollas, J. (2013). Is crowdfunding now a threat to traditional finance? *Corporate Finance Review*, 18(1), 27-31.
- Iyer, R., Khwaja, A., Luttmer, E. & Shue, K. (2009). Screening in New Credit Markets: Can Individual Lenders Infer Borrower Creditworthiness in Peer-to-Peer Lending? *HKS Faculty Research Working Paper Series*, No RWP09-031. John F. Kennedy School of Government, Harvard University Publishing.
- Iyer, R., Khwaja, A. I., Luttmer, E. F., & Shue, K. (2016). Screening peers softly: Inferring the quality of small borrowers. *Management Science*, 62(6), 1554–1577.
- Khandani, A., Kim, A. & Lo, A. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34, 2767-2787.
- Klafft, M. (2008a, 6. märts). *Peer to Peer Lending: Auctioning Microcredits over the Internet*. Proceedings of the International Conference on Information Systems, Technology and Management, A. Agarwal, R. Khurana, eds., IMT, Dubai. Kättesaadav: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1352383 11. november 2021.
- Klafft, M. (2008b, 14. juuli). Online Peer-to-Peer Lending: A Lenders' Perspective. *Proceedings of the International Conference on E-Learning, E-Business, Enterprise Information Systems, and E-Government, EEE 2008, H. R. Arabnia and A. Bahrami, eds., pp. 371-375, CSREA Press, Las Vegas*. Kättesaadav: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1352352 11. november 2021.
- Knief, U. & Forstmeier, W. (2021). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, 1-15.
- Laht, P. (2016). *Platvormist sõltumatu voogedastussüsteem partnervõrgus telesarjade näitel*. (Magistritöö) Taltech infotehnoloogia teaduskond, Tallinn.
- Lambrecht, A. & Tucker, C. (2018). Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads. *Management Science*, 65(7), 2966-2981.
- Lee, E. & Lee, B. (2012). Herding behavior in online P2P lending: An empirical investigation. *Electronic Commerce Research and Applications, Elsevier B.V.*, 11(5), 495–503.
- Lei, X. (2016). Discussion of the Risks and Risk Control of P2P in China. *Modern Economy*, 7(4), 399.
- Li, J. & Hu, J. (2019). Does university reputation matter? Evidence from peer-to-peer lending. *Finance Research Letters, Elsevier*, 31(C), 66-77.

- Lin, M., Prabhala, N.R. & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, 59(1), 17–35.
- Lin, X., Li, X. & Zheng, Z. (2017). Evaluating Borrower's Default Risk in Peer-to-Peer Lending: Evidence from a Lending Platform in China. *Applied Economics*, 49(35), 3538–3545.
- Liu, D., Brass, D., & Chen, D. (2015). Friendships in online peer-to-peer lending: pipes, prisms, and relational herding. *MIS Quarterly*, 39(3), 729–742.
- Lumley, T., Diehr, P., Emerson, S. & Chen, L. (2002). The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual review of public health*, 23, 151-69.
- Petkantchin, V. EU anti-discrimination policy's impact on insurance risk management: A parallel with the US sub-prime crisis. *Pensions Int J*, 15, 155–160.
- Philippon, T. (2012). Has the U.S. Finance Industry Become Less Efficient? on the Theory and Measurement of Financial Intermediation. *The American Economic Review*, 105(4), 1408-1438.
- Pokorná, M. & Sponer, M. (2016). Social Lending and Its Risks. *Procedia - Social and Behavioral Sciences*, 220, 330-337.
- Pope, D. & Sydnor, J. (2011). What's in a Picture? Evidence of Discrimination from Prosper.com. *Journal of Human Resources*, 46(1), 53–92.
- Quillian, L., Lee, J. J., & Honoré, B. (2020). Racial Discrimination in the U.S. Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016. *Race and Social Problems*, 12(1), 13-28.
- Ravina, E. (2012). Love & Loans: The Effect of Beauty and Personal Characteristics in Credit Markets. *Columbia University*.
Kättesaadav:
<https://law.yale.edu/sites/default/files/area/workshop/leo/document/E.Ravina2.pdf>
11. november 2021.
- de Roure, C., Pelizzon, L. & Tasca, P. (2016). How does P2P lending fit into the consumer credit market? *Discussion Papers*, No 30/2016. Deutsche Bundesbank.
- Mach, T., Carter, C. & Slattery, C. (2014). Peer-to-peer lending to small businesses. *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.*
- Malekipirbazari, M. & Aksakalli, V. (2015). Risk Assessment in Social Lending via Random Forests. *Expert Systems with Applications*, 42, 4621-4631.
- Milne, A. & Parboteeah, P. (2016). The Business Models and Economics of Peer-to-Peer Lending. *ECRI Papers*, No 11594. Centre for European Policy Studies Publishing.

- Mitt, M. (2019). *Eesti elamukinnisvara investeerimisobjektina vs. üürielamispind ja alternatiivinvesteeringud, 2003.-2011. aastate lõikes*. (Magistritöö) Estonian Business School, Tallinn.
- Moenninghoff, S. & Wieandt, A. (2012). The Future of Peer-to-Peer Finance. *Zeitschrift für Betriebswirtschaftliche Forschung*, August/September 2013, 466-487.
- Montoya, P., Soli, A. & Undurraga, R. (2020). Bad Taste: Gender Discrimination in Consumer Credit Markets. *IDB Working Paper Series*, No. 1053. IDB Publishing.
- Moro, A., Wisniewski, T.P. & Mantovani, G.M. (2017). Does a manager's gender matter when accessing credit? Evidence from European data. *Journal of Banking & Finance*, 80, 119–134.
- Muravyev, A., Talavera, O. & Schafer, D. (2009). Entrepreneurs' gender and financial constraints: Evidence from international data. *Journal of Comparative Economics*, 37, 270–286
- Santoso, W., Trinugroho, I. & Risfandy, T. (2020). What Determine Loan Rate and Default Status in Financial Technology Online Direct Lending? Evidence from Indonesia. *Emerging Markets Finance and Trade, Taylor & Francis Journals*, 56(2), 351-369.
- Sauga, A. (2005). *Kvantitatiivsed meetodid majanduses*. Audentese Ülikool.
- Serrano-Cinca, C., Gutiérrez-Nieto, B. & López-Palacios, L. (2015) Determinants of Default in P2P Lending. *PLoS ONE*, 10(10).
- Stein, J. (2002). Information Production and Capital Allocation: Decentralized versus Hierarchical Firms. *Journal of Finance LVII*, (5), 1891-1921.
- Stefan, M., Holzmeister, F., Müllauer, A. & Kirchler, M. (2018). Ethnical discrimination in Europe: Field evidence from the finance industry. *PLoS ONE*, 13(1).
- Stiglitz, J. & Weiss, A. (1981). Credit Rationing in Markets With Imperfect Information. *American Economic Review*, 71, 393-410.
- Sun, H. & Gao, L. (2019). Lending practices to same-sex borrowers. *Proceedings of the National Academy of Sciences*, 116(19), 9293-9302.
- Sundén, A. & Surette, B. (1998). Gender differences in the allocation of assets in retirement savings plans. *American Economic Review*, 88, 207–211.
- Tao, Q., Dong, Y., & Lin, Z. (2017). Who Can Get Money? Evidence from the Chinese Peer-to-Peer Lending Platform. *Information Systems Frontiers*, 19, 425-441.
- Thakor, Anjan V. (2020). Fintech and banking: What do we know? *Journal of Financial Intermediation, Elsevier*, 41(C).
- United Nations. (2007). *Definition of Youth*.
Kättesaadav: <https://www.un.org/esa/socdev/documents/youth/fact-sheets/youth-definition.pdf> 11. november 2021.

- Vainu, J. (2006). *Ökonomeetria. Lihtsad mudelid*. Tallinn: Külim.
- Vallée, B. & Yao, Z. (2019). Marketplace Lending: A New Banking Paradigm? *The Review of Financial Studies*, 32(5), 1939–1982.
- Wang, C. (2018). Technology, the Nature of Information, and Fintech Marketplace Lending. *Federal Reserve Bank of Boston Research Paper Series Current Policy Perspectives Paper, No 18-3*. Federal Reserve Bank of Boston Publishing.
- Wei, S. (2015). Internet Lending in China: Status Quo, Potential Risks and Regulatory Options. *Computer Law & Security Review*, 31(6), 793–809.
- Weizsäcker, G. & Zankiewicz, C. (2017). Measuring Applicant Quality to Detect Discrimination In Peer-to-Peer Lending. *Rationality and Competition Discussion Paper Series*, No 13. Berlin: CRC TRR 190 Publishing.
- White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817–838.
- Williams, B. A., Brooks, C. & Shmargad, Y. (2018). How Algorithms Discriminate Based on Data They Lack: Challenges, Solutions, and Policy Implications. *Journal of Information Policy*, 8, 78-115.
- Winter, J. (2015). Algorithmic Discrimination: Big Data Analytics and the Future of the Internet. *Springer, The Future Internet: Alternative Visions*, 125-140.
- Wiswall, M. & Zafar, B. (2018). Preference for the workplace, investment in human capital, and gender. *Quarterly Journal of Economics*, 133, 457–507.
- World Bank. (2018). Womens Financial Inclusion and The Law. Kättesaadav: <http://pubdocs.worldbank.org/en/610311522241094348/Financial-Inclusion.pdf> 11. november 2021.
- Xu, J., Hilliard, J., & Barth, J. (2018). On Education Level and Terms in Obtaining P2P Funding: New Evidence from China: Peer-to-Peer Funding. *International Review of Finance*, 20(4), 801-826.
- Zhongming, T., Baah, A. & Guoping, D. (2019). Modelling Default Risk of Borrowers: Evidence from Online Peer to Peer Lending Platforms in Australia. *Research Journal of Finance and Accounting*, 10, 1 - 10.
- Ziegler, T., Shneor, R., Wenzlaff, K., Wang, B., Kim, J., Odorovic, A., Paes, F., Suresh, K., Zhang, B., Johanson, D., López, C., Mammadova, L., Adams, N. & Luo, D. (2020). The Global Alternative Finance Market Benchmarking Report. *Cambridge Centre for Alternative Finance*.

LISAD

Lisa 1. Korrelatsioonimaatriks (Gretl)

Correlation coefficients, using the observations 1 – 69902. 5% critical value (two-tailed) = 0.0074 for n = 69902

Vahemusrahv	Vanus	Sugu	Haridus	Toostaaz	
1.0000	0.0401	0.0340	0.0143	-0.0017	Vahemusrahv
	1.0000	0.0209	0.1249	0.2862	Vanus
		1.0000	-0.0434	-0.0007	Sugu
			1.0000	0.0811	Haridus
				1.0000	Toostaaz
Kinnisvara	Laenureiting	EsmakLaenaj	1_LaenuSuuru	1_Sissetulek	
-0.0240	0.0358	0.0765	-0.0558	-0.0899	Vahemusrahv
0.2671	0.1494	0.1414	-0.1226	-0.1493	Vanus
-0.0382	-0.0186	0.0844	-0.0033	-0.1691	Sugu
0.1350	0.1422	0.0727	-0.0383	-0.1798	Haridus
0.1941	0.0928	0.0969	-0.0450	-0.0994	Toostaaz
1.0000	0.1413	0.0804	-0.0536	-0.1028	Kinnisvara
	1.0000	0.1336	0.0209	-0.0274	Laenureiting
		1.0000	0.0761	-0.0298	EsmakLaenaj
			1.0000	0.2279	1_LaenuSuuru
				1.0000	1_Sissetulek
					1_LaenuPikku
				0.0097	Vahemusrahv
				0.0244	Vanus
				0.0602	Sugu
				0.0561	Haridus
				-0.0086	Toostaaz
				0.0102	Kinnisvara
				0.1911	Laenureiting
				-0.0238	EsmakLaenaj
				0.2477	1_LaenuSuuru
				-0.0396	1_Sissetulek
				1.0000	1_LaenuPikku

Lisa 2. Regressioonanalüüsi mudel I (Gretl)

Model 1: Heteroskedasticity-corrected, using observations 1-69902
Dependent variable: Intress

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	-6.21014	0.274823	-22.60	<0.0001	***
Sugu	1.17012	0.0319699	36.60	<0.0001	***
Vahemusrahvus	0.334123	0.0373869	8.937	<0.0001	***
Vanus	0.400798	0.0516746	7.756	<0.0001	***
Haridus	0.389362	0.0394401	9.872	<0.0001	***
Kinnisvara	0.329243	0.0320882	10.26	<0.0001	***
Toostaz	0.161081	0.0357392	4.507	<0.0001	***
l_LaenuPikkus	-0.545070	0.0366003	-14.89	<0.0001	***
l_LaenuSuurus	-0.146793	0.0178971	-8.202	<0.0001	***
Laenureiting	6.73599	0.0126853	531.0	<0.0001	***
EsmakLaenaja	0.943927	0.0326705	28.89	<0.0001	***
l_Sissetulek	0.812520	0.0330193	24.61	<0.0001	***

Statistics based on the weighted data:

Sum squared resid	201291.8	S.E. of regression	1.697093
R-squared	0.815524	Adjusted R-squared	0.815495
F(11, 69890)	28087.90	P-value(F)	0.000000
Log-likelihood	-136153.0	Akaike criterion	272329.9
Schwarz criterion	272439.8	Hannan-Quinn	272363.8

Statistics based on the original data:

Mean dependent var	26.10322	S.D. dependent var	10.87772
Sum squared resid	1431054	S.E. of regression	4.525020

Lisa 3. Regressioonanalüüsi mudel II (Gretl)

Model 2: Heteroskedasticity-corrected, using observations 1-69902
 Dependent variable: Intress

	<i>Coefficient</i>	<i>Std. Error</i>	<i>t-ratio</i>	<i>p-value</i>	
const	0.0441682	0.635177	0.06954	0.9446	
Sugu	0.537911	0.0791651	6.795	<0.0001	***
Vahemusrahvus	0.876062	0.0904500	9.686	<0.0001	***
Vanus	2.60814	0.123687	21.09	<0.0001	***
Haridus	2.47187	0.0942143	26.24	<0.0001	***
Kinnisvara	1.90350	0.0792973	24.00	<0.0001	***
Toostaaz	0.796463	0.0881994	9.030	<0.0001	***
l_LaenuPikkus	4.34865	0.0657632	66.13	<0.0001	***
l_LaenuSuurus	-0.517430	0.0448630	-11.53	<0.0001	***
EsmakLaenaja	2.75593	0.0784681	35.12	<0.0001	***
l_Sissetulek	1.21409	0.0797630	15.22	<0.0001	***

Statistics based on the weighted data:

Sum squared resid	184528.7	S.E. of regression	1.624880
R-squared	0.120440	Adjusted R-squared	0.120314
F(10, 69891)	957.0320	P-value(F)	0.000000
Log-likelihood	-133114.0	Akaike criterion	266249.9
Schwarz criterion	266350.6	Hannan-Quinn	266281.0

Statistics based on the original data:

Mean dependent var	26.10322	S.D. dependent var	10.87772
Sum squared resid	7511440	S.E. of regression	10.36695

Lisa 4. VIF (Gretl)

Variance Inflation Factors

Minimum possible value = 1.0

Values > 10.0 may indicate a collinearity problem

Sugu 1.052

Vahemusrahvus 1.019

Vanus 1.202

Haridus 1.078

Kinnisvara 1.123

Toostaaz 1.116

l_LaenuPikkus 1.130

l_LaenuSuurus 1.163

Laenureiting 1.105

EsmakLaenaja 1.070

l_Sissetulek 1.158

Lisa 5. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina Marko Mitt

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

SOTSIAALSETEL TUNNUSTEL PÕHINEV DISKRIMINEERIMINE BONDORA
LAENUPLATVORMIL,

mille juhendaja on Mari-Liis Kukk,

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

1.2 üldsusele kättesaadavaks tegemiseks TalTechi veebikeskkonna kaudu, sealhulgas TalTechi 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.

¹*Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil.*