

TALLINNA TEHNIKAÜLIKOOL
Majandusteaduskond
Majandusanalüüsi ja rahanduse instituut

Erle Leppikson

**KÕRGHARIDUSE JA KESKMISE PALGA VASTASTIKUNE
SEOS OECD RIIKIDE NÄITEL**

Bakalaureusetöö

Õppekava TAAB, peeriala rakenduslik majandusteadus

Juhendaja: Jelena Matina, Lektor

Tallinn 2021

Deklareerin, et olen koostanud bakalaureusetöö 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 8031 sõna sissejuhatusest kuni kokkuvõtte lõpuni.

Erle Leppikson

(allkiri, kuupäev)

Üliõpilase kood: 185053TAAB

Üliõpilase e-posti aadress: erle.leppikson@gmail.com

Juhendaja: Jelena Matina, Lektor:

Töö vastab kehtivatele nõuetele

.....

(allkiri, kuupäev)

Kaitsmiskomisjoni esimees:

Lubatud kaitsmisele

.....

(nimi, allkiri, kuupäev)

SISUKORD

LÜHIKOKKUVÕTE	5
SISSEJUHATUS	6
1. KÕRGHARIDUSE JA KESKMISE PALGA VAHELINE SEOS	8
1.1. Haridus kui inimkapitali komponent	8
1.2. Hariduse ja sissetulekute vaheline seos	10
1.2.1. Hariduse ja sissetulekute vaheline seos makro tasandil	10
1.2.2. Hariduse ja sissetulekute vaheline seos mikro tasandil	13
2. METOODIKA JA ANDMED	15
2.1. Kasutatavad andmed ja allikad	15
2.2. Meetod ja mudeli püstitus.	17
2.3. Haridustasemed ja keskmine töötasu OECD riikides	18
3. ANALÜÜSI TULEMUSED JA JÄRELDUSED	26
3.1. Keskmise palga seos haridusega	26
3.1.1. Hajuvusdiagrammi ja korrelatsioonimaatriks	26
3.1.2. Keskmise palga seos kõrghariduse osatähtsusega vanuserühmade järgi.	28
3.2. Kõrghariduse seos keskmise palgaga	31
3.2.2. Kõrghariduse seos keskmise palgaga.	31
3.3. Järeldused ja hüpoteeside paikapidavus	34
KOKKUVÕTE	36
SUMMARY	38
KASUTATUD ALLIKATE LOETELU	40
LISAD	43
Lisa 1. OECD riike kirjeldav statistika	43
Lisa 2. Sõltuv muutuja keskmine töötasu, algmudel (16-24)	46
Lisa 3. Sõltuv muutuja keskmine töötasu, lõplik mudel (16-24)	47
Lisa 4. Sõltuv muutuja keskmine töötasu, algmudel (25-54)	48
Lisa 5. Sõltuv muutuja keskmine töötasu, lõplik mudel (25-54)	49
Lisa 6. Sõltuv muutuja keskmine töötasu, algmudel (55-65)	50
Lisa 7. Sõltuv muutuja keskmine töötasu, lõplik mudel (55-65)	51
Lisa 8. Sõltuv muutuja kõrghariduse osakaal, algmudel (16-24)	52
Lisa 9. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (16-24)	53
Lisa 10. Sõltuv muutuja kõrghariduse osakaal, algmudel (25-54)	54
Lisa 11. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (25-54)	55
Lisa 12. Sõltuv muutuja kõrghariduse osakaal, algmudel (55-65)	56
Lisa 13. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (55-65)	57

LÜHIKOKKUVÕTE

Käesoleva bakalaureusetöö eesmärgiks on leida, kas kõrghariduse ning sissetulekute vahel esineb vastastikune seos. Töös koostatud analüüs on tehtud 2018 aasta andmete põhjal ning valimis oli 22 OECD riiki. Autor uurib seose olemasolu erinevates vanuserühmades.

Töö eesmärgi saavutamiseks analüüsib autor OECD riikide ristandmeid. Antud töös kasutab autor analüüsi läbiviimisel makroandmeid. Valimisse osutasid riigid, mille kohta olid andmed antud ajaperioodil kättesaadavad. Autor viib läbi analüüsi ökonomeetriapakettis Gretl, kus viiakse läbi vajaminevad testid ning leitakse mudelite lõplikud kujud. Töös on esitatud andmed kolmes erinevas vanuserühmas, milleks on 16-24, 25-54 ning 55-65. Muutujateks töös on keskmine palk, tööhõive määr, tervise hinnang „hea või väga hea“, põhihariduse-, keskhariduse-, ning kõrghariduse näitajad. Sõltuvad muutujad antud töös on vastavalt keskmine palk ning kõrghariduse näitaja.

Autor viis läbi korrelatsioon-, ja regressioonanalüüsi, mille tulemusena selgus, et eksisteerib ühesuunaline keskmise palga ja kõrghariduse vaheline seos vanuserühmades 25-54 ja 55-64 aastat. Seost keskmise palga ja kõrghariduse vahel ei õnnestunud tuvastada.

Võtmesõnad: keskmine palk, kõrgharidus, OECD riigid, PIAAC uuring, ristandmed

SISSEJUHATUS

Haridust on pidanud inimesed alati tähtsaks, kuid eriti oluliseks on see saanud just tänapäevases globaliseerunud maailmas. Haridusele ja teadusele kulutatavad summad moodustavad kokku suure osa ühe riigi eelarvest, seega on oluline, et iga kulutus oleks maksimaalselt efektiivne. Tänapäevases maailmas on samuti esiplaanile seatud sissetulekud, sealhulgas majandus ehk mida edukam on riik, seda kõrgem on ka inimeste töötasu. Kõrgem haridustase riigis mõjutab otseselt ka inimkapitali hulka. Selle tõttu on oluline, et rahvas oleks võimalikult efektiivne ehk omaks võimalikult palju inimkapitali, et suurendada ka riigi majanduslikku võimekust ja töötasu. Antud teema on oluline, eriti Euroopas, kus paljudes riikides on kehtestatud tasuta kõrgharidus, mis tõstab samuti nii kõrgharitud inimeste arvu kui ka kulutusi antud valdkonnas. Aktuaalne on see ka Eestis, kus on tekkinud suur avalik debatt teemal, kas tasuta kõrgharidus on õigustatud või mitte.

Inimese haridustase näitab, kui palju inimkapitali ta omab. Mida kõrgem on inimkapital, seda suurem on võimalus leida kõrgemapalgalist tööd. Sama oluline on kõrgema inimkapitali olemine ka ühiskonna tasandil. See aitab kaasa näiteks elutaseme tõusule ja tööhõive kasvule. Maailm areneb pidevalt ning kuna tehnoloogia tase iga aastaga ainult paraneb, siis selle tõttu on ka palgaerinevused erinevate haridustasemetega vahel suurenenud. Sellise arengu tõttu vajavad tööandjad kvalifitseeritud töötajad ning see tähendab omakorda, et nõudmised töötajate haridustasemele on kasvanud. (Rõõm 2007)

Töö eesmärgiks on uurida kõrghariduse ja keskmise palga vahelise vastastikuse seose tugevust OECD riikide, erinevate vanusegruppide vahel. Autor eeldab, et kõrgharidusel on keskmisele palgale positiivne mõju ning samuti on kõrgemal keskmisel palgal positiivne mõju indiviidide haridustasemele.

Töös kasutuses olevad andmed on võetud OECD riikide põhjal. Kasutusele võetakse andmed aastal 2018. Piirangud valimile seab andmete puudulikkus mõne OECD riigi kohta. Samuti võetakse kasutusele 2018. aasta keskmine töötasu OECD riikides. Andmed on töös analüüsitud vanusegruppide lõikes.

Bakalaureusetöös on püstitatud järgmised hüpoteesid:

1. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 16-24 esineb positiivne vastastikune seos.
2. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 25-54 esineb positiivne vastastikune seos
3. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 55-65 esineb positiivne vastastikune seos
4. Seose tugevus kõrghariduse ja keskmise sissetuleku vahel on erinevat vanusegruppide puhul erinev.

Töö on jaotatud kolme peatükki. Esimene peatükk annab ülevaate varasemast teoreetilisest ning empiirilisest kirjandusest. Teoreetiline osa omakorda on jaotatud kaheks alapeatükiks, kus antakse ülevaade sissetulekutest ning haridusest. Teises peatükis kirjeldatakse antud töös kasutatavat analüüsimeetodit ning empiirilist uuringut. Viimases peatükis kirjeldatakse töös läbi viidud analüüsi tulemusi ning tehakse nende põhjal järeldusi.

1. KÕRGHARIDUSE JA KESKMISE PALGA VAHELINE SEOS

Käesolevas peatükis antakse varasemale kirjandusele tuginev teoreetiline ja empiiriline ülevaade haridusest, sissetulekust ning hariduse ja sissetulekute vahelise seose kohta avaldatud seisukohtadest.

1.1. Haridus kui inimkapitali komponent

Uurimuses (Blundell et al. 2005) on välja toodud, et inimkapital koosneb kolmest komponendist- esimeseks komponendiks on kaasasündinud omadused ja oskused, teiseks on erinevad koolitused ja programmid, kus omandad oskused õppimise käigus ning viimaseks on oskused ja teadmised, mida omandatakse töö tegemise käigus. Lanzi (2007) toob samuti oma uurimuses välja, et inimkapital koosneb erinevatest komponentidest ning, kui on soov inimkapitali kasvatada tuleb sellesse investeerida. Woodhall (1987) väidab, et investeeringuks inimkapitali loetakse näiteks erinevaid koolitusi, investeringuid haridusse ning samuti läheb sinna arvesse kõik, mis võib tulevikus suurendada sissetulekuid.

Kõrgem hariduse tase tagab tööturul madalama töötuse ohu. Selle põhjusena võib välja tuua selle, et ettevõtted on haritud töötajate koolitamisele kulutanud rohkem raha ja on selle kaudu töötajaga rohkem seotud. Samuti on leitud, et haritumate inimeste tööotsingu aeg on lühem. Üheks põhjuseks võib olla see, et sellistel töökohtadel on vaja kindlate oskustega inimesi ning ettevõtted on antud töötajate otsimisel initsiatiivikamad. Kõrgema haridustasemega töötajad otsivad ka ise tööd põhjalikumalt ning loodavad leida pikaajalisi ning enda jaoks sobivaid töökohti. (Mincer 1991)

Haridus on aastatega omandamas järjest olulisemat rolli ühiskonnas, kuna haridus on kasulik nii indiviidile endale, kui ka ühiskonnale. Haridust saab vaadelda ka, kui investeringut inimkapitali. Uurimuses on välja toodud, et panustamine inimkapitali küll kasvatab töötajate produktiivsust, kuid väga oluline on ka, et oleks olemas töökohad rohkem kvalifitseeritud töötajatele. (Levin, Kelley 1994) Teises uurimuses on autorid välja toonud, et kiiremini arenevas tehnoloogia

tööstuses tasub üha enam ära parema haridustasemega töötajad. On välja toodud, et kui töötajate haridustase on kõrgem, seda paremini kohanevad nad tehnoloogia arenguga ning on selle tõttu ka produktiivsemad. (Blundell et al. 2005)

Haridusel on väga oluline roll inimkapitalis. Hariduse kaudu koguvad inimesed erinevaid teadmisi ning oskusi, mida saab vaadelda, kui investeringut ettevõtte produktiivsusesse. Mida haritum on inimene, seda produktiivsem on ta ettevõttele ja selle kaudu on võimalik inimesel teenida kõrgemat palka. Haridustaseme kaudu on võimalik hinnata võimaliku majandusliku tulu. (Robeyns 2006)

Uuringus on välja toodud, et inimesed erinevad kahte varianti võimete poolest. Ühelt poolt inimene pärib võimed ja oskused, aga teisalt on oskused, mida omandatakse näiteks tööl käies. Inimkapital aga käsitleb, just neid elu jooksul omandatavaid oskuseid ja võimeid, mida arendatakse näiteks hariduse, koolituste või tööl omandatud kogemuste kaudu. Kuid nagu akumulatsioon teatud aja möödudes kapital, siis akumulatsioon ka inimkapital. Vananedes tervis halveneb ning seeläbi väheneb inimkapital, mingil määral saab seda parandada, kuid mitte lõpmatuseni. (Mincer 1981)

Uurimuses on välja toodud, et järjest olulisem on põhjalikult tunda sissetuleku tähendust. Hewett (1925) on leidnud, et nii palju, kui majandusteadlased on kirjutanud sissetuleku teemal, võib nad jagada kahte üldgruppi, ühed on need, kes määratlevad sissetulekut kui teenuse voogu rikkusest ja inimestelt, ning teised, kes kasutavad seda mõistet kaupade ja teenuste voona.

Töös liigitati sissetulekute saajad vastavalt vanusegruppidesse, kuna iga vanuserühma töölased oskused on seoses vanusega. Uuringus leiti, et sissetulekute erinevused kasvavad koos vanusega. Samuti on leitud, et mida vähem oskusi on tööl vaja omandada, seda väiksem on ka töötasu. Vanusega saab mõõta kogemuste omandamist ning samuti bioloogilist arengut kuni aga selle languseni. Vanus peegeldab tootlikkuse kasvu aga nii kaua, kuni jõutakse hetkeni, kus bioloogiline langus hakkab tootlikkust mõjutama negatiivselt. (Mincer 1958)

Uuringus on välja toodud, et hariduse ja majanduskasvu vahel on olemas seos. Kui riigis on rohkelt hea haridustasemega inimesi, siis majandus õitseb. Kui riik aga soovib majanduskasvu kiirendada, siis on mõttekas teha investeringuid haridusse ehk haridusel on väga oluline roll majanduskasvus. (Gillies 2017)

Ridell ja Song (2011) on oma uurimuses, mis on läbi viidud USA näitel, leidnud, et hariduse ja töötuse vahele esineb positiivne seos. Kui hariduse tase riigisiselt kasvab, väheneb antud riigis ka töötuse tase. Autorid tõid välja, et iga täiendav hariduse aasta võib suurendada tööle asumise tõenäosust kuskil 4,7%. Antud järeldustele jõuti kasutades regressioonanalüüsi. Hariduse mõju töötuse tasemele uurisid ka Núñez ja Livanos (2010) ning nemad uurisid seost Euroopa näitel. Autorid uurisid nii akadeemilise kraadi, kui ka teadusvaldkonna mõju töötuse määrale. Nad jõudsid samuti järeldusele, et kõrgharidus suurendab võimalusi ning kasvatab selle kaudu tööhõive võimalusi. Nad vaatlesid hariduse mõju töötusele nii pikaajaliselt kui ka lühiajaliselt.

1.2. Hariduse ja sissetulekute vaheline seos

1.2.1. Hariduse ja sissetulekute vaheline seos makro tasandil

Canton (2007) toob oma töös, mis on läbi viidud makroandmete põhjal, välja, et kui tööjõu keskmises haridustasemes toimub kasv aasta võrra, siis inimkapitali tase võib kasvada 11% kuni 15%. Samuti on autor seostanud hariduse mõju ka tehnoloogia arenguga. Töös kasutati andmeid SKT elaniku kohta, SKT töötaja kohta, SKT tunni kohta, töötatud tunnid, investeeringuid, rahvastiku kasvu ning hariduse taset kahes erinevas vanuse grupis.

Makroandmetel viidi läbi ka järgnev uuring, kus uuriti, kuidas mõjutab haridustaseme tõus palgainflatsiooni ning töötuse määra tasakaalu. Töös kasutati analüüsi läbiviimiseks töötuse määra erinevates vanusegruppides, tööjõu osakaalu erinevate haridustasemete järgi ning samuti vaadeldi töötuse määra erinevate haridustasemete järgi. Uuringus jõuti järeldusele, et kui haridustase kasvab võib see suurendada palgasurvet sellele rühmale, kelle palk on suuremas sõltuvuses töötuse määraga, kuid samas võib haridustaseme kasv vähendada palgasurvet ning viia tööjõu koosseisu madalama töötuse määra tasakaalupunkti. (Daly *et al.* 2007)

Uuringus võrreldi Eesti andmeid Läti, Leedu, Soome, Rootsi ning Taani andmetega. Töös hinnati, millised on olulised komponendid, et madala haridustasemega täiskasvanud suunduksid tagasi tasemeõppesse, ning kuidas tuleks korraldada põhi- ja üldhariduse korraldamist. Välisriikide andmete kohta tehakse järeldused ning nende kaudu tehakse töös soovitusel, et arendada Eesti täiskasvanute üldhariduse taset. Töös kasutati antud soovitudele jõudmiseks riikide makroandmeid. Näiteks kasutati töös aastaseid kulusi keskharidusele õpilase kohta ning samuti kasutati andmeid põhiharidusega inimeste osakaalu kohta riigis. Veel vaadeldi riikide tööhõive määra ning bruto

palka. Töös kasutatud andmete põhjal koostasid töö autorid mitmeid soovitusi. Näiteks tuleks edendada koostöid erinevate institutsioonide vahel, mille kaudu oleks täiskasvanutel rohkem võimalusi haridusse tagasi pöördumiseks. Samuti aitaks sellele kaasa stabiilne riigipoolne rahastus. (Müürisepp *et al.* 2017)

Kõrgharidus on oluline komponent, et kasvatada inimkapitali ning selle kaudu samuti sissetulekuid. Väga oluline on aga jälgida iga haridustaseme mõju eraldi. Jälgides haridustasemeid eraldi on võimalus hariduspoliitikat kujundada vastavalt oludele riigis. Aafrika põhjal viidi läbi aastatel 1960-2000 paneeländmete analüüs ning jõuti järeldusele, et Aafrikas suunatakse päris suur osa ressursse kõrgharidusse, kuid see on endaga kaasa toonud selle, et kõrgharitud rändavad aga riigist välja. Aafrika peaks astuma erinevaid samme, et seda väljarännet peatada, kuna nende kõrgema inimkapitaliga indiviidide lahkumine mõjutab üldpildis ka majandust ning majanduskasvu. Antud töös kasutati majanduskasvu näitajana SKP reaalkasvu elaniku kohta ning muutujateks olid veel rahvastiku kasvumäär, amortisatsioonimäär ning tehnika areng, kodusõja esinemine riigis ning haridus oli esitatud kolmes erinevas haridustasemes. (Gyimah-Brempong *et al.* 2006)

Majanduskasvu seost inimkapitaliga on uuritud ka Türgi näitel. Uuringus viidi läbi mitmeid analüüse, kasutades andmeid aastatest 1937-1996, et leida antud seose olemasolu. Analüüsides kasutati makroandmeid. Majanduskasvu näitajana kasutati reaalsest SKP-d ning haridust kirjeldati sellega, palju inimesi suundub erinevaid haridustasemeid omandama ehk kui palju inimesi läheb näiteks ülikooli. Antud uuringus jõuti järeldusele, et investeerimine kesk- ja põhiharidusse mõjub positiivselt sissetulekute suurenemisse. (Sari, Soytaş 2006)

Uuringus, mis on läbi viidud Ameerika Ühendriikide näitel, on välja toodud, et haridustase on seotud sissetulekute ja samuti säästmisega. Kõrgema haridustasemega indiviidid teenivad suuremat sissetulekut ning samuti on neil rohkem likviidseid varasid, mis on toeks raskematel aegadel. Nad hoiavad võla taset madalamal ning selle kaudu suudavad nad kindlustada tuleviku. Antud töös kasutati andmeid inimeste säästmise kohta, inimeste sissetulekut erinevate haridustasemete järgi ning töötust. (Wolla, Sullivan 2017) Säästmise puhul on aga väga oluline roll finantsharidusel. Varasemate uuringute ja kirjanduse põhjal on leitud, et kui puuduvad teadmised ei pruugi säästmisel olla kasu tuleviku kindlustamisel. Sääste on võimalik kasvatada läbi investeerimise, kuid oluline on selle puhul investeerimiskäitumine. Et säästudest saada maksimaalne kasulikkus on oluline arendada finantsharidust. (Atkinson *et al.* 2015)

Haridustase ühtlustab palgalõhet rikaste ja vaeste vahel. Uurimuses, mis põhineb makro andmetel, on leitud, et OECD riikide 10% rikkaima elanikkonna sissetulek on lausa 9,5 korda kõrgem, kui vaeseima 10% . Võrreldes seda lõhet 1980. aasta andmetega, on see kasvanud. Haridus on aga investering ning selle tõttu sõltub haridus inimese praegusest sissetulekust. Töös on välja toodud, et vaesematel inimestel ei pruugi olla piisavalt raha, et võimaldada endale hariduse omandamist. Selle tõttu on vaestel raskem endale head haridust hankida ning nende tulevased sissetulekud võivad kannatada. Oluline on riigis pöörata tähelepanu võrdsetele võimalustele hariduse omandamisel nii rikaste, kui ka vaeste vahel. Selline lähenemine aitab kaasa majanduskasvule ning edendab vaesemate inimeste tööhõive määra. Töös kasutati ebavõrdsuse näitajaks Gini koefitsienti ning inimkapitali näitajaks on 15-64 aastaste inimeste keskmine kooliaastate arv. (Cingano 2014)

Lähtudes varasematest uuringutest on leitud, et hariduse ja sissetulekute vahel esineb seos. Inimkapitali tasemel on väga oluline roll sissetulekute määratlemisel. Samuti on leitud, et sissetulekud mõjutavad ka hariduskulutusi ehk hariduse tase mõjutab sissetulekute suurust ning sissetulekute suurus mõjutab hariduse taset. Uuringus vaadeldi riigi sissetulekut kogutoodangu funktsiooni järgi. Antud funktsiooni tulemuseks oli seos, kus riigi sissetulekud sisaldasid ka haridus kulusid, sest antud muutujal on oluline roll inimkapitali kvaliteedis. (Tolley, Olson 1971) Peamiselt saab vaadelda inimkapitali taset, kui jälgida riiklike kulutusi haridusele. Mida kõrgemad on riiklikud kulutused haridusele, seda kõrgem on riigis inimkapitali tase. Oluline on pöörata tähelepanu hariduspoliitikale, kuna inimkapitali areng on oluline, sest see mõjutab indiviidide sissetulekuid. SKT elaniku kohta ning noorte osakaal kogu riigi elanikkonnast on näitajad, mille kaudu saab vaadelda kulutusi haridusele. Autorid leidsid, et kui sissetulekud kasvavad, siis kasvavad ka kulutused haridusele. Antud järeldustele jõuti võrrandi kaudu, mis koosnes kolmest muutujast. (Balázs *et al.* 2020)

Uuringus on leitud, et haridustaseme ja sissetulekute vahel on seos, kuid sissetulekuid võib mõjutada ka sugu või hoopiski rass. 1979-2000 aastal kogutud andmete põhjal, milleks oli näiteks nädala mediaan palk, läbiviidud analüüsis kasvas iga rühma haridustase, kuid sissetuleku kasvud nende rühmade vahel olid erinevad. Mustanahaliste naiste, meeste ja Hispaania naiste sissetuleku kasvutempo oli aeglasem, kui valgenahalistel. Põhjuseid sellele võib olla mitmeid, näiteks diskrimineerimine või siis erinevate inimrühmade hariduse kvaliteedi erinevus. (Bradbury 2002)

Mida haritum on inimene seda kõrgemat palka on võimalik tal teenida. Mõningad uuringud näitavad, et keskharidus on selle lõhe vähendamisel üks olulisemaid haridustasemeid ehk kui riigisisesele oleks soov vähendada rikaste ja vaeste vahe olevat lõhet tuleks suunata vahendeid keskharidusse. Uuringus koostati ökonomeetriline mudel, mis koosnes ebavõrdsuse näitajast, haridust mõõtvast näitajast, teiste muutujate selgitavast vektorist ning veaterminist. Uurimuse tulemusena leiti, et haridusel on oluline roll sissetulekute kujunemisel. Rikastel oli hariduse mõju sissetulekutele aga väiksem kui vaestel. (Abduallah *et al.* 2013)

1.2.2. Hariduse ja sissetulekute vaheline seos mikro tasandil

Haridus on omandamas väga olulist rolli majanduskasvus. Mida rohkem üksikisikud haridusse investeerivad, seda edukamad on nad ise, kuid samuti edendab see ka majandust üldiselt. (Gillies 2017) Riigi majanduskasvuks on vaja erinevate oskuste ja teadmistega inimkapitali. Kui riigis on inimesi, kes omavad piisavalt oskusi erinevates valdkondades, tagab see riigi majanduspoliitika edukuse, seda on kinnitanud uuringud nii arenenud riikides kui ka arenguriikides. Malaisia näitel on leitud, et haridusel ja inimkapitali tasemel on oluline roll ka majanduskasvus. Uuringus sai kinnitust, et hariduse ja majanduskasvu vahel esineb positiivne seos ning samuti on positiivne seos inimkapitali ning majanduskasvu vahel. Sellele tulemusle jõuti kasutades Pearsoni korrelatsioonanalüüsi ning samuti viidi läbi ka regressioonanalüüs. Töös kasutatud andmed koguti küsimustiku vormis. (Islam *et al.* 2016)

Inimkapitali tase ning kõrgkoolide olemasolu mõjutavad positiivselt elukvaliteeti. On leitud, et kui linnas on olemas kõrgkool kasvatab see selles piirkonnas inimkapitali osakaalu, kuna see võimaldab kohalikele head ligipääsu kõrgharidusele ning samuti suurendab võimalust, et keskkooli lõpetaja suundub edasi kõrgkooli. Teiseks toob kõrgkooli olemasolu väljastpoolt piirkonda inimesi, kes samuti soovivad kõrgkoolis õppida. Pärast kooli lõpetamist jäävad linna elama ka mõned sisserännanud ning selle kaudu kasvab inimkapitali osakaal ning samuti paraneb elukvaliteet. (Winters 2011)

Inimeste tervis ja haridus on eduka majanduse alustalad. See on nii, sest need loovad väärtust iseenesest, kui ka tagavad inimese võimaluse töötada ja makse maksta, mis omakorda tagab vähem vaesust. Kui hea haridus ja tervis saavad normaalsuseks antud riigis, hakkab see õitsema ja tagab pikaajalise ja kestva tugeva majanduse. Töös eeldati, et eksogeensed muutujad mõjutavad lineaarselt vaesuse määra ning selle tulemusena koostati statistiline võrrand. Algsesse mudelisse toodi sisse kuus muutujat, milleks olid näiteks elanikkonna protsent, kes elavad alla vaesusepiiri,

sünnituse ja rasedusega seotud põhjustel surnud naiste arv, rahvamajanduse kogutulu elaniku kohta ostujõu pariteet. (Dao 2008)

Kui indiviid teenib perioodi jooksul rohkem, kui ta tarbib, tekib ülejääk. Seda ülejääki saab defineerida, kui säästu, kuid on väga erinevaid arvamusi, kas säästmine liigitub sissetuleku alla. Uurimuses välja toonud, et antud teemal jagunevad majandusteadlased kahte leeri. Ühed väidavad, et säästud ei kuulu sissetulekute hulka, kuna säästmine ei anna praegusel perioodil kasulikkust vaid aitab kindlustada tulevikku. Antud teadlased usuvad, et selline ülejääk kvalifitseerub kapitali alla, sest kui tarbimist ei toimu, siis ei ole see tulu. Teisel poolel on aga majandusteadlased, kes leiavad, et selline ülejääk ehk säästmine on sissetulek. Nad väidavad, et see raha küll säästetakse, aga kunagi see tarbitakse ikkagi ehk seda saab käsitleda kui säästu. Säästmine saab aset leida ainult siis, kui tekib tulu ehk sissetulek. (Hewett 1925)

Sissetulekud ja haridus mõjutavad tervist. Schnittker (2004) toob oma töös välja, et kõrgemal hariduse tasemel on positiivne mõju tervisele ning seda mõju on kõige rohkem tunda madalama sissetulekutega inimeste hulgas. Sellest saab järeldada, kellel on kõrgem hariduse tase, sel on ka suurem sissetulek, mis kokkuvõttes põhjustab väiksemat mõju tervisele. Uuringus leiti, et sissetulekute ja hariduse vahel esineb positiivne seos, kuid iga haridustaseme vahel on selle seose tugevus aga erinev. Muutujateks töös olid haridus, mis oli esitatud kolmes erinevas haridustasemes, perekonna sissetulek ning toodud oli ka tervise näitaja, kus inimesed hindasid enda tervist viie punkti skaalal. Järgnevas uuringus (Herd *et al.* 2007) on samuti leitud, et hariduse ja sissetuleku vahel on seos, mis mõjutab tervist. Antud tulemusele jõuti küsitluse käigus kogutud andmete põhjal.

Uuringus, mis viidi läbi Suurbritannia näitel, koostati koondpaneeli andmekogum, mis sisaldas, kas 16-aastane isik suitsetas või ei, haridust ning tunnipalga määra. Kasutatavad andmed olid täiskohal töötavate meeste kohta vanusevahemikus 19-65. Samuti uuriti suitsetamise seost haridustasemega. Uuringus leiti, et täiendav kooliaasta suurendas palka 4,6%. Töös välja toodud tõendite näitel saab väita, et investeringud haridusse on kasumlikud ning hariduse ja töötasu vahel esineb positiivne seos. (Dickson 2012)

2. METOODIKA JA ANDMED

Bakalaureusetöö teises pooles antakse ülevaade töös kasutatavatest andmetest ning analüüsimeetoditest. Samuti pannakse paika töös kasutatav ökonomeetriline mudel. Lisaks valitakse välja töös kasutatavad muutujad, et näha mudeli tugevust ning hinnata erinevate tegurite vahelisi seoseid. Andmete töötlemisel ja mudeli koostamisel kasutatakse programme Microsoft Excel ja Gretl. Kolmandas osas esitatakse mudeli tulemused ning tehakse antud tulemuste põhjal järeldused.

2.1. Kasutatavad andmed ja allikad

Antud peatükis tuuakse välja andmeanalüüsis kasutatavad andmed. Töös kasutatavateks teguriteks on sissetulekud ja haridustase. Sissetulekute näitajaks valis autor keskmise palga ning haridustaseme näitajaks valis autor näitaja, mis näitab täiskasvanute haridustaseme osakaalu protsentides erinevates vanusegruppides. Vanusegruppideks on 16-24, 25-54 ja 55-65. Autor koostab antud töös kaks mudelit, et hinnata haridustaseme ja sissetulekute vastastikust seost. Esimeses mudelis on sõltuvaks muutujaks keskmine sissetulek vastavalt vanusegrupile ning sõltumatuteks muutujateks on antud vanusegrupi tööhõive määr, täiskasvanute haridustaseme protsent nii põhi-, kesk-, kui kõrghariduses ning tervisenäitaja „hea või väga hea“ vastavalt mudelis olevale haridustasemele. Töös kasutatakse andmeid OECD riikide kohta aastal 2018, kuna töös kasutatakse andmeid OECD statistika andmebaasi PIAAC uuringust, mis on tehtud aasta 2018 andmete põhjal. Andmed hariduse, tööhõive määra, ning tervise kohta on pärit OECD statistika andmebaasist ning andmed keskmise töötasu kohta on pärit Eurostat statistika andmebaasist. Autori poolt töös kasutatavad andmed on ristanudmed.

Haridustasemeid tuleks vaadelda eraldi, kuna erinevate haridustasemete seos sissetulekule varieerub ning siis on võimalik näiteks riigisiselt näha, millisesse haridustasemesse oleks vaja teha täiendavaid investeeringuid. (Gymah-Brempong *et al.* 2006) Autor valis haridust kirjeldavaks näitajaks OECD riikide põhjal läbi viidud PIAAC uuringust näitaja, kus on esitatud andmed

protsentides rahvastikust, kes omavad vastavat haridustaset. Haridustasemed on esitatud kolmes tasemes ning erinevates vanusegruppides. Haridustasemed on vastavalt põhiharidus, keskharidus ning kõrgharidus. Vanusegruppe algsetes andmetes oli viis. Analüüsiks valis autor vanusegruppid, 16-24, 25-54 ja 55-56. Kuna andmed ülejäänud töös kasutatavate näitajate kohta on kolmes erinevas vanuserühmas, leidis autor vastavalt kolme vanuserühma, 25-34, 35-44 ning 45-54 aritmeetilise keskmise ning sai vanuserühma 25-54.

Sissetulekute näitajaks antud töös on OECD riikide ekvivalent sissetuleku mediaan tötötasu dollarites aastal 2018. Andmed on võetud Eurostat statistika andmebaasist. Sissetuleku näitaja hõlmab tööst saadavat sissetulekut, investeerimisest ning kinnisvarast teenitavat tulu, kodumajapidamiste vahelistest ülekannetest ning sularahas saadavaid sotsiaalseid toetusi, näiteks vanaduspension. Bradbury (2002) on oma töös leidnud, et hariduse ja sissetulekute vahel esineb positiivne seos. Antud töös kasutati sissetuleku näitajana nädalast mediaanpalka ning seost uuriti vaadeldes indiviidide sugu kui ka rassi.

Töös on üheks sõltumatuks muutujaks ka inimese tervise näitaja. Schnittker (2004) on oma töös leidnud, et haridustasemel ja sissetulekul on mõju tervisele. Mida kõrgem on indiviidi haridustase, seda rohkem mõjutab see positiivselt inimese tervist. Kui inimene omab kõrgemat ja paremat haridust, on tal suurem võimalus teenida kõrgemat sissetulekut ning samuti mõjub see positiivselt indiviidi tervisele. Schnittker võttis oma töös tervist seletavaks näitajaks nii füüsilise kui vaimse tervise. Tervist kirjeldavaks näitajaks antud töös on andmed, kus inimesed vanuses 15 ja vanemad on hinnanud oma tervist „hea või väga hea“. Andmed on esitatud kolmes erinevat haridustaset omavates gruppides, milleks on põhiharidus, keskharidus ning kõrgharidus. Andmed on esitatud protsent elanikkonnast.

Ridell ja Song (2011) on leidnud, et hariduse ja tööhõive vahel esineb seos ning antud seos on positiivne. Mida kõrgem on inimese haridustase seda paremad on indiviidi võimalused töö leidmiseks. Antud töös kasutab autor sõltumatu muutujana veel tööhõive määra. Tööhõive määra näitaja antud töös näitab tööga hõivatud protsenti vanuserühmas olevate inimeste koguarvust. Kasutatavateks vanuserühmadeks on 15-24, 25-54 ning 55-65. Näitaja on pärit OECD statistika andmebaasist.

Vaadeldes kirjeldatud andmeid eemaldas autor OECD riikide valimist riigid, mille kohta olid andmed osaliselt puudulikud. Andmeid vaadeldi aastatel 2018 ning vastavalt andmete

kättesaadavusele jäi valimisse 22 OECD riiki. Valimisse jäid järgmised riigid: Austria, Belgia, Eesti, Hispaania, Holland, Iirimaa, Itaalia, Kreeka, Leedu, Norra, Poola, Prantsusmaa, Rootsi, Saksamaa, Sloveenia, Soome, Taani, Tšehhi vabariik, Türgi, Ungari ning Ühendkuningriik. OECD on majandusliku koostöö ja arengu organisatsioon kuhu kuulub 38 liiget ning antud organisatsiooni eesmärgiks on tugevdada demokraatiat ning soodustada vabaturumajandust kõikjal maailmas.

2.2. Meetod ja mudeli püstitus.

Bakalaureusetöö eesmärgiks on leida hariduse ja sissetulekute vastastikuse seose olemasolu OECD riikide näitel. Antud töö analüüsiks valis autor muutujad ning koostas nende põhjal regressioonanalüüsi ning korrelatsioonanalüüsi. Regressioonanalüüsi käigus koostatakse lineaarne mudel. Antud mudeli põhjal on võimalik selgeks teha hariduse ja sissetulekute vastastikuse seose tugevus ja mõju. Kuna töös uuritakse näitajate vahelist vastastiku seost koostatakse kaks mudelit. Esimesena koostatakse lineaarne mudel, kus sõltuvaks muutujaks on keskmine sissetulek vastavalt vanuserühmale ning sõltumatuteks muutujateks on vastava vanuserühma tööhõive määr, põhi-, kesk-, ja kõrghariduse osakaal ning tervisenäitaja „hea või väga hea“. Teises mudelis on sõltuvaks muutujaks indiviidide kõrghariduse osakaal vastavalt vanuserühmale ning sõltumatuteks muutujateks on vastava vanuserühma keskmine töötasu, tööhõive määr, ning tervisenäitaja „hea või väga hea“. Autor kasutab töös OLS-i meetodid, ehk hariliku vähimruutude meetod. Autor koostas alloleva tabeli 1, kus tõi välja erinevate muutujate lühendid.

Tabel 1. Andmete lühendid

Näitaja	Lühend
15-24 (Tööhõive tase %)	TH1
25-54 (Tööhõive tase %)	TH2
55-65 (Tööhõive tase%)	TH3
16-24 (Kesk töötasu EUR)	KTT1
25-54 (Kesk töötasu EUR)	KTT2
55-65 (Kesk töötasu EUR)	KTT3
Hea, väga hea (madal haridusetase)	HAR1TER

Hea, väga hea (keskmine haridusetase)	HAR2TER
Hea, väga hea (kõrge haridusetase)	HAR3TER
16-24 (Põhi haridustase %)	PHAR1
25-54 (Põhi haridustase %)	PHAR2
55-65 (Põhi haridustase %)	PHAR3
16-24 (Kesk haridustase %)	KHAR1
25-54 (Kesk haridustase %)	KHAR2
55-65 (Kesk haridustase %)	KHAR3
16-24 (Kõrg haridustase %)	KOHAR1
25-54 (Kõrg haridustase %)	KOHAR2
55-65 (Kõrg haridustase %)	KOHAR3

Allikas: OECD statistika andmebaas, Eurostat (2021), autori koostatud lisa 1 põhjal.

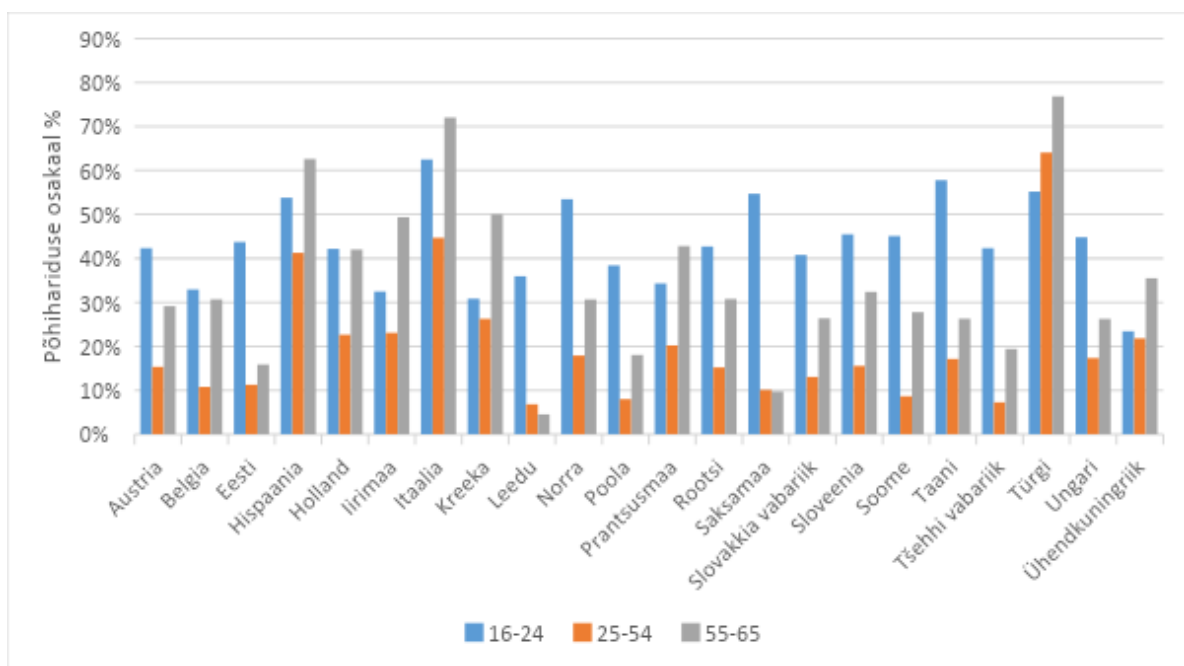
Töös kasutatavad andmed on ristanndmed ning sellest tulenevalt töös statsionaarsust ei kontrollita. Statsionaarsust saaks kontrollida, kui kasutatavad andmed oleksid aegread. Järgnevalt toob autor välja probleemid, mis võivad tekkida ristanndmete analüüsimisel:

1. Multikollineaarsus. Kui regressioonmudelis kasutatavate muutujate vahel esineb tugev multikollineaarsus, siis võivad tunnused olla statistiliselt mitteolulised, sest nende standradravigade hinnangud on suured. Tegelikult aga antud muutujad on mudelis vajalikud. (Sauga 2020)
2. Heteroskedastiivsus. Heteroskedastiivsuse leidmiseks kasutatakse töös White'i testi. Lui mudelis esineb heteroskedastiivsus siis juhuslike liikmete dispersioon mudelis ei ole konstantne (Sauga 2020).
3. Jääkliikmete normaaljaotus. Kui jääkliikmed alluvad normaaljaotusele, siis nende väärtus peaks olema väga nulli lähedane.
4. Mudeli kuju. Lõpuks viib autor läbi RESET testi, et hinnata mudeli õigsust.

2.3. Haridustasemed ja keskmine töötasu OECD riikides

Lõputöös kasutatakse haridustasemete näitajana vanuserühmades esitatud vastava haridustaseme osakaalu protsentides. Haridustasemeid on kolm, nendeks on põhi-, kesk- ning kõrgharidus. Vanuserühmad on vastavalt 16-24, 25-54 ja 55-65. Haridustase 25-54 on leitud autori poolt

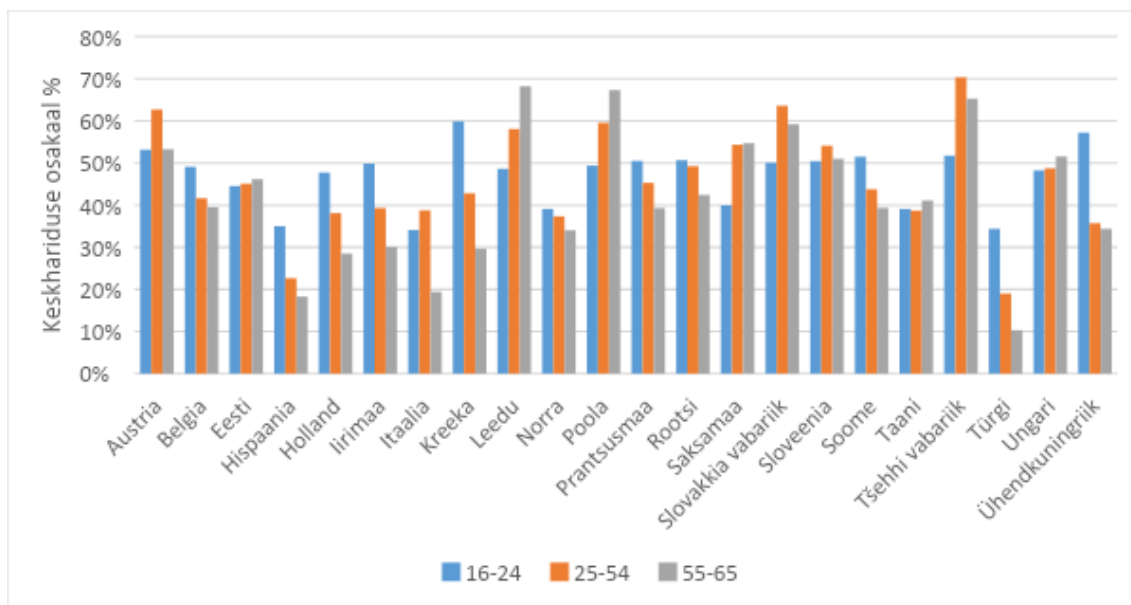
vastavalt kolme (25-34, 35-44, 45-54) vanuserühma aritmeetiline keskmine. Antud joonisel on näha põhihariduse osakaalu kolmes erinevas vanuserühmas. Kolme erinevat vanuserühma on kujutatud kolme erineva värviga. Vanuserühm 16-24 on kujutatud sinisega, 25-54 on kujutatud oranži värviga ning 55-65 on kujutatud halli värviga. Jooniselt 1 on näha, valimisse jäänud OECD riikide põhihariduse osakaalu vastavalt vanuserühmale. Põhiliselt on põhihariduse osakaal kõige suurem vanuserühmas 16-24. Põhihariduse tase vanuserühmas 16-24 on kõige kõrgem Itaalias vanuserühmas 25-54 ning vanuserühmas 55-65 aga kõige kõrgem Türgis. Jooniselt on näha, et põhihariduse osakaal on suuremas osas riikides kõige kõrgem vanuserühmas 16-24, selle põhjuseks võib olla see, et haridust hakatakse aastatega üha enam hindama.



Joonis 1. Põhihariduse osakaal vastavalt vanuserühmale OECD riikide põhjal läbi viidud PIAAC uuringu näitel aastal 2018

Allikas: OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetele tuginedes.

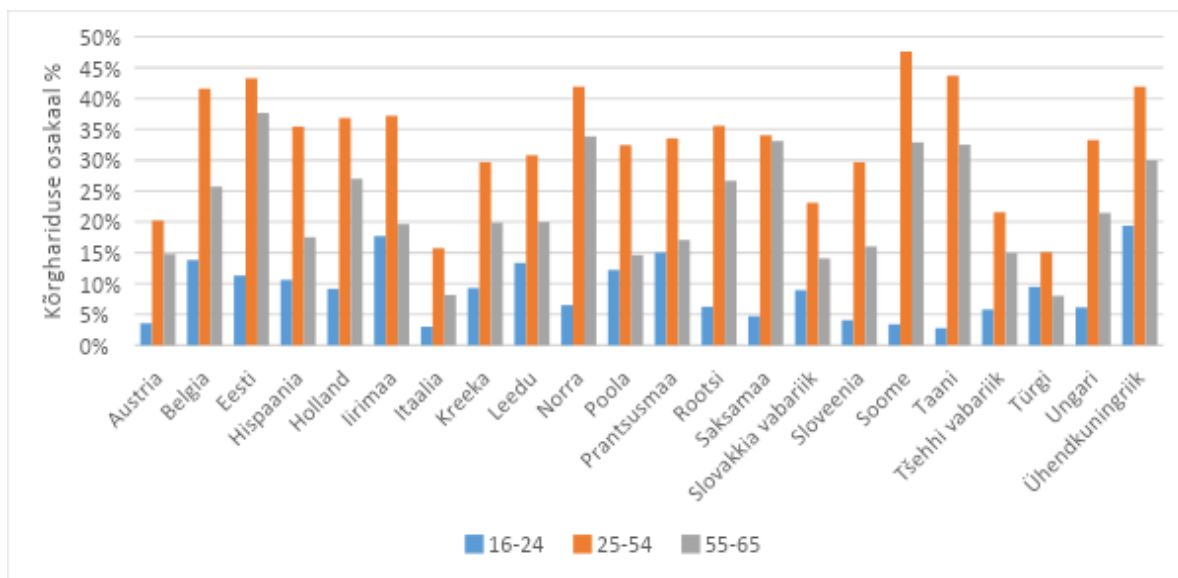
Joonisel 2 on näha keskhariduse osakaalu samuti kolmes erinevas vanuserühmas. Keskhariduse osakaal valimisse jäänud OECD riikides on erinevate vanuserühmade vahel ühtlasem, kui see oli põhihariduse puhul. Kõige kõrgem on keskharidust omavate inimeste osakaal vanuserühma 16-24 puhul Kreekas, vanuserühmas 25-54 on osakaal kõige suurem Tšehhi vabariigis ning vanuserühmas 55-65 puhul Leedus. Kõige väiksem on keskhariduse osakaal kõikides vanuserühmades aga Türgis. Keskharidust omavate inimeste osakaal vastavalt vanuserühmale, valimisse jäänud OECD riikide puhul on maksimaalselt 70%



Joonis 2. Keskhariduse osakaal vastavalt vanuserühmale OECD riikide põhjal läbi viidud PIAAC uuringul näitel aastal 2018

Allikas: OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetele tuginedes.

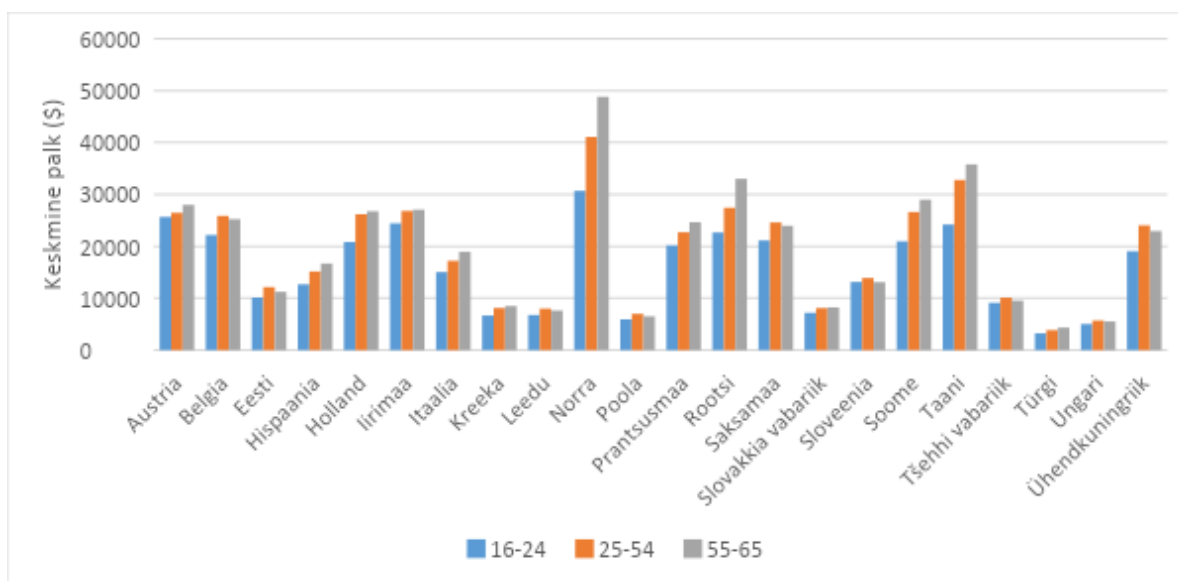
Allpool oleval joonisel 3 on näha kõrghariduse osakaalu kolmes erinevas vanuserühmas. Kõrghariduse kohta on andmed samuti esitatud protsentides. Kõige suurem osa kõrgharidust omavaid inimesi valimisse jäänud riikide puhul on vanuserühmas 25-54. Selle põhjuseks võib olla see, et haridus on varasema ajaga saanud inimestele kättesaadavaks ning samuti soovib rohkem inimesi haridust omandama minna. Kõige väiksem on kõrgharitud inimeste osakaal vanuserühmas 16-24. Autor usub, et selle üheks oluliseks põhjuseks on vanus, mille tõttu ei ole veel suur osa noori jõudnud haridust omandama minna või ei ole veel leidnud enda jaoks õiget eriala, mida õppima minna. Kõige kõrgem on kõrghariduse osakaal, valimisse jäänud riikides Soomes vanuserühmas 25-54. Kõige väiksem kõrghariduse osakaal on Itaalias ja Taanis, kus vanuserühmas 16-24 on kõrgharitud inimeste osakaal võrreldes teiste valimisse jäänud riikidega kõige väiksem.



Joonis 3. Kõrghariduse osakaal vastavalt vanuserühmale OECD riikide põhjal läbi viidud PIAAC uuringul näitel aastal 2018

Allikas: OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetele tuginedes.

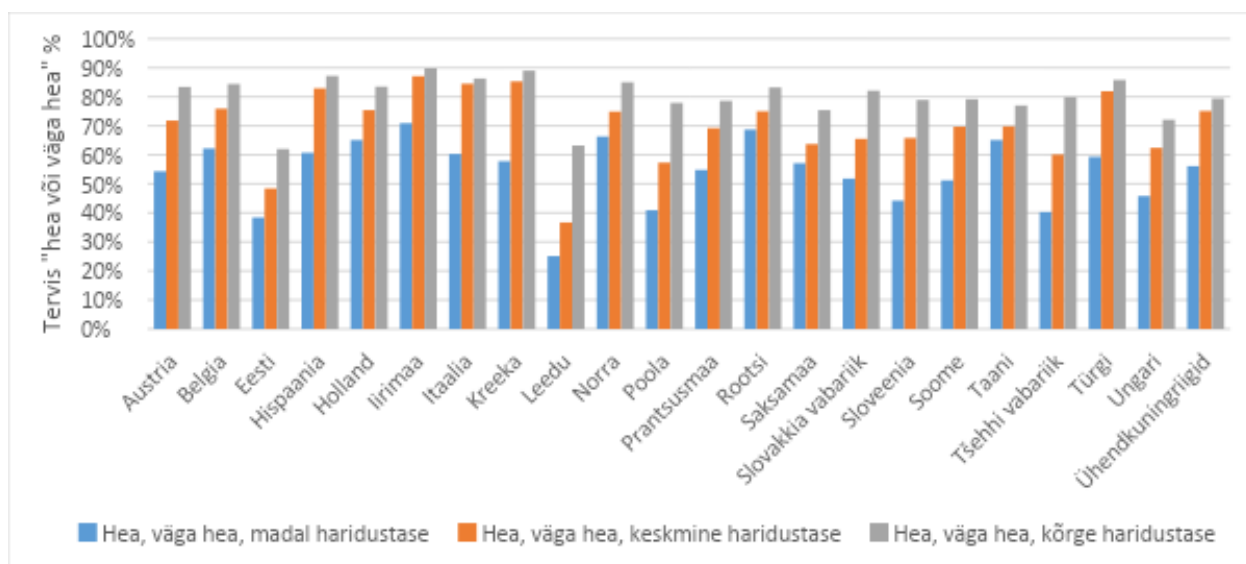
Joonisel 4 on näha OECD riikide keskmist palka, mis on vastavalt riig mediaan palk. Keskmise palga tase riigiti on väga erinev. Kõige väiksem on keskmine palk Türgis, Ungaris ning Poolas kõige kõrgem aga Norras, Taanis ning Rootsis. Jooniselt on samuti näha, kuidas erineb keskmise palga tase vastavalt vanuserühmale. Keskmise palga suurus vanuserühmas 25-54 ja 55-65 on suhteliselt sarnane. Keskmine palk on kõikide riikide puhul kõige väiksem vanuserühmas 16-24. Autor usub, et seda võib mõjutada nii omandatud kõrgeim haridustase, kui ka varasem töökogemus. Kindlasti on ka riigiti erinevaid faktoreid, mis riigi keskmist palka mõjutavad, milleks võivad olla erinevad riigisisised kriisid jne.



Joonis 3. Keskmine palk OECD riikides aastal 2018

Allikas: Eurostat andmebaas (2021), autori koostatud lisa 1 andmetele tuginedes.

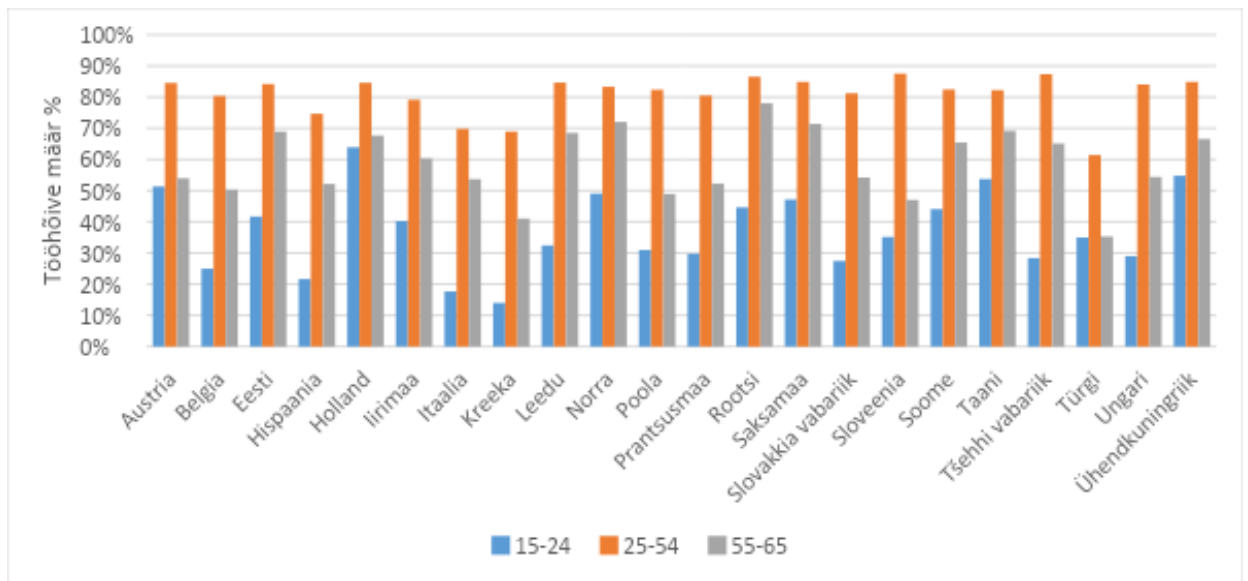
Allpool oleval joonisel 4 on kujutatud inimeste hinnangut „hea või väga hea“ enda tervisele valimisse jäänud OECD riikides. Jooniselt on näha, et kõige kõrgemalt hindavad oma tervist inimesed, kes omavad kõrgharidust ning kõige väiksem osakaal inimesi, kes hindavad oma tervist „hea või väga hea“ omavad aga põhiharidust. Samuti on jooniselt näha hinnangut oma tervisele riigiti. Valimisse jäänud OECD riikide hinnang „hea või väga hea“ oma tervisele on riigiti suhteliselt sarnane. Kõige väiksem osakaal on aga Leedus ning Eestis, kus võrreldes teistesse valimisse jäänud riikidega hindab oma tervist „hea või väga hea“ kõige väiksem osakaal inimesi. Selle põhjuseks võib olla kliima, kuna antud riikides on jahedam periood aastas pikem ning selle tõttu haigestub ka rohkem inimesi külmetushaigustesse, mis võib nende hinnangut oma tervisele muuta. Kõige suurem osakaal inimesi kes hindab oma tervist „hea või väga hea“ on Iirimaa.



Joonis 4. Inimeste hinnang enda tervisele „hea või väga hea“ protsentides OECD riikides, aastal 2018

Allikas: OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetele tuginedes.

Järgneval joonisel 5 on kujutatud tööhõive määra kolmes erinevas vanuserühmas OECD riikide näitel aastal 2018. Tööhõive määr valimisse jäänud OECD riikides on sarnane. Kõige kõrgem on tööhõive vanuserühmas 25-54 ning kõige väiksem vanuserühmas 16-24. Selle põhjuseid võib olla mitmeid, milleks võib olla omandatud kõrgeim haridustase kui ka töökogemus. Veel üheks põhjuseks, miks tööhõive määra osakaal on kõige kõrgem vanuserühmas 25-54, võib olla see, et juba suur osa noori siirdunud elama iseseisvalt ning enda elatamiseks läinud ka tööle. Vanuserühmas 55-65 võib tööhõive määra mõjutada näiteks tervis.



Joonis 5. Tööhõive määr OECD riikides vanuserühmade järgi, aastal 2018
Allikas: OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetel tuginedes.

Jooniste põhjal usub autor, et seos hariduse ja sissetulekute vahel on olemas, kuid antud jooniste põhjal ei saa järeldada, kas antud seos on vastastikune. Samuti usub autor, et kuna riigisisest on hariduspoliitika erinevad, siis sellel on ka oma roll haridusele ning hariduse tasemele OECD riikides. Hariduspoliitika riigis võib kaudselt mõjutada sissetulekuid.

Tabel 2. OECD riike kirjeldav statistika

Näitaja	Aritmeetiline keskmine	Mediaan	Standardhälve	Miinum	Maksimum
TH1 (%)	37,14	35,08	12,70	14,00	63,93
TH2 (%)	80,91	82,89	6,44	61,43	87,50
TH3 (%)	58,94	57,39	10,84	35,30	78,03
KTT1 (EUR)	15769	15044	7882	3235	30694
KTT2 (EUR)	18811	19959	9932	3859	41081
KTT3 (EUR)	19788	20946	11415	4337	48783
HAR1TER (%)	54,38	56,6	11,12	25,1	70,9
HAR2TER (%)	69,93	70,8	11,98	36,7	87,1
HAR3TER (%)	80,15	81	7,05	62,00	89,8
PHAR1 (%)	43,42	42,52	9,59	23,41	62,56
PHAR2 (%)	19,91	16,32	13,60	6,80	64,07
PHAR3 (%)	34,50	30,67	18,17	4,50	76,85
KHAR1 (%)	47,02	49,25	6,98	34,10	59,90
KHAR2 (%)	45,87	44,43	12,37	19,00	70,34
KHAR3 (%)	41,96	40,30	15,58	10,32	68,26
KOHAR1 (%)	8,93	9,04	4,72	2,77	19,37
KOHAR2 (%)	32,92	33,79	8,93	15,09	47,61
KOHAR3 (%)	22,07	19,93	8,43	8,00	37,66

Allikas: OECD statistika andmebaas, Eurostat (2021), autori koostatud lisa 1 põhjal.

Tabel 2 kujutab töös kasutatavate andmete kirjeldavat statistikat. Tabelist on näha aritmeetilist keskmist, mediaani, standardhälvet, miinimumi ja maksimumi. Tabelist on näha, et aastal 2018 oli kõige kõrgem keskmine palk vanusegrupis 55-65, kus palk oli 48783 dollarit ning kõige madalam keskmine palk oli vanuserühmas 16-24, kus keskmine töötasu oli 3235 dollarit. Selle põhjuseks võib olla, et noored inimesed ei ole jõudnud veel omandada piisavalt haridust ning samuti pole neil veel vajaliku töökogemust, mis võiks nende potentsiaalset teenitavat tulu suurendada. Samuti on näha, et keskmine töötasu on kõige väiksem vanuserühmas 16-24 ning kõige kõrgem vanuserühmas 55-65.

Tabelist on näha kolme erinevat haridustaset kolmes erinevas vanuserühmas. Andmed on protsentides ning näitavad valimisse jäänud OECD riikide vanuserühma kaupa, kui suur osakaal on kindla haridustaseme omandanud. Maksimaalselt põhiharidust omavate inimeste protsent kolmes vanuserühmas jääb 62,56%-76,85% vahele. Kõige väiksem põhiharidust omavate inimeste osakaal kolmes erinevas vanuserühmas jääb aga 4,5%-23,41% vahele. Keskkharidust omavate inimeste osakaal keskmiselt kolmes vanuserühmas jääb vahemikku 41,96%-47,02%. Kõige suurem on põhiharidust omavate inimeste osakaal vanuserühmas 55-65. Keskkhariduse ja kõrghariduse puhul on osakaalud kõige suuremad vanuserühmas 25-54.

Tabelis on välja toodud tervise näitaja, kus on andmed esitatud protsentides. Andmed on esitatud vanuses 15+ ning inimesed, kes omavad, kas põhi-, kesk-, ja kõrgharidust hindavad oma tervist „hea või väga hea“. 89,8 % inimestest, kes omavad kõrgharidust hindavad oma tervist „hea või väga hea“ ning kõige väiksem osa 25,1% inimestest hindavad oma tervist „hea või väga hea“ omavad põhiharidust. Keskmiselt on oma tervist hinnanud „hea või väga hea“ kõige suuremas osakaalus kõrgharidusega inimesed ning kõige väiksemas osakaalus põhiharidusega inimesed.

Esimesel kolmel real tabelis on näha tööhõive määra kolmes erinevas vanuserühmas. Tööhõivemäär on kõige kõrgem vanuserühmas 25-54, kus hõivatuid on 87,5%. Kõige väiksem on aga tööhõive vanuserühmas 16-24, kus hõivatus on 14%. Põhjuseks, miks tööhõive määr on kõige väiksem just vanuserühmas 16-24 võib olla see, et noored käivad alles koolis ning sõltuvad oma vanematest, mis võib mõjutada tööl käivate noorte osakaalu. Tööhõive määra keskmine, mediaan, miinimum ning maksimum on kõigi nende näitajate puhul suurim vanuserühmas 25-54. Seda võib seletada sellega, et selles vanuses olevad inimesed on kõik töörealised ning väga suures osas elavad

iseseisvalt, mille tõttu käivad ka suuremas osas tööl. Samuti on väga suur osa inimesi selles eas ka lõpetanud enda koolitee ehk selle võrra on neil rohkem aega ning suurem osa neist saab käia ka tööl.

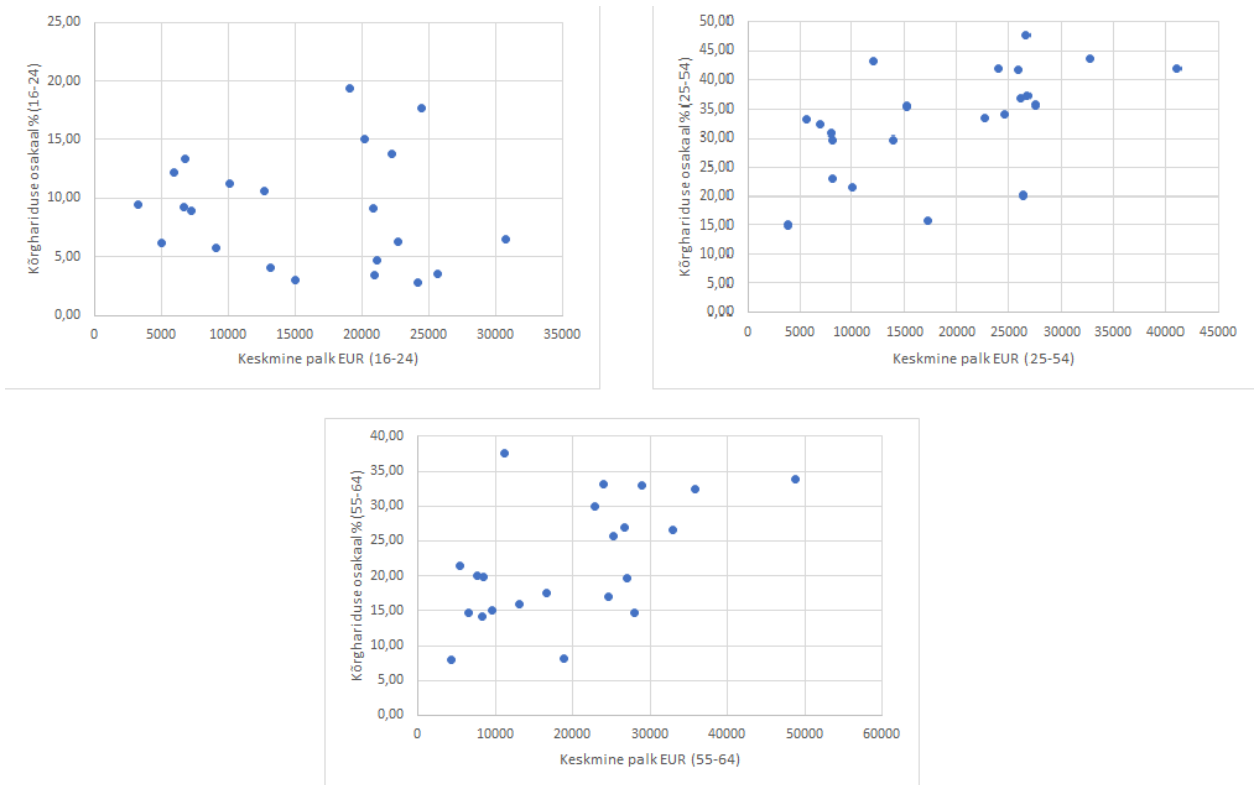
3. ANALÜÜSI TULEMUSED JA JÄRELDUSED

Kolmandas peatükis teostatakse andmeanalüüs ning tuuakse välja selle tulemused. Andmete analüüsimiseks kasutatakse ökonomeetriapaketti Gretl. Andmete ettevalmistamiseks kasutas autor Microsoft Excelit. Sissejuhatuses püstitatud hüpoteesidele vastamiseks on teostatud andmeanalüüs, kus uuritakse seost kõrghariduse ja keskmise palga vahel ning mudelisse lisatakse veel sõltumatuid muutujaid, milleks on näiteks tööhõive määr. Autor koostab kaks mudelit, et leida, kas seos on vastastikune. Antud peatükk jagatakse kolmeks alapeatükiks, kus tuuakse välja tehtud testid mõlema mudeli puhul ning tulemused ja järeldused.

3.1. Keskmise palga seos haridusega

3.1.1. Hajuvusdiagrammi ja korrelatsioonimaatriks

Hajuvusdiagramm ehk korrelatsiooniväli näitab kahe muutuja vahelist seost graafiliselt. Allpool asuval joonisel 6 on näha keskmise palga ja kolme erineva haridustaseme hajuvusdiagrammid. Y-teljel on näha kõrgharitud inimeste osakaalu ning x-teljel on näha keskmise palga suurust. Haridustaset mõõdetakse protsentides põhi-, gümnaasium- ja kõrgharidustasemega inimesed riigi rahvastikust. Hajusdiagramm näitab, et põhihariduse hajuvusdiagrammil pole konkreetset suunda, mis viitab korrelatsiooni puudumisele või nullilähedasele korrelatsioonile. Positiivset suhet kõrghariduse ja keskmise palga vahel kolmes erinevas vanuserühmas võib jälgida allpool joonisel 6 asuvalt hajuvusdiagrammilt.



Joonis 6. Keskmise palga ja kõrghariduse hajuvusdiagrammid vastavalt vanuserühmale. Allikas: Eurostat ja OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetel tuginedes.

Allpool olevas tabelis 3 on esitatud kõrghariduse ja keskmise palga korrelatsioonikordajad. Antud töö korrelatsioonimaatriksilt on näha, et mis suunaga ning kui tugev on seos erinevate vanuserühmade puhul kõrghariduse ning keskmise palga vahel. Vanuserühmas 55-65 on keskmise palga ning kõrghariduse seos olemas ning on näha, et seos on keskmise tugevusega ($r=0,551$). Antud tabelist on võimalik välja tuua, et keskmise palga ning kõrghariduse vahel esineb positiivne seos. Järgnevalt saab vaadelda keskmise palga ja kõrghariduse seost vanuserühmas 25-54, kus seos on eelnevale vanuserühmale väga sarnane ehk esineb keskmine positiivne seos antud muutujate vahel ($r=0,547$). Kõige väiksem ehk peaaegu olematu seos on kõrghariduse ja keskmise palga vahel vanuserühmas 16-24 ($r=-0,090$). Selle põhjuseid võib olla mitmeid. Autor usub, et üheks faktoriks on kindlasti see, et selles vanuserühmas on kõrgharidute arv kindlasti väiksem, kui teistes vanuserühmades ning samuti on oluline roll töötasus ka varasemal kogemusel. Vanuserühmas 16-24 olevatel indiviididel on varasem töökogemus teise vanuserühmadega võrreldes kindlasti väiksem.

Tabel 3. Korrelatsioonimaatriks

	<i>16-24 (Kesk töötasu EUR)</i>	<i>25-54 (Kesk töötasu EUR)</i>	<i>55-64 (Kesk töötasu EUR)</i>	<i>16-24 (Kõrg har %)</i>	<i>25-54 (Kõrg har %)</i>	<i>55-65 (Kõrg har %)</i>
16-24 (Kesk töötasu EUR)	1					
25-54 (Kesk töötasu EUR)	0,984103	1				
55-64 (Kesk töötasu EUR)	0,964047	0,989521	1			
16-24 (Kõrg har %)	-0,09015	-0,10632	-0,15819	1		
25-54 (Kõrg har %)	0,473536	0,547003	0,513302	0,242755	1	
55-65 (Kõrg har %)	0,501394	0,580587	0,550542	0,000448	0,868294	1

Allikas: Eurostat ja OECD statistika andmebaas (2021), autori koostatud lisa 1 andmetel tuginedes.

3.1.2. Keskmise palga seos kõrghariduse osatähtsusega vanuserühmade järgi.

Järgnevalt viib autor läbi regressioonanalüüsi, kus sõltuvaks muutujaks on keskmine palk ning sõltumatuteks muutujateks on tööhõive määr, tervisehinnang „hea või väga hea“, põhi-, kesk-, ja kõrghariduse osakaal vastavalt vanuserühmale. Algmudel on esitatud lisades (Lisa 2). Töö autor kontrollis, kas mudelis esineb multikollineaarsus. Multikollineaarsuse puhul on sõltumatud muutujad omavahel tugevamas seoses, kui sõltuv muutuja ning selle tõttu eemaldas autor mudelist muutuja või muutujad, mille multikollineaarsuse näitaja oli üle 10. Kui multikollineaarsuse koefitsient on üle kümne, siis see võib mõjutada seost statistiliselt oluliste muutujate vahel. Järgnevalt vaatas autor, millised muutujad mudelis olid statistiliselt olulised ning eemaldas mudelist muutujad, mis polnud statistiliselt olulised nivool 0,05. Kui muutuja oli statistiliselt oluline jäi see mudelisse. Esimese mudeli lõplikuks tulemuseks vanuserühmas 16-24 olid statistiliselt olulised nivool 0,05 tööhõive määr ning tervise näitaja, kus põhiharidusega inividid hindasid oma tervist „hea või väga hea“. Antud mudeli tulemused on esitatud lisades (Lisa 3). Järgnevalt viis autor läbi White'i testi heteroskedastiivsuse kontrolliks, Ramsey reset testi mudeli kuju kontrolliks ning kontrollis jääkliikmete normaaljaotust. Mudelis heteroskedastiivsust ei esinenud, Ramsey testi puhul jäi kehtima nullhüpotees ehk mudeli kuju oli õige samuti allusid jäägid antud mudelis normaaljaotusele.

Tabel 4. Seos 1 sõltuv muutuja keskmine palk vanuserühmas 16-24

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	-14940,5	5978,16	-2,499	0,0218*
TH1	291,390	92,0253	3,166	0,0051***
HAR1TER	365,683	105,123	3,479	0,0025***

Allikas: autori koostatud lisa 3 andmetel tuginedes

Kommentaar: *** oluline 0,01nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Ülevaloleval tabelis 4 on näha keskmise palga ja hariduse vahelise seose tulemusi. Mudel tuli statistiliselt oluline nivool 0,05 ning mudeli seletusvõime ehk R^2 tuli antud mudelis 60,91%. Edasi kasutati mudelis spetsifikatsioonivigade testi ehk Ramsey reset testi. Ramsey testi puhul püstitatud hüpoteesid on järgmised: $H_0: p > 0,05$ siis mudeli kuju on õige; $H_1: p < 0,05$ siis mudeli kuju on vale. Antud mudelis Ramsey reset testi tulemusel jäi kehtima null hüpotees ehk mudeli kuju oli õige. White'i testi tulemusena leidis autor, et mudelis heteroskedastiivsust ei esinenud ning samuti allusid jääkliikmed antud mudelis normaaljaotusele. Tabelist 4 on näha, et kui tööhõive määr suureneb 1% võrra kasvab keskmine töötasu 291,39\$ ning, kui madalama haridusega elanikkonnast veel 1% hakkab hindama oma tervist „hea või väga hea“, kasvab keskmine palka 365,683\$. Antud mudeli tulemusena saab väita, et keskmisel töötasul vanuserühmas 16-24 ei ole haridusega seost.

Järgnevalt viis autor samuti läbi regressioonanalüüsi, kuid seekord vanuserühmas 25-54. Antud mudelis oli sõltuvaks muutujaks keskmine töötasu ning sõltumatuteks muutujateks tööhõive määr, tervisehinnang „hea või väga hea“, põhi-, kesk-, ja kõrghariduse osakaal vastavalt vanuserühmale. Antud vanuserühma algmudel on esitatud lisades (Lisa 4). Autor kontrollis, kas mudelis esineb multikollineaarsust ja vajadusel eemaldas mudelist muutujad, mille koefitsient oli kümnest suurem. Lõpuks testis autor, kas mudelis esineb heteroskedastiivsust, mudeli kuju õigsust ning samuti testis, kas jääkliikmed mudelis alluvad normaaljaotusele. Antud mudeli tulemused on näha allolevas tabelis 5.

Tabel 5. Seos 1 sõltuv muutuja keskmine töötasu vanuserühmas 25-54

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	-101616	30588,4	-3,322	0,0038***

HAR2TER	559,311	151,717	3,687	0,0017***
TH2	833,648	313,073	2,663	0,0159**
KOHAR2	421,252	187,236	2,250	0,0372**

Allikas: autori koostatud lisa 5 andmetel tuginedes

Kommentaar: *** oluline 0,01nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Mudeli tulemuste kohaselt olid statistiliselt olulised muutujad tervise näitaja, kus keskharidust omavad inividid hindasid enda tervist „hea või väga hea“, tööhõive määr ning kõrgharitud inimeste osakaal antud vanuserühmas. Mudel oli statistiliselt oluline nivool 0,05 ning mudeli seletusvõime oli 60,57%. Ramsey testi tulemusena oli mudeli kuju õige, heteroskedastiivsust ei esinenud ning samuti allusid jäägid normaaljaotusele. Tabelist 5 on näha, et kui kõrgharidusega inimeste osakaal kasvab 1% siis keskmine töötasu kasvab 421,252\$ võrra. Tööhõive määra 1% kasvu korral keskmine töötasu aga kasvab 833,648\$ võrra ning antud vanuserühma kõrghariduse näitaja 1% kasvu korral kasvab keskmine töötasu 559,648\$. Antud mudeli tulemusena saab väita, et vanuserühmas 25-54 sõltub keskmine töötasu haridusest, täpsemalt siis kõrgharidusest. Mudeli tulemusi on võimalik vaadelda lisades (Lisa 5).

Autor viis läbi ka kolmanda analüüsi vanuserühmas 55-65, kus kasutati eelnevate mudelitega samu muutujaid, kui antud mudelis olid andmed vanuserühmas 55-65. Kolmanda vanuserühma algmudel on esitatud lisades (Lisa 6) Autor kontrollis, et kas mudelis esineb multikollineaarsust ning viis läbi White'i testi, Ramsey reset testi ning kontrollis, kas mudelis alluvad jäägid normaaljaotusele.

Tabel 6. Seos 1 sõltuv muutuja keskmine töötasu vanuserühmas 55-65

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	-107007	22678,1	-4,719	0,0002***
TH3	524,225	197,428	2,655	0,0161**
HAR3TER	1036,23	226,713	4,571	0,0002***
KOHAR3	581,922	253,386	2,297	0,0339**

Allikas: autori koostatud lisa 7 andmetel tuginedes

Kommentaar: *** oluline 0,01nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Mudelis osutusid statistiliselt olulisteks näitajateks tööhõivemäär, kõrgharidust omavate inimeste tervisenäitaja „hea või väga hea“ ning kõrgharitud inimeste osakaal. Mudel oli statistiliselt oluline nivool 0,05 ning mudeli enda seletusvõime oli 70,51%. Testide tulemuste põhjal saab väita, et

mudeli kuju oli õige, heteroskedastiivsust mudelis ei esinenud ning jäägid allusid normaaljaotusele. Ülaltoodud tabelis 6 on näha sõltuvate muutujate kordajaid, millest saab välja lugeda, et 1% kõrgharidusega inimeste osakaalu kasv toob kaasa 581,922\$ suuruse keskmise töötasu kasvu. Samuti toimub keskmise töötasu kasv 524,225\$, kui tööhõivemäär kasvab antud vanuserühmas 1% ning kui mudelis kasutatud tervisenäitaja, mis näitas kõrgharitud inimeste hinnangut oma tervisele „hea või väga hea“ indiviidide osakaal kasvab 1% siis kasvab keskmine palk aga 1036,23\$. Antud mudeli tulemusena saab väita, et töötasu ja hariduse vahel esineb positiivne seos. Antud seose mudelit on võimalik vaadelda lisades (Lisa 7).

3.2. Kõrghariduse seos keskmise palgaga

3.2.2. Kõrghariduse seos keskmise palgaga.

Järgnevalt viib autor läbi järgmise regressioonanalüüsi, kus sõltuvaks muutujaks on kõrgharidust omavate inimeste osatähtsus vanusegrupis 16-24 ning sõltumatuteks muutujateks on tööhõive määr, keskmine töötasu, tervisehinnang ning põhihariduse ja keskhariduse muutujad vastavalt vanuserühmale. Vanuserühma 16-24 algmudel on esitatud lisades (Lisa 8). Mudel tuli statistiliselt oluline nivool 0,05. Järgnevalt eemaldas autor mudelist ükshaaval muutujad, mille multikollineaarsuse näitaja oli kõige kõrgem ning vähemalt üle kümne. Näitajaid eemaldas autor mudelist ükshaaval selle hetkeni, kuni mudelisse jäänud näitajate multikollineaarsuse näitajad olid igal muutujal vähem kui 10. Seose regressioonanalüüsi mudeli tulemused on esitatud lisades (Lisa 9). Järgnevalt viis autor läbi mudeli sobivuse kontrolliks White'i heteroskedastiivsuse testi, Ramsey reset testi ning samuti kontrollis, kas jäägid mudelis alluvad normaaljaotusele.

Tabel 7. Seos 2 sõltuvaks muutujaks kõrgharidus vanuserühmas 16-24

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	91,4740	4,24503	21,55	2,65e-014***
PHAR1	-0,924940	0,0405988	-22,78	1,01e-014***
KHAR1	-0,896434	0,0557202	-16,09	3,97e-012***
KTT1	-1,48023e-05	2,45787e-05	-0,6022	0,5545

Allikas: autori koostatud lisa 9 andmetel tuginedes

Kommentaar: *** oluline 0,01 nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Mudelid olid statistiliselt olulised näitajad nivool 0,05 nii põhihariduse, kui ka keskhariduse näitaja. Antud mudeli seletusvõime ehk R^2 oli 96,98%, mis on väga kõrge. Ramsey reset testi tulemusena sai võtta vastu nullhüpoteesi ehk mudeli kuju oli õige, kuna p väärtus oli suurem kui 0,05. Samuti ei esinenud mudelis heteroskedastiivsust, kuna selle näitaja p väärtus oli samuti suurem kui 0,05. Antud mudelis jäägid aga normaaljaotusele ei allunud. Tabelist 7 on näha, et kui põhihariduse osakaal suureneb 1% kahaneb aga kõrgharidute osakaal 0,925%. Sarnaselt põhiharidusega kahaneb keskharidusega inimeste osakaalu 1% kasvu puhul kõrgharitud inimeste osakaal aga 0,896%. Keskmise palga regressioonikordaja ei osutunud mudelis statistiliselt oluliseks, seega ei õnnestunud tuvastada seost kõrghariduse ja keskmise palga vahel vanuserühmas 16-24.

Järgnevalt viis autor läbi regressioonanalüüsi vanuserühma 25-54 andmete põhjal. Vastava mudeli algmudel on esitatud lisades (Lisa 10). Antud mudelis oli sõltuvaks muutujaks kõrghariduse osakaalu näitaja ning sõltumatud muutujad tööhõive määr, tervise näitaja ning põhihariduse ja keskhariduse osakaal vastavalt vanuserühmale. Autor kontrollis mudelis ka multikollineaarsust ning mudel tuli statistiliselt oluline nivool 0,05. Autor viis läbi mudeli kontrolliks ka White'i testi, Ramsey reset testi ning samuti kontrollis, et kas jäägid alluvad normaaljaotusele. Regressioonanalüüsi tulemused on näha allpool asuvas tabelis 6.

Tabel 8. Seos 2 sõltuvaks muutujaks kõrgharidus vanuserühmas 25-54

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	97,1356	3,53017	27,52	3,68e-016***
KTT2	-8,34671e-06	4,05952e-05	-0,2056	0,8394
PHAR2	-0,975258	0,0446330	-21,85	2,08e-014***
KHAR2	-0,973276	0,0490772	-19,83	1,11e-013***

Allikas: autori koostatud lisa 11 andmetel tuginedes

Kommentaari: *** oluline 0,01 nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Statistiliselt olulised näitajad antud mudelis olid põhihariduse näitaja ning samuti keskhariduse näitaja. Mudeli seletusvõime aga tuli mudelis 97,50%. Testide tulemused aga näitasid, et mudeli kuju oli õige ning heteroskedastiivsust mudelis ei esinenud. Sarnaselt eelnevale mudelile ei allunud jäägid normaaljaotusele ka selles mudelis. Ülaltoodud tabelis 8 on näha, kuidas mõjutab põhihariduse ning keskhariduse osakaalu kasv kõrgharitud inimeste osakaalu. Kui põhihariduse osakaal kasvab 1% kahaneb kõrgharitud inimeste osakaal 0,975% ning keskhariduse puhul

kahaneb kõrgharitud inimeste osakaal 0,973%. Vanuserühmas 25-54 ei osutunud keskmise palga näitaja samuti statistiliselt oluliseks. Sellest saab järeldada, et antud vanuserühmas ei õnnestunud tuvastada seost kõrghariduse ning keskmise palga vahel. Antud mudeli tulemused on esitatud lisades (Lisa 11).

Viimasena viis autor läbi analüüsi ka kolmanda töös kasutatud vanuserühma andmetega. Viimase mudeli algmudel on esitatud lisades (Lisa 12). Sõltuvaks muutujaks kõrghariduse osakaalu näitaja vanuserühmas 55-65. Sõltumatud muutujad sarnaselt eelmisele kahele mudelile on tööhõive määr, keskmine palk, hariduse näitaja ning põhihariduse ja keskhariduse osakaalu näitajad vastavalt analüüsitava vanuserühmale. Analüüsi puhul kontrollib autor multikollineaarsust ning viib läbi White'i testi, Ramsey reset testi ning kontrollib jääkliikmete normaaljaotust. Analüüsi muutujate tulemused on näha allpool asuvas tabelis 8.

Tabel 9. Seos 2 sõltuvaks muutujaks kõrgharidus vanuserühmas 55-66

	<i>Coefficient</i>	<i>Std. Error</i>	<i>T-ratio</i>	<i>P-value</i>
const	93,8716	5,64293	16,64	2,26e-012***
KTT2	4,39776e-05	4,54847e-05	0,9669	0,3464
PHAR3	-0,954715	0,0605465	-15,77	5,57e-012***
KHAR3	-0,947049	0,0727866	-13,01	1,36e-010***

Allikas: autori koostatud lisa 13 andmetel tuginedes

Kommentaar: *** oluline 0,01 nivool, ** oluline 0,05 nivool, * oluline 0,1 nivool.

Antud mudelis, sarnaselt kahele eelneva osutusid statistiliselt olulisteks muutujateks konstant, põhihariduse osakaalu näitaja ning keskhariduse osakaalu näitaja. Mudeli seletusvõime analüüsi puhul oli 95,58%. Testidest tuli välja, et mudeli kuju oli õige ning samuti ei esinenud heteroskedastiivsust, kuid jääkliikmed antud mudelis normaaljaotusele ei allunud. Tabelis 9 on näha, kuidas erinevate haridustasemete kasv mõjutab kõrgharitud inimeste osakaalu. Kui põhiharidus kasvab 1% kahaneb kõrgharitud inimeste osakaal 0,955% ning keskhariduse puhul kahaneb kõrgharitud inimeste osakaal 0,947%. Sarnaselt kahele eelnevale mudelile ei osutunud keskmise palga näitaja ka selles mudelis statistiliselt oluliseks. Sellest saab järeldada, et vanuserühmas 55-65 ei suudetud tuvastada seost kõrghariduse ning keskmise palga vahel. Analüüsi tulemused on esitatud lisades (Lisa 13).

Kõigi kolme läbiviidud mudeli puhul jäägid normaaljaotusele ei allunud. Autor viis kõigi kolme mudeli puhul läbi Ramsey reset testi ning selle tulemusena oli kõigi kolme mudeli kuju korrektne. Antud testi tulemused mudeli kuju muutumist ei toeta ning selle tõttu on antud mudelid lõplikud vaatamata sellele, et jäägid ei allu normaaljaotusele.

3.3. Järeldused ja hüpoteeside paikapidavus

Käesoleva bakalaureusetöö eesmärgiks oli uurida kõrghariduse ning keskmise palga vastastikust seost. Seose uurimiseks valis autor andmed, et läbi viia andmeanalüüs. Töös kasutas autor haridusenäitajaks OECD riikide kohta läbiviidud PIAAC uuringut, kust kasutati andmeid kolme erineva haridustaseme kohta, milleks olid põhi-, kesk-, ning kõrgharidus. Samuti kasutati keskmist töötasu, tööhõivemäära ning tervise näitajat, kus invidiidid hindasid enda tervist „hea või väga hea“. Analüüs viidi läbi kolmes vanuserühmas, milleks olid 16-24, 25-54 ning 55-65. Töö viidi läbi 22 OECD riigi andmete põhjal.

Käesoleva töö eesmärk oli kontrollida nelja hüpoteesi:

1. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 16-24 esineb positiivne vastastikune seos.
2. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 25-54 esineb positiivne vastastikune seos
3. Kõrghariduse ja keskmise sissetuleku vahel vanusegrupis 55-65 esineb positiivne vastastikune seos
4. Seose tugevus kõrghariduse ja keskmise sissetuleku vahel on erinevat vanusegruppide puhul erinev.

Autor püstitas töös neli hüpoteesi ning viis läbi nii korrelatsioonanalüüsi, kui ka regressioonanalüüsi, et kontrollida hüpoteeside paikapidavust. Esimene töös püstitatud hüpotees ei pidanud paika. Korrelatsioonanalüüsi tulemusena oli näha, et kõrghariduse ning keskmise palga vaheline seos vanuserühmas 16-24 on põhimõtteliselt olematu. Sama tulemuseni jõudis autor ka regressioonanalüüsis, kus keskmise töötasu näitaja ei osutunud mudelis statistiliselt oluliseks. Töös püstitatud teine ja kolmas hüpotees pidasid aga paika osaliselt. Positiivse seose olemasolu vanuserühmas 25-54 ning 55-65 oli näha korrelatsioonimaatriksis ning samuti näitas seda ka

regressioonanalüüs, mille tulemusena oli näha, et kõrghariduse määra tõus toob kaasa keskmise palga kasvu. Vastupidises regressioonmudelil, kus sõltuvaks muutujaks oli kõrgharidus ei suudetud aga vanuserühmas 25-54 ning 55-65 seost tuvastada. Töös oli püstitatud ka neljas hüpotees, mis pidas samuti paika osaliselt, kuna maatriksilt oli näha, et vanuserühmas 16-24 seos põhimõtteliselt puudus, vanuserühmas 25-54 oli seos keskmise tugevusega ning viimases vanuserühmas 55-65 oli seos natukene tugevam kui vanuserühmas 25-54 aga samuti keskmise tugevusega. Selle põhjal saab väita, et kõrghariduse ning keskmise palga seos on erinevates vanuserühmades erineva tugevusega. Vastupidises seoses, kus oli sõltuv muutuja kõrgharidus aga keskmine palk sõltuvaks muutujaks ei osutunud ehk vastupidist seost ei suudetud tuvastada ning selle tõttu ei saa hinnata ka seose tugevust.

Põhinedes uuringus kasutatud empiirilisele kirjandusele usub autor, et seos keskmise palga ning kõrghariduse vahel võib olla olemas, kuid mudelisse tuleks juurde lisada veel mõni tunnus ning samuti võib seose leidmisele kaasa aidata see, kui analüüsis kasutatavat valimi mahtu suurendada ehk vaadelda rohkemaid riike või lisada valimisse andmeid rohkemate aastate kohta.

KOKKUVÕTE

Käesoleva bakalaureusetöö eesmärgiks oli leida kõrghariduse ja keskmise palga vastastikuse seose olemasolu. Seose olemasolu uurimiseks tegi autor ülevaate varasemast empiirilisest kirjandusest ning viis samuti läbi endapoolse andmeanalüüsi.

Varasemast empiirilisest kirjandusest selgus, et töötasu ning hariduse vahel on positiivne seos. Mida kõrgem on indiviidi haridustase, seda suurem võimalus on tal teenida kõrgemat sissetulekut. Kuid, mitte ainult haridus ei mõjuta sissetulekute taset. Varasemast kirjandusest selgus, et sissetulekute ja hariduse vahelist seose uurimisel on näiteks kasutatud ka töötuse- ning tööhõive määra, riigipoolseid investeeringuid haridusse, kui ka inimeste tervisliku seisundi näitajaid. Samuti selgus, et seose tugevus ning olemasolu võib sõltuda nii soost, rassist, vanusest ning võib olla riigiti erinev.

Sissetulekute ning kõrghariduse seose uurimiseks kasutas OLS-i meetodit ehk hariliku vähimruutude meetodit. Autor koostas hajuvusdiagrammi ning viis läbi korrelatsioon-, ja regressioonanalüüsi. Töös koostati kaks regressioonmudelit, kus sõltuvaks muutujaks oli keskmine palk või kõrghariduse osakaal, et uurida sissetulekute ning kõrghariduse vastastikust seost. Antud analüüsid viis autor, mõlemal puhul, läbi kolmes erinevas vanusegrupis. Töös kasutati 2018 aasta andmeid 22 OECD riigi kohta. Andmed olid pärit Eurostat-, ning OECD statistika andmebaasist. Haridust kirjeldavad näitajad olid pärit OECD statistika andmebaasi PIAAC uuringust.

Esimesena uuris autor, kuidas mõjutab keskmine palk hariduse osakaalu. Regressioon-, ning korrelatsioonanalüüsi tulemusena selgus, et kõrgharitud inimeste osatähtsus mõjutab keskmist palka vanuserühmas 25-54 ning 55-65. Statistiliselt oluliseks vanuserühmades 25-54 ja 55-65 osutust kõrgharidus, ehk analüüsi tulemusena võib väita, et keskmine palk mõjutab kõrgharitud inimeste osakaalu. Samuti osutusid mudelites statistiliselt oluliseks tööhõive määr ning tervise näitaja, kus individid hindasid oma tervist „hea või väga hea“.

Järgnevalt uuris autor kõrghariduse osakaalu mõju keskmisele palgale. Sõltumatuteks muutujateks kõigis kolmes vanuserühmas osutusid põhihariduse ning keskhariduse osakaalu näitajad. Autori poolt läbiviidud testide tulemusena jäägid normaaljaotusele ei allunud. Regressioonanalüüsi tulemusena kõrghariduse puhul mõju keskmisele palgale kinnitust ei saanud.

Töö eesmärgiks oli leida kõrghariduse ning keskmise palga vastastikuse seose olemasolu ning selleks püstitas autor töös neli hüpoteesi. Esimene hüpotees paika ei pidanud, kuna vanuserühmas 16-24 oli seos põhimõtteliselt olematu. Teine hüpotees aga pidas osaliselt paika. Analüüsi tulemustena vanuserühmas 25-54 mõjutas keskmine palk küll kõrgharitud inimeste osakaalu, kuid kõrghariduse mõju keskmisele palgale kinnitust ei saanud. Kolmas hüpotees pidas paika samuti osaliselt. Vanuserühma 55-65 puhul oli keskmisel palgal kõrgharidusele mõju olemas, kuid vastupidine seos aga kinnitust ei saanud. Töös püstitatud neljas hüpotees pidas paika keskmise palga ning kõrghariduse seose puhul. Analüüsi tulemusena oli näha, et erinevate vanuserühmade vahel oli antud seos erineva tugevusega. Vastupidises seoses, kus sõltuvaks muutujaks oli kõrgharidus aga keskmine palk statistiliselt oluline muutuja ei olnud ehk seos puudus. Selle tõttu seose tugevust erinevates vanuserühmades vaadelda ei saanud.

Autor usub, et antud seose uurimine on kasulik ning analüüsi tulemuste põhjal on võimalik riigiti korraldada hariduspoliitikat, mille kaudu on võimalik edendada ka riigi majandust. Mudelit tuleks kindlasti täiustada, et mudeli põhjal saaks teha paremaid järeldusi. Näiteks võiks suurendada valimis olevate riikide arvu ning kasutada andmeid riikide kohta pikemas ajavahemikus. Samuti soovitab autor hariduse mõju puhul palgale lisada mudelisse veel muutujaid, et saaks teha järeldusi selle kohta, kuidas mõjutab haridus indiviidide sissetuleku suurust.

SUMMARY

THE RELATIONSHIP BETWEEN HIGHER EDUCATION AND AVERAGE WAGES IN THE EXAMPLE OF OECD COUNTRIES.

Erle Leppikson

The aim of this bachelor's thesis was to find the relationship between higher education and the average salary. To investigate the existence of the relationship, the author reviewed the previous empirical literature and also conducted her own data analysis.

Previous empirical literature has shown that there is a positive relationship between salary and education. The higher the level of education of an individual, the greater are his or her chances of earning a higher income. However education is not the only factor that determines the income level. Previous literature has shown that, for example, unemployment and employment rates, public investment in education, and human health indicators have been used to examine the links between income and education. It was also found that the strength and existence of the link may depend on gender, race, age and may vary from country to country.

To study the relationship between income and higher education the author used the OLS method or the least squares method. The author compiled a scatterplot and performed correlation and regression analysis. Two regression models were developed, in which the dependent variable was the average salary or the share of higher education in order to study the relationship between income and higher education. The author performed these analyzes in both cases in three different age groups. The study used 2018 data from 22 OECD countries. The data came from Eurostat and the OECD statistical database. The indicators describing education were from the PIAAC survey of the OECD statistical database.

Firstly, the author studied how the average salary affects the share of education. As a result of the regression and correlation analysis, the share of highly educated people affects the average wage

in the age groups 25-54 and 55-65. Higher education proved to be statistically significant in the age groups 25-54 and 55-65, as a result of the analysis it can be stated that the average salary affects the share of highly educated people. The employment rate and the health indicator, where individuals rated their health as “good or very good”, also proved to be statistically significant in the models.

Next, the author examined the effect of the share of higher education on the average salary. The indicators of the share of basic education and secondary education turned out to be dependent variables in all three age groups. As a result of the tests performed by the author, the residues did not undergo normal distribution. As a result of the regression analysis, the effect on higher wages in higher education was not confirmed.

The aim of the work was to find the relationship between higher education and the average salary, and for this purpose the author put forward four hypotheses. The first hypothesis was not valid, as the relationship was basically non-existent in the 16-24 age group. However, the second hypothesis was partially correct. As a result of the analysis in the age group 25-54, the average salary affected the share of highly educated people, but the effect of higher education on the average salary was not confirmed. The third hypothesis was also partially true. For the 55-65 age group, the average salary had an effect on higher education, but the opposite relationship was not confirmed. The fourth hypothesis put forward in the work was true in the case of the relationship between the average salary and higher education. The analysis showed that there was a different relationship between the different age groups. In the opposite relationship, where the dependent variable was higher education, the average salary was not statistically significant variable, so there was no correlation. As a result, the strength of the association in different age groups could not be observed.

The author believes that the study of this connection is useful and based on the results of the analysis, it is possible to organize education policy in each country, through which it is also possible to promote the national economy. The model should definitely be improved so that better conclusions can be drawn from it. For example, the number of countries in the sample could be increased and country-by-country data could be used over the longer term, and additional variables on the impact of education on wages should be added to the model to draw conclusions about how education affects an individual's income.

KASUTATUD ALLIKATE LOETELU

- Abduallah, A., Doucouliagos, H., Manning, E. Turner. (2013). Does education reduce income inequality? A meta-regression analysis. *Journal of Economic Surveys*, 29 (2), 301-316.
- Atkinson, A., Messy, F. A., Rabinovich, L., Yoong, J. (2015). Financial education for long-term savings and investments: Review of research and literature. *OECD Working Papers on Finance, Insurance and Private Pensions*, No. 39.
- Balázs, É., Jaroslava, B., David, T. (2020). The contribution of human capital and its policies to per capita income in Europe and the OECD. *European Economic Review*, 129, 103560.
- Blundell, R., Dearden L., Meghir, C., Sianesi, B. (2005). Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy. *Fiscal Studies*, 20 (1), 1-23.
- Bradbury, K. L. (2002). Education and wages in the 1980s and 1990s: are all groups moving up together?. *New England Economic Review*, 19-46.
- Canton, E. (2007). Social returns to education: Macro-evidence. *De Economist*, 155(4), 449-468.
- Cingano, F. (2014). Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Papers*, No. 163.
- Daly, M. C., Osborne, J., Valletta, R. G. (2007). Educational Attainment, Unemployment, and Wage Inflation*. *Economic Review-Federal Reserve Bank of San Francisco*, 49-61.
- Dao, M. Q. (2008). Human capital, poverty, and income distribution on developing countries. *Journal of Economic Studies*, 35 (4), 294-303.
- Dickson, M. (2012). The Casual Effect of Education on Wages Revisited. *Oxford Bulletin of Economics and Statistics*, 75 (4), 477-498.
- Gillies, D., (2017). Human Capital Theory in Education. *Encyclopedia of educational philosophy and theory*, 1-5.
- Gyimah-Brempong, K., Paddison, O., Mitiku, W. (2006). Higher education and economic growth in Afrika. *The Journal of Development Studies*, 42 (3), 509-529.
- Herd, P., Goesling, B., House, J. S. (2007). Socioeconomic Position and Health: The Differential Effects of Education versus Income on the Onset versus Progression of Health Problems. *Journal of Health and Social Behavior*, 48 (3), 223-238.

- Hewett, W. W. (1925). The Definition of Income. *The American Economic Review*, 15 (2), 239-246.
- Islam, R., Ghani, A. B. A., Kusuma, B., Theseira, B. B. (2016). Education and Human Capital Effect on Malaysian Economic Growth. *International Journal of Economics and Financial Issues*, 6 (4), 1722-1728.
- Lanzi, D. (2007). Capabilities, human capital and education. *The Journal of Socio-Economics*, 36, 424-435.
- Levin, H. M., Kelley, C. (1994). Can Education Do It Alone? *Economics of Education Review*, 13 (2), 97-108.
- Mincer, J. (1958). Investment in Human Capital and Personal Income Distribution. *Journal of Political Economy*, 66 (4), 281-302.
- Mincer, J. (1981). Human capital and economic growth. *NBER Working Paper*, No. 803.
- Mincer, J. (1991). Education and unemployment. *NBER Working Paper*, No. 3838.
- Müürisepp, K., Nausédaitė, R., Madubuko, T., Taal, K. E. (2017). Välisriikide kogemus madala haridustasemega täiskasvanute tasemekoolitusse tagasitoomisel ja nn teise võimaluse hariduse korraldamisel. Kättesaadav: https://dspace.ut.ee/bitstream/handle/10062/56421/Raport_2017.pdf?sequence=1&isAllowed=y 30. aprill 2021.
- Núñez, I., Livanos, I. (2010). Higher education and unemployment in Europe: an analysis of the academic subject and national effects. *Higher Education*, 59(4), 475-487.
- Ridell, W. C., Song X. (2011). The impact of education on unemployment incidence and re-employment success: Evidence from US labour market. *Labour Economics*, 18(4), 453-463.
- Robeyns, I. (2006). Three models of education. *School Field*, 4 (1), 69-84.
- Rõõm, T. (2007). Haridus ja tööturg Eestis. Tallinn: Eesti Pank, 66 lk.
- Sari, R., Soytaş, U. (2006). Income and Education in Turkey: Multivariate Analysis. *Education Economic*, 14 (2), 181-196.
- Sauga, A. (2017). Statistika õpik majanduseriala üliõpilastele. Tallinn: TTÜ Kirjastus.
- Schnittker, J. (2004). Education and the Changing Shape of the Income Gradient in Health. *Journal of Health and Social Behavior*, 45 (3), 286-305.
- Tolley, G. S., Olson, E. (1971). The Interdependence between Income and Education. *Journal of Political Economy*, 79(3), 460-480.
- Winters, J. V. (2011). Human capital, higher education institutions, and quality of life. *Regional Science and Urban Economics*, 41 (5), 446-454.

Wolla, S. A., Sullivan, J. (2017). Education, income, and Wealth. *Page One Economics*®.
Kättesaadav: https://files.stlouisfed.org/files/htdocs/publications/page1-econ/2017-01-03/education-income-and-wealth_SE.pdf. 16. märts 2021.

Woodhall, M. (1987). Human Capital Concepts. *Economics of Education*, 21-24.

LISAD

Lisa 1. OECD riike kirjeldav statistika

Riik	1(%)	2(%)	3(%)	4(\$)	5(\$)	6(\$)	7(%)	8(%)	9(%)	10(%)	11(%)	12(%)	13(%)	14(%)	15(%)	16(%)	17(%)	18(%)
Austria	51,35	84,55	53,98	25,67 0	26,45 4	27,98 4	54,30	71,80	83,40	42,33	15,32	29,13	53,17	62,70	53,28	3,56	20,21	14,79
Belgia	24,98	80,40	50,35	22,18 1	25,86 6	25,23 5	62,20	75,90	84,40	32,93	10,76	30,69	49,12	41,65	39,51	13,82	41,61	25,71
Eesti	41,70	84,20	68,95	10,13 2	12,14 0	11,22 0	38,40	48,40	62,00	43,73	11,24	15,84	44,55	45,10	46,17	11,29	43,27	37,66
Hispania	21,65	74,68	52,15	12,68 0	15,15 2	16,68 2	60,70	82,90	87,20	53,80	41,27	62,63	35,03	22,66	18,29	10,60	35,44	17,51
Holland	63,93	84,60	67,68	20,83 2	26,18 0	26,75 1	65,10	75,40	83,50	42,17	22,62	42,02	47,74	38,11	28,50	9,15	36,84	26,98
Iirimaa	40,28	79,18	60,38	24,43 3	26,81 8	27,06 7	70,90	87,10	89,80	32,45	23,05	49,37	49,86	39,32	30,09	17,69	37,20	19,64
Itaalia	17,68	69,83	53,70	15,04 4	17,20 9	18,95 3	60,30	84,50	86,30	62,56	44,69	72,01	34,10	38,77	19,43	3,02	15,73	8,16
Kreeka	14,00	68,95	41,05	66,63 66	81,81 13	84,84 93	57,80	85,30	89,00	30,83	26,24	50,04	59,90	42,82	29,64	9,28	29,69	19,87
Leedu	32,43	84,65	68,53	67,67 66	79,79 95	76,76 25	25,10	36,70	63,20	35,94	6,80	4,50	48,65	58,14	68,26	13,34	30,80	19,99
Norra	49,08	83,33	72,05	30,69 4	41,08 1	48,78 3	66,30	74,90	85,00	53,48	17,88	30,65	39,11	37,34	34,04	6,51	41,94	33,86

Lisa 1 järg

Riik	1(%)	2(%)	3(%)	4(\$)	5(\$)	6(\$)	7(%)	8(%)	9(%)	10(%)	11(%)	12(%)	13(%)	14(%)	15(%)	16(%)	17(%)	18(%)
Poola	30,98	82,38	48,95	5932	6992	6507	40,90	57,30	77,90	38,40	7,97	18,03	49,38	59,59	67,27	12,21	32,44	14,62
Prantsusmaa	29,75	80,53	52,28	20182	22708	24638	54,80	69,20	78,60	34,30	20,20	42,80	50,51	45,29	39,32	15,03	33,54	17,07
Rootsi	44,65	86,58	78,03	22664	27427	33008	68,70	75,00	83,20	42,71	15,18	30,77	50,68	49,26	42,42	6,24	35,56	26,65
Saksamaa	47,18	84,85	71,43	21134	24599	23959	57,10	63,70	75,40	54,71	10,07	9,70	39,96	54,34	54,72	4,72	34,03	33,10
Slovakkia vabariik	27,45	81,25	54,23	7186	8107	8240	51,80	65,50	82,10	40,78	13,03	26,35	50,07	63,59	59,27	8,93	23,09	14,09
Sloveenia	35,15	87,50	47,05	13135	13898	13086	44,20	65,80	78,90	45,48	15,53	32,38	50,45	54,11	51,00	4,07	29,68	16,01
Soome	44,08	82,45	65,43	20959	26613	28985	51,20	69,70	79,20	45,09	8,63	27,76	51,51	43,76	39,35	3,40	47,61	32,89
Taani	53,75	82,25	69,18	24201	32778	35789	65,10	69,80	77,00	57,79	17,12	26,27	39,11	38,70	41,10	2,77	43,71	32,51
Tšehhi vabariik	28,38	87,45	65,13	9117	10106	9545	40,30	60,10	79,90	42,33	7,24	19,40	51,75	70,34	65,26	5,80	21,55	14,97
Türgi	35,00	61,43	35,30	3235	3859	4337	59,30	81,90	85,80	55,20	64,07	76,85	34,39	19,00	10,32	9,46	15,09	8,00
Ungari	28,98	84,08	54,40	5037	5702	5517	45,80	62,40	72,10	44,81	17,32	26,23	48,27	48,76	51,55	6,13	33,26	21,45
Ühendkuningriik	54,79	84,89	66,58	19050	24042	22938	56,10	75,10	79,40	23,41	21,80	35,49	57,21	35,73	34,39	19,37	41,95	29,98

Allikas: OECD statistika andmebaas, Eurostat (2021), autori koostatud.

Märkused:

1. Näitajate pikkade nimetuste tõttu asendas autor nimetused numbritega ning toob järgnevalt välja nende sisu: 1 – 15-24 (Tööhõive tase %), 2 – 25-54 (Tööhõive tase %), 3 – 55-65 (Tööhõive tase %), 4 – 16-24 (Kesk töötasu \$), 5 – 25-54 (Kesk töötasu \$), 6 – 55-65 (Kesk töötasu \$), 7 – Hea, väga hea, madal haridustase (ISCED 0 kuni 2), 8 – Hea, väga hea, keskmine haridustase (ISCED 3

ja 4), 9 – Hea, väga hea, kõrge haridustase (ISCED 5 kuni 8), 10 – 16-24 (Põhi har %), 11 – 25-54 (Põhi har %), 12 - 55-65 (Põhi har %), 13 - 16-24 (Kesk har %), 14 - 25-54 (Kesk har %), 15 - 55-65 (Kesk har %), 16 - 16-24 (Kõrg har %), 17 - 25-54 (Kõrg har %), 18- 55-65 (Kõrg har %)

Lisa 2. Sõltuv muutuja keskmine töötasu, algmudel (16-24)

```
Model 38: OLS, using observations 1-22
Dependent variable: KTT1

      coefficient   std. error   t-ratio   p-value
-----
const    145951      130168      1,121     0,2787
TH1      289,918        95,0559     3,050     0,0076 ***
HAR1TER  389,096         111,311     3,496     0,0030 ***
PHAR1    -1689,81        1319,48     -1,281     0,2186
KHAR1    -1528,48        1300,13     -1,176     0,2569
KOHAR1   -1889,28        1399,17     -1,350     0,1957

Mean dependent var  15769,41   S.D. dependent var  8067,901
Sum squared resid  4,70e+08   S.E. of regression  5421,306
R-squared           0,655977   Adjusted R-squared  0,548470
F(5, 16)           6,101705   P-value(F)          0,002402
Log-likelihood      -216,8717   Akaike criterion    445,7434
Schwarz criterion   452,2896   Hannan-Quinn        447,2855

Excluding the constant, p-value was highest for variable 13 (KHAR1)
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 3. Sõltuv muutuja keskmine töötasu, lõplik mudel (16-24)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model8
Lõplik noored:
OLS, using observations 1-22
Dependent variable: KTT1

      coefficient  std. error  t-ratio  p-value
-----
const    -14940,5    5978,16   -2,499   0,0218  **
TH1       291,390     92,0253   3,166   0,0051  ***
HARITER    365,683     105,123   3,479   0,0025  ***

Mean dependent var  15769,41  S.D. dependent var  8067,901
Sum squared resid   5,34e+08  S.E. of regression  5302,686
R-squared            0,609155  Adjusted R-squared  0,568013
F(2, 19)            14,80629  P-value (F)         0,000133
Log-likelihood      -218,2753  Akaike criterion    442,5506
Schwarz criterion   445,8238  Hannan-Quinn        443,3217

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 1,79885
with p-value = P(Chi-square(5) > 1,79885) = 0,876219

RESET test for specification -
Null hypothesis: specification is adequate
Test statistic: F(2, 17) = 1,25802
with p-value = P(F(2, 17) > 1,25802) = 0,309376

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 3,18369
with p-value = 0,203549
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 4. Sõltuv muutuja keskmine töötasu, algmudel (25-54)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 15 X model 17 X
Mudel12 (alg) :
OLS, using observations 1-22
Dependent variable: KTT2

      coefficient   std. error   t-ratio   p-value
-----
const      43246,3      101317      0,4268    0,6752
HAR2TER     571,787          163,426     3,499     0,0030 ***
PHAR2     -1568,55          1022,55    -1,534     0,1446
KHAR2     -1670,00          1027,85    -1,625     0,1237
TH2        1046,35          449,542     2,328     0,0334 **
KOHAR2    -1253,00          1054,74    -1,188     0,2522

Mean dependent var  18810,86   S.D. dependent var  10165,93
Sum squared resid  7,33e+08   S.E. of regression  6766,479
R-squared           0,662455   Adjusted R-squared  0,556972
F(5, 16)           6,280207   P-value(F)         0,002088
Log-likelihood      -221,7479   Akaike criterion    455,4957
Schwarz criterion   462,0420   Hannan-Quinn       457,0378

Excluding the constant, p-value was highest for variable 17 (KOHAR2)
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 5. Sõltuv muutuja keskmine töötasu, lõplik mudel (25-54)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 12
loplikkeskealised:
OLS, using observations 1-22
Dependent variable: KTT2

      coefficient  std. error  t-ratio  p-value
-----
const    -101616      30588,4    -3,322    0,0038 ***
HAR2TER    559,311      151,717     3,687    0,0017 ***
TH2        833,648      313,073     2,663    0,0159 **
KOHAR2     421,252      187,236     2,250    0,0372 **

Mean dependent var  18810,86  S.D. dependent var  10165,93
Sum squared resid   8,56e+08  S.E. of regression  6894,623
R-squared            0,605742  Adjusted R-squared  0,540033
F(3, 18)            9,218467  F-value(F)          0,000651
Log-likelihood       -223,4562  Akaike criterion    454,9124
Schwarz criterion    459,2766  Hannan-Quinn        455,9405

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 1,83013
with p-value = P(Chi-square(9) > 1,83013) = 0,993878

RESET test for specification -
Null hypothesis: specification is adequate
Test statistic: F(2, 16) = 0,905033
with p-value = P(F(2, 16) > 0,905033) = 0,424264

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 0,393946
with p-value = 0,821213
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 6. Sõltuv muutuja keskmine töötasu, algmudel (55-65)

```

gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 17
Mudel13 (alg):
OLS, using observations 1-22
Dependent variable: KTT3

      coefficient   std. error   t-ratio   p-value
-----
const    -47748,4      83443,3    -0,5722   0,5751
TH3       539,338        215,412     2,504     0,0235 **
HAR3TER   1060,28        314,623     3,370     0,0039 ***
PHAR3    -626,583        895,539    -0,6997   0,4942
KHAR3    -643,643        881,275    -0,7304   0,4757
KOHAR3   -27,7252        880,996    -0,03147  0,9753

Mean dependent var   19788,27   S.D. dependent var   11683,30
Sum squared resid    8,17e+08   S.E. of regression    7145,000
R-squared             0,715047   Adjusted R-squared    0,625999
F(5, 16)             8,029909   P-value(F)            0,000593
Log-likelihood        -222,9454   Akaike criterion      457,8907
Schwarz criterion     464,4370   Hannan-Quinn          459,4328

Excluding the constant, p-value was highest for variable 18 (KOHAR3)

```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 7. Sõltuv muutuja keskmine töötasu, lõplik mudel (55-65)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 13
lopplik kolmas vanus:
OLS, using observations 1-22
Dependent variable: KTT3

      coefficient  std. error  t-ratio  p-value
-----
const    -107007    22678,1    -4,719    0,0002 ***
TH3       524,225     197,428     2,655    0,0161 **
HAR3TER   1036,23       226,713     4,571    0,0002 ***
KOHAR3    581,922       253,386     2,297    0,0339 **

Mean dependent var  19789,27  S.D. dependent var  11683,30
Sum squared resid   8,45e+08  S.E. of regression  6852,527
R-squared            0,705135  Adjusted R-squared  0,655991
F(3, 18)            14,34829  P-value(F)          0,000051
Log-likelihood       -223,3215  Akaike criterion    454,6429
Schwarz criterion    459,0071  Hannan-Quinn        455,6710

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 1,99901
with p-value = P(Chi-square(9) > 1,99901) = 0,991483

RESET test for specification -
Null hypothesis: specification is adequate
Test statistic: F(2, 16) = 1,21059
with p-value = P(F(2, 16) > 1,21059) = 0,323906

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 2,35461
with p-value = 0,308107
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 8. Sõltuv muutuja kõrghariduse osakaal, algmudel (16-24)

```

Model 39: OLS, using observations 1-22
Dependent variable: KOHAR1

      coefficient      std. error      t-ratio      p-value
-----
const      89,6302          4,65366        19,26        1,71e-012 ***
TH1         0,0183556          0,0197093        0,9313        0,3655
HAR1TER     0,0242791          0,0242783        1,000         0,3322
PHAR1      -0,923499          0,0413342       -22,34        1,72e-013 ***
KTI1       -5,41461e-05       4,00997e-05      -1,350         0,1957
KHAR1      -0,887936          0,0578978       -15,34        5,48e-011 ***

Mean dependent var      8,925877      S.D. dependent var      4,831825
Sum squared resid      13,47712      S.E. of regression      0,917780
R-squared                0,972511      Adjusted R-squared      0,963921
F(5, 16)                113,2111      P-value (F)             6,64e-12
Log-likelihood          -25,82611      Akaike criterion        63,65222
Schwarz criterion       70,19848      Hannan-Quinn           65,19433

Excluding the constant, p-value was highest for variable 1 (TH1)

```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 9. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (16-24)

```

gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 12
lopliknoored2(haridus):
OLS, using observations 1-22
Dependent variable: KOHAR1

      coefficient    std. error    t-ratio    p-value
-----
const      91,4740         4,24503      21,55      2,65e-014 ***
PHAR1     -0,924940         0,0405988    -22,78     1,01e-014 ***
KHAR1     -0,896434         0,0557202    -16,09     3,97e-012 ***
KTT1      -1,48023e-05       2,45787e-05    -0,6022    0,5545

Mean dependent var    8,925877    S.D. dependent var    4,831825
Sum squared resid    14,79571    S.E. of regression    0,906633
R-squared              0,969822    Adjusted R-squared    0,964792
F(3, 18)              192,8186    P-value(F)            7,22e-14
Log-likelihood        -26,85289    Akaike criterion      61,70579
Schwarz criterion     66,06996    Hannan-Quinn          62,73385

Excluding the constant, p-value was highest for variable 4 (KTT1)

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 13,0176
with p-value = P(Chi-square(9) > 13,0176) = 0,16181

RESET test for specification -
Null hypothesis: specification is adequate
Test statistic: F(2, 16) = 0,0135845
with p-value = P(F(2, 16) > 0,0135845) = 0,986519

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 10,7288
with p-value = 0,0046802

```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 10. Sõltuv muutuja kõrghariduse osakaal, algmudel (25-54)

gretl: models

File Edit Tests Save Graphs Analysis LaTeX

Mudel5(alg) X Mudel6(alg) X

Model 40: OLS, using observations 1-22
Dependent variable: KOHAR2

	coefficient	std. error	t-ratio	p-value	
const	82,3525	10,5888	7,777	7,98e-07	***
TH2	0,151664	0,111930	1,355	0,1942	
HAR2TER	0,0583397	0,0471282	1,238	0,2336	
PHAR2	-0,964993	0,0609873	-15,82	3,42e-011	***
KIT2	-6,46888e-05	5,44532e-05	-1,188	0,2522	
KHAR2	-0,988794	0,0493443	-20,04	9,29e-013	***

Mean dependent var	32,92003	S.D. dependent var	9,136205
Sum squared resid	37,82010	S.E. of regression	1,537451
R-squared	0,978424	Adjusted R-squared	0,971681
F(5, 16)	145,1127	P-value(F)	9,65e-13
Log-likelihood	-37,17643	Akaike criterion	86,35285
Schwarz criterion	92,89911	Hannan-Quinn	87,89496

Excluding the constant, p-value was highest for variable 5 (KIT2)

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 11. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (25-54)

```

gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 16
loplikkeskvanus2.1 (korg) :
OLS, using observations 1-22
Dependent variable: KOHAR2

      coefficient   std. error   t-ratio   p-value
-----
const      97,1356      3,53017   27,52     3,68e-016 ***
KIT2       -8,34671e-06     4,05952e-05  -0,2056   0,8394
PHAR2      -0,975258       0,0446330  -21,85    2,08e-014 ***
KHAR2      -0,973276       0,0490772  -19,83    1,11e-013 ***

Mean dependent var   32,92003   S.D. dependent var   9,136205
Sum squared resid    43,80093   S.E. of regression    1,559931
R-squared             0,975012   Adjusted R-squared    0,970847
F(3, 18)              234,1148   P-value(F)            1,32e-14
Log-likelihood        -38,79139   Akaike criterion      85,58277
Schwarz criterion     89,94694   Hannan-Quinn          86,61084

Excluding the constant, p-value was highest for variable 5 (KIT2)

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 8,54384
with p-value = P(Chi-square(9) > 8,54384) = 0,4804

RESET test for specification -
Null hypothesis: specification is adequate
Test statistic: F(2, 16) = 0,590691
with p-value = P(F(2, 16) > 0,590691) = 0,565581

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 11,1768
with p-value = 0,00374098

```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 12. Sõltuv muutuja kõrghariduse osakaal, algmudel (55-65)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
Mudel6(alg)
Model 41: OLS, using observations 1-22
Dependent variable: KOHAR3

      coefficient      std. error      t-ratio      p-value
-----
const      86,1772          10,3904          8,294         3,46e-07 ***
TH3         0,0546757          0,0708052        0,7722        0,4513
HAR3TER     0,0755769          0,115200         0,6560        0,5211
PHAR3      -0,964248          0,0918768       -10,50        1,40e-08 ***
KHAR3      -0,955212          0,0871422       -10,96        7,55e-09 ***
KIT3       -2,23244e-06       7,09380e-05      -0,03147      0,9753

Mean dependent var  22,06788  S.D. dependent var  8,627034
Sum squared resid  65,77027  S.E. of regression  2,027472
R-squared           0,957919  Adjusted R-squared  0,944769
F(5, 16)           72,84359  P-value(F)          1,96e-10
Log-likelihood      -43,26303  Akaike criterion    98,52605
Schwarz criterion   105,0723  Hannan-Quinn        100,0682

Excluding the constant, p-value was highest for variable 6 (KIT3)
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 13. Sõltuv muutuja kõrghariduse osakaal, lõplik mudel (55-65)

```
gretl: models
File Edit Tests Save Graphs Analysis LaTeX
model 20
lopikkorg2.1(korg3):
OLS, using observations 1-22
Dependent variable: KOHAR3

      coefficient    std. error    t-ratio    p-value
-----
const    93,8716         5,64293      16,64      2,26e-012 ***
KIT3     4,39776e-05        4,54847e-05    0,9669     0,3464
PHAR3   -0,954715          0,0605465    -15,77     5,57e-012 ***
KHAR3   -0,947049          0,0727866    -13,01     1,36e-010 ***

Mean dependent var    22,06788    S.D. dependent var    8,627034
Sum squared resid    69,06690    S.E. of regression    1,958839
R-squared              0,955810    Adjusted R-squared    0,948445
F(3, 18)              129,7762    P-value(F)            2,22e-12
Log-likelihood        -43,80101    Akaike criterion      95,60203
Schwarz criterion     99,96620    Hannan-Quinn          96,63009

Excluding the constant, p-value was highest for variable 6 (KIT3)

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 15,9248
with p-value = P(Chi-square(9) > 15,9248) = 0,0684696

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 10,0136
with p-value = 0,00669213
```

Allikas: Autori arvutused programmis Gretl lisas 1 toodud andmete põhjal.

Lisa 14. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks¹

Mina Erle Leppikson (*autori nimi*)

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose Kõrghariduse ja keskmise palga vastastikune seos OECD riikide näitel.
(*lõputöö pealkiri*)

mille juhendaja on Jelena Matina,

(*juhendaja nimi*)

- 1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
- 1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.
3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

(kuupäev)

¹ Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtjaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. jq 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.