

LUT School of Business and Management

Kauppätieteiden kandidaatintutkielma

Talousjohtaminen

Tutkimus- ja kehitysmenoihin vaikuttavat sisäiset tekijät
The internal determinants affecting R&D expenditures

15.1.2018

Tekijä: Marko Tikkanen

Ohjaaja: Maija Hujala

TIIVISTELMÄ

Tekijä: Marko Tikkanen
Opiskelijanumero:
Tutkielman nimi: Tutkimus- ja kehitysmenoihin vaikuttavat sisäiset tekijät
Akateeminen yksikkö: School of Business and Management
Koulutusohjelma: Kauppatiede / Talousjohtaminen
Ohjaaja: Maija Hujala
Hakusanat: Tutkimus ja kehitys, paneelidata, regressioanalyysi

Tällä tutkielmalla pyritään selvittämään, mitkä yrityksen sisäiset tekijät vaikuttavat sen tutkimukseen ja kehitykseen kohdistettujen varojen määrään. Teoriaosassa sisäiset tekijät jaetaan taloudellisiin, fyysisiin ja aineettomiin tekijöihin, joiden näkökulmasta aihetta tarkastellaan kirjallisuuden sekä aikaisempien tutkimusten avulla. Teoriaan nojaten jokaisesta näistä kolmesta kategoriasta valitaan yksi muuttuja, jolla tutkimus- ja kehitysmenojen suuruutta pyritään selittämään.

Tutkimusaineistona käytetään paneelidataa 78:sta Euroopan teollisuusyrityksestä, joista on kerätty tilinpäätöstietoja viimeisen 10 vuoden ajalta. Tutkimus toteutetaan tilastollisena tutkimuksena, ja tutkimusmenetelmänä käytetään paneelidatan regressioanalyysiä. Tutkimuksen laatua parantaakseen tehdään muuttujien logaritimuunnoksia sekä ääriarvojen suodattamista. Saatujen tulosten luotettavuutta arvioidaan tutkimalla multikollinearisuutta sekä estimoitujen residuaalien normaalijakautuneisuutta.

Tutkimustulokset paljastavat, että liikevaihdolla on tilastollisesti merkitsevä, heikko negatiivinen yhteys yrityksen kokoon suhteutettujen tutkimus- ja kehitysmenoihin määrään. Nettovelkaantumisasteella ja keskimääräisillä henkilöstökuluilla ei yhteyttä huomattu olevan.

ABSTRACT

Author: Marko Tikkanen
Student number:
Title: The internal determinants affecting R&D expenditures
School: School of Business and Management
Degree programme: Business Administration / Financial Management
Supervisor: Maija Hujala
Keywords: Research and development, panel data, regression analysis

This thesis pursues to find out which internal determinants of a company affect its amount of funds targeted towards research and development. In the theory section the internal determinants are divided into financial, physical and intangible determinants from whose point of view the topic is studied through literature and previous researches. Based on the theory section one variable is chosen from each of these three categories to explain the amount of research and development expenditures.

A panel data containing 78 European manufacturing companies' financial statements from last 10 years is used as research material. The research is executed as statistical research and panel data regression analysis is used as the research method. To improve the quality of the research, logarithmic transformations and outlier filtering is executed. The reliability of the findings is evaluated by studying multicollinearity and distribution of the estimated residuals.

The findings reveal that sales has statistically significant weak negative relation with the research and development expenditures of a company in proportion to its size. With gearing and average cost of employee a corresponding relation is not found.

SISÄLLYSLUETTELO

1. JOHDANTO	1
1.1 Tutkimuksen aihe, tavoitteet ja tutkimusongelmat	1
1.2 Teoreettinen viitekehys	2
1.3 Tutkimuksen rajaukset	4
1.4 Tutkimusaineisto ja menetelmät.....	5
1.5 Tutkimuksen rakenne.....	5
2. TUTKIMUKSEN JA KEHITYKSEN TEORIAA.....	6
2.1 Tutkimus- ja kehitysmenot	6
2.2 Fyysiset tekijät	7
2.3 Taloudelliset tekijät	9
2.3.1 Tutkimus ja kehitys investointina.....	9
2.3.2 Ulkoisen rahoituksen ongelmat	10
2.4 Aineettomat tekijät	12
3. TUTKIMUSMENETELMÄ JA -AINEISTO.....	13
3.1 Muuttujien kuvailu	14
3.2 Paneelidata.....	15
3.3 Paneelidatan analysointi	16
4. TULOKSET	20
4.1 Aineiston kuvailu ja muokkaus	20
4.2 Estimointimenetelmän valinta ja tulokset.....	23
4.3 Tutkimuksen luotettavuuden arviointi	27
5. JOHTOPÄÄTÖKSET	29
LÄHDELUETTELO	31

LIITTEET

Liite 1. Aineiston hakuperusteet

Liite 2. Tutkimusaineiston yritykset

Liite 3. Katsaus muunnetusta aineistosta

Liite 4. Aineiston muunnokseen käytetty R-koodi

Liite 5. Muuttujien jakauma

Liite 6. Logaritmisten muuttujien jakauma

Liite 7. Regressioanalyysin tulokset

Liite 8. Muuttujien laatikko-jana-kuviot

Liite 9. Outlier-havaintojen poiston rajat

Liite 10. Regressioanalyysin tulokset outlier-havaintojen suodattamisen jälkeen

Liite 11. Mallin residuaalien kvantiili- ja jakaumakuviot

KUVIOLUETTELO

Kuvio 1. T&k -menojen jakautuminen ulkoisiin ja sisäisiin tekijöihin

TAULUKKOLUETTELO

Taulukko 1. Käytettävän mallin valinta (Park 2010)

Taulukko 2. Muuttujien rdtosales, sales, gearing ja wage perustiedot

Taulukko 3. Muuttujien vinous ja huipukkuus

Taulukko 4. Logaritmisten muuttujien vinous ja huipukkuus

Taulukko 5. Estimointimenetelmän valintatesti, kiinteiden vaikutusten mallin parametriestimaatit ja muut perustiedot

Taulukko 6. Pearsonin korrelaatiokertoimet ja p-arvot

Taulukko 7. Suodatetun aineiston estimointimenetelmän valintatesti, satunnaisten vaikutusten mukaiset parametriestimaatit ja muut perustiedot

1. JOHDANTO

Nykyajan globaaleilla ja kilpailullisilla markkinoilla innovaatio on eilinehto yrityksen selviytymiselle. Vaikka innovaatioon johtava oivallus ei aina vaadikaan rahallista panosta, monet yritysjohtajat pitävät tutkimus- ja kehitysmenoja (t&k -menoja) tulevaisuuden teknologisten innovaatioiden mittarina. T&k -menoille on ominaista se, että niistä saatavia tulevaisuuden tuottoja on vaikeaa tai lähes mahdotonta ennustaa. Tämä johtuu siitä, että tulevaisuus on epävarmaa ja tuotetun asian arvo on kontekstisidonnaista. (Germeraad 2003) Jos optimaalista tutkimukseen ja kehitykseen allokoitua rahamäärää ei voida määrittää niistä saatavien tuottojen perusteella, kuinka yritykset päättävät t&k -menojensa suuruudesta? Coad ja Rao-Nicholson (2007) esittävät, että koska investoiminen tutkimukseen ja kehitykseen on lähes enemmän taidetta kuin tiedettä, on kohtuullista olettaa, että yritykset käyttävät nyrkkisääntöä sitomalla t&k -menonsa johonkiin tiettyyn tunnuslukuun, kuten liikevaihtoon, työntekijöiden määrään tai omaan pääomaan.

1.1 Tutkimuksen aihe, tavoitteet ja tutkimusongelmat

Tämän tutkimuksen aiheena on t&k -menoihin vaikuttavat yrityksen sisäiset tekijät. Työssä pyritään etsimään mitattavissa olevia tunnuslukuja, joihin yritykset sitovat t&k -menonsa. Lisäksi tutkitaan, kuinka ison osan t&k -menoista löydetyt tekijät selittävät, ja vastaavasti kuinka ison osan selittävät muut tekijät, joita tässä tutkimuksessa ei huomioida. Teorian pohjalta käydään läpi syitä, miksi tiettyjen tekijöiden voidaan olettaa toimivan hyvin tai huonosti selittävinä tekijöinä. Lisäksi pohditaan sekä syitä, jotka voivat vaikuttaa tutkimuksen luotettavuuteen, että tekijöitä, joita tämä tutkimus ei ota huomioon.

Aihetta on tutkittu aiemminkin hyvin vaihtelevin tuloksin. Hall et al. (2009) havaitsivat kassavirralla olevan vahva positiivinen vaikutus t&k -menoihin Yhdysvalloissa ja Ranskassa. Bhagat ja Welch (1995) eivät kuitenkaan tutkimuksessaan havainneet näiden kahden välistä yhteyttä Yhdysvalloissa. Tuoreemmassa tutkimuksessa Coad ja Rao-Nicholson (2007) huomasivat t&k -menojen seuraavan liikevaihtoa ja työntekijöiden määrää. Tietääkseni aiheesta ei ole tuoretta Eurooppalaisiin yrityksiin kohdistuvaa tutkimusta, joten tämän tutkimuksen tekeminen on mielestäni hyvin perusteltua.

Tutkimuksen tavoitteena on tuoreen aineiston avulla kartoittaa Euroopan teollisuusyritysten t&k -menoihin vaikuttavia sisäisiä tekijöitä. Jos merkitseviä tekijöitä ei löydy, pyritään teorian avulla selittämään syitä tälle sekä tarjoamaan ehdotuksia, mitä tulevien tutkimusten tulee ottaa huomioon. Näin ollen tutkimuksen päätutkimusongelmaksi muodostuu:

”Mitkä yrityksen sisäiset tekijät vaikuttavat tutkimus- ja kehitysmenoihin?”

Alatutkimusongelmana selvitetään, kuinka ison osan t&k -menoista löydetty tekijät selittävät, ja kuinka iso osa jää residuaaliin, eli tekijöihin, joita tässä tutkimuksessa ei huomioida. Alatutkimusongelma on:

”Miten ja kuinka paljon löydetty tekijät vaikuttavat tutkimus- ja kehitysmenoihin?”

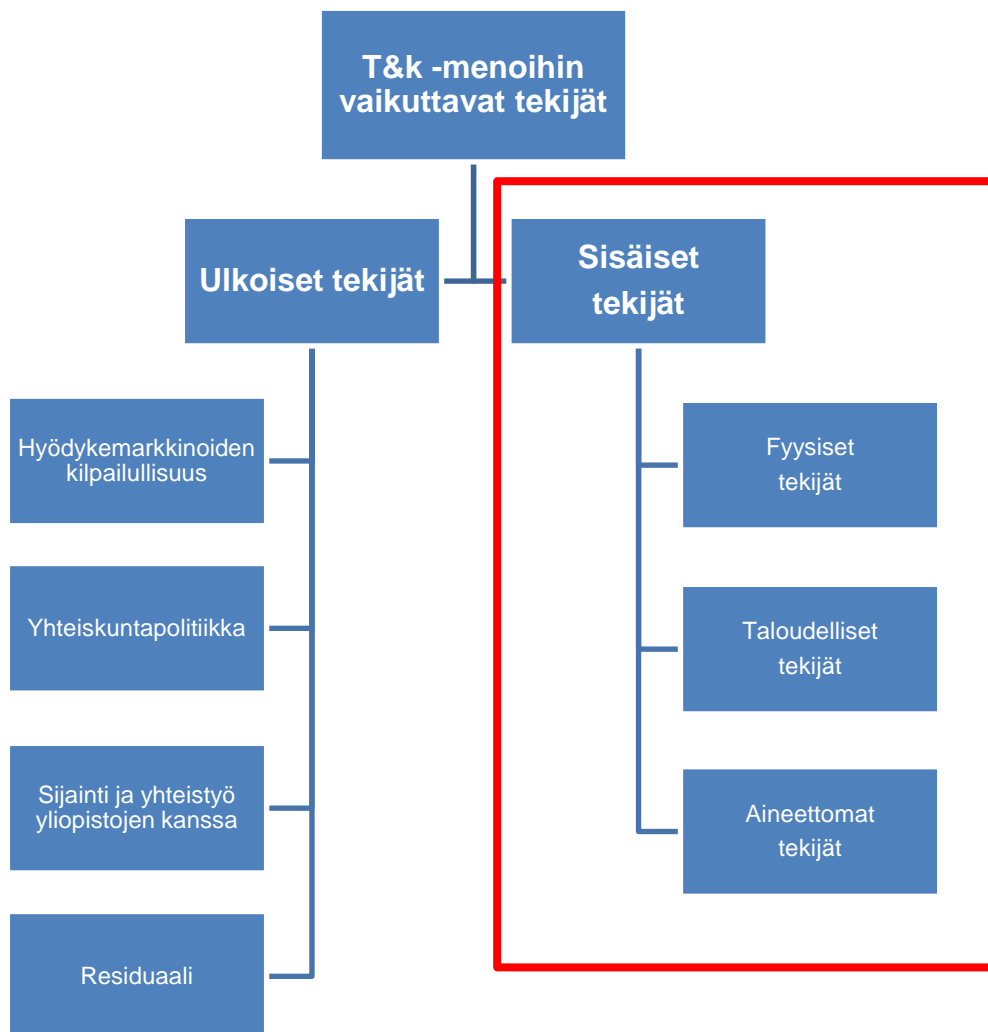
1.2 Teoreettinen viitekehys

Tutkimuksen lähtökohtana on oletus siitä, että yritykset eivät voi luotettavasti ennustaa t&k -menojensa tuottoja neoklassiseen tyyliin voittolisän kautta. Perustelut tälle väitteelle esitetään luvussa 2.3.1.

Tämä tutkimus pohjautuu teoriaan, jossa tutkimukseen ja kehitykseen vaikuttavat tekijät jaetaan viiteen eri kategoriaan, joista neljä koostuu ulkoisista tekijöistä ja yksi

sisäisistä. Ulkoisia tekijöitä ovat hyödykemarkkinoiden kilpailullisuus, yhteiskuntapolitiikka, sijainti ja yhteistyö yliopistojen kanssa sekä residuaali, johon kuuluu mm. tiedon positiivinen ulkoisvaikutus. Sisäisiä tekijöitä ovat yrityksen luonteenomaiset ominaisuudet, kuten liikevaihto tai sisäinen rahoitus. (Becker & Pain 2003)

Del Canto ja González (1999) jakavat nämä sisäiset tekijät vielä kolmeen osaan: taloudelliset, fyysiset sekä aineettomat tekijät. Taloudelliset tekijät sisältävät mm. oman pääoman, velkaisuuden ja kertyneet voittovarot. Fyysisiä tekijöitä ovat esimerkiksi yrityksen koko mitattuna jollain asteikolla tai yrityksen pääomaintensiteetti. Aineettomia tekijöitä ovat "näkymättömät" varat, kuten työntekijöiden tieteellinen ja teknologinen osaaminen sekä yrityksen imago. Näitä tekijöitä on huomattavasti vaikeampi mitata kuin kahta ensimmäistä. Kuvio 1 havainnollistaa t&k -menoihin vaikuttavien tekijöiden jakautumista ulkoisiin ja sisäisiin tekijöihin.



Kuvio 1. T&k -menojen jakautuminen ulkoisiin ja sisäisiin tekijöihin (Becker & Pain 2003; Del Canto & González 1999). Tämä tutkimus keskittyy sisäisiin tekijöihin

1.3 Tutkimuksen rajaukset

Tutkimuksen rajauksen tavoite on tehdä aineiston yrityksistä mahdollisimman vertailukelpoisia pitäen kuitenkin otoskoon niin isona, että tuloksia voidaan pitää luotettavana. Koska tutkimus käsittelee yrityksiä sisäisiä tekijöitä, aineiston on hyvä koostua yrityksistä, joiden ulkoiset tekijät ovat mahdollisimman yhteneväiset. Näin residuaaliin jää mahdollisimman vähän näitä ulkoisten tekijöiden vaikutuksia, ja tutkimuksen tulokset ovat mahdollisimman tarkat.

Nämä tavoitteet huomioiden tutkimus rajataan maantieteellisen sijainnin, toimialan sekä saatavilla olevan datan määrän perusteella. Maantieteellisesti tutkimus rajataan koskemaan Euroopan valtioita, sillä näissä valtioissa lainsäädännön ja taloudellisten olosuhteiden voidaan olettaa tämän tutkimuksen tarvitsemissa rajoissa olevan yhtenäiset. Toimialan suhteen tutkimus rajataan koskemaan teollisuusalan yrityksiä. Ala on riittävän suuri tutkittavaksi, ja yritykset riittävän samankaltaisia, että tutkimus on mielekästä toteuttaa. Eteen aineiston saamiseksi tutkimukseen otetaan mukaan yritykset, joista löydetään tarvittavat tilinpäätöstiedot ainakin 10 vuoden ajalta.

1.4 Tutkimusaineisto ja menetelmät

Määriteltyihin tutkimusongelmiin pyritään vastaamaan kvantitatiivisella, eli määrällisellä, tutkimuksella. Tärkeimpänä analyysimenetelmänä käytetään paneelidatan regressioanalyysiä. Tämän lisäksi tehdään testejä tutkimuksen luotettavuuden arvioimiseksi. Analyysi toteutetaan RStudio- ja SAS Enterprise Guide 6.1-ohjelmistoilla. Tutkimusmenetelmästä ja sen valinnasta kerrotaan lisää luvussa 3.

Tutkimuksen aineisto on Amadeus-tietokannasta haettu paneelidata. Paneelidatalla tarkoitetaan dataa, jossa yksiköistä on useampia havaintoja. Paneelidatan käytössä on useita etuja aikasarja- tai poikkileikkausdatan käyttöön verrattuna. Näitä etuja käydään läpi kolmannessa luvussa.

1.5 Tutkimuksen rakenne

Tutkimus koostuu viidestä pääluvusta. Ensimmäisen luvun, eli johdannon, jälkeen käsitellään tutkimus- ja kehitysmenojen teoriaa. Luvussa käydään aiheesta kirjoitetun teorian avulla läpi mahdolliset t&k -menoihin vaikuttavat tekijät, minkä jälkeen voidaan päättää tässä tutkimuksessa käytettävät muuttujat. Kolmannessa luvussa käydään läpi aineistoon ja tutkimusmenetelmiin liittyvä teoria sekä käytettävien muuttujien kuvailu.

Neljännessä luvussa esitellään empirian eli itse tutkimuksen tulokset sekä arvioidaan sen luotettavuutta. Viimeisessä luvussa tehdään yhteenveto tutkimuksesta, esitetään tutkimuksen johtopäätökset sekä esitetään mahdollisia jatkotutkimuskohteita.

2. TUTKIMUKSEN JA KEHITYKSEN TEORIAA

Tulevissa alaluvuissa käydään läpi t&k -menoihin vaikuttavien tekijöiden teoreettista viitekehystä entistä laajemmin. Luvussa määritellään tutkimukselle tärkeät termit sekä vertaillaan eriäviä teorioita ja aiempia tutkimuksia.

2.1 Tutkimus- ja kehitysmenot

Tutkimus ja kehitys käsittää kaiken luovan ja systemaattisen toiminnan, jonka tarkoituksena on lisätä tietämyksen määrää tai luoda uusia sovellutuksia olemassa olevasta tietämyksestä (OECD 2015, 28). Erilaiset tutkimus- ja kehitysprojektit voidaan lajitella usealla eri tavalla. Ehkä yleisin tapa on jakaa tutkimus ja kehitys kolmeen osaan, jotka ovat perustutkimus, soveltava tutkimus ja kokeellinen kehittäminen. Perustutkimuksessa kokeellisella tai teoreettisella työllä pyritään saavuttamaan uutta tietoa jostain ilmiöstä ilman tavoitetta soveltaa tätä tietoa tiettyyn tarkoitukseen. Myös soveltavan tutkimuksen tavoitteena on uuden tiedon saavuttaminen, mutta tutkimus on suuntautunut jonkin tietyn ongelman ratkaisuun tai tavoitteen saavuttamiseen. Kokeellisen kehittämisen tarkoitus on tutkimuksesta ja kokemuksesta saadun tiedon avulla joko luoda uusi tuote tai prosessi, tai kehittää olemassa olevaa tuotetta tai prosessia. (Hall et al. 2009)

Yksinkertaisin tapa mitata yrityksen t&k -menoja on käyttää yrityksen tilinpäätöksessä ilmoittamaa lukua. Tämän tunnusluvun käyttö ei ole täysin ongelmatonta, sillä yrityksillä on huomattavasti liikkumatilaa siinä, mitä ne luokittelevat t&k -menoksi (Becker & Pain 2003). Otetaan esimerkiksi pieni yritys, jossa yksikään työntekijä ei ole

palkattu ensisijaisesti tekemään tutkimusta ja kehitystä, mutta osa kuitenkin osallistuu oman työnsä ohella t&k -projekteihin. Tällaisella yrityksellä ei välttämättä ole kannustinta pitää kirjaa ja merkitä tilinpäätökseen tutkimukseen ja kehitykseen käytettyjen työtuntien määrä, joten tilinpäätöksessä oleva luku voi olla mitä tahansa nollan ja todellisen luvun välistä.

Myös laskennallisia eroja on. Cohen ja Mowery (1984) tutkivat t&k -menoja kahdesta eri lähteestä samoista yrityksistä samoina vuosina, ja huomasivat t&k -menojen olevan keskimäärin 12% korkeampia toisen lähteen mukaan laskennallisista eroista johtuen. Tämän tutkimuksen aineisto on kerätty kokonaisuudessaan samasta lähteestä, joten laskennallisia eroja ei pääse syntymään.

Aihetta tutkittaessa täytyy myös muistaa, että tutkimus ja kehitys eivät ole yrityksen ainoat keinot luoda uusia tuotteita tai prosesseja. Yrityksen muita innovatiivisia toimintoja ovat esimerkiksi yristysostot, muotoilun parantaminen, laadunvalvonta ja tekemällä oppiminen (Del Canto & González 1999). Nämä toiminnot eivät näy yrityksen tilinpäätöksessä, eikä niitä näin ollen oteta huomioon tässä tutkimuksessa. Tämä tutkimus ei pyri löytämään selittäjiä innovaatiolle tai innovaation tuotoksen määrälle, vaan yhdelle innovaatioon vaikuttavan panoksen määrälle.

2.2 Fyysiset tekijät

Yrityksen koon vaikutusta sen t&k -toimintaan on tutkittu pitkään. Jo vuonna 1942 aiheen oppi-isä Schumpeter esitti väitteensä siitä, että innovaation tasolla on positiivinen korrelaatio yrityksen koon ja siihen liittyvän kilpailuvoiman kanssa. Schumpeterin mukaan isot yritykset tarvitsevat juridisen lyhytaikaisen suojan innovaatioilleen, jotta näiden tuoma kilpailuetu olisi riittävä houkuttelemaan yrityksiä investoimaan tutkimukseen ja kehitykseen. (Schumpeter 1942) Schumpeterin (1961) hypoteesin mukaan isot yritykset kuluttavat suhteessa enemmän innovaatioon kuin pienet yritykset. Hypoteesin perimmäisenä tekijänä on skaalautuvuus: isot yritykset voivat odottaa isompia tuottoja yksittäiselle innovaatiolle kuin pienet yritykset, ja suuri

osuus markkinoista suojaa isoja yrityksiä mahdollisilta matkijoilta. Vastaavasti pienillä yrityksillä ei ole kykyä sijoittaa tehokkaasti tutkimukseen ja kehitykseen, sillä kilpailullisilla markkinoilla riski on liian suuri odotettuun tuottoon (Majumdar 2009).

Galbraith (1952, 85-90) on samoilla linjoilla sanoen isojen yritysten olevan täydellisiä innovaation lähteitä. Hän myös lisää, että isot yritykset voivat hajauttaa riskinsä suureen määrään t&k -projekteja. Arrow (1962) jatkaa Schumpeterin (1942, 1961) ja Galbraithin (1952) ajatuksia selittämällä, että pienet yritykset ali-investoivat tutkimukseen ja kehitykseen, koska ne ovat haluttomia ottamaan riskejä, taloudellisesti heikompia ja kykenemättömiä hyödyntämään innovaation tuloksia optimaalisesti. Arrow (1962) käsittelee aihetta artikkelissaan painottaen aineettomia oikeuksia ja riskejä. Hän väittää, että pienten yritysten on vaikeampi puolustaa t&k -projekteista syntyneitä aineettomia oikeuksia, minkä takia innovaatioista tulee julkisia hyödykkeitä. Tämän pohjalta Arrow (1962) esittää, että koska yksikään yritys ei halua tuottaa julkisia hyödykkeitä ja tällä tavalla ajaa itseään konkurssiin, vain vahvan markkina-aseman omaavilla isoilla yrityksillä on kannustin innovaatioon.

Aiheesta ei kuitenkaan olla yksimielisiä, sillä esimerkiksi Demsetz (1969) väittää, että Arrowin (1962) teoria ei ole uskottava. Demsetz ei kyseenalaista johtopäätöstä siitä, että t&k sopii paremmin isoille yrityksille, vaan kritisoi Arrowin päättelyketjua. Demsetzin mukaan Arrowin perustelua sille, miksi pienet yritykset ali-investoivat tutkimukseen ja kehitykseen, voitaisiin aivan yhtä hyvin käyttää perusteluna sille, miksi isot yritykset ali-investoivat. Näin ollen Arrowin väitteellä, että isoilla yrityksillä on paremmat kyvykkyydet t&k -toimintaan pienemmän riskin takia, ei ole pitävää pohjaa. Demsetz ei ota kuitenkaan itse kantaa, onko joko pienet tai suuret yritykset etulyöntiasemassa t&k -toiminnassa, vaan nimeää kolme tavoitetta, jotka innovaatiotoimintaa harjoittavan yrityksen tulisi saavuttaa: olla mukana laajasti erilaisissa kokeiluissa, investoida lupaaviin hankkeisiin sekä olla kykenevä käyttämään kokonaisuudessaan hyväkseen tieto, joka innovaatioprosessissa saavutetaan. (Arrow 1962; Demsetz 1969)

T&k -menoihin vaikuttavia tekijöitä tutkittaessa Schumpeterin (1961) hypoteesia on vaikea olla huomioimatta. Intuition mukaan ja loogisesti ajateltuna on perusteltua sanoa, että innovaatio vaatii huomattavia resursseja, uhka innovaation matkimisesta vähentää houkutusta investoida siihen ja iso yritys voi hyödyntää yksittäistä innovaatiota laajemmilla markkinoilla kuin pieni yritys. (Goodwin 1998) Kuitenkin empiirisesti tutkittuna aiheesta on saatu ristiriitaisia tuloksia. Esimerkiksi Scherer (1965) ja Comarov (1967) eivät tutkimuksissaan havainneet t&k -menojen ja yrityksen koon välistä yhteyttä. Tuoreemmassa tutkimuksessa Coad & Rao-Nicholson (2007) taas havaitsivat merkittävän yhteyden t&k -menojen ja liikevaihdon sekä työntekijöiden määrän välillä.

2.3 Taloudelliset tekijät

Taloudellisilla tekijöillä tarkoitetaan tässä yhteydessä yrityksen eri varoja sekä sen pääomarakennetta, eli esimerkiksi omaa pääomaa, velkaisuusastetta tai kertyneitä voittovaroja. Taloudellisten tekijöiden, ja varsinkin sisäisen rahoituksen määrän, on laajalti pidetty tärkeimpinä t&k -menoihin vaikuttavina tekijöinä (esimerkiksi Himmelberg & Petersen 1994, Bougheas et al. 2001, Hall & Lerner 2009). Kamien & Schwartz (1972) sanovat, että t&k -investoinnin huomattavin ominaisuus on tarve rahoittaa se sisäisesti yrityksen omilla tuotoilla ja kertyneillä varoilla. Pääasiallinen perustelu tälle väitteelle on se, että tutkimuksen ja kehityksen tuloksena syntyy ei-kilpailullista tietoa, eli tiedon käyttäminen ei estä toista yritystä käyttämästä samaa tietoa (Hall & Lerner 2009).

2.3.1 Tutkimus ja kehitys investointina

T&k -investoinnilla on useita ominaispiirteitä, jotka erottavat sen normaaleista investoinneista. Ensinnäkin käytännössä 50 prosenttia tai enemmän t&k -menoista koostuu korkeasti koulutettujen tutkijoiden ja insinöörien palkoista. Näiden

asiantuntijoiden työ luo yritykseen aineetonta pääomaa, josta yrityksen tulevaisuuden innovaatiot muodostuvat. Asiantuntijan lähtiessä yrityksestä katoaa myös osa aineettomasta pääomasta. Jotta yritysten ei tarvitsisi irtisanoa osaavia asiantuntijoita, yritysten on tavanomaista tasata t&k -menonsa pitkälle ajanjaksolle. Tästä johtuen t&k -investoinneilla on usein korkeat sopeuttamiskustannukset (Hall & Lerner 2009). Koska osa investoinnin tuotoista kuluu sopeuttamiskustannuksiin, tarvittavan tuotto-odotuksen täytyy olla normaalia investointia korkeampi.

Toinen t&k -investoinnin ominaisuus on lopputuotoksen suuri epävarmuus (Hall & Lerner 2009). Tämän takia t&k -investoinnin kannattavuutta on hyvin vaikea arvioida neoklassiseen tyyliin voittolisän kautta. Jos tämä voitaisiin ennustaa, yritykset seuraisivat Modigliani-Millerin irrelevanttiusteoreemaa. Teorian mukaan yrityksen optimaaliseen investointitasoon ei vaikuta sen pääomarakenne (Modigliani & Miller 1958). Toisin sanoen yritykset asettaisivat t&k -menonsa sille tasolle, että niistä saadut tuotot ovat samalla tasolla kaikkien muiden investointien tuottojen kanssa. Mm. tuottojen epävarmuuden takia moni asiantuntija on pitänyt yrityksen sisäisen rahoituksen määrää tärkeimpänä t&k -menoja selittävänä tekijänä (Himmelberg & Petersen 1994).

2.3.2 Ulkoisen rahoituksen ongelmat

Seuraavaksi käydään läpi syitä, joiden takia ulkoinen rahoitus usein ei ole mahdollista t&k -projektin rahoitusmuotona. Nämä syyt ovat asymmetrinen informaatio, moraalinen riski sekä tutkimuksen ja kehityksen aineettomuus.

Asymmetrisellä informaatiolla tarkoitetaan tilannetta, jossa yrityksellä on parempi kuva innovaatioprojektista ja innovaation onnistumisen todennäköisyydestä kuin ulkoisella rahoittajalla (Hall & Lerner 2009). Tämän takia markkinat innovaatioille muistuttavat Akerlof'n (1970) mallintamaa "sitruunamarkkinaa". Sitruunamarkkinat tässä tapauksessa tarkoittavat sitä, että riskipreemio t&k -investoinneille on tavallista suurempi, koska pitkäkestoisissa t&k -projekteissa sijoittajien on vaikeampi erottaa

tuottavat projektit huonoista kuin lyhytaikaisissa ja pieniriskisissä sijoituksissa (Myers ja Majluf 1984).

Asymmetrinen informaatio siis nostaa ulkoisen rahoituksen hintaa t&k -investoinneille, ja sen poistaminen olisi molemmille osapuolille hyödyllistä. Asymmetrisyyden vähentäminen on kuitenkin vaikeaa innovaation matkimisen helppouden takia (Hall & Lerner 2009). Yritys ei halua kertoa tarkkaa tietoa innovaatioprojekteistaan rahoitusmarkkinoille, koska tieto päätyisi helposti myös kilpailijoille, jolloin yritys menettäisi entistäkin nopeammin kilpailuedun, jonka innovaatio voisi tuoda (Bhattacharya & Ritter 1985). Tätä ongelmaa lieventää tosin se, että uuden innovaation matkiminen ei missään nimessä ole ilmaista, vaan voi kustantaa jopa 50-70% alkuperäisen innovaation kustannuksista (Levin et al. 1987).

Moraalinen riski on asymmetriseen informaatioon liittyvä ongelma, jossa lainaava osapuoli vastaa varojen allokoinnista projektissa, ja sijoittajalla ei ole tietoa siitä, kuinka varat käytetään t&k -projektissa. Tällöin voi syntyä tilanne, jossa rahoituksen saaja ohjaa varat oman etunsa ajamiseen itse projektin sijasta. (Arslan-Ayaydin et al. 2014) Ongelmaa pahentaa se, että vaikka yritys ei aikoisikaan käyttää ulkoista rahoitusta epämoraalisesti, se ei silti halua paljastaa liikaa projektista kilpailijoiden matkinnan pelossa (Bhattacharya & Ritter 1985). Arrow (1962, 153) väittää moraalisen riskin olevan erityisen suuri t&k -projekteissa, sillä projektin tulosta ei ikinä voi ennustaa täydellisesti käytetyistä panoksista. Oman edun ajamisen lisäksi Stiglitz ja Weiss (1981) esittävät, että koron kasvaessa seuraamattomilla lainaajilla on kannustin käyttää laina korkeamman riskin projekteihin, jotka lisäävät konkurssin mahdollisuutta, mutta ei tuo lisäarvoa lainaajalle.

Kuten sanottu, yksi t&k -menojen ominaisuuksista on se, että iso osa siitä koostuu korkeasti koulutettujen asiantuntijoiden palkoista (Hall & Lerner 2009). Lisäksi epäonnistuneesta innovaatioprojektista usein jää jäljelle hyvin vähän aineellista omaisuutta (Kamien & Schwartz 1978). Näistä johtuen varsinkin pienillä yrityksillä voi olla vaikeaa saada ulkoista rahoitusta vakuuksien puuttumisen takia.

Williamson (1988) sanoo, että päätös projektin rahoittamisesta joko velalla tai pääomalla riippuu hankittavien kohteiden ominaisuuksista. Pääomarahoitusta tulisi käyttää, kun sijoitetaan hyvin erikoistuneisiin kohteisiin, kun taas velkaa käytetään tavallisempiin kohteisiin. Normaalien sijoitusten omaisuus on helposti uudelleenmyytävissä, joten yrityksen omistajilla on aina mahdollisuus myydä nämä projektin epäonnistuessa. Mitä erikoistuneempi sijoituskohte on, sitä vaikeampi sitä on enää myydä, ja sitä suurempi on sijoituksen riski. Tämän takia pääomarahoitusta on käytetympi rahoituskeino t&k -projekteissa, jotka usein ovat luonteeltaan hyvin erikoistuneita. Samasta syystä osakeyrityksissä, joissa osakkeenomistajilla on paljon neuvotteluvoimaa ja näin ollen mahdollisuus painostaa yritysjohto pieniriskisiin sijoituksiin, t&k -menojen voidaan odottaa olevan pienet.

2.4 Aineettomat tekijät

Aineettomat tekijät ovat yrityksen näkymätöntä pääomaa, kuten työntekijöiden tieteellistä ja teknologista osaamista sekä yrityksen imagoa (Foss 1996). Strategisesta näkökulmasta ne usein ovat yrityksen tärkeimpiä varoja, sillä ne tuovat yhteen lähtökohdat kilpailuedun saavuttamiselle: arvokkuus, harvinaisuus ja vaikeus imitoida (Barney 1991). Samalla ne ovat kuitenkin yrityksen ulkopuoliselle toimijalle vaikeimpia varoja arvioida, sillä aineettoman pääoman arvoa ei voi nähdä suoraan tilinpäätöksestä. Vaikean mitattavuuden takia aineettomat tekijät ovat usein jätetty vähemmälle huomiolle t&k -menoihin vaikuttavia tekijöitä tutkittaessa, vaikka niiden merkitys yrityksen t&k -toiminnalle voidaankin olettaa olevan merkittävä.

Karkeasti aineeton pääoma voidaan jakaa inhimillisiin- ja kaupallisiin resursseihin. Inhimilliset resurssit pitävät sisällään mm. yrityksessä työskentelevien henkilöiden tiedon, taidon, kokemuksen, taidot sekä taipumuksen riskinottoon. (Barney 1991) Nämä kaikki ominaisuudet ovat tarpeellisia tutkimuksen ja kehityksen onnistumiselle, minkä takia voidaan tehdä hypoteesi, että yritys, jolla on paljon inhimillistä pääomaa, panostaa myös paljon tutkimukseen ja kehitykseen. Inhimilliselle pääomalle ei ole

olemassa yksiselittäistä mittaria, joten tässä tutkimuksessa tullaan käyttämään proxy-muuttujaa inhimillisen pääoman arvioimiseksi.

Kaupalliset resurssit, kuten imago, kansainvälisyys ja sidosryhmäverkosto, ovat yrityksen toimintaa täydentäviä resursseja, jotka ovat tarpeellisia kehitetyn innovaation tehokkaaseen kaupalliseen hyödyntämiseen. Jos yrityksellä ei ole aineettomia kaupallisia resursseja hyödyntää uusia innovaatioita, ei sillä myöskään ole kannusteita harjoittaa tutkimusta ja kehitystä, koska todennäköisyys tuotoille ovat pienemmät ja todennäköisyys kilpailijoiden matkinnalle ovat suuremmat kuin korkeiden kaupallisten resurssien yrityksellä. (Teece 1987)

Mahdollisuudesta hyödyntää innovaatioita puhuttiin jo kappaleessa 2.2. Yrityksen koon ja sen kaupallisten resurssien määrällä välillä voidaan olettaa olevan vahva positiivinen korrelaatio. Multikollinearisuuden välttämiseksi sekä datan vaikean saatavuuden takia kaupallisia resursseja ei käytetä selittävinä tekijöinä tässä tutkimuksessa. Del Canto ja González (1999) tutkivat kaupallisten resurssien vaikutusta siihen, onko yrityksellä t&k –toimintaa ottamalla tutkimukseen mukaan dummy-muuttujan, jolla kuvattiin, onko yrityksellä vientitoimintaa. Muuttujalla pyrittiin kuvaamaan yrityksen maineen ja kaupallisen toiminnan laajuutta. Tutkimus havaitsi positiivisen yhteyden viennin ja t&k –toiminnan välillä.

3. TUTKIMUSMENETELMÄ JA -AINEISTO

Tässä luvussa käydään läpi käytettävät selittävät muuttujat, käytettävän aineiston luonne sekä tutkimusmenetelmän teoriaa. Muuttujat ovat valittu edellisen teorialuvun sekä aineiston saatavuuden perusteella. Muuttujat ovat yritysten tilinpäätöksissään esittämiä arvoja viimeisen kymmenen vuoden ajalta. Selitettävänä muuttujana käytetään vastaavasti tilinpäätöstiedoista saatuja t&k -menoja suhteutettuna yrityksen liikevaihtoon. Muuttuja siis pyrkii kuvailemaan t&k -menojen määrää yrityksen kokoon suhteutettuna.

3.1 Muuttujien kuvailu

Aiemmin tehtyjen tutkimusten (esimerkiksi Scherer 1965, Coad & Rao-Nicholson 2007) mukaisesti fyysisissä tekijöissä keskitytään yrityksen kokoon. Tavalliset tunnusluvut yrityksen koon mittaamiseksi ovat esimerkiksi liikevaihto, tasearvo sekä henkilöstömäärä. Liikevaihto on varsinaisesta toiminnasta syntyneitä myyntituottoja alennusten ja arvonalisäveron jälkeen. Tasearvolla tai taseen loppusummalla tarkoitetaan yritykseen sitoutuneen pääoman määrää. Toisin sanoen se mittaa yhtiön varoja. Henkilöstömäärä tarkoittaa nimensä mukaisesti yrityksessä työskentelevien henkilöiden määrää vuotuisena keskiarvona ilmoitettuna.

Multikollinearisuuden välttämiseksi tutkimuksessa käytetään vain yhtä yrityksen kokoa mittaavaa fyysistä tekijää. Henkilöstömäärää ei käytetä, koska yritykset voivat käyttää huomattavan paljon erilaisia henkilöstöpalveluita omien työntekijöiden sijaan, minkä takia se ei ole niin tarkka yrityskoon mittari kuin muut esitetyt mittarit. Tämän takia tutkimuksessa yrityksen kokoa mittaavana selittävänä tekijänä käytetään liikevaihtoa. Ristiriitaisen teorian takia on vaikea ennustaa, onko tällä muuttujalla yhteyttä t&k -menoihin, ja jos on, minkä suuntaista. Tämän takia liikevaihdosta muodostetaan kaksisuuntainen hypoteesi:

H₁: ”Liikevaihto ei vaikuta t&k -menojen suuruuteen.”

Taloudellisista tekijöistä käsittelevässä luvussa esille nousi se, että tutkimus ja kehitys joudutaan usein rahoittamaan sisäisesti. Tätä ilmiötä tutkitaan ottamalla mukaan muuttuja nettovelkaantumisasaste-%. Nettovelkaantumisasaste-%, eli gearing, kuvaa yrityksen korollisen nettovelan ja oman pääoman suhdetta. Toisin sanoen on kyse omistajien yritykseen sijoittamien omien pääomien ja rahoittajilta lainattujen korollisten velkojen suhteesta. Nettovelkaantumisasaste-% lasketaan kaavalla: (Ikäheimo et al. 2011, 72)

$$\text{Nettovelkaantumisasaste} - \% = \frac{\text{korollinen vieras pääoma} - \text{rahat ja arvopaperit}}{\text{oma pääoma}}. (1)$$

Teorian pohjalta voidaan tehdä hypoteesi:

H₂: "Velkaantumisasaste prosentien noustessa yrityksen t&k -menot pienenevät."

Kuten sanottu, aineettomien tekijöiden mittaaminen on huomattavasti haastavampaa kuin taloudellisten tai fyysisten tekijöiden mittaaminen. Aineettomissa tekijöissä keskitytään aiempien tutkimusten mukaisesti henkiseen pääomaan. Mittauksen vaikeuden takia tutkimuksessa käytetään proxy-muuttujaa keskimääräiset henkilöstökulut. Vaikka korkea palkka ei aina tarkoita taidokkaampia työntekijöitä, on kohtuullista olettaa korkeamman keskipalkan tarkoittavan korkeampaa kertynyttä henkistä pääomaa yrityksessä (Wakelin 1998). Keskimääräiset henkilöstökulut lasketaan kaavalla:

$$\text{Keskimääräiset henkilöstökulut} = \frac{\text{Henkilöstökulut}}{\text{Työntekijöiden määrä}} \quad (2)$$

Teorian pohjalta on syytä olettaa, että yritys, jolla on paljon henkistä pääomaa, investoi myös innovaatioon eli tutkimukseen ja kehitykseen enemmän. Näin voidaan muodostaa hypoteesi:

H₃: "Keskimääräisten henkilöstökulujen noustessa yrityksen t&k -menot kasvavat."

Keskipalkan käyttämisessä on kuitenkin ongelmansa. Kuten aiemmin sanottu, yrityksillä on hyvin paljon vapautta siinä, minkä ne luokittelevat t&k -menoksi. Näin ollen osalla yrityksistä yhteys voi olla jopa negatiivinen, jos kaikkia t&k -menoja ei ole erikseen ilmoitettu, vaan ne näkyvät vain korkeissa henkilöstökuluissa.

3.2 Paneelidata

Kuten sanottu, paneelidatalla tarkoitetaan aineistoa, jossa tarkastellaan useaa poikkileikkausyksikköä useana eri ajanhetkenä (Hill et al. 2012, 9). Siinä on siis sekä

aikasarja- että poikkileikkausaineistolle tyypillisiä elementtejä. Paneelidatassa on useita hyviä puolia, jotka tekevät tutkimuksesta kattavamman ja luotettavamman. Siinä on yleisesti huomattavasti enemmän datapisteitä kuin poikkileikkaus- tai aikasarjadataassa, jolloin vapausaste on korkeampi ja selitettävien muuttujien kollineaarisuus pienempi (Hsiao 2003, 4-5).

Paneelidata ratkaisee myös suurimmaksi osaksi yritysten eriävien laskentatapojen aiheuttaman ongelman. Koska käytännöt yritysten sisällä harvemmin muuttuvat, suurin ero aineistossa on yritysten välillä yksittäisen yrityksen yli ajan tapahtuvien muutoksien sijaan. Näin ollen regressioanalyysi selittää pääasiassa yritysten välisiä eroja tai yhtäläisyyksiä selittävissä muuttujissa. (Grabowski 1968)

Tässä tutkimuksessa käytettävä aineisto koostuu 78 yrityksestä, joista tietoa on kerätty kymmenen vuoden ajalta. Aineistoa voidaan siis kuvailla leveäksi ja lyhyeksi, mikä nostaa aineiston poikkileikkausominaisuudet entistäkin korostetumpaan asemaan. Aineisto on myös taspainoinen, koska kaikista yhtiöistä on tiedossa kaikki muuttajat koko kymmenen vuoden ajalta (Hill et al. 2012, 538-539).

3.3 Paneelidatan analysointi

Tutkimuksessa halutaan selvittää, kuinka johonkin tiettyyn muuttujaan y vaikuttaa useampi muu muuttuja x . Tällaisia muuttujien välisiä yhteyksiä voidaan tarkastella paneelidatan regressioanalyysillä. Kyseistä analyysiä tehdessä täytyy ensin valita estimointimenetelmä, joka käyttää paneelidatan tietoa mahdollisimman tehokkaasti sekä antaa mahdollisimman merkityksellisiä ja harhattomia tuloksia. Seuraavaksi käydään läpi kolme paneelidatan regressioanalyysin lähestymistapaa sekä perusteet, joilla estimointimenetelmä valitaan.

Ensimmäiseksi on syytä tarkastella, onko aineiston yksiköillä yksilöllisiä eroja, jotka voisivat aiheuttaa eroavia kertoimia. Jos eroja yksiköiden välillä ei ole, mallina voidaan käyttää yhdistettyä OLS:ää (pooled ordinary least squares). Yhdistetty OLS ei itse

asiassa ole paneelidatan analysointimalli, vaan siinä jätetään huomiotta aineiston paneelimaisuus ja suoritetaan tavallinen usean selittävän muuttujan regressioanalyysi. Näin ollen yhdistetyn OLS:n kaava on: (Hill et al. 2012, 540)

$$y_{it} = \beta_1 + \beta_2 x_{2it} + \beta_3 x_{3it} + \dots + e_{it}. \quad (3)$$

Koska tässä tutkimuksessa tutkitaan tekijöitä, jotka voisivat selittää yksiköiden käyttäytymistä, on hyvin epätodennäköistä, että yksiköillä ei olisi eroja ja yhdistettyä OLS:ia voitaisiin käyttää. Vaihtoehtoja tälle ovat kiinteiden vaikutusten malli sekä satunnaisten vaikutusten malli. Kiinteiden vaikutusten mallissa kaikki erot yksiköiden välillä oletetaan kuuluvan vakiotermiin. Mallilla on kaksi eri laskutapaa, dummy muuttuja -malli sekä kiinteiden vaikutusten estimaattori. Dummy muuttuja -mallia ei tässä tutkimuksessa käytetä, sillä sitä on luontevaa käyttää vain, kun tutkittavia yksiköitä on vähän. Kiinteiden vaikutusten estimaattorissa käytetään yksikkökohtaisia yli ajan otettuja keskiarvoja. Muuttujina toimivat näiden lukujen poikkeamat keskiarvosta, jolloin kaavaksi muodostuu:

$$\tilde{y}_{it} = \beta_2 \tilde{x}_{2it} + \beta_3 \tilde{x}_{3it} + \dots + \tilde{e}_{it}. \quad (4)$$

Estimaattorin heikkous on se, että malliin ei voida sisällyttää muuttujia, jotka säilyvät vakioina yli ajan. Tällaisia muuttujia ei kuitenkaan tässä tutkimuksessa käytetä, joten ongelmaa ei ole.

Kiinteiden vaikutusten mallissa tutkitaan vain niitä yksiköitä, joita aineisto koskee. Satunnaisten vaikutusten malli taas olettaa aineiston olevan satunnaisotos jostain populaatiosta, ja tutkimus kohdistuu koko populaatioon. Myös satunnaisten vaikutusten mallissa kaikki erot yksiköiden välillä oletetaan kuuluvan vakiotermiin, mutta tässä mallissa huomioidaan yksiköiden satunnaiset muutokset jakamalla vakio-termi kahteen osaan, populaation keskiarvoon sekä satunnaiseen yksikkökohtaiseen vaihteluun:

$$\beta_{1i} = \bar{\beta}_1 + u_i. \quad (5)$$

Yksikkökohtainen vaihtelu ja mallin residuaali niputetaan usein yhteen, jolloin satunnaisten vaikutusten malli yhtälömuodossa on:

$$y_{it} = \bar{\beta}_1 + \beta_2 x_{2it} + \beta_3 x_{3it} + \dots + v_{it}. \quad (6)$$

Koska tässä tutkimuksessa aineistona toimii kohtuullisen pieni otos kaikista Euroopan teollisuusyrityksistä, olisi helppoa olettaa satunnaisten vaikutusten mallin olevan paras malli. Ensin täytyy kuitenkin tutkia, onko asia näin, ja voidaanko kyseistä mallia käyttää. Taulukko 1 esittää mallinvalintaprosessin. Seuraavana käydään läpi käytettävät testit sekä niiden tuloksien tulkinta.

Taulukko 1. Käytettävän mallin valinta (Park 2010)

F-testi*	Breusch-Pagan testi**	Käytettävä malli
H ₀ jää voimaan	H ₀ jää voimaan	Yhdistetty OLS
H ₀ hylätään	H ₀ jää voimaan	Kiinteiden vaikutusten malli
H ₀ jää voimaan	H ₀ hylätään	Satunnaisten vaikutusten malli
H ₀ hylätään	H ₀ hylätään	1) Satunnaisten vaikutusten malli, jos Hausman-testin H ₀ jää voimaan, muulloin kiinteiden vaikutusten malli, tai 2) Sekä satunnaisten- että kiinteiden vaikutusten malli

* F-testillä tutkitaan kiinteiden vaikutusten läsnäoloa mallissa nollahypoteesinaan se, että kiinteitä vaikutuksia ei ole.

** Breusch-Pagan testillä tutkitaan satunnaisten vaikutusten läsnäoloa mallissa nollahypoteesinaan se, että satunnaisia vaikutuksia ei ole.

Yhdistetyn OLS:in käytettävyyttä testataan kahdella testillä, joilla tutkitaan, onko mallissa satunnaisia tai kiinteitä vaikutuksia. Ensimmäisenä testinä on Breusch-Pagan, jolla testataan heteroskedastisuutta. Testin nollahypoteesi on, että satunnaisten vaikutusten varianssi yksiköiden välillä on nolla:

$$H_0: \sigma_u^2 = 0. \quad (7)$$

Testin nollahypoteesin jäädessä voimaan mallissa ei ole satunnaisia vaikutuksia, ja yhdistettyä OLS:ää voidaan käyttää, jos myöskään kiinteitä vaikutuksia ei löydy. Kiinteitä vaikutuksia voidaan verrata yhdistettyyn OLS:ään kiinteiden vaikutusten F-testillä. Testi vertailee mallien sopivuutta nollahypoteesinaan se, että molemmat mallit toimivat aineistolle yhtä hyvin. Jos nollahypoteesi jää voimaan, kiinteitä vaikutuksia ei ole, eikä näin ollen kiinteiden vaikutusten mallin käyttö ole perusteltua.

Tilanteessa, jossa sekä Breusch-Pagan -testin että F-testin nollahypoteesit hylätään, vertaillaan kiinteiden vaikutusten mallia satunnaisten vaikutusten malliin. Vertailu tehdään Hausman-testillä, jolla tutkitaan yksikkökohtaisten virhetermien korrelaatiota selittävien muuttujien kanssa. Jos korrelaatiota ei ole, sekä kiinteiden vaikutusten malli että satunnaisten vaikutusten malli ovat konsistentteja. Tässä tapauksessa käytetään satunnaisten vaikutusten mallia, koska GLS-estimaattorina sen varianssi on pienempi kuin OLS-estimaattoreilla suurissa otoksissa (Hill et al. 2012, 557). Vaihtoehtona Hausman-testille on arvioida sekä satunnaisten vaikutusten mallin että kiinteiden vaikutusten mallin tuloksia (Park 2010).

Käytettäessä kiinteiden- tai satunnaisten vaikutusten mallia täytyy lisäksi päättää, käytetäänkö yksi- (one-way) vai kaksisuuntaista (two-way) menetelmää. Yksisuuntaisessa menetelmässä residuaaliin kuuluu kiinteiden vaikutusten mallissa vain yksiköiden ja ajan välillä vaihteleva e_{it} . satunnaisten vaikutusten yksisuuntaisessa mallissa residuaaliin kuuluu yksiköiden välillä vaihteleva u_i sekä yksiköiden ja ajan välillä vaihteleva e_{it} . Kaksisuuntainen menetelmä ottaa näiden lisäksi residuaalissaan huomioon mallin ulkopuoliset tekijät λ_t , jotka vaikuttavat kaikkiin yksiköihin samalla tavalla tiettyä aikana. (Mátyás & Sevestre 2008, 50, 64) Jotta kaikkiin yksiköihin tiettyä aikana vaikuttavat tekijät, joita voisi olla esimerkiksi koko toimialaan vaikuttava lama tai raaka-aineiden tarjontashokki, tulisi huomioitua, tässä tutkimuksessa käytetään kaksisuuntaista menetelmää.

4. TULOKSET

Tässä luvussa käsitellään tutkimuksen empiriaa. Luku koostuu ensin aineiston sekä sen muokkaamisen kuvailusta, minkä jälkeen perustellaan käytettävän estimointimenetelmän valintaa sekä esitellään regressioanalyysin tuloksia. Lopuksi arvioidaan tutkimuksen oikeellisuutta ja luotettavuutta.

4.1 Aineiston kuvailu ja muokkaus

Tutkimuksessa käytettävä aineisto on Amadeus-tietokannasta haettu leveä ja lyhyt paneelidata. Tietokannasta etsitään toiminnassa olevat tai toimintatilanteeltaan tuntemattomat yritykset, jotka sijaitsevat Itä- tai Länsi-Euroopassa ja toimivat NAICS 2017 -lajittelujärjestelmän mukaisesti lajiteltuna teollisuusalalla. Täydelliset Amadeus-tietokannan käyttämät hakuperusteet löytyvät liitteestä 1. Näiden lisäksi yrityksestä täytyy olla käytettävissä täysin eheät tiedot kaikkien tutkimuksessa käytettävien muuttujien osalta viimeisen 10 vuoden ajalta. Tutkimuksessa päätettiin käyttää tasapainotettua aineistoa, koska näin päästään eroon mahdollisuudesta, että aineisto on tasapainottoman jostain tietystä tutkimukseen vaikuttavasta syystä, mikä näin ollen voisi muuttaa tutkimuksen tuloksia huomattavasti. Tasapainotettua aineistoa on käytössä niin paljon, että tutkimuksesta saadaan riittävä, joten riskiä ei ole tarpeellista ottaa. Toisaalta tämä tarkoittaa myös sitä, että tässä tutkimuksessa huomioidaan vain yli 10 vuotta vanhat yritykset. Lopulta kaikki hakukriteerit täyttävät 78 yritystä, joiden nimet ovat listattu liitteessä 2.

Valituista yrityksistä kerätään nimen ja valtion lisäksi tutkimus- ja kehitysmenot, liikevaihto, velkaantumisasentti sekä keskimääräiset henkilöstökulut kymmenen vuoden ajalta. Ensin Amadeuksesta ladattu leveä paneelidata muunnetaan tilastollisen tutkimuksen ohjelmistoille käytettävään pitkään muotoon, jossa yhdellä rivillä esitetään yhden yrityksen yhden vuoden tiedot. Muunnos tehdään RStudio-

ohjelmointiympäristössä, ja muunnoksen tulos ja käytetty koodi on esitetty liitteissä 3 ja 4.

Tutkimuksessa käytettävien muuttujien perustiedot esitetään taulukossa 2. Muuttuja *rdtosales* on tehty jakamalla muuttuja *rd* muuttujalla *sales*. Taulukosta nähdään, että aineistossa on hyvin erikokoisia yhtiöitä, pienimmän liikevaihdon ollessa 9 miljoonaa euroa ja suurimman 79,6 miljardia euroa. Muuttujan *gearing* luvuista nähdään, että myös pääomarakenne on hyvin erilainen yritysten välillä. Keskiarvona yrityksillä on yhtä paljon korollisia velkoja kuin omia varoja. Muuttuja *wage* on muuttujista selvästi tasaisemmin jakautunut. Tämän voi huomata siitä, että se on ainoa muuttuja, jonka keskihajonta on keskiarvoa pienempi. Selitettävästä muuttujasta *rdtosales* nähdään jälleen kuinka erilaisia yrityksiä tutkimuksessa on, sillä t&k –menojen suhde liikevaihtoon vaihtelee nolasta lähes viiteenkymmeneen prosenttiin.

Taulukko 2. Muuttujien rdtosales, sales, gearing ja wage perustiedot. rdtosales sekä gearing ovat esitetty prosentteina ja sales sekä wage ovat esitetty tuhansina euroina

Muuttuja	Keskiarvo	Keskihajonta	Minimi	Maksimi	N
rdtosales	5,86	6,15	0	47,4	780
sales	4320000	12500000	9000	79600000	780
gearing	99,9	117	2,75	930	780
wage	60,3	17,9	8,38	158	780

Regressioanalyysissä muuttujien tulisi olla normaalijakautuneita. Vahvasti vinoutuneita jakaumia voidaan muuttaa normaalijakautuneiksi erilaisten muunnosten, kuten logaritmi- tai neliöjuurimuunnosten, avulla. Muuttujamuunnosten huono puoli on se, että ne vaikeuttavat tulosten luettavuutta sekä ääriarvojen painoarvoa. (Hill et al. 2012, 71-73)

Liitteessä 5 esitetään muuttujien jakautuma. Vaikka jo paljaalla silmällä on helppoa nähdä muuttujien *rdtosales*, *sales*, ja *gearing* vahva vinoutuneisuus sekä huipukkuus, on hyvä todeta asia myös tunnuslukujen avulla. Taulukko 3 esittää muuttujien

vinouden ja huipukkuuden. Molempia tunnuslukuja luetaan niin, että lähelle nollaa sijoittuvat arvot tarkoittavat muuttujan olevan lähes normaalijakautunut. Taulukosta havaitaan, että kaikki muuttujat *wage*:a lukuunottamatta ovat hyvin vinoutuneita ja huipukkaita.

Taulukko 3. Muuttujien vinous ja huipukkuus

Muuttuja	Vinous	Huipukkuus
rdtosales	3,06	13,12
sales	4,36	19,69
gearing	3,50	16,81
wage	0,58	1,59

Oikealle vinoutuneen muuttujan voi muuttaa lähemmäs normaalia logaritimuunnoksella (Hill et al. 2012, 71). Muuttujista *rdtosales*, *sales* ja *gearing* otetaan kymmenkantainen logaritmi, minkä jälkeen muuttujia tarkastellaan uudestaan. Taulukosta 4 nähdään uusien muuttujien *lgrdtosales*, *lgsales* ja *lggearing* vinous ja huipukkuus.

Taulukko 4. Logaritmisten muuttujien vinous ja huipukkuus

Muuttuja	Vinous	Huipukkuus
lgrdtosales	-1,02	2,25
lgsales	0,53	-0,51
lggearing	-0,34	0,10

Kaikkien muuttujien vinous- ja huipukkuusarvot ovat selvästi lähempänä nollaa, joten tutkimuksessa päädytään käyttämään näitä logaritmisia muuttujia. Myös liitteessä 6 esitetyistä jakaumakuvaajista voidaan nähdä, että logaritmiset muuttujat ovat huomattavasti paremmin normaalijakautuneet kuin alkuperäiset muuttujat.

Liitteen 5 jakaumakuvaajista nähdään myös, että muuttujissa on paljon mahdollisia outlier-havaintoja eli ääriarvoja. Näitä ei kuitenkaan tässä vaiheessa tutkimuksessa poisteta, sillä kyseessä on ennemminkin vaikutusvaltainen havainto kuin ääriarvo. Ääriarvolla tarkoitetaan havaintoa, joka on numeerisesti kauempana odotetuista arvoista kuin on kohtuullista olettaa. Vaikutusvaltainen havainto taas on havainto, joka vaikuttaa huomattavasti analyysin tulokseen. (Andersen & Skovgaard 2010, 175) Ensimmäisen tutkimuksen jälkeen voidaan kuitenkin kokeilla mallin parantamista näitä mahdollisia ääriarvoja poistamalla.

4.2 Estimointimenetelmän valinta ja tulokset

Muuttujamuunnosten jälkeen tutkimukseen ollaan saatu selitettävä muuttuja *lgrdtosales* sekä selittävät muuttujat *lgsales*, *lggearing* ja *wage*. Seuraavaksi valitaan luvun 3.3 mukaisesti mallille sopivin estimointimenetelmä ja tulkitaan itse tutkimuksen tulokset.

Estimointimenetelmän valintaan liittyvien testien tulokset esitetään taulukossa 5. Breusch-Paganin testin p-arvo, joka on alle 0,001, osoittaa, että mallissa on satunnaisia vaikutuksia. Tämä tarkoittaa sitä, että yhdistettyä OLS:ää ei voida käyttää. Myös F-testin nollihypoteesi hylätään, eli mallissa on myös kiinteitä vaikutuksia. Tässä tilanteessa voidaan joko tulkita sekä kiinteiden- että satunnaisten vaikutusten mallien tuloksia tai valita jompikumpi Hausman-testin avulla. Hausman-testin p-arvo on 0,001, mikä viittaa siihen, että satunnaisten vaikutusten malli ei ole konsistentti, ja on parempi käyttää kiinteiden vaikutusten mallia.

Luvussa 3.1 tehtiin teorian pohjalta hypoteeseja, millä tavalla valitut selittävät muuttujat vaikuttavat t&k –menojen suuruuteen. Hypoteesit olivat, että nettovelkaantumisaste vaikuttaa negatiivisesti, keskimääräiset henkilöstökulut positiivisesti ja liikevaihdon vaikutuksen suuntaa ei pystytty ennustamaan. Taulukkoon 5 on kerätty tehdyn kiinteiden vaikutusten mallin parametriestimaatit sekä tilastolliset merkitsevyydet. Muuttujilla *lggearing* ja *wage* p-arvot ovat erittäin korkeat ja parametriestimaatit

lähenevät nollaa. Tästä voidaan sanoa, että näillä muuttujilla ei ole yhteyttä selitettävään muuttujaan *lgrdtosales*.

Taulukko 5. Estimointimenetelmän valintatesti, kiinteiden vaikutusten mallin parametriestimaatit ja muut perustiedot. Selitettävä muuttuja on lgrdtosales, joka on logaritmuunnos liikevaihtoon suhteutetuista t&k -menoista.

Testi	P-arvo
Breusch-Paganin testi	<0,0001
F-testi	<0,001
Hausman-testi	0,0004

Muuttuja	Kerroin	Keskivirhe	P-arvo
vakiotermi	2,30	0,38	<0,001
lgsales	-0,37	0,07	<0,001
lggearing	-0,00	0,03	0,98
wage	0,001	0,001	0,60

Estimointimenetelmä	Fixed Two Way Estimate
Havaintojen lukumäärä	742
Selitysaste	0,88

Muuttujalla *lgsales* on mallin mukaan negatiivinen, yhden prosentin riskitasolla tilastollisesti merkitsevä vaikutus t&k-menojen suhteelliseen suuruuteen. Tulos on siis ristiriidassa Schumpeterin (1961) hypoteesin kanssa. Tämä ei ole kuitenkaan mikään yllätys, sillä monet aikaisemmat tutkimukset (esimerkiksi Scherer 1965, Comarov 1967) ovat saaneet ristiriitaisia tuloksia Schumpeterin hypoteesin kanssa. Logaritmuunnosten takia parametriestimaattia on haastava tulkita. Tulos tulee tulkita niin, että kun *sales* kasvaa yhden prosentin, *rdtosales* pienenee 0,37 prosenttia. Esimerkiksi jos yrityksen t&k –menot ovat 5% liikevaihdosta ja liikevaihto kasvaa 10%, uudet t&k –menot näin ollen olisivat 4,815%.

Muutos ei ole kovinkaan suuri, mikä nostattaa kysymyksen, miksi analyysin selitysaste on niinkin korkea kuin 0,88. Korkean selitysasteen voi selittää muuttujien *lgrdtosales* ja *lgsales* välinen osittaiskorrelaatio, joka tarkoittaa korrelaatiota sen jälkeen, kun yhden tai useamman muuttujan vaikutus on poistettu. Syynä voi olla myös multikollinearisuus selittävien muuttujien välillä. Lineaarista multikollinearisuutta voidaan tutkia Pearsonin korrelaatiokertoimien avulla. Taulukkoon 6 on koottu selittävien muuttujien Pearsonin korrelaatiokertoimet ja p-arvot. Taulukosta nähdään, että ainoastaan muuttujien *lgsales* ja *lggearing* välillä voidaan sanoa olevan tilastollisesti merkitsevää korrelaatiota. Tämäkin korrelaatio on ainoastaan 0,31, joten multikollinearisuudesta ei missään nimessä voida puhua. Tältä osin muuttujien valinta on siis onnistunut hyvin. Tuloksia tulkittaessa on kuitenkin hyvä muistaa, että selitysaste on harhainen. Täydellinen SAS-tuloste analyysille löytyy liitteestä 7.

Taulukko 6. Pearsonin korrelaatiokertoimet ja p-arvot

	wage	lgsales	lggearing
wage	1		
lgsales	0,05570 0,1213	1	
lggearing	0,00557 0,8769	0,31391	1 <0,0001

Seuraavaksi pyritään parantamaan mallia poistamalla mahdollisia ääriarvoja. Liitteen 8 laatikko-jana-kuvioista nähdään, että kaikki mallin muuttujat saavat paljon huomattavan isoja arvoja. Näiden arvojen poistaminen voi paljastaa yhteyksiä selitettävän ja selittävien muuttujien välillä yrityksissä, jotka ovat kyseisiltä arvoiltaan lähempänä keskiarvoa.

Ääriarvojen poistamiseen käytetään Interquartile range- eli kvartiiliväli-menetelmää. Menetelmän ideana on poistaa aineistosta kaikki havainnot, joiden arvot eivät ole välillä: (Hubert & Van der Veecken 2007)

$$[Q_1 - 1,5 * IQR, Q_3 + 1,5 * IQR]. \quad (8)$$

Yhtälössä Q_1 on alakvartaali, Q_3 on yläkvartaali ja IQR on yläkvartaalin ja alakvartaalin erotus. Normaalijakautuneella aineistolla kyseinen menetelmä poistaa n. 0,7% havainnoista. Vinoutuneella datalla poistettavien havaintojen määrä on suurempi. (Hubert & Van der Veecken 2007)

Liitteessä 9 esitetään käytetyt ylä- ja alarajat havaintojen suodattamiselle. Toimenpiteen jälkeen aineistoon jää 586 havaintoa, joka on yhä tarpeeksi suuri tutkimuksen toteuttamiselle. Aineistolle toistetaan aiemmat testit, ja tulokset esitetään Taulukossa 7.

Taulukko 7. Suodatetun aineiston estimointimenetelmän valintatesti, satunnaisten vaikutusten mukaiset parametriestimaatit ja muut perustiedot. Selitettävänä muuttujana on lgrdtosales.

Testi	P-arvo
Breusch-Paganin testi	<0,0001
F-testi	<0,001
Hausman-testi	0,0984

Muuttuja	Kerroin	Keskivirhe	P-arvo
vakiotermi	1,12	0,24	<0,001
lgsales	-0,09	0,04	0,0436
lggearing	-0,04	0,03	0,1456
wage	0,001	0,001	0,7440

Estimointimenetelmä	Random Two Way Estimate
Havaintojen lukumäärä	560
Selitysaste	0,0126

Taulukosta 7 nähdään, että edelleen Breusch-Paganin testi osoittaa, että mallissa on satunnaisia vaikutuksia. Mallissa on myös kiinteitä vaikutuksia, mikä nähdään F-testin nollihypoteesin hylkäämisestä. Erona tässä analyysissä edelliseen on se, että suodatetulla aineistolla Hausman-testin nollihypoteesi jää voimaan. Tämä tarkoittaa sitä, että molemmat mallit ovat konsistentteja, ja satunnaisten vaikutusten mallia voidaan käyttää. GLS-estimaattorina sen varianssi on pienempi kuin OLS-estimaattoreilla suurissa otoksissa (Hill et al. 2012, 557).

Parametriestimaatit näyttävät hyvin samanlaisilta verrattuna ensimmäiseen analyysiin. Muuttujat *lggearing* ja *wage* eivät edelleenkään saa merkitseviä tuloksia. Muuttuja *lgsales* saa tilastollisesti merkitseviä tuloksia viiden prosentin riskitasolla. Parametriestimaatti tälle muuttujalle on -0,09, mikä on hieman pienempi kuin ensimmäisessä analyysissä. Tämä viittaa siihen, että isoimmilla yrityksillä suhteelliset t&k –menot laskevat nopeammin liikevaihdon kasvaessa kuin muilla yrityksillä. Mallin selitysaste laski estimointimenelmää muuttamalla uskottavammalle tasolle 1,26 prosenttiin. Täydellinen SAS-tuloste analyysille löytyy liitteestä 10.

Sekä kiinteiden vaikutusten malli että suodatetulle aineistolle tehty satunnaisten vaikutusten malli antoivat saman tuloksen, mikä lisää tutkimustulosten uskottavuutta. Molemmilla malleilla havaittiin, että liikevaihdolla on negatiivinen yhteys yrityskokoon suhteutettujen t&k –menojen suuruuteen, tosin parametriestimaatti on hyvin lähellä nollaa, eli löydetty yhteys on hyvin pieni eikä selitä kovinkaan merkittävää osaa t&k –menojen suuruudesta. Näin ollen kaikki tutkimuksen hypoteesit H_1 , H_2 ja H_3 hylätään.

4.3 Tutkimuksen luotettavuuden arviointi

Seuraavaksi pohditaan paneelidatan regressioanalyysin taustaoletusten täyttymistä sekä tutkimuksen luotettavuutta. Kandidaatintutkielman vaatiman osaamistason sekä aineiston luonteen huomioiden kaikkia taustaoletuksia ei tässä käydä läpi. Esimerkiksi koska kaikkien mallissa käytettävien muuttujien arvot ovat vahvasti riippuvaisia

edellisten vuosien arvoista, mallissa on varmasti autokorrelaatiota, eikä asiaa tutkita testin avulla.

Liitteen 11 kvantiili- ja jakaumakuviosta voidaan tulkita mallien residuaalien normaalijakautuneisuutta sekä keskiarvoja. Normaalijakautuneisuus on tärkeä ominaisuus, jotta mallille voidaan tehdä hypoteeseja, kuten eroavatko muuttujien parametrit merkitsevästi nolasta (Wooldridge 2009, 172). Kuvasta nähdään, että kiinteiden vaikutusten mallin residuaalit (ylemmät kuvaajat) ovat kohtuullisen normaalijakautuneita pientä huipukkuutta lukuun ottamatta. Myös residuaalien keskiarvo on lähellä nolaa. Satunnaisten vaikutusten mallilla (alemmat kuvaajat) oletus normaalijakautuneisuudesta ei toteudu, vaan jakauma on vinoutunut vasemmalle, ja residuaalien keskiarvo on selvästi nolaa suurempi. Tämä on kuitenkin kohtuullisen normaalia taloustieteessä tehdyille malleille (Brooks 2008, 162).

On erittäin mahdollista, että muuttujalla *wage* ei tässä tutkimuksessa saatu merkitseviä tuloksia, koska se on huono proxy-muuttuja tutkittavalle asialle, joka on siis henkisen pääoman määrä. Tulevissa tutkimuksissa henkistä pääomaa voitaisiin kuvata jollain muulla mittareilla, esimerkiksi patenttien määrällä tai tuotettujen patenttien määrä suhteessa työntekijöiden määrään. Voidaan siis sanoa, että tämän muuttujan osalta tutkimus ei ole kovin validi, eikä tutkimuksen pohjalta voida tehdä johtopäätöstä, että aineettoman pääoman määrä ei vaikuta t&k –menoihin.

Muuttujan *lggearing* toimimattomuutta on vaikeampi selittää. Teoria tuki vahvasti velkaisuuden vaikutusta t&k –menojen suuruuteen. Voi olla, että yhteys olisi löytynyt, jos olisi käytetty rahallisesti mitattua mittaria joko oman pääoman määrästä tai veloista. Toisin sanoen näiden kahden suhdeluvulla ei välttämättä ole vaikutusta t&k -menoihin, vaan ainoastaan absoluuttisten varojen määrällä. Tämänkään muuttujan puolesta ei siis voida tehdä lopullista johtopäätöstä siitä, vaikuttaako yrityksen pääomarakenne sen t&k –aktiivisuuteen.

5. JOHTOPÄÄTÖKSET

Tutkimuksessa tarkasteltiin yritysten sisäisten tekijöiden vaikutusta niiden kokoon suhteutettujen tutkimus- ja kehitysmenojen suuruuteen. Vaikutusta tutkittiin paneelidatan regressioanalyysillä, jolla tutkittiin teorian pohjalta valittujen muuttujien yhteyttä suhteellisiin t&k –menoihin. Tutkimus kohdistui 78:aan Euroopan teollisuusalan yritykseen 10 vuoden ajanjaksolla. Maantieteellisellä ja alakohtaisella rajauksella ei ollut teoriasta noussutta syytä, vaan rajaus oli tehty aineiston saatavuutta ajatellen, sekä sillä tavoitteella, että tutkittavilla yrityksillä olisi mahdollisimman yhtenevät ulkoiset olosuhteet.

Teoriaosuudessa aihetta lähestyttiin näkökulmasta, jossa t&k –menoihin vaikuttavat tekijät jaoteltiin ulkoisiin ja sisäisiin tekijöihin. Sisäiset tekijät jaettiin vielä fyysisiin, taloudellisiin ja aineettomiin tekijöihin. Näistä kolmesta kategoriasta etsittiin teoriakirjallisuudesta sekä aiemmista tutkimuksista tekijöitä, joilla t&k –menojen suuruutta on pyritty selittämään. Jokaisesta kategoriasta valittiin yksi selittävä muuttuja, minkä jälkeen jokaisesta muodostettiin teorian pohjalta hypoteesi, miten kyseinen muuttuja vaikuttaa. Valitut muuttujat olivat liikevaihto, nettovelkaantumisasaste prosentti sekä keskimääräiset henkilöstökulut.

Tutkimustuloksista huomattiin, että liikevaihdolla on negatiivinen yhteys suhteellisiin t&k –menoihin, tosin yhteys on kohtuullisen pieni. Nettovelkaantumisasasteella tai keskimääräisillä henkilöstökuluilla ei huomattu olevan vaikutusta t&k –menoihin. Näin ollen tutkimuskysymyksiin saatiin vastaukset:

”Mitkä yrityksen sisäiset tekijät vaikuttavat tutkimus- ja kehitysmenoihin?”

”Liikevaihto.”

”Miten ja kuinka paljon löydetyt tekijät vaikuttavat tutkimus- ja kehitysmenoihin?”

”Liikevaihdon vaikutus on negatiivinen ja pieni.”

Tutkimusta tehdessä heräsi muutamia ehdotuksia jatkotutkimukselle. Aineiston sen salliessa tutkimus olisi hyvä rajata entistäkin pienemmälle alueelle ja toimialalle. Tutkimuksen luonteen huomioon ottaen voisi olla järkevää tutkia vain muutaman teorian pohjalta potentiaalisimman tekijän sijaan kymmeniä eri muuttujia. Näitä muuttujia voisivat olla esimerkiksi taseen loppusumma, oman pääoman määrä, vieraan pääoman määrä, patenttien lukumäärä suhteessa työntekijöiden määrään, uusien patenttien määrä vuodessa, current ratio, suhteellinen velkaantuneisuusprosentti ja liikevaihdon kasvuprosentti. Jatkotutkimusaiheina ulkoisten tekijöiden vaikutuksen tutkiminen olisi iso ja mielenkiintoinen aihe, vaikkakin varmasti vaikea toteuttaa.

LÄHDELUETTELO

Akerlof, G. (1970) The Market for 'Lemons': Quality, Uncertainty, and the Market Mechanism. *Quarterly Journal of Economics*, vol. 84, s. 488-500

Andersen, P. & Skovgaard, L. (2010) Regression with Linear Predictors. Springer, London

Arrow, K. (1962) Economic Welfare and the Allocation of Resources for Invention. Princeton Univ. Press, Princeton

Arslan-Ayaydin, O., Barnum, D., Karan, M. & Ozdemir, A. (2014) How is Moral Hazard Related to Financing R&D and Innovations? *European Research Studies*, vol. 17, s. 111-131

Becker, B. & Pain, N. (2003) What Determines Industrial R&D Expenditure in the UK? *NIESR Discussion Papers*, vol. 211

Barney, J. (1991) Firm resources and sustained competitive advantage. *Journal of Management*, vol. 17, s. 99–120

Bhattacharya S. & Ritter, J. (1985) Innovation and Communication: Signalling with Partial Disclosure. *Review of Economic Studies*, vol. 52, s. 331-46.

Bougheas, S., Görg, H. & Strobl, E. (2001) Is R&D financially restrained? Theory and evidence from Irish manufacturing. *Leverhulme Centre for Research on Globalisation and Economic Policy*, vol. 16

Brooks, C. (2008) Introductory Econometrics for Finance. 2. p. Cambridge University Press, New York

Coad, A. & Rao-Nicholson, R. (2007) Firm Growth and R&D Expenditure. *Papers on Economics and Evolution*, vol. 19

Cohen, W., Mowery, D. (1984): The internal characteristics of the firm and the level and composition of research & development spending. Interim report NSF, Carnegie-Mellon University

Del Canto, J. & González, I. (1999) A resource-based analysis of the factors determining a firm's R&D activities. *Research Policy*, vol 28, s. 891-905

Demsetz, H. (1969) Information and Efficiency: Another Viewpoint. Palgrave, London

Foss, K. (1996) Transaction cost and technological development: the case of the Danish fruit and vegetable industry. *Research Policy*, vol. 25, s. 531–547

Galbraith, J. (1952) American Capitalism: The Concept of Countervailing Power. Houghton Mifflin Company, Boston

Germeraad, P. (2003) Measuring R&D in 2003. *Research-Technology Management*, vol 46, s. 47-56

Goodwin, M (1998) Firm Size and R&D; Testing the Schumpeterian Hypothesis. *University Avenue Undergraduate Journal of Economics*, vol. 2, artikkeli 4

Grabowski, H. (1968) The Determinants of Industrial Research and Development: A Study of the Chemical, Drug, and Petroleum Industries. *Journal of Political Economy*, vol. 76, s. 292-306

Hall, B. H., Mairesse, J. & Mohnen, P. (2009) Measuring the Returns to R&D. *Handbook of the Economics of Innovation*, vol. 2, s. 1033-1082

Hall, B. & Lerner, J. (2009) The Financing of R&D and Innovation. *Handbook of the Economics of Innovation*, vol. 1, s. 609-639

Hill, C., Griffiths, W. & Lim. G. (2012) Principles of Econometrics, Fourth Edition. Wiley, Hoboken (NJ)

Himmelberg, C. & Petersen, B. (1994) R & D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries. *The Review of Economics and Statistics*, vol 76, s. 38-51

Hsiao, C. (2014) Analysis of Panel Data, Third Edition. Cambridge University Press, Cambridge

Ikäheimo, S., Laitinen, E., Laitinen, T. & Puttonen, V. (2011) Laskentatoimi ja rahoitus. Vaasan Yritysinformaatio Oy, Vaasa

Kamien, M. & Schwartz, N. (1978) Self-Financing of an R&D Project. *American Economic Review*, vol. 68, s. 252-261

Levin, R., Klevorick, A., Nelson, R. & Winter, S. (1987) Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, s. 783-832

Majumdar S. (2009) Scalability versus flexibility: firm size and R&D in Indian industry. Springer US, New York

Mátyás, L. & Sevestre, P. (2008) The Econometrics of Panel Data, Third Edition. Springer, Berlin

Modigliani, F. & Miller, M. (1958) The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review*, vol. 48, s. 261-297

Myers S. & N. Majluf (1984) Corporate Financing and Investment Decisions when Firms Have Information that Investors Do Not. *Journal of Financial Economics*, vol. 13, s. 187-221

OECD (2015) Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development. OECD Publishing, Paris

Park, H. (2010) Practical Guides to Panel Data Analysis [verkkodokumentti] [viitattu 26.11.2017] Saatavilla:
http://www.iuj.ac.jp/faculty/kucc625/writing/panel_guidelines.pdf

Scherer, F. (1965) Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions. *American Economic Review*, vol. 55, 1097-1125

Schumpeter, J. (1942) *Capitalism, Socialism, and Democracy*. Harper and Row, New York

Schumpeter, J. (1961) *Theory of Economic Development*. Oxford Univ. Press, New York

Stiglitz, J. & Weiss, A. (1981) Credit Rationing in Markets with Imperfect Information. *American Economic Review*, vol. 71, s. 393-410

Teece, D. (1987) Profiting From Technological Innovation: Implications of Integration, Collaboration, Licensing and Public Policy, Teoksessa: Teece, D. *The Competitive Challenge: Strategies for Industrial Innovation and Renewal*. Harper & Row, New York, s. 185–221

Wakelin, K. (1998) Innovation and Export Behaviour at the Firm Level. *Research Policy*, vol. 26, s, 829–841

Williamson, O. (1988) Corporate finance and corporate governance. *Journal of Finance*, vol. 43, s. 567–591

Wooldridge, J. (2009) *Introductory Econometrics: A Modern Approach*. 4. p. Mason, South-Western Cengage Learning

Liitteet

Liite 1. Aineiston hakuperusteet

Product name	Amadeus		
Update number	278		
Software version	15.01		
Data update	17/11/2017 (n° 2782)		
Username	Lappeenranta university -1402		
Export date	27/11/2017		
Cut off date	31/03		
		Step result	Search result
1. All active companies and companies with unknown situation		21,023,076	21,023,076
2. Region/Country/region in country: Western Europe, Eastern Europe		24,475,947	21,023,076
3. NAICS 2017 (Primary codes only): 31 - Manufacturing, 32 - Manufacturing, 33 - Manufacturing		1,880,102	1,644,749
Boolean search : 1 And 2 And 3			
		TOTAL	1,644,749

Liite 2. Tutkimusaineiston yritykset

ADIDAS AG
ADVANCED NUCLEAR FUELS GMBH
AIXTRON SE
ALFMEIER PRÄZISION SE
ANGIOMED GMBH & CO. MEDIZINTECHNIK KG
APCA AUTOMOTIVE PLASTIC COMPONENTS ASSETS GMBH & CO. KG
B. BRAUN MELSUNGEN AKTIENGESELLSCHAFT
BASELL POLYOLEFINE GMBH
BASF SE
BAYER AKTIENGESELLSCHAFT
BEIERSDORF AKTIENGESELLSCHAFT
BOLLORE
BORCHERS GMBH
BRAINLAB AG
BRUKER AXS GMBH
BRUKER BIOSPIN GMBH
BRUKER BIOSPIN MRI GMBH
CARL ZEISS MEDITEC AG
CEOTRONICS AKTIENGESELLSCHAFT AUDIO. VIDEO. DATA COMMUNICATION
CLARIANT PLASTICS & COATINGS (DEUTSCHLAND) GMBH
CONDUCTIX-WAMPFLER GMBH
CONTINENTAL AKTIENGESELLSCHAFT

DASSAULT AVIATION
DIAGNOSTIC MEDICAL SYSTEMS
DILAS DIODENLASER GMBH
DRÄGERWERK AG & CO. KGAA
DÜRKOPP ADLER AKTIENGESELLSCHAFT
ELMA ELECTRONIC AG
ELMOS SEMICONDUCTOR AKTIENGESELLSCHAFT
ELRINGKLINGER AG
EPPENDORF AG
ESSILOR INTERNATIONAL
EUROGERM
FLUKE PROCESS INSTRUMENTS GMBH
GARDENA MANUFACTURING GMBH
GLEASON-HURTH TOOLING GMBH
GLEASON-PFAUTER MASCHINENFABRIK GMBH
GRÜNENTHAL GMBH
GUERBET
HAWKER GMBH
HENKEL AG & CO. KGAA
HILTI AKTIENGESELLSCHAFT
I2S
INFINEON TECHNOLOGIES AG
ITRON GMBH
JOHNSON & JOHNSON MEDICAL GMBH
KARL STORZ SE & CO. KG
KOENIG & BAUER AG
LANXESS AKTIENGESELLSCHAFT
LEIFHEIT AKTIENGESELLSCHAFT
LEONHARD KURZ STIFTUNG & CO. KG
LUKAS HYDRAULIK GMBH
MEDIGENE AG
MGI COUTIER
MTU AERO ENGINES AG
NATUREX
ORAPI
PRISMAFLEX INTERNATIONAL
PULSION MEDICAL SYSTEMS SE
RATIONAL AKTIENGESELLSCHAFT
ROFIN-BAASEL LASERTECH GMBH & CO. KG
ROFIN-SINAR LASER GMBH
ROGERS GERMANY GMBH
SCHOELLERSHAMMER GMBH & CO. KG
SFC ENERGY AG
SGL CARBON SE
SIEMENS AKTIENGESELLSCHAFT
SIGNAUX GIROD
SMA SOLAR TECHNOLOGY AG
SMART-BRABUS GMBH
SOCIETE QUANTEL

SYMRISE AG
 SÜSS MICROTEC SE
 WACKER CHEMIE AG
 WACKER NEUSON SE
 VAILLANT GMBH
 VETOQUINOL SA
 VIRBAC
 ZF FRIEDRICHSHAFEN AG

Liite 3. Katsaus muunnetusta aineistosta

Name	id	year	Country ISO code	rd	sales	gearing	wage
SIEMENS AKTIENGESELLSCHAFT	1	1	DE	3399000	72448000	83,505	68,407561
SIEMENS AKTIENGESELLSCHAFT	1	2	DE	3784000	77327000	120,771	60,945817
SIEMENS AKTIENGESELLSCHAFT	1	3	DE	3900000	76651000	134,965	59,637375
SIEMENS AKTIENGESELLSCHAFT	1	4	DE	3558000	68978000	148,725	54,34815
SIEMENS AKTIENGESELLSCHAFT	1	5	DE	3925000	73515000	119,81	56,305825
SIEMENS AKTIENGESELLSCHAFT	1	6	DE	4245000	77395000	141,471	60,888617
SIEMENS AKTIENGESELLSCHAFT	1	7	DE	4291000	75882000	152,475	66,880223
SIEMENS AKTIENGESELLSCHAFT	1	8	DE	4020000	71227000	138,532	71,271386
SIEMENS AKTIENGESELLSCHAFT	1	9	DE	4483000	75636000	151,923	78,776812
SIEMENS AKTIENGESELLSCHAFT	1	10	DE	4732000	79644000	165,226	80,833811
BAYER AKTIENGESELLSCHAFT	2	1	DE	2578000	32385000	124,195	73,511278
BAYER AKTIENGESELLSCHAFT	2	2	DE	2653000	32918000	187,589	71,563682
BAYER AKTIENGESELLSCHAFT	2	3	DE	2746000	31168000	129,735	73,406967
BAYER AKTIENGESELLSCHAFT	2	4	DE	3053000	35088000	126,798	72,758144
BAYER AKTIENGESELLSCHAFT	2	5	DE	2932000	36528000	141,664	77,592723

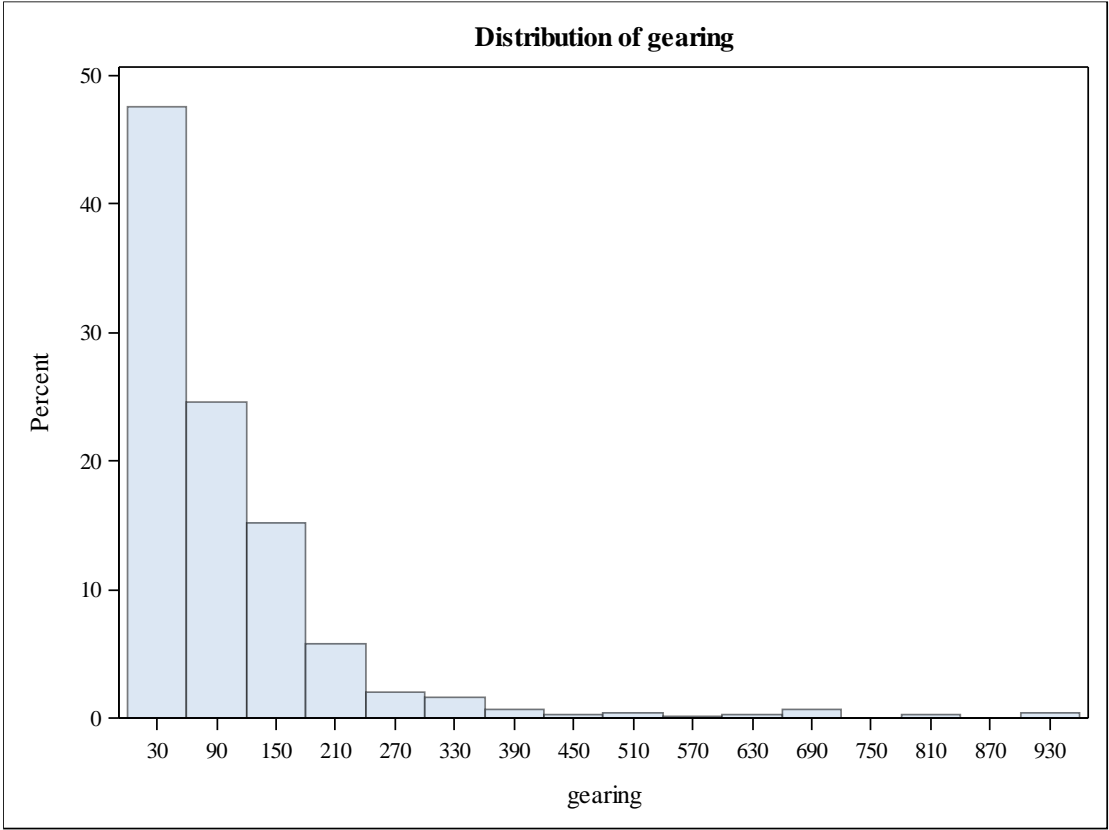
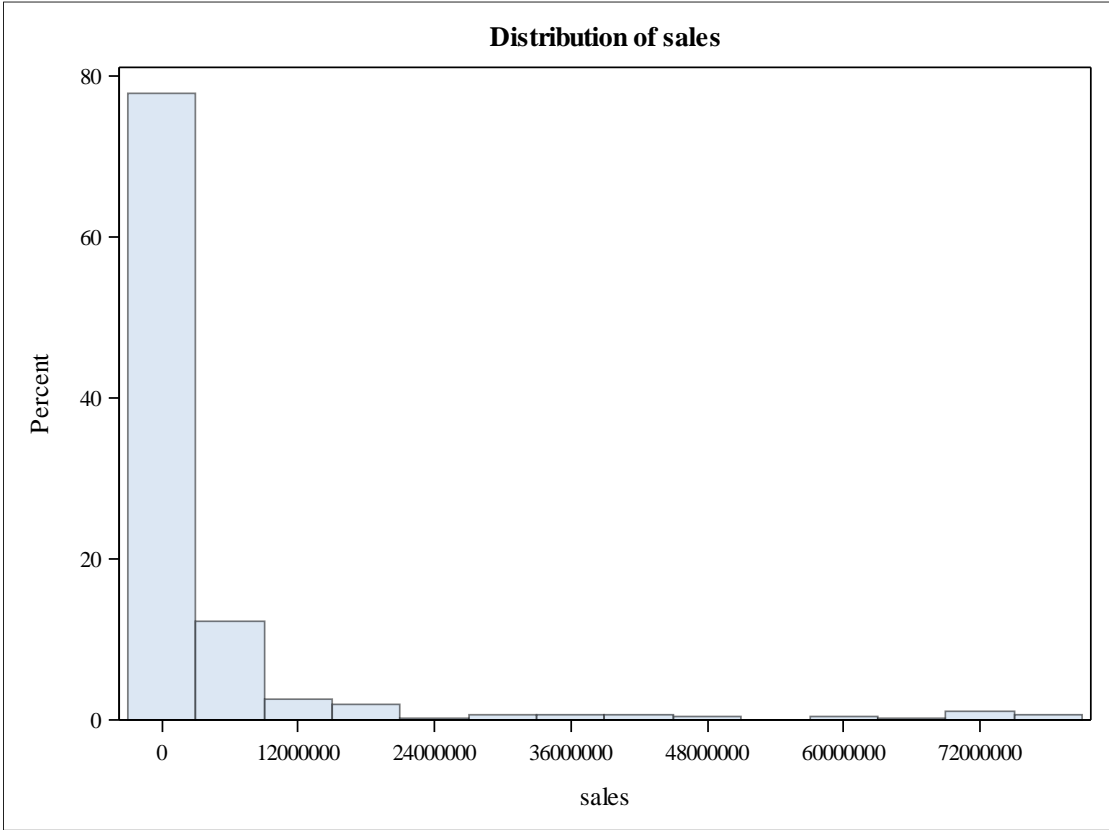
Liite 4. Aineiston muunnokseen käytetty R-koodi

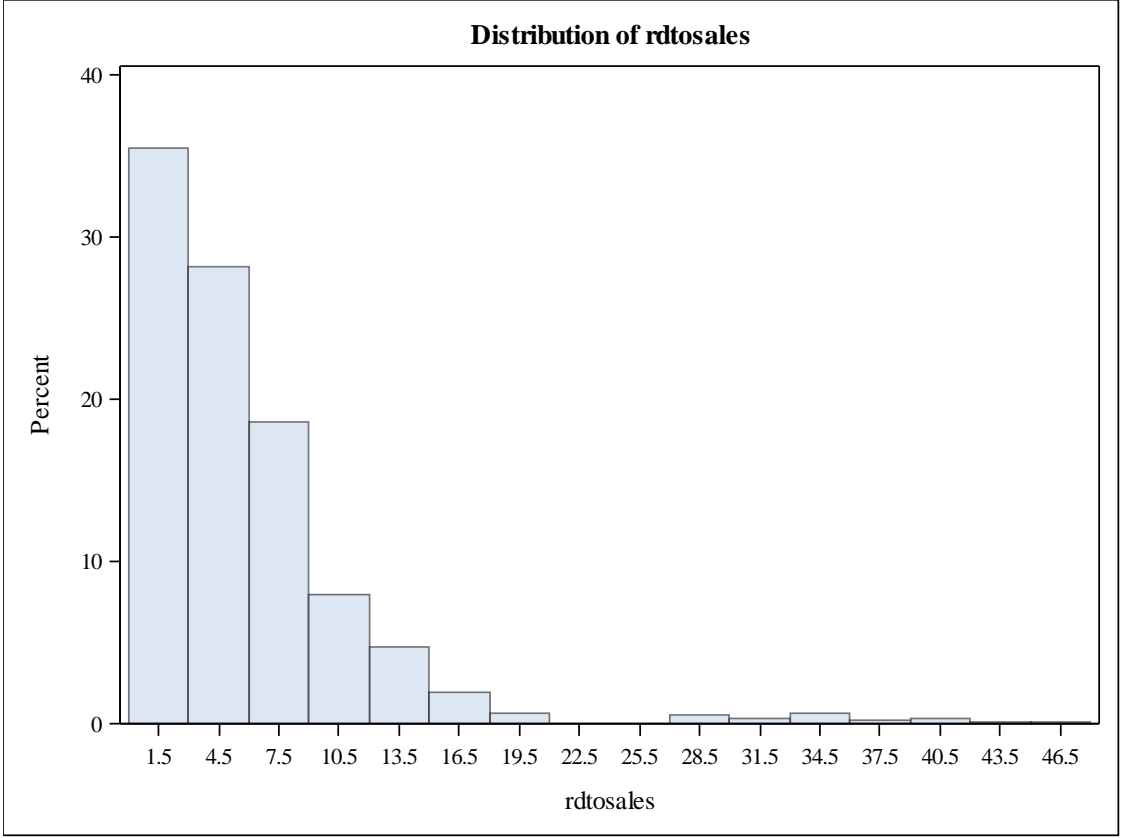
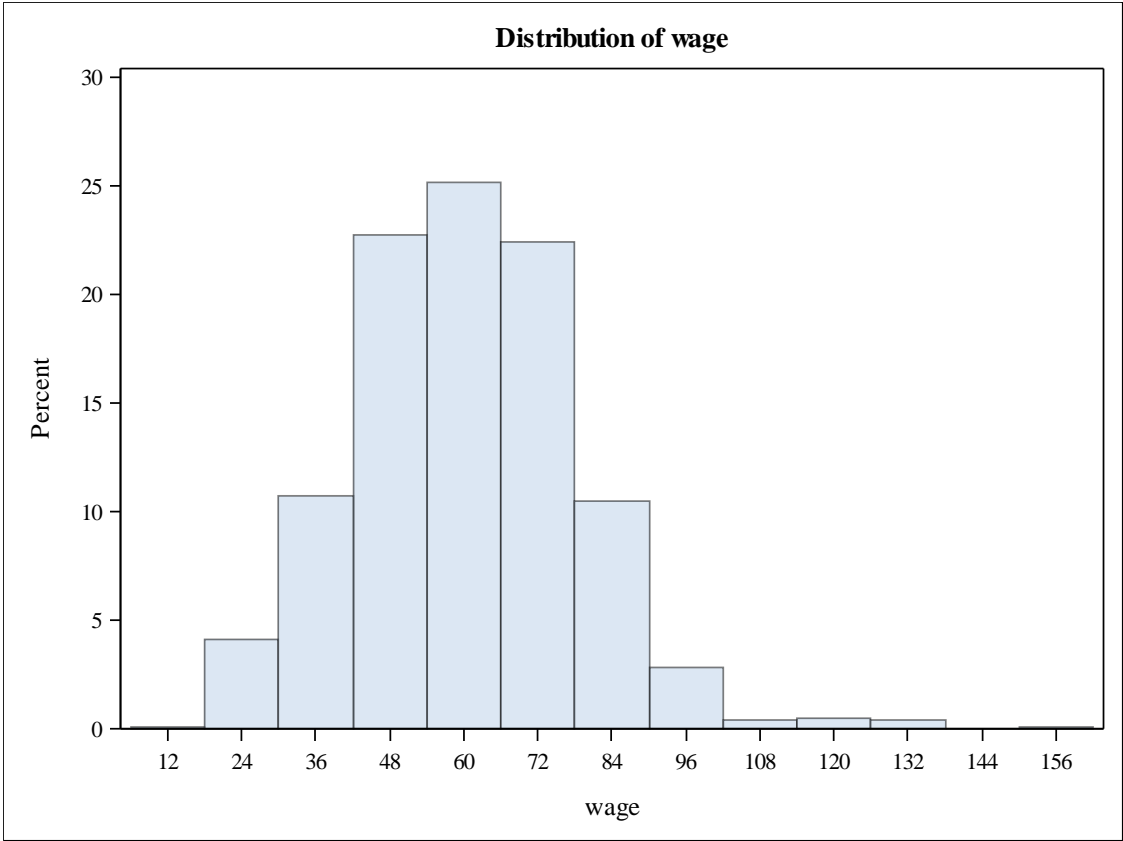
```

dfwide <- open.table(file.choose())
dflong <- reshape(dfwide, direction="long",
  varying= list(c(3:12),c(13:22),c(23:32), c(33:42)),
  v.names= c("rd", "sales", "gearing", "wage"),
  idvar= "id",
  timevar="year", times= c(1:10))

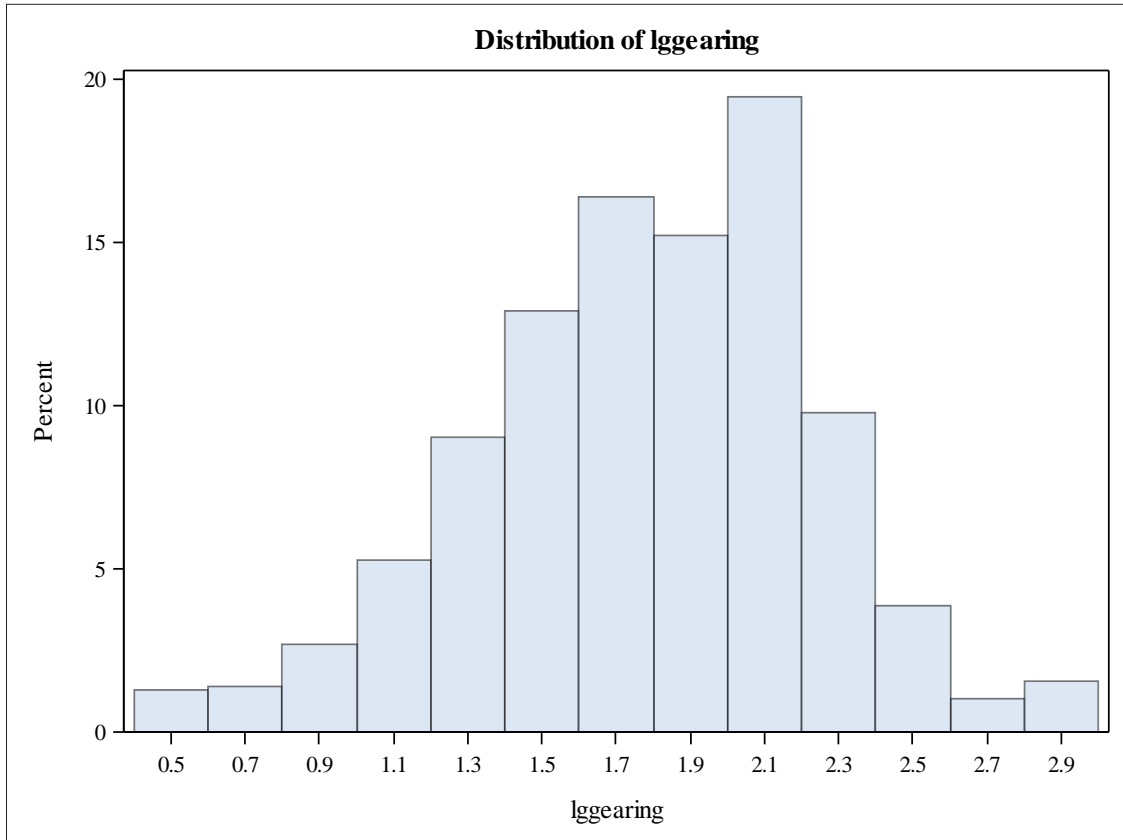
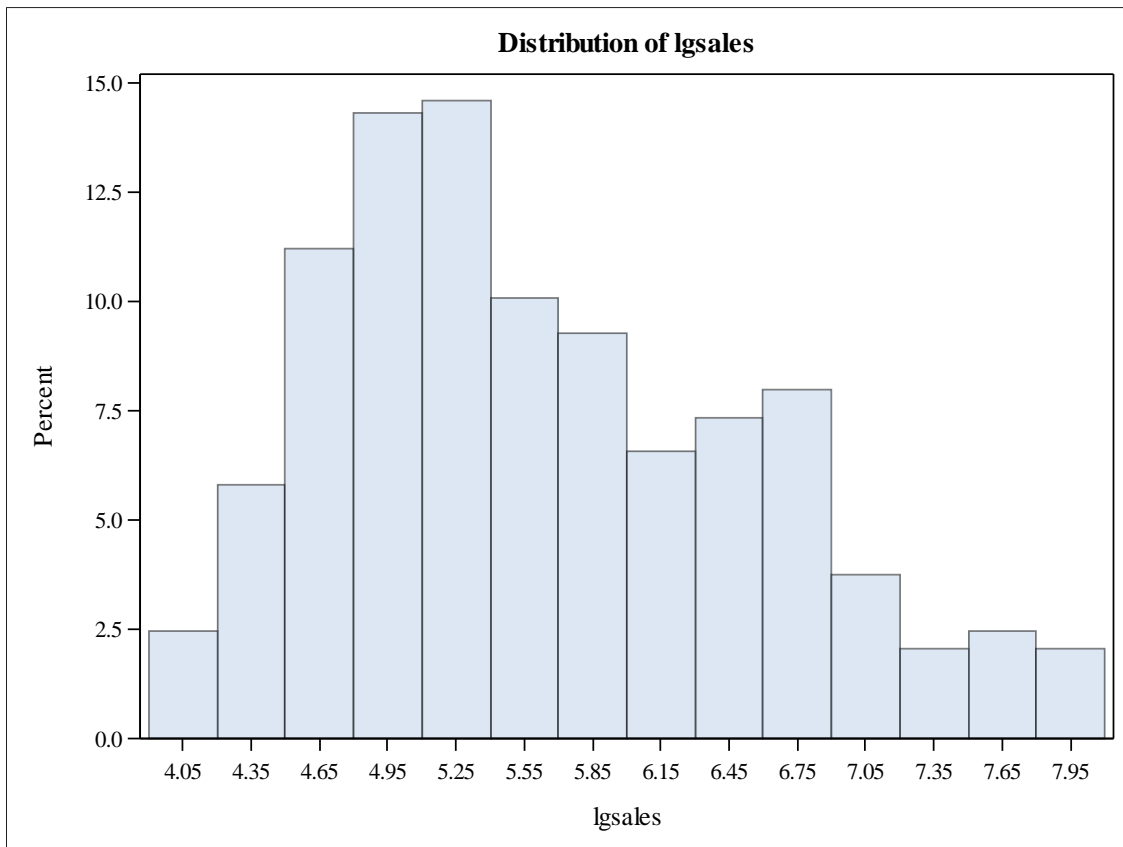
```

Liite 5. Muuttujien jakaumat

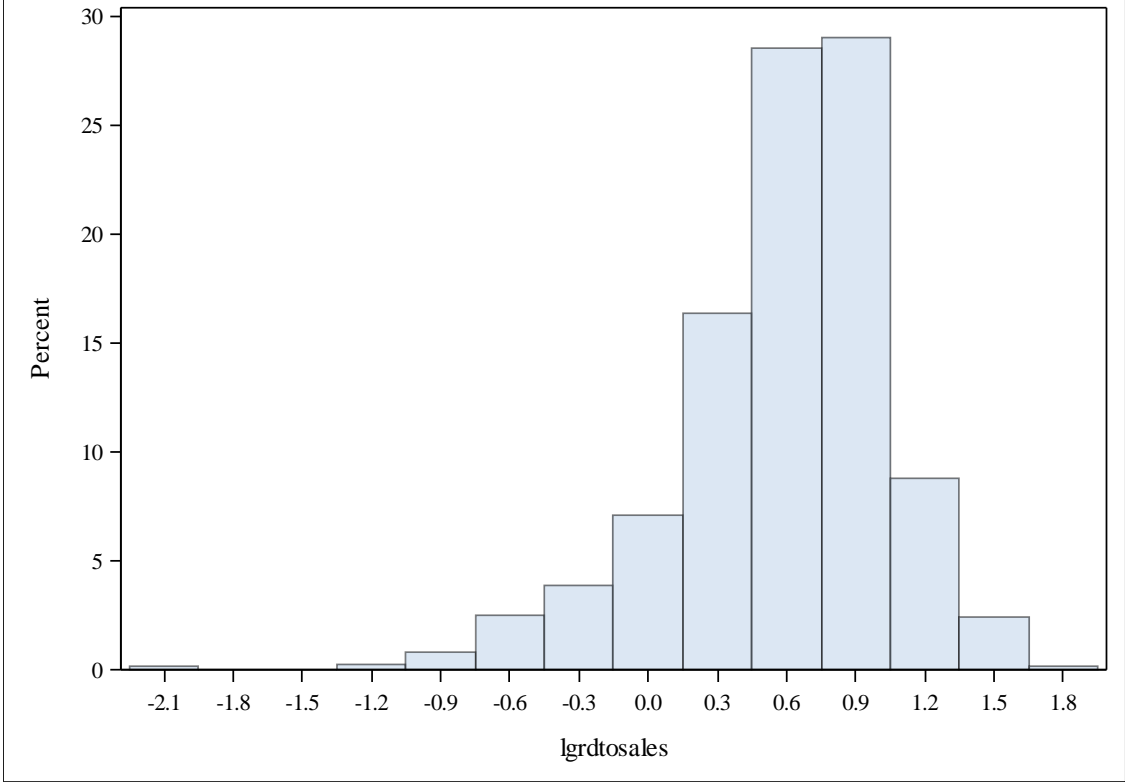




Liite 6. Logaritmisten muuttujien jakaumat



Distribution of lgrdtosales



Liite 7. Regressioanalyysin tulokset

Dependent Variable: lgrdtosales

Model Description	
Estimation Method	FixTwo
Number of Cross Sections	78
Time Series Length	10

Fit Statistics			
SSE	18.7159	DFE	657
MSE	0.0285	Root MSE	0.1688
R-Square	0.8841		

F Test for No Fixed Effects			
Num DF	Den DF	F Value	Pr > F
86	657	48.82	<.0001

Breusch Pagan Test for Random Effects (One Way)		
DF	m Value	Pr > m
1	1782.41	<.0001

Parameter Estimates						
Variable	DF	Estimate	Standard Error	t Value	Pr > t	Label
Intercept	1	2.296265	0.3829	6.00	<.0001	Intercept
lgsales	1	-0.37212	0.0658	-5.65	<.0001	
lggearing	1	-0.00069	0.0334	-0.02	0.9835	
wage	1	0.000402	0.000764	0.53	0.5991	

Dependent Variable: lgrdtosales

Model Description	
Estimation Method	RanTwo
Number of Cross Sections	78
Time Series Length	10

Fit Statistics			
SSE	20.9507	DFE	743
MSE	0.0282	Root MSE	0.1679
R-Square	0.0295		

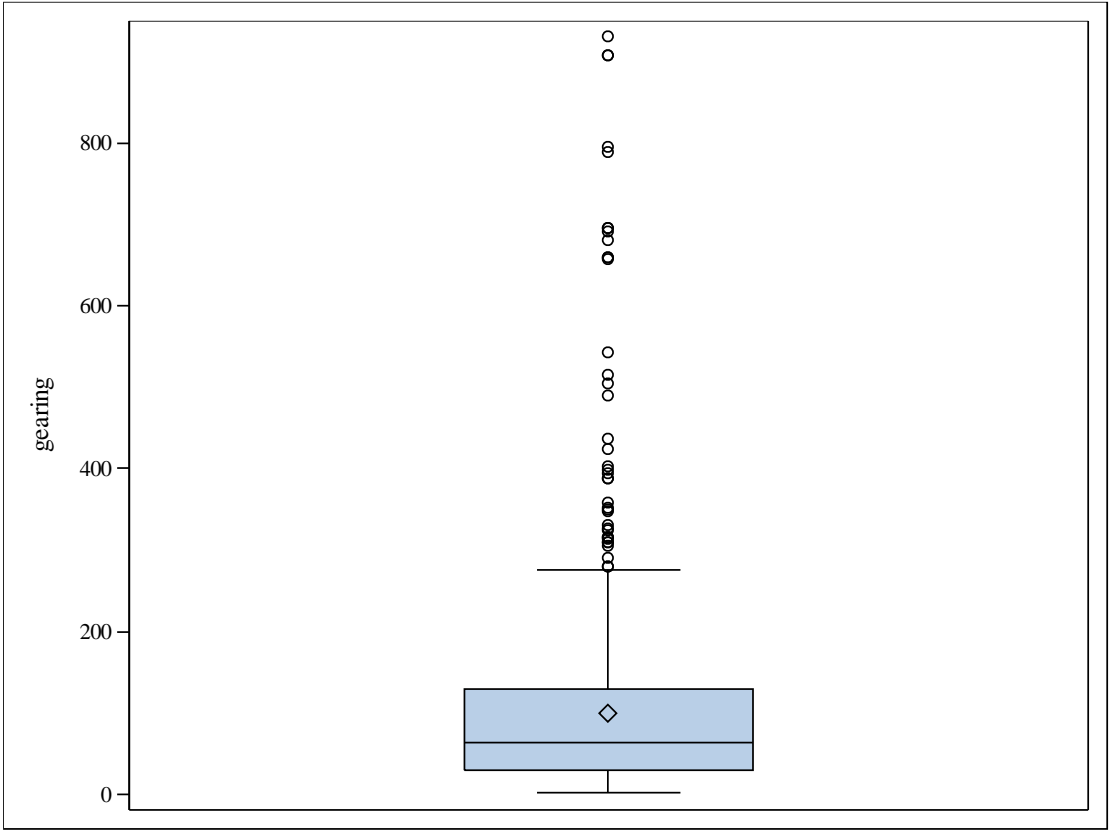
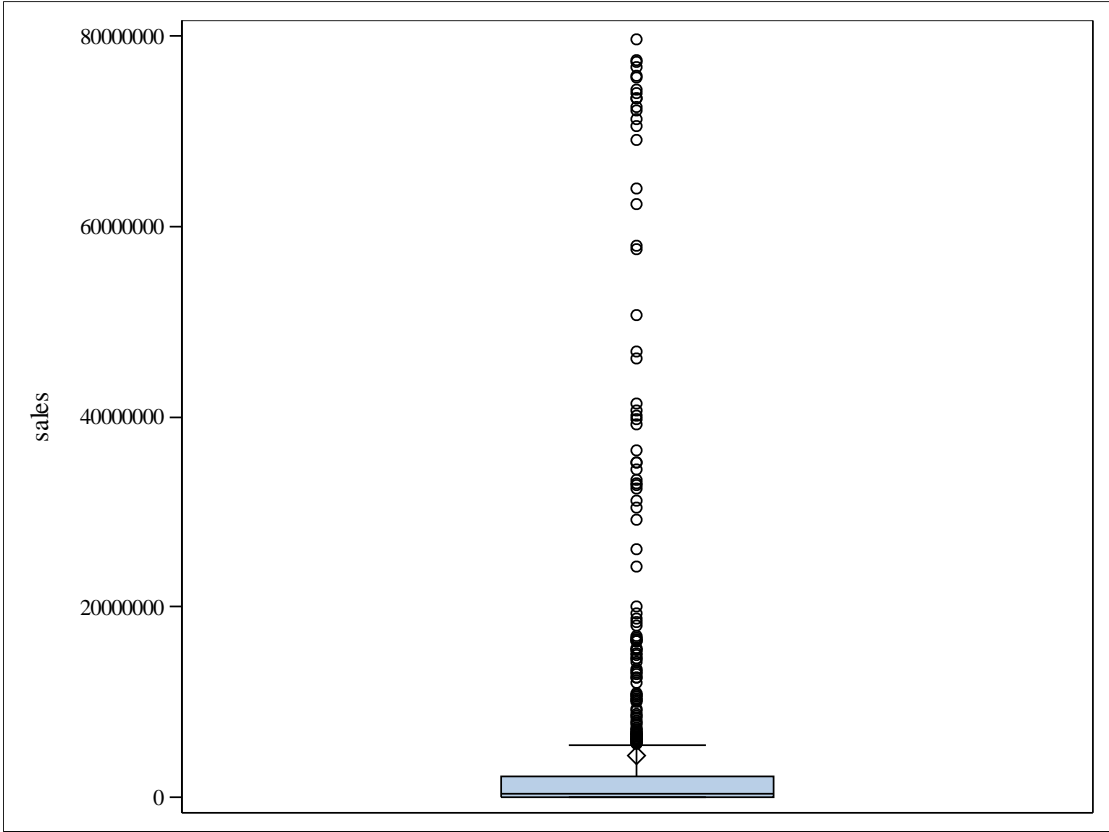
Variance Component Estimates	
Variance Component for Cross Sections	0.276446
Variance Component for Time Series	0
Variance Component for Error	0.028487

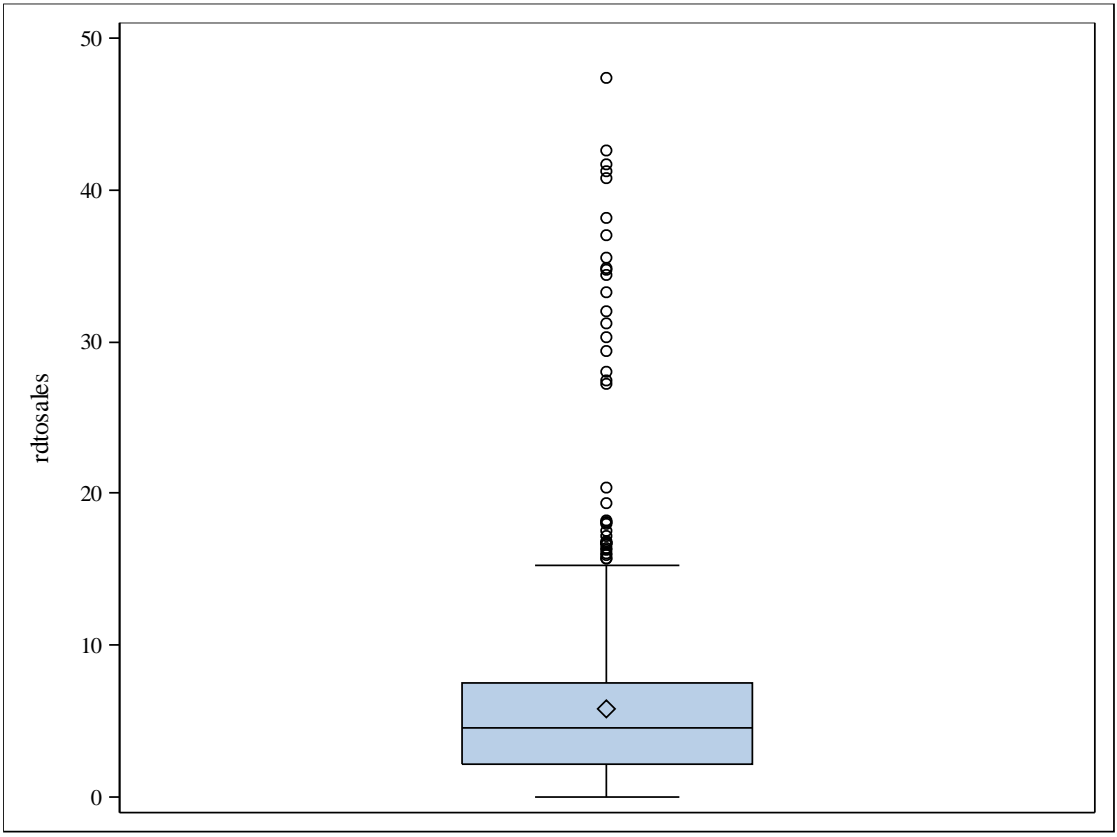
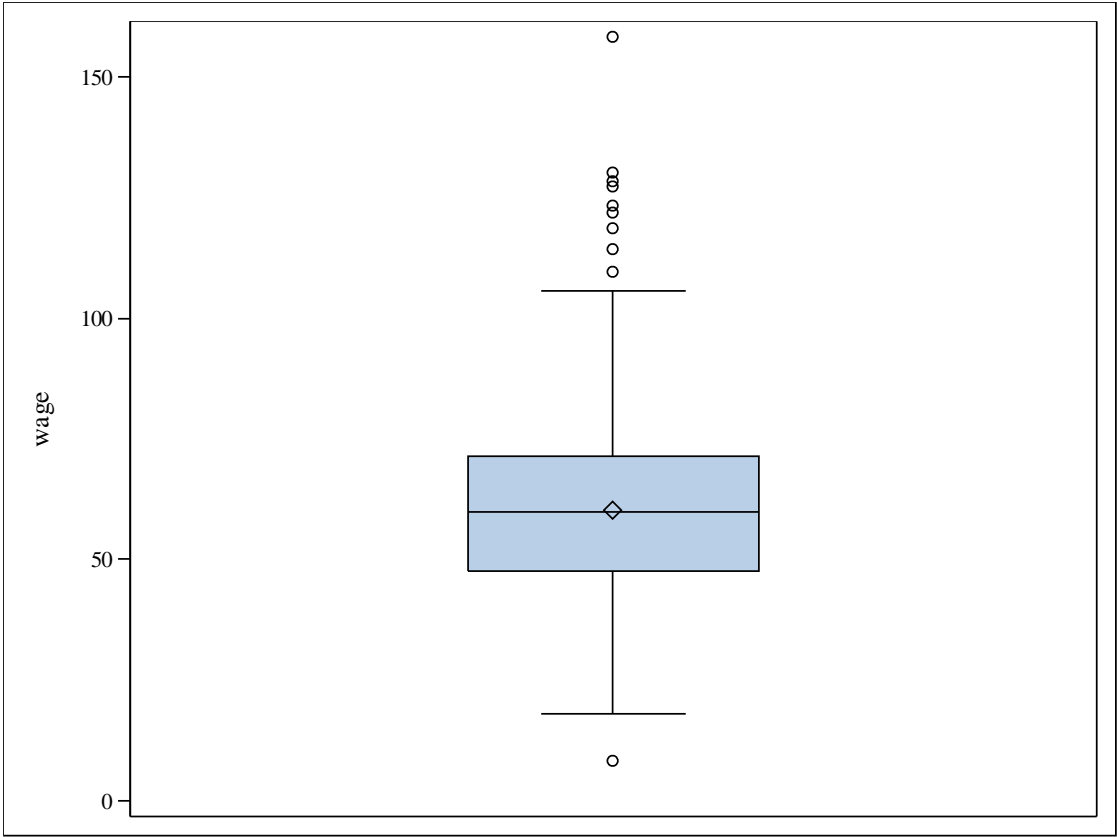
Hausman Test for Random Effects		
DF	m Value	Pr > m
3	18.39	0.0004

Breusch Pagan Test for Random Effects (One Way)		
DF	m Value	Pr > m
1	1782.41	<.0001

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	1.737801	0.2533	6.86	<.0001
lgsales	1	-0.19881	0.0435	-4.57	<.0001
lggearing	1	-0.03203	0.0315	-1.02	0.3099
wage	1	0.00038	0.000701	0.54	0.5882

Liite 8. Muuttujien laatikko-jana-kuviot





Liite 9. Outlier-havaintojen poiston rajat

Variable	Lower Quartile	Upper Quartile	IQR	yläraja	alaraja
sales	85662,76	2275900,00	2190237,24	5561255,86	-5475593,10
gearing	30,2910000	128.8400000	98,55	276,66	-246,37
wage	47,5071040	71,4285710	23,92	107,31	-59,80
rdtosales	2,1837626	7,4886894	5,30	15,45	-13,26

Liite 10. Regressioanalyysin tulokset outlier-havaintojen suodattamisen jälkeen

Dependent Variable: lgrdtosales

Model Description	
Estimation Method	FixTwo
Number of Cross Sections	66
Time Series Length	10

Fit Statistics			
SSE	8.3551	DFE	482
MSE	0.0173	Root MSE	0.1317
R-Square	0.9143		

F Test for No Fixed Effects			
Num DF	Den DF	F Value	Pr > F
74	482	62.32	<.0001

Breusch Pagan Test for Random Effects (One Way)		
DF	m Value	Pr > m
1	1358.42	<.0001

Parameter Estimates						
Variable	DF	Estimate	Standard Error	t Value	Pr > t	Label
Intercept	1	1.236848	0.3428	3.61	0.0003	Intercept
lgsales	1	-0.13018	0.0601	-2.17	0.0308	
lggearing	1	-0.02504	0.0315	-0.80	0.4268	
wage	1	0.00041	0.000963	0.43	0.6703	

Dependent Variable: lgrdtosales

Model Description	
Estimation Method	RanTwo
Number of Cross Sections	66
Time Series Length	10

Fit Statistics			
SSE	558.7141	DFE	556
MSE	1.0049	Root MSE	1.0024
R-Square	0.0126		

Variance Component Estimates	
Variance Component for Cross Sections	0.161824
Variance Component for Time Series	0.000285
Variance Component for Error	0.017334

Hausman Test for Random Effects		
DF	m Value	Pr > m
3	6.29	0.0984

Breusch Pagan Test for Random Effects (One Way)		
DF	m Value	Pr > m
1	1358.42	<.0001

Parameter Estimates					
Variable	DF	Estimate	Standard Error	t Value	Pr > t
Intercept	1	1.121514	0.2423	4.63	<.0001
lgsales	1	-0.09041	0.0447	-2.02	0.0436
lggearing	1	-0.04416	0.0303	-1.46	0.1456
wage	1	0.000288	0.000883	0.33	0.7440

Liite 11. Mallien residuaalien kvantili- ja jakaumakuviot

