

AARHUS UNIVERSITET  
SCHOOL OF BUSINESS AND SOCIAL SCIENCE  
INSTITUT FOR ØKONOMI

---

# EJERSKABSTRUKTUR & RENTABILITET

EN ANALYSE AF DEN GLOBALE FARMACEUTISKE INDUSTRI

---

BACHELORAFHANDLING

1. MAJ 2019

Kristian Rumph Frederiksen  
(Stud.nr. 201610744)



Antal Tegn: 99.730 (ekskl. blanktegn)

Grafer og figurer: 5 af 4.000 tegn

Samlet antal tegn: 103.730

Vejleder:  
Jan Bartholdy  
Lektor



SCHOOL OF BUSINESS AND SOCIAL SCIENCES  
AARHUS UNIVERSITET

## Abstract

This thesis examines the relationship between ownership structure and firm performance, with an emphasis on the global pharmaceutical industry. This relationship is investigated both theoretically, using theories from the corporate governance literature, and empirically using fixed effects and random effects statistical models.

As mentioned, the theoretical starting point of the thesis is agency theory and incomplete contracts theory, both corporate governance theories which focuses on potential conflicts between agents and principals. Using the terminology from Thomsen & Conyon (2012), the focus of this thesis is type-1 agency problems: which is conflicts between owners and managers. Incomplete contracts theory is a broad theory concerning the contractual agreements between units and/or agents. As managers are contractually employed by the board of directors (representing the owners), incomplete contracts theory can be applied to this area as well.

These focal theories underline the underlying assumption, that a firm's ownership structure should have a causal effect on its financial performance. This theoretical relationship stems from the likely correlation between ownership concentration and the degree of control a specific parent company can exercise on the management of its subsidiary.

Included in the thesis is a description of the financial performance measures used in the analysis. The applied performance measures are *Return on Assets*, *Return on Equity* and *Tobin's q*. The first two measures are accounting measures of company profitability. Tobin's q however, is a market measure of firm performance.

The empirical study consists of several advanced statistical models, estimated from a panel dataset with 10 years of data (2008-2017), containing more than 5.000 observations distributed on ~1.500 unique companies. The dataset is constructed from several sub-datasets extracted from Orbis Historical archives, using Visual Basic Application. Unfortunately, the regression results imply underlying inconsistencies with the dataset, most likely caused by the continuous updating and expansion of the Orbis database.

The applied statistical models are fixed effects and random effects, and part of the empirical analysis is investigating which model is most efficient. This is studied using a Hausman test to compare the consistency of the random effects estimator to the less efficient, fixed effects estimator. The validity of the Hausman test depends on whether the assumptions, for the fixed effects estimator, hold, as this infers consistency of the estimator.

The procedure of the empirical study is: First establishing an initial fixed effects model, then discussing whether its assumptions are held, and lastly deciding on whether to establish a random effects model as well. Following said procedure, both estimators are applied to compute several models. The validity of both models is constrained considerably by the failure to comply with several critical assumptions. Violation of these assumptions implies that the models are not exogenous, indicating that the coefficients suffer from bias and the estimators from inconsistency.

In the thesis' discussion section, the robustness of the findings is investigated, and it is concluded that the results are not very robust, and far from the findings of other studies. This underlines the other complications and indicates that the results are not very feasible.

The main conclusion of the empirical study is, even considering the models imperfections, that the explanatory power is very low and almost exclusively controlled for by age and size control variables. Therefore, the empirical results are not valid for concluding any causal relationship between firm performance and ownership structure or concentration.

# Indholdsfortegnelse

---

<b>1. UNDERSØGELSENS DESIGN .....</b>	<b>1</b>
<b>1.1. INDLEDNING .....</b>	<b>1</b>
<b>1.2. PROBLEMFORMULERING.....</b>	<b>2</b>
1.2.1 DET CENTRALE FORSKNINGSSPØRGSMÅL .....	2
<b>1.3. AFGRÆNSNING.....</b>	<b>3</b>
<b>1.4. VIDENSKABSTEORETISKE OVERVEJELSER .....</b>	<b>3</b>
<b>1.5. AFHANDLINGENS OPBYGNING.....</b>	<b>4</b>
<b>2. TEORI.....</b>	<b>5</b>
<b>2.1. LITTERATUR REVIEW.....</b>	<b>5</b>
<b>2.2. CORPORATE GOVERNANCE .....</b>	<b>6</b>
2.2.1 PRINCIPAL-AGENT-TEORI .....	7
2.2.2 KONTRAKTTEORI.....	8
2.2.3 OPSUMMERING .....	9
<b>2.3. PERFORMANCEMÅL.....</b>	<b>10</b>
2.3.1 EGEN- OG TOTALKAPITALFORRENTNING (ROE OG ROA).....	10
2.3.2 TOBIN'S Q .....	11
<b>3. METODE .....</b>	<b>12</b>
<b>3.1. FIXED EFFECTS MODEL .....</b>	<b>12</b>
3.1.1 FORUDSÆTNINGER .....	14
<b>3.2. RANDOM EFFECTS MODEL .....</b>	<b>18</b>
3.2.1 FORUDSÆTNINGER .....	18
<b>3.3. VALG AF MODEL .....</b>	<b>20</b>
<b>3.4. FREMGANGSMÅDE FOR ANALYSEN.....</b>	<b>20</b>
<b>4. DATA .....</b>	<b>21</b>
<b>4.1. DATAGRUNDLAG .....</b>	<b>21</b>
<b>4.2. DATAUDVÆLGELSE.....</b>	<b>22</b>
4.2.1 SÆRLIGT FOR ÅR 2008 - 2010.....	22
<b>4.3. BEHANDLING AF DATA.....</b>	<b>23</b>
4.3.1 KONSOLIDERING AF DATASÆTTET MED VBA.....	24
4.3.2 IMPORT OG SORTERING I STATA.....	24

4.3.3	EJERSKABSVARIABLER.....	25
4.3.4	FILTRERING AF DATASÆTTET.....	26
<b>4.4.</b>	<b>BEGRÆNSNINGER.....</b>	<b>27</b>
<b>5.</b>	<b><u>RESULTATER.....</u></b>	<b>29</b>
<b>5.1.</b>	<b>FIXED EFFECTS MODELLEN.....</b>	<b>29</b>
5.1.2	RESULTATER .....	31
5.1.3	FORUDSÆTNINGER .....	31
5.1.4	OPSUMMERING .....	35
<b>5.2.</b>	<b>HAUSMAN TEST.....</b>	<b>37</b>
<b>5.3.</b>	<b>RANDOM EFFECTS MODELLEN.....</b>	<b>39</b>
5.3.2	FORUDSÆTNINGER .....	39
5.3.3	OPSUMMERING .....	40
<b>6.</b>	<b><u>DISKUSSION.....</u></b>	<b>41</b>
<b>6.1.</b>	<b>RESULTATERNES VALIDITET.....</b>	<b>41</b>
<b>6.2.</b>	<b>SAMMENLIGNING MED LIGNENDE STUDIER.....</b>	<b>44</b>
6.2.1	OPLÆG TIL VIDERE ANALYSE .....	44
<b>7.</b>	<b><u>KONKLUSION .....</u></b>	<b>45</b>
<b>8.</b>	<b><u>BIBLIOGRAFI.....</u></b>	<b>46</b>

## Oversigt over figurer

<b>FIGUR 1 – AFHANDLINGENS STRUKTUR .....</b>	4
<b>FIGUR 2 – DET UDVIDEDE AGENTPROBLEM .....</b>	8
<b>FIGUR 3 – OVERSIGT OVER SØGEKRITERIER .....</b>	21
<b>FIGUR 4 – ORBIS’ KLASSEFIKATION AF DIREKTE OG TOTALT EJERSKAB .....</b>	23
<b>FIGUR 5 – QQ-PLOT OVER FIXED EFFECTS RESIDUALER (MODEL 1) .....</b>	35

## Oversigt over tabeller

<b>TABEL 1 – OVERSIGT OVER PERFORMANCEMÅL .....</b>	10
<b>TABEL 2 – FORUDSÆTNINGER FOR REGRESSIONER.....</b>	12
<b>TABEL 2 – EJERTYPEFORDELINGEN FOR DE TRE STØRSTE EJERE .....</b>	25
<b>TABEL 4 – SORTERING AF DATASÆT .....</b>	26
<b>TABEL 5 – BESKRIVENDE STATISTIKKER FOR INKLUDEREDE VARIABLER .....</b>	28
<b>TABEL 6 – RETURN ON ASSET, FØRSTE MODELLER .....</b>	30
<b>TABEL 7 – ÆNDRINGER I DUMMY VARIABLE .....</b>	32
<b>TABEL 8 – VIF FOR VARIABLERNE I FIXED EFFECTS MODELLEN .....</b>	33
<b>TABEL 9 – FIXED- &amp; RANDOM EFFECTS KOEFFICIENTER .....</b>	36
<b>TABEL 10 – RANDOM EFFECTS MODELLER .....</b>	38
<b>TABEL 11 – VIF FOR VARIABLERNE I RANDOM EFFECTS MODELLEN .....</b>	40
<b>TABEL 12 – SENSITIVITETSANALYSER.....</b>	43

## 1. Undersøgelsens design

---

Afhandlingens første kapitel vil redegøre for problemformuleringen, de grundlæggende videnskabsteoretiske overvejelser samt afhandlingens struktur og opbygning. Det er forfatterens ønske at kapitlet vil tydeliggøre motivationen bag det overordnede emne- og metodevalg og samtidig anskueliggøre hvorfor (og for hvem) problemstillingen er relevant.

### 1.1. Indledning

Ligeså længe som der har været en adskillelse mellem kapitalsejere og de ledelsesagenter som er bemyndiget til at forvalte ejernes kapitalapparat, har man været opmærksom på interessekonflikter imellem de to parter. Den berømte britisk-skotske økonom Adam Smith, satte, i hans berømte værk *Wealth of Nations* (1776), ord på den interessekonflikt, som afgestadkommer af ansvarsdelingen mellem *ejer* og *leder*:

*“The directors (...), being the managers rather of other people’s money than of their own, it cannot well be expected that they should watch over it with the same anxious vigilance with which the partners in a private copartnery frequently watch over their own. (...)"*  
(Smith, 1776)

Adam Smith argumenterer i citatet for, hvordan direktører, som forvalter *af andre folks penge*, ikke må forventes at udvise samme årvågenhed som hvis de administrerede deres egne penge. Denne konflikt er udgangspunktet for corporate governance fagområdet, som netop beskæftiger sig med kontrollen af og adfærdens iblandt selskabsledere. En stor del af corporate governance-litteraturen omhandler derfor hvordan selskabsejere i praksis kontrollerer og sikrer sig at ledelsen handler i overensstemmelse med deres interesser.

I det moderne kapitalistiske system, som kendetegnes af åbne, liberale kapitalmarkeder, vælger mange selskaber at børsnoteres, for at sikre let adgang til finansiering. Store børsnoterede virksomheder kendetegnes ofte af et fragmenteret ejerskab, med mange mindre aktionærer. Det skaber en risiko for at afstanden mellem ejere og ledelse bliver for stor, til at ejerne på tilstrækkelig vis kan kontrollere at ledelsen ikke handler på egennyttige interesser.

Udgangspunktet for denne afhandling er ikke at afdække de egentlige mekanismér bag interessekonflikterne, men snarere at undersøge om konsekvenserne heraf kan testes empirisk. Konkret vil afhandlingen fokusere på om ejerskabsstrukturen, og koncentrationen, påvirker rentabiliteten af moderne børsnoterede selskaber. Effekten af ejerskabskoncentrationen på virksomhedernes performance er velundersøgt (Short, 1994), men det står stadig uklart om der reelt eksisterer et observerbart forhold mellem selskabernes ejerskabsstruktur og deres rentabilitet/performance (Clifford G. Holderness, 2003). Håbet er at afhandlingen formår at bygge videre på de eksisterende studier, ved at anvende paneldata samt mere komplekse statistiske modeller, end det har været tilfældet i tidligere studier på området. Den primære empiriske problemstilling er at måle hvor stor forklaringskraft, og foranlediget heraf, hvor stor påvirkning ejerskabsstrukturen har på virksomhedernes rentabilitet.

Hvis der påvises en statistisk signifikant, og praktisk betydelig, sammenhæng, har det relevans for investorer; Uanset om det er mindre private investorer, som overvejer at investere i børsnoterede selskaber eller om det er større investeringsselskaber, som træffer beslutninger om hvor store andel de skal eje i bestemte virksomheder.

Det primære omdrejningspunkt, for afhandlingen, vil være den empiriske analyse, derfor vil metode- og resultatafsnittene indeholde udfyldende beskriver, af de anvendte statistiske modeller. Problemformuleringen nedenfor, vil redegøre for afhandlingens konkrete forskningsdesign samt de centrale forskningsspørgsmål, som ønskes besvaret. Sidst i afhandlingen, vil der følges op på disse spørgsmål i ét samlet konklusionsafsnit (afsnit 7). De øvrige afsnit i indeværende kapitel, vil redegøre for studiets afgrænsning, overordnede videnskabs-teoretiske overvejelser samt afhandlingens opbygning.

## 1.2. Problemformulering

Afhandlingens overordnede formål er at undersøge sammenhængen mellem børsnoterede farmaceutiske virksomheters ejerskabsstruktur og performance over en 10-årig periode (2007-2017). Undersøgelsen vil gennemføres som et empirisk studie, på grundlag af panel-data fra Orbis Historical-databasen, med anvendelse af fixed- og random effects modeller. Afhandlingens teoretiske udgangspunkt er principal-agent-teori og kontraktteori, som begge ligger indenfor den overordnede corporate governance litteratur.

### 1.2.1 Det centrale forskningsspørgsmål

*Eksisterer der en kausal sammenhæng mellem ejerskabsstrukturens karakteristika og virksomhedernes langsigtede performance, indenfor den globale farmaceutiske industri?*

Det centrale forskningsspørgsmål leder naturligt videre til en række delspørgsmål, som det bestræbes at finde afklaring af, i løbet af afhandlingen:

1. *Hvordan måles ejerskab og hvordan kan ejerskabsstrukturerne karakteriseres?*
2. *Hvilken model er mest fordelagtig til estimation heraf: fixed- eller random effects?*
3. *Hvordan relaterer studiets resultater sig til tidligere empiri?*

### 1.3. Afgrænsning

Som formuleringen af det centrale forskningsspørgsmål afslører, så afgrænses opgavens omfang til selskaber indenfor den farmaceutiske industri. Årsagen skyldes det langsigtede perspektiv som er et vilkår indenfor denne branche. Dette iboende vilkår, skyldes de lange udviklingstider indenfor industrien: Farmaceutiske produkter udvikles over mange år og den gennemsnitlige udviklingstid er på mere end 10 år (DiMasi, 1995). Det langsigtede perspektiv gør den farmaceutiske industri særlig relevant, når ejerskabsstrukturens effekt på virksomhedernes langsigtede værdiskabelse skal undersøges.

Udover afgrænsningen på den farmaceutiske industri, begrænser undersøgelsen sig også til kun at omfatte børsnoterede selskaber med en markedsværdi over USD 100 mio. Dette valg er truffet for at sikre størst mulig datatilgængelighed og transparens, så analysens inkluderede variabler er så pålidelige som muligt, både de forklarende variabler samt de anvendte performance mål<sup>1</sup>. Afgrænsningen foretages helt konkret gennem søgekriterier, filtrering og konstruktion af datasættet, som er beskrives i dataafsnittet (afsnit 4).

Foranlediget af disse to markante afgrænsninger, så ender analysens udgangspunkt med at være ret snævert og eksterne validitet af resultaterne, må derfor anses som ret lav – Det betyder altså at resultaterne ikke er generaliserbare udenfor den farmaceutiske industri samt børsmarkedet. Forhåbentligt bidrager den snævre afgrænsning til at analysen derimod har en høj intern validitet. Konkret forventes det at øgede datakvaliteten vil sikre resultaternes pålidelighed og gyldighed.

### 1.4. Videnskabsteoretiske overvejelser

Afhandlingen videnskabsteoretiske afsæt er post-positivismen, hvor det ontologiske udgangspunkt er kritisk realisme. Et virkelighedssyn som kendetegnes af troen på én sand virkelighed, samtidig med erkendelsen af at denne virkelighed udelukkende kan forstås i brudstykker og gennem approksimationer (Guba, 1990). Post-positivismens videnssyn (epistemologi) er modifieret objektivistisk og kendetegnes af bestræbelsen på objektivitet, men erkendelsen af at det er usandsynligt at opnå fuld objektivitet.

Det post-positivistiske syn, vil have implikationer for hele afhandlingen, men særligt har det indflydelse på metodevalget. Her vil analysen, som beskrevet i indledningen (1.1.), baseres på et kvantitativt studie af sammenhængen mellem ejerskabsstruktur og virksomhedsperformance. Studiet vil inkludere en række variabler, hvis udvælgelse sker med henblik på at sikre størst mulig objektivitet. For at mindske bias, udvælges variablerne på baggrund af tidligere studier samt datatilgængelighed. Formålet med analysen er derfor ikke at skabe et fuldendt billede af faktorerne bag virksomhedernes værdiskabelse, men blot en forståelse af påvirkningen fra virksomhedernes ejerskabsstruktur – det accepteres at den sande effekt på værdiskabelsen kun tilnærmedsesvist kan forklares.

---

<sup>1</sup> Performance mål beskriver de, i analysen, uafhængige variabler, som benyttes til at måle selskabets indtjening og værdiskabelse: *ROE, ROA, Tobin's q etc.*

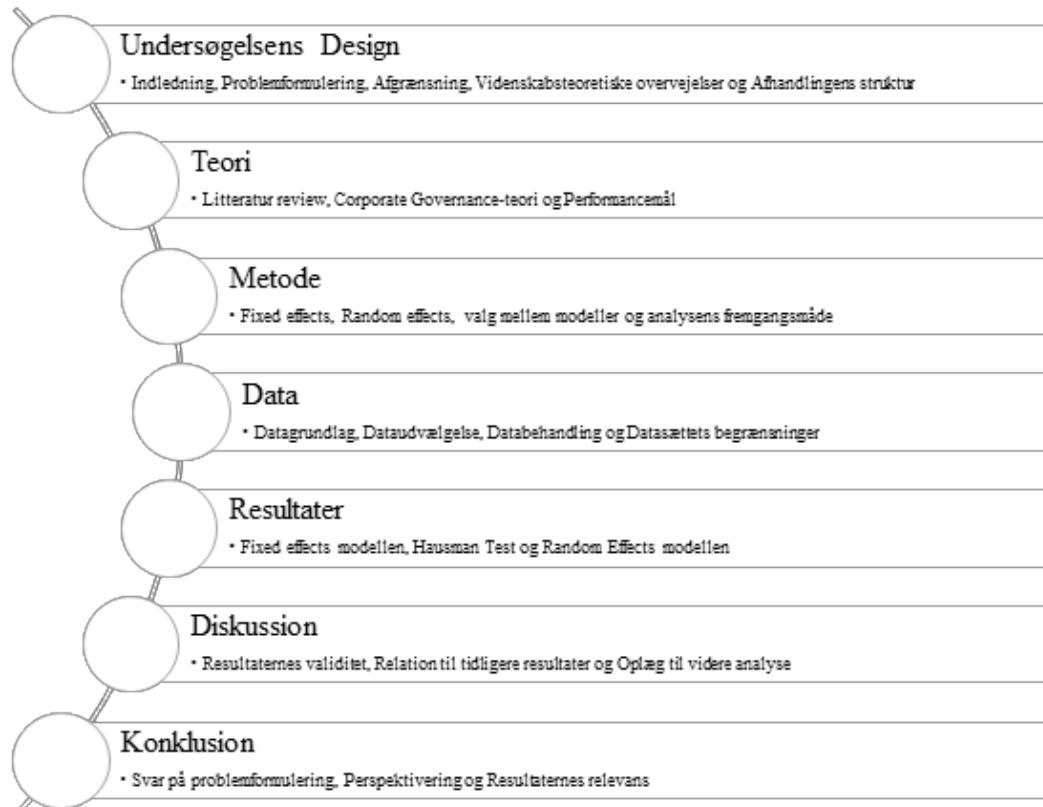
## 1.5. Afhandlingens opbygning

Afhandlingen er inddelt i syv overordnede kapitler:

1. *Undersøgelsens design* (indeværende kapitel) omfatter den underliggende motivation for emnevalget, afhandlingens problemformulering og en beskrivelse af de videnskabsteoretiske overvejelser.
2. *Teori-kapitlet* redegør for afhandlingens overordnede litterære og teoretiske referenceramme med udgangspunkt i et litteratur review af den eksisterende empiri indenfor corporate governance generelt, samt ejerskabsstruktur i særdeleshed.
3. *Metode-kapitlet* beskriver afhandlingens statistiske modeller: Fixed- og random effects og forudsætningerne for deres gyldighed.
4. *Data-kapitlet* indeholder en udfyldende beskrivelse af det endelige datasæt og en diskussion af datasættets begrænsninger.
5. *Resultat-kapitlet* indeholder beskrivelser af resultaterne fra de statistiske modeller, samt en vurdering af om forudsætningerne for deres gyldighed er opfyldt.
6. *Diskussion-kapitlet* vil vurdere analysens resultater i et bredere perspektiv, og vil sammenholde dem med resultaterne fra lignende studier.
7. *Konklusionen* opsummerer pointerne fra de øvrige afsnit, i afhandlingen, og bestræber sig på at besvare spørgsmålene fra problemformuleringen.

Afhandlingens struktur er illustreret i figur 1 nedenfor:

**Figur 1 – Afhandlingens Struktur**



Note: Egen udarbejdelse

## 2. Teori

I dette kapitel er der redegjort for afhandlingens litterære og teoretiske udgangspunkt. Kapitlet begynder med en generel redegørelse for den eksisterende litteratur indenfor corporate governance, ejerskabsstrukturer og virksomhedsperformance. De øvrige kapitler vil bygge videre på denne redegørelse, ved blandt andet at redegøre for principal-agent-teori, kontraktteori samt de tre performance-mål som anvendes i analysen.

### 2.1. Litteratur Review

*"The bulk of corporate governance theory examines the agency problems that arise from two extreme ownership structures: 100% small shareholders or one large, controlling owner combined with small shareholders."*

Laeven and Levine (2007, p. 579)

Der eksisterer, som nævnt i indledningen (afsnit 1.1), allerede en betydelig mængde litteratur omkring corporate governance og ejerskabets betydning. Størstedelen af den ældre litteratur bestræber sig på teoretisk at beskrive de potentielle omkostninger, som kan opstå ved en ikke ugunstig ejerskabsstruktur. Særligt har Jensen og Mecklings indflydelsesrige artikel: 'Theory of the firm: Managerial behavior, agency costs and ownership structures' (Jensen & Meckling, 1976) haft stor betydning. Artiklen introducerede konceptet *Agency Costs*, som definerer de mulige omkostninger af manglende kongruens mellem to parters motiver (Se afsnit 2.2.1 om agentproblemer).

Senere studier har empirisk undersøgt sammenhængen mellem ejerskabsstrukturen og virksomhedernes finansielle performance. De empiriske studier har dog, jf. Short (1994, pp. 244-245), primært undersøgt få variabler i simple tværsnitsanalyser og stort set kun på data fra amerikanske virksomheder. Siden Short-artiklen udkom i 1994, er litteraturen dog udvidet til også at omfatte europæiske virksomheder, en gruppe som er kendetegnes ved at have særligt komplekse ejerskabsstrukturer (Faccio & Lang, 2002). Yderligere har litteraturen dog udviklet sig til også at undersøge effekten af de enkelte ejeres karakteristika – og ikke kun ejerkoncentrationen – på virksomhedernes finansielle performance (Thomsen & Pedersen, 2000).

Selvom litteraturen, de senere år, har udviklet sig markant og mange interessante aspekter er blevet undersøgt, så er der endnu uafdækkede forhold omkring sammenhængen mellem ejerskab og performance. Mange studier medtager ikke nuancerne mellem én kontrollerende ejer og fuldstændig spredt ejer. Et studie som bestræber sig på netop at undersøge disse nuancer er Laeven and Levine (2007), hvis inddeling af ejerskabet vil have stor indflydelse på denne afhandling (Se afsnit 4.3.3). En yderligere mangelfuldhed i den eksisterende litteratur er den meget generel tilgang, hvor enkelte specifikke brancher ikke er undersøgt: Søgninger på både Google Scholar og AU Library giver kun sparsomme resultater når der søges på 'Ownership' og 'Industries', og majoriteten af de branche-specifikke artikler omhandler alene banksektoren. Der er kun fundet enkelte artikler, som vedrører den farmaceutiske-branche og disse er afgrænset til asiatiske selskaber og har derfor ringe relevans for denne afhandling.

## 2.2. Corporate governance

*"Corporate governance is the system by which companies are directed and controlled"*

(Cadbury, 1992, p. 14)

Det primære teoretiske udgangspunkt, for afhandlingen, ligger indenfor corporate governance. Eller hvad Gyldendals *Den Store Danske* omtaler som ”principper for god selskabsledelse” (Gyldendal, 2017), en definition som ikke er helt dækker fuldstændigt over hvad corporate governance er i den akademiske verden. Det er således nødvendigt med en klar definition af hvad corporate governance er – og ligeså vigtigt: hvad det *ikke* er.

Corporate governance betragtes i litteraturen ud fra to perspektiver: Ét afkast/investorperspektiv, som måske bedst beskrives af følgende citat: *“Corporate Governance deals with the ways in which suppliers of finance to corporations assure themselves of getting a return on their investment.”* (Shleifer & Vishny, 1997, p. 737). Dette perspektiv lægger altså vægt på aktionærernes afkastet og deres kontrol herover. Corporate governance faget beskæftiger sig, indenfor denne ”skole”, med hvordan aktionærerne kan øge deres afkast. At alene se på corporate governance med et afkastperspektiv, er dog noget begrænsende, da corporate governance, i bredeste forstand, ikke alene omhandler profitsøgende virksomhed, men også virksomheder med andre motiver, f.eks. statsejede selskaber, hjælpe- og nødhjælpsorganisationer eller familieejede virksomheder. Afkastperspektivet forholder sig heller ikke til kompleksiteten af virkelige corporate governance problemstillinger, da andre aktører end ejerne har interesse i virksomheden, f.eks. banker, forsikringsselskaber eller stater.

For at definere corporate governance med den nødvendige bredde, er det nødvendigt at forholde sig til mere end blot ejere og afkast. Derfor ses ofte, i litteraturen, at en mere omfattende definition af corporate governance anvendes. Dette perspektiv beskrives bedst af et citat fra den britiske Cadbury-rapport: *“Corporate governance is the system by which companies are directed and controlled”* (Cadbury, 1992, p. 14). Definitionen her er mere fokuseret på kontrol og indflydelse af selskaberne, end alene aktionærernes afkast.

De to definitioner er ikke indbyrdes udelukkende og det afkastorienterede perspektiv på corporate governance er, i ligeså høj grad, relevant for denne afhandling, da udgangspunktet for analysen er selskabernes rentabilitet. Men Cadbury-definitionen, som fokuserer mere på kontrol, er langt mere relevant for indeværende teoriafsnit, da den i højere grad forholder sig til kompleksiteten af corporate governance problemstillinger.

Netop fordi at corporate governance er så kompleks en størrelse og definitionen er så bred, som den er, så er samtidig nødvendigt at definere nogle grænser for hvad corporate governance *ikke* er. Her tages udgangspunkt i de centrale punkter fra afsnit 1.2 i Steen Thomsen og Martin Conyon’s lærebog *Corporate Governance: Mechanisms and Systems*, som også anvender Cadbury-definitionen af corporate governance (Thomsen & Conyon, 2012, pp. 4-5). Jf. disse punkter, så er corporate governance ikke synonym med lovgivning, da også uregulerede virksomheder har governance-mekanismer samt governance i højere grad omhandler fænomener og adfærd. Corporate governance er heller ikke et synonym med

selskabsledelse<sup>2</sup>, da governance ikke direkte omhandler de operationelle forhold, men alene kontrollen af ledelsen, som selvstændigt faciliterer disse forhold. Sidstnævnte skilning er særlig vigtig, da det netop er i denne relation at der kan opstå agentproblemer, hvilket er omdrejningspunktet for næste afsnit.

### 2.2.1 Principal-agent-teori

Siden udgivelsen af Jensen og Meckling (1976), har Principal-agent-teori<sup>3</sup>, været den mest indflydelsesrige teori indenfor corporate governance. Udgangspunktet for teorien er at undersøge konsekvensen af motivforskelle blandt aktører, som har et principal-agent-forhold. Definitionen af en sådan relation, er i den oprindelige Jensen & Meckling-artikel beskrevet som: ”(...) a contract under which one or more persons (the principal(s)) engage another person (the agent) to perform some service on their behalf which involves delegating some decision making authority to the agent.” (Jensen & Meckling, 1976, p. 5). Principal-agent-forhold baserer sig således på delegation af formel autoritet fra principalen til agenten. Sådanne forhold forekommer i flere situationer og kan resultere i interessekonflikter mellem principalen og agenten. Thomsen og Conyon (2012, p. 20) skelner mellem tre overordnede principal-agent-problemer:

- *Type 1 (Ejer- og ledelseskonflikter)*: Omhandler interessekonflikter mellem aktionærerne (principal) og direktionen (agent).
- *Type 2 (Minoritet og majoritetskonflikter)*: Omhandler interessekonflikter mellem majoritetsejere med kontrollerende ejerskab (agent) og de øvrige mindre ejere (principal). Her fungerer majoritetsejerne som agent for minoritetsaktionærerne, da sidstnævnte er uden kontrollerende medindflydelse og derfor afhængig af majoritetsejernes kontrol med virksomheden og ledelsen.
- *Type 3 (Aktionær og interessent konflikter)*<sup>4</sup>: Omhandler interessekonflikter mellem selskabsejjerne (agent) og øvrige interesser (principal), som kan være stater, banker eller samfundet som helhed. Et eksempel kan være overdreven risikovillighed blandt systemisk vigtige virksomheder.

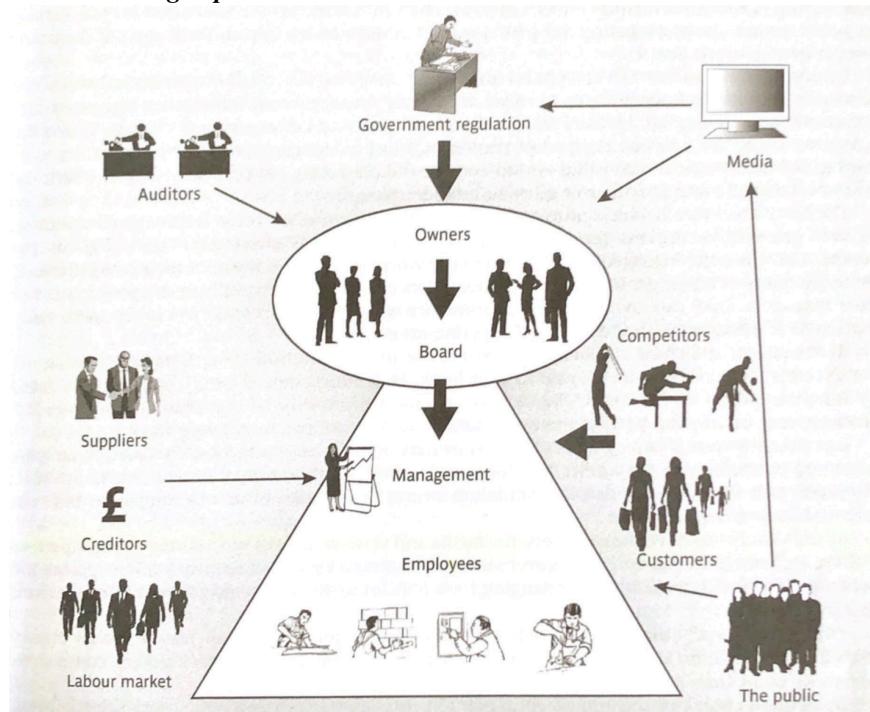
Denne afhandlings primære fokus er type-1 og type-2 agentproblemer. Type-1 problemer opstår på grund af adskillelsen af det kontrollerende ejerskab og den styrende ledelse, som kendtegner både private såvel som offentlige virksomheder. Gennem denne adskillelse risikerer ejerne at alene være udbydere af kapital, og agentproblemerne opstår derfor hvis ledelsen forvalter kapitalen med egennyttige motiver og ikke i overensstemmelse med ejernes ønsker. Foranlediget af principal-agent-teorien, må det forventes at forskellige ejerskabsstrukturer, er bedre eller dårligere til at kontrollere at ledelsen handler i overensstemmelse med sine interesser.

<sup>2</sup> Thomson & Conyon (2012) anvender her det engelske begreb: *Corporate Management*

<sup>3</sup> *Agency Theory* på originalsprog (Jensen & Meckling, 1976)

<sup>4</sup> Denne type af agentproblemer, relaterer sig til *Corporate Social Responsibility* (CSR).

**Figur 2 – Det udvidede agentproblem**



Kilde: Thomsen and Conyon (2012) figure 1.1, *The Extended Agency Problem*

Note: Det udvidede agentproblem illustrerer kompleksiteten i virkelige scenarier, hvor principal-agent-forholdet ikke kan begrænses til kun to parter.

## 2.2.2 Kontraktteori

Efter at, i forrige afsnit, have slået fast at ejernes opgave er at sikre at ledelsen handler i overensstemmelse med deres interesser, så må ejernes kontrol af ledelsen vurderes som et vigtigt udgangspunkt for at rentabilitet og værdiskabelse. Ejernes kontrol opnås gennem to forhold: formelt ejerskab og kontrakter. Det formelle ejerskab, som er udgangspunktet for den senere analyse, beskrives i afsnit 2.3. Dette korte afsnit, vil omhandle kontrahering mellem ejere og ledelse, med udgangspunkt i kontraktteori.

Kontraktteori<sup>5</sup> gør op med den tidligere forståelse af kontrakter som fuldkomne og tager i stedet det mere realistiske udgangspunkt at kontrakter er grundlæggende ufuldkomne og umuligt kan medtage alle eventuelle nutidige samt fremtidige forhold (Grossman & Hart, 1986). Kontraktteori beskæftiger sig således med hvem der har beføjelserne over de ikke-kontraktuelle forhold, som indirekte er en del af parternes indbyrdes aftale. Disse implicitte beføjelser beskrives indenfor teorien som residual-kontrollen.

Inddragelsen af kontraktteori har flere implikationer for ejerskab. Eksempelvis kan det udledes af teorien at incitament-baseret aflønning – en kontrakt mellem ejere og ledelse – aldrig vil kunne tage højde for alle tænkelige forhold og derfor vil der alene kunne opnås ”korrekte” incitamenter for ledelsen, hvis ledelsen ligeledes er medejer af virksomheden (Hart, 2016). Kontraktteori underbygger desuden vigtigheden af at forskellige ejerskabsstrukturer. Én bagvedliggende konsekvens af teorien er at fuldkommen ejerskab er mere

<sup>5</sup> Incomplete Contracts Theory på originalsprog (Hart, 2016)

favorabelt end simpel medejerskab og desuden at direkte ejerskab er mere favorabelt end indirekte ejerskab, da flere kontraktuelle ”led” giver flere fejlkilder (Thomsen & Conyon, 2012, p. 24). Implicit i teorien ligger derfor én grundlæggende kritik af hvordan mange virkelige selskab- og ejerskabsstrukturer er organiseret. Årsagerne til denne diskrepans mellem det teoretiske ”ideal” og virkeligheden er mange. Bl.a. adgangen til finansiering og ledelsens evner og kompetencer.

### 2.2.3 Opsummering

*”The ‘best owner’ of a company, or an asset, is the owner which can create most value with it, which involves weighing costs and benefits of ownership”*

(Thomsen & Conyon, 2012, p. 133)

Med baggrund i principal-agent-teori og kontraktteori er der et teoretisk belæg for at undersøge sammenhængen, mellem ejerskabsstruktur og virksomhedernes performance, empirisk. Adskillelsen af ejerskab og ledelse har medført at det er ejernes primære ansvar at sikre at ledelsen handler i deres interesse, den opgave løftes gennem formelt ejerskab, som er udgangspunkt for denne afhandling, og kontrahering med ledelsen. Kontraktteori siger at disse kontrakter sjældent er fuldkomne og selv i kontraktuelle forhold, så er det umuligt at skabe perfekt overensstemmelse mellem ledelsen og ejernes interesser.

Det forhold at fuldstændig overensstemmende interesser ikke er teoretisk muligt, gør at fuldstændigt ejerskab, og ideelt set med ejeren i en ejerleder-rolle, bør anses som et ideal indenfor corporate governance. Denne påstand understøttes også empirisk, af studier som har undersøgt ejerlederskab og familieejerskab (Anderson & Reeb, 2003; Andres, 2008; Clifford G Holderness & Sheehan, 1988). Virkeligheden er dog for mange virksomheder mere kompleks og finansieringsbegrænsninger, manglende ledelseskompentence samt forskellige investeringshorisonter er alle forhold som gør at langt de fleste virksomheder fraviger fra idealet om ejerledelse.

Efter at have redejort for principal-agent- og kontraktteori, som begge er teorier som forklarer de underliggende årsager bag corporate governance problemstillinger, så vil der i næste afsnit (2.3) redegøres for de performancemål som vil anvendes i analysen.

### 2.3. Performancemål

Hvor de forrige afsnit omkring principal-agent- og kontraktteori, redegjorde for det teoreti- ske udgangspunkt for afhandling, så er formålet med indeværende afsnit at redegøre for de performancemål som anvendes i analysen.

Udgangspunkt er de tre performancemål som tidligere lignende studier har anvendt: Totalkapitalforrentningen (ROA), Egenkapitalforretningen (ROE) og Tobin's q (Anderson & Reeb, 2003; Laeven & Levine, 2007). Tabel 1 nedenfor indeholder en overordnet beskrivelse af de tre mål og de kommende delafsnit vil beskrive dem i detaljer:

**Tabel 1 – Oversigt over performancemål**

Performancemål	Beskrivelse
ROA (Return On Assets)	ROA mäter selskabets rentabilitet med udgangspunkt i forrentningen af selskabets totale aktiver og mäter derfor både egen- og fremmedkapitalsforrentningen.
ROE (Return On Equity)	ROE mäter selskabets egenkapitalforrentning. I modsætning til ROA, så mäter ROE forrentningen af selskabet egenkapital og forholder sig ikke til andelen af fremmedkapital.
Tobin's q	Tobin's q forholder sig i modsætning til de øvrige performancemål, ikke alene til regnskabsværdien af selskabet, men også markeds- værdien heraf. Tobin's q er et mål for forholdet mellem selskabets markedsværdi (aktiekapital samt gæld) og dets totale aktiver.

#### 2.3.1 Egen- og totalkapitalforrentning (ROE og ROA)

Både egen- og totalkapitalforrentningen er regnskabsmål for virksomhedens rentabiliteten. Totalkapitalforrentningen (ROA) er et mål for forrentningen af selskabets totale aktiver og beregnes som resultat af primær drift (EBIT) ift. balancens samlede aktiver (Bodie, Kane, & Marcus, 2014):

$$ROA = \frac{EBIT}{Aktiver} \quad (2.1)$$

ROA (Totalkapitalforretningen/Return On Assets) er et forrentningsmål for selskabets totale aktiver og skelner dermed ikke mellem egen- og fremmedkapital. Ved at inkludere fremmedkapital i nævneren, så kendtegner ROA sig ved at mæle virksomhedens forretning af hele kapitalapparatet. Modsætningsvist forholder ROE (Egenkapitalforrentningen/Return On Equity) sig kun til ejernes egenkapital og ROE mäter derved selskabets rentabilitet ud fra et investeringsøjemed.

$$ROE = \frac{Nettoindkomst}{Egenkapital} \quad (2.2)$$

Den primære forskel på ROA og ROE er således om gælden inkluderes i nævneren og observerede forskelle mellem de to mål, er en indikation for selskabets finansielle politik, altså om selskabet har meget eller lidt gæld. ROE adskiller sig også ved at inkludere renter og skattekostninger i tælleren (Bodie et al., 2014).

### 2.3.2 Tobin's q

Tobin's q adskiller sig fra både ROA og ROE, ved ikke alene at være et regnskabsmål, men også se på virksomhedens børsværdi. Fordelen ved at inddrage markedsværdien, er at den i højere grad afspejler aktivernes nuværende værdi, i modsætning til regnskabstal som alene er et mål for aktivernes historiske værdi: Aktivets initiale anskaffelsesomkostninger justeret for afskrivninger.

$$Tobin's q = \frac{\text{Aktivets Markedsværdi}}{\text{Aktivets Erstatningsomkostninger}} \quad (2.3)$$

Tobin's q bygger på det teoretiske forhold at et aktivs markedsværdi og dets erstatnings-/reproduktionsomkostninger bør være 1:1 på lang sigt. Denne teori er første gang fremstillet af James Tobin (deraf navnet Tobin's q) og var oprindeligt en tilgang til at analysere makroøkonomiske forhold, ved at se på børsmarkedet som helhed (Tobin & Brainard, 1976), senere studier har dog vidst at børsmarkedets q-værdi kan afvige markant fra 1 i længere perioder (Bodie et al., 2014, p. 593).

Tobin's q har senere vundet indflydelse i corporate finance-litteraturen som et mål for enkelte selskabers værdiansættelse – q-målet er anvendt i både Anderson & Reeb (2003) og Laeven & Levine (2007), som er to lignende artikler, der ser på hhv. familieejerskab og ejerskabsstrukturer generelt – Tobin's q er ligesom ROA et mål for totalkapitalforrentningen, da selskabets totale aktiver tages i betragtning.

En væsentlig udfordring ved Tobin's q er estimation af erstatningsomkostningerne. Her anvender tidlige studier forskellige approksimationer heraf (Anderson & Reeb, 2003; Bharadwaj, Bharadwaj, & Konsynski, 1999; Laeven & Levine, 2007). To typiske antagelser er at erstatningsomkostningerne og gældens markedsverdi er identiske med hhv. aktivernes og gældens bogførte værdier, hvorfor det approksimerede  $q^*$ -forhold kan udledes:

$$Tobin's q^* = \frac{\text{Markedsverdi} + \text{Gæld}}{\text{Aktiver}} \quad (2.4)$$

Hvor tæller inkluderer egenkapitalens markedsverdi samt gældens bogførte værdi (som proxy for gældens markedsverdi) og nævneren omfatter alle selskabets aktiver. Nævneren indeholder altså både selskabets bogførte egenkapital og dets forpligtelser, som dermed er en proxy for erstatningsomkostningerne.

### 3. Metode

I dette kapitel vil de to anvendte modeller: Fixed Effects (FE) og Random Effects (RE) beskrives. Desuden vil der redegøres for forudsætningerne for de to modellers validitet. Modellerne vil anvendes til analysen og forudsætningerne testes i resultatafsnittet.

**Tabel 2 – Forudsætninger for regressioner**

Forudsætninger for Fixed Effects (FE) og Random Effects (RE) modeller		
<b>FE.1</b>	Linearitet i parametrene	
<b>FE.2</b>	Simpel tilfældig udvælgelse	
<b>FE.3</b>	Periodeafhængige variabler og ingen perfekt multikollinearitet	
<b>FE.4</b>	Nulforventning for fejlleddet	$E(u_{it} X_i, a_i) = 0, \forall t$
<b>FE.5</b>	Homoskedasticitet	$Var(u_{it} X_i, a_i) = Var(u_{it}) = \sigma_u^2, \forall t$
<b>FE.6</b>	Ingen seriekorrelation	$Cov(u_{it}, u_{is} X_i, a_i) = 0, \forall t \neq s$
<b>FE.7</b>	Normalitet af fejlleddet	$u_{it} X_i, a_i \sim i. i. d N(0, \sigma_u^2)$
<b>RE.1</b>	Ingen perfekt multikollinearitet	
<b>RE.2</b>	Ingen korrelation mellem $a_i$ og $X_i$	$E(a_i X_i) = \beta_0$
<b>RE.3</b>	Homoskedasticitet	$Var(a_{it} X_i) = \sigma_a^2$

Kilde: Wooldridge (2015, pp. 457-480)

#### 3.1. Fixed effects model

For at opstille en statistisk model med flere perioder, vil der i analysen anvendes to modeller: fixed effects og random effects. Afsnittet her tager udgangspunkt i fixed effects modellen og dens forudsætninger.

Opstillingen af en generel fixed effects model findes i formel 3.1 nedenfor:

$$y_{it} = X_{it} \cdot \beta + \alpha_i + u_{it}: \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (3.1)$$

hvor:

- $y_{it}$  er den afhængige variabel for selskab  $i$  for periode  $t$ .
- $X_{it}$  er den periodeafhængige  $1 \times k$  regressionsvektor ( $k$  er antallet af uafhængige variabler)
- $\beta$  er den  $k \times 1$  matrix for parametrene.
- $\alpha_i$  er den uboserverede periodekonstante effekt.
- $u_{it}$  er fejlledet.

Fixed effects dækker over en tilgang til at håndtere periodeuafhængige<sup>6</sup> effekter i et panel-datasæt, uden at antage at den specifikke effekt  $a_i$  er ukorreleret med de observerede variabler:  $Corr(a_i, X_{it}) \neq 0$ .

<sup>6</sup> Begreberne: periodekonstante, periodeuafhængige og tidsinvariante anvendes om samme effekter – Altså variabler som er konstante over tid. Modsætningsvist dækker begreberne periodeafhængige og tidsvariante effekter over variabler som ændrer sig over tid.

I praksis håndteres dette, ved at lave en *within*-transformation af model 3.1, ved at beregne gennemsnittene af modellens variable ved:

$$\bar{y}_i = T^{-1} \sum_{t=1}^T y_{it} \quad (3.2)$$

Og ligeledes for  $\bar{u}_i$  og de øvrige variabler:  $\bar{x}_{i1}, \dots, \bar{x}_{ik}$ . Ved at fratrække disse gennemsnit, i den oprindelige model, fås den transformerede model:

$$\ddot{y}_{it} = \ddot{X}_{it} \cdot \beta + \ddot{u}_{it}: \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (3.3)$$

Hvor  $\ddot{y}_{it} = y_{it} - \bar{y}_i$  osv. for de øvrige variabler. Det bør bemærkes at den uobserverede effekt  $\alpha_i$  er konstant over tid og forsvinder i ligning 3.3, da:  $\alpha_i - \bar{\alpha}_i = 0$ .

Heldigvis beregnes denne transformation automatisk i Stata. Stata laver dog en mere kompleks beregning, hvor  $\ddot{y}_{it}$  samtidig korrigeres for den samlede gennemsnit:

$$\ddot{y}_{it} = y_{it} - \bar{y}_i + \bar{\bar{y}} \quad (3.4)$$

Hvor  $\bar{\bar{y}}$  angiver det samlede gennemsnit beregnet ved:

$$\bar{\bar{y}} = \frac{\sum_{i=1}^N \sum_{t=1}^T y_{it}}{N} \quad (3.5)$$

Stata transformation er derfor beregningsmæssig mere kompleks og afledt heraf er at Stata i praksis estimerer et konstantled *\_cons*, når programmet estimerer en fixed effects model (Kommando: *xtreg y x1 ... xk, fe*). Konstantleddet er dog arbitraert og ikke egnet til tolkning, men angiver blot den gennemsnitlige værdi af de konstante effekter (Gould, 2019).

Modellen angivet i ligning 3.3 estimeres med *pooled OLS*<sup>7</sup> og derved beregnes den estimerede fixed effects model<sup>8</sup>. Forudsætningerne for at estimere modellen er ikke ulig forudsætningerne for almindelig *pooled OLS*. Pooled OLS vil dog ikke berøres yderligere, men der henvises i stedet til Wooldridge (2015).

I næste afsnit (3.1.1) vil forudsætningerne for fixed effects model beskrives.

---

<sup>7</sup> Ordinary Least Squares, eller på dansk: ”Mindste kvadraters metode”.

<sup>8</sup> Antallet af frihedgrader for modellen er:  $df = N(T - 1) - k$ , hvor  $N$  er det totale antal observationer,  $T$  er det totale antal periode og  $k$  er det totale antal parametre (Wooldridge, 2015).

### 3.1.1 Forudsætninger

#### FE.1 – LINEARITET I PARAMETRENE

Linearitet i parametrene indebærer at modellen skal kunne specificeres i en funktionel form, som i ligning 3.1, henover alle  $t$  perioder:

$$y_{it} = X_{it} \cdot \beta + \alpha_i + u_{it}: \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (3.1)$$

Et brud på linearitet opstår hvis den lineære model forsøges estimeret på ikke-lineært data. Det er i praksis svært at afklare om forholdet mellem den uafhængige og de forklarende variabler er lineært, men ved at tillade *ln*-transformationer af de forklarende variabler, bør det være muligt at estimere en model hvor parametrene i det mindste er lineære.

En yderligere forudsætning for paneldata, er ligesom ved tidsserieanalyser at den stokastiske proces skal være stationær. En stokastisk proces er strengt stationær, hvis hver  $k$  uafhængig variabel  $x_{ik}$  er identisk fordelt for alle tidsperioder. Stationæritet ift. paneldata diskuteres ikke uddybende i hverken Wooldridge (2010) eller Wooldridge (2015), årsagen er formentligt at paneldata ofte har kortere tidshorisonter og pludselige forskydninger i parametrene (*unit-roots*) sjældent anses som et problem, desuden er tests for stationæritet sjældent pålidelige for korte perioder.

#### FE.2 – SIMPEL TILFÆLDIG UDVÆLGELSE

Den gængse OLS-forudsætning om en tilfældigt udvalgt stikprøve, gælder også for fixed effects modellen. Simpel tilfældig udvælgelse indebærer at stikprøven skal udvalgt tilfældigt fra populationen. Paneldata er per definition konstrueret så de samme  $N$  virksomheder følges gennem hele perioden, men forudsætningen om tilfældig udvælgelse indebærer her at disse  $N$  antal virksomheder er tilfældig udvalgt.

Et yderligere forhold omhandler om datasættet er balanceret eller ej. Et datasæt som mangler observationer for enkelte selskaber, i enkelte perioder, kaldes ubalanceret. Et ubalanceret datasæt er ikke et umiddelbart problem, da *within*-transformationen i Stata automatisk korrigerer for dette forhold ved at anvende de  $T_i$  antal perioder for hvert  $i$  selskab, hvor:  $T_i < T$ . Det eneste problem er hvis  $T_i = 1$ , hvilket indebærer at selskabet ingen periodeafhængig variation har, og dermed bryder den med næste forudsætning (FE.3).

Forudsætningen om simpel tilfældig udvælgelse kan ikke testes, men vurderes på baggrund af ens datakilde. I resultatetafsnittet (5.1.3:FE.2) vurderes denne forudsætning på baggrund af Orbis-datasættet.

#### FE.3 – PERIODEAFHÆNGIGE VARIABLER OG INGEN PERFEKT MULTIKOLLINEARITET

Årsagen til at alle  $T_i > 1$ , for hvert selskab  $i$ , skyldes at *within*-transformationen fjerner alle konstanter effekter og da selskaber med kun én observation per definition er konstante, så forsvinder disse i den transformerede model.

Foruden forudsætningen om at variablerne ikke må være konstante over tid, så forudsætter FE.3 også at der ikke må være perfekt multikollinearitet – altså at ingen variabler må forklare nogen af de øvrige variabler perfekt.

Perfekt multikollinearitet indtræffer når to eller flere af de forklarende variabler, fra den underliggende stikprøve, korrelerer perfekt:

$$\text{Corr}(x_j, x_k) = -1 \vee 1, \quad \forall j \neq k \quad (3.6)$$

Multikollinearitet kan ofte påvises i datasættet, hvis der enten inddrages flere dummy-variabler som forklarer hinanden perfekt (f.eks. hvis begge køn er repræsenteret af dummy-variabler) eller hvis én variabel indgår med forskellige faktorer<sup>9</sup>.

Stata opdager og fjerner automatisk variabler med fuldstændig perfekt korrelation, men et værktøj til manuelt at bedømme graden af multikollinearitet er variansinflationsfaktoren (VIF), som bestemmes ved:

$$VIF_i = \frac{1}{\text{Tolerancen}} \quad (3.7)$$

Hvor tolerancen bestemmes som:  $1 - R_i^2$ .  $VIF_i$  angiver den  $i$ 'te uafhængige variabels forklaringskraft af de øvrige forklarende variabler i modellen.

VIF har dog sine begrænsninger, da der ikke er konsensus omkring en tærskel for VIF (O'brien, 2007) og VIF-tærsklen vælges ofte arbitraert. En tommelfingerregel om at  $VIF_i < 10$  anvendes dog ofte. En værdi af  $VIF_i$  over 10 indikerer at den  $i$ 'te variabel har en forklaringskraft på  $R_i^2 = 0,9$  ift. de øvrige variabler.

#### FE.4 – NULFORVENTNING FOR FEJLLEDET

FE.4 forudsætter at for alle  $t$  perioder, er den forventede værdi af fejlledet nul, betinget af konstantleddet og samtlige uafhængige variabler i samtlige perioder:

$$E(u_{it} | X_i, a_i) = 0, \quad \forall t \quad (3.8)$$

Hvis denne forudsætning er opfyldt, kan modellen karakteriseres som strengt eksogen. Med andre ord er fejlledet  $u_{it}$  uafhængigt af  $X_i$  og  $a_i$  i *samtlige* perioder. Er  $u_{it}$  alene uafhængig i én periode altså:  $E(u_{it} | X_{it}, a_{it}) = 0$ , anses modellen som svagt eksogen. Der skal alene svag eksogenitet til for at OLS er konsistent, mens det er nødvendigt med streng eksogenitet for at OLS også er unbiased.

Et brud på denne forudsætning skyldes ofte misspecifikation af modellen, enten i form af manglende variabler (*Omitted Variable Bias*) eller ved at ikke-lineære effekter, f.eks. logaritmiske eller højere ordens effekter, ikke er medtaget.

Der findes flere tests for misspecifikation, men i resultatafsnittet vil Ramsey's RESET<sup>10</sup> blive anvendt.

---

<sup>9</sup> Et eksempel kan være en variabel angivet i procent fra [0;1] og samme variabel angivet i intervallet [0;100].

<sup>10</sup> Ramsey's Regression Specification Error Test (Wooldridge, 2015, p. 277).

RESET-testen følger den naturligere konsekvens at hvis den oprindelige model:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u \quad (3.9)$$

Opfylder forudsætningen om nulforventning for fejlledet, så vil ingen ikke-lineære funktioner af de forklarende variabler tilføje yderligere signifikans til modellen, hvis de tilføjes. RESET testes indebærer derfor at modellen tilføjes ikke-lineære effekter, f.eks. kvadratiske og kubiske effekter:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \delta_1 y^2 + \delta_2 y^3 + u \quad (3.10)$$

Og dernæst gennemføres som en F-test, for fælles signifikans, med  $H_0: \delta_1 = \delta_2 = 0$ . Der er ingen klar regel for hvilke ikke-lineære effekter som bør testes, men specifikationstesten i analysen vil tage udgangspunkt i kvadratiske og kubiske effekter.

Hvis forudsætningerne FE.1-FE.4 er opfyldt, er OLS estimatoren unbiased og konsistent når  $N \rightarrow \infty$ , for et fast  $T$  antal perioder. Forudsætningen om streng eksogenitet (FE.4) er den mest begrænsende og hvis den ikke er opfyldt, så er modellen biased.

De næste forudsætninger FE.5-FE.7 kræves for at de F- og t-teststatistikkerne og dermed de underliggende standardfejl er gyldige:

## FE.5 – HOMOSKEDASTICITET

Homoskedasticitetsforudsætningen indebærer at variansen af fejlledet skal være konstant for alle  $t$  perioder, betinget af  $X_i$  og  $a_i$ :

$$\text{Var}(u_{it}|X_i, a_i) = \text{Var}(u_{it}) = \sigma_u^2, \quad \forall t \quad (3.11)$$

Hvis modellen viser tegn på heteroskedasticitet, altså at fejlledet har en periodeafhængig varians, så er almindelige teststatistikker ikke længere gyldige og  $p$ -værdierne for de inkluderede variabler og modellen i helhed er upålidelige.

I finansielt data er homoskedasticitet forventeligt, da variationen i selskabernes indtjening typisk ændres over tid, da visse virksomheder af natur er mere konjunkturafhængige end andre, det gælder for så vidt også indenfor samme industri, da farmaceutiske selskaber, da disse kan præges af produktcyklusser el.lign.

Der kan testes for homoskedasticitet med b.la. en Breush-Pagan-test (Wooldridge, 2015), men i analysen, vil der i stedet anvendes *cluster*-robuste standardfejl i analysen, som korrigerer for både heteroskedasticitet, klyngeffekter og seriekorrelation, så der ikke er noget godt argument for at undlade at anvende dem. Ulempen ved beregningen af de robuste standardfejl, er alene fremtrædende når stikprøven er relativt lille ift. antallet af klynger. Anvendelsen af de robuste standardfejl medfører at almindelig statistisk inferens kan gennemføres på de beregnede teststatistikker.

## FE.6 – INGEN SERIEKORRELATION

Forudsætningen om ingen seriekorrelation medfører at for alle  $t \neq s$  er det idiosynkratiske fejlled ukorreleret, betinget af alle forklarende variabler  $X_i$  og  $a_i$ :

$$\text{Cov}(u_{it}, u_{is} | X_i, a_i) = 0, \quad \forall t \neq s \quad (3.12)$$

Dvs. at fejlleddene ikke skal afhænge af hinanden på tværs af perioden. Hvis der er korrelation mellem fejlleddene, vil de OLS-estimerede  $\beta$ -parametre være biased. Modsætningsvist vil OLS være den bedste lineære unbiased estimator (BLUE), når der ikke forekommer seriekorrelation.

For at teste for 1. ordens seriekorrelation kan der anvendes en Durbin-Watson test:

$$DW = \frac{\sum_{t=2}^T (\epsilon_t - \epsilon_{t-1})^2}{\sum_{t=1}^T \epsilon_t^2} \quad (3.13)$$

Hvor  $\epsilon_t$  er residualleddet for perioden  $t$ . Durbin-Watson testes med nulhypotesen:  $\rho = 0$ , og  $DW$ -teststørrelsen skal sammenlignes med nogle specifikke tærskelværdier i intervallet [0;4] for hhv. negativ og positiv seriekorrelation, nulhypotesen bekræftes omkring 2.

Som nævnt under FE.5, så er *cluster*-robuste standardfejl også robuste overfor seriekorrelation og eftersom Durbin-Watson testen kræver at FE forudsætningen om streng eksogenitet er opfyldt, samt at den kun tester for 1. ordens seriekorrelation (Altså korrelation én tidsperiode tilbage), så er Durbin-Watson af ringe relevans for analysen.

## FE.7 – NORMALITET AF FEJLLEDET

Den sidste fixed effects forudsætning er normalitet af fejlleddet, betinget af alle forklarende variabler  $X_i$  og  $a_i$ :

$$u_{it} | X_i, a_i \sim \text{i.i.d. } N(0, \sigma_u^2) \quad (3.14)$$

Hvis fejlleddet er normalfordelt og forudsætningen er opfyldt, indebærer det per definition at FE.4-FE.6 også er opfyldt, dermed er FE.7 en meget stærk forudsætning. Hvis FE.7 er opfyldt kan det med rette siges at fixed effects modellen er konsistent og at alle  $F$  og  $t$ -statisistikker følger nøjagtige  $F$  og  $t$ -fordelinger, hvilket sikrer korrekt statistisk inferens.

Der findes en række tests som kan anvendes til undersøge for normalitet, som udgangspunkt kan der dog antages normalitet på baggrund af OLS-estimatorens asymptotiske egenskaber, hvilket medfører at  $u_{it}$  tilnærmelsesvist bliver normalfordelt når  $N \rightarrow \infty$ , hvilket følger af den centrale grænseværdi sætning. Specifikt for panedata gælder desuden at  $T$  ikke må være for stor relativt til stikprøven, men med over 1.000 unikke selskaber i stikprøven og  $T = 10$ , så bør det ikke være et problem at antage normalitet på baggrund af OLS-estimatorens asymptotiske egenskaber.

### 3.2. Random effects model

I forrige afsnit afklarede at udgangspunktet for en fixed effects model er transformation af en *unobserved effects* model, så at den uobserverede effekt  $\alpha_i$  forsvinder, ræsonnementet er at den uobserverede effekt kan være korreleret med én eller flere forklarende variabler:  $\text{Corr}(\alpha_i, X_{it}) \neq 0$ .

Ved at forudsætte at  $\alpha_i$  er ukorreleret med de forklarende variabler:  $\text{Corr}(\alpha_i, X_{it}) = 0$ , er det muligt at fastholde de periodekonstante effekter i modellen. Ved at beholde de periodekonstante variabler i modellen, kan man foranlediges til at tro at almindelig OLS-estimation af den generelle model (3.1) er tilstrækkelig, men hvis det sammensatte fejlsomme defineres som:  $v_{it} = \alpha_i + u_{it}$ , kan model 3.1 opstilles som:

$$y_{it} = X_{it} \cdot \beta + v_{it}: \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (3.15)$$

Da  $\alpha_i$  indgår i det sammensatte fejlsomme  $v_{it}$  indebærer det at fejlsomme har seriekorrelation:

$$\text{Corr}(v_{it}, v_{is}) = \frac{\sigma_a^2}{(\sigma_a^2 + \sigma_u^2)}, \quad t \neq s \quad (3.16)$$

Afhængig af størrelsesforholdet mellem  $a_i$  og  $u_{it}$ , så indebærer det en betydelig seriekorrelation, hvilket bryder med OLS forudsætningen herom. Af denne grund, er det nødvendigt at anvende en estimator som fjerner seriekorrelation i fejlsomme, denne estimator er GLS<sup>11</sup>. GLS er en OLS regression på en transformert model. GLS-transformationen, vil ikke beskrives yderligere i denne opgave, da den kræver kompliceret algebra, men der henvises til Wooldridge (2010), Wooldridge (2015) og StataCorp (2013b). GLS estimatoren og random effects modellen beregnes automatisk i Stata med kommandoen: `xtset y x1 ... xk, re`.

#### 3.2.1 Forudsætninger

Forudsætningerne for random effects modellen bygger videre på fixed effects forudsætningerne FE.1, FE.2, FE.4, FE.5 og FE.6. Da random effects modellen tillader inklusion af periodekonstante variabler, så justeres FE.3 til alene at omhandle multikollinearitet (Se RE.1 nedenfor), af samme årsag tilføjes en yderligere forudsætning (RE.2) som omhandler sammenhængen mellem  $a_i$  og modellens variabler. Homoskedasticitetsforudsætningen FE.5 udvides desuden så der også kræves homoskedasticitet af det uobserverede konstante fejlsomme  $a_i$ , hvilket eksplickeres i RE.3 nedenfor:

#### **RE.1 – INGEN PERFEKT MULTIKOLLINEARITET**

For random effects, gælder i modsætning til FE.3, at der gerne må inkluderes periodekonstante effekter i modellen. Derfor erstatter denne forudsætning FE.3, ved alene at forudsætte at modellen ikke har perfekt multikollinearitet mellem nogen variabler og derfor at ligning (3.6) er opfyldt.

---

<sup>11</sup> Generalized Least Squares

## RE.2 – DET UOBSERVEREDE FEJLLED ER UKORRELERET MED DE ØVRIGE VARIABLER

Som introduceret ovenfor, er det en nødvendig forudsætning at  $a_i$  er ukorrelert med modellens øvrige variabler  $X_{it}$ :  $\text{Corr}(a_i, X_{it}) = 0$ . Denne forudsætning indikerer at den forventede værdi af det uboserverede fejlled  $a_i$  er konstant, betinget af alle øvrige variabler:

$$E(a_i|X_i) = \beta_0 \quad (3.17)$$

Konsekvensen af denne forudsætning er, foruden at  $a_i$  og  $X_{it}$  er ukorrelerede, at  $a_i \neq 0$  og modellen beholdet således et konstantled  $\beta_0$ .

Hvis denne forudsætning er opfyldt, anses random effects estimatoren som konsistent og asymptotisk normalfordelt når  $N \rightarrow \infty$ , med et fast  $T$  antal perioder. Desuden er random effects estimatoren mere efficient end pooled OLS, når RE.2 er opfyldt og foretrækkes derfor over fixed effects estimatoren. Denne kritiske random effects forudsætning, testes ift. en allerede kendt konsistent model med en Hausman test (Se afsnit 3.3).

## RE.3 – HOMOSKEDASTICITET

Foruden homoskedasticitetsforudsætningen FE.5, så er det desuden nødvendigt, for en random effects model, at forudsætte at det uboserverede fejlled  $a_i$  også udviser homoskedasticitet, betinget af alle  $X_i$  variabler:

$$\text{Var}(a_i|X_i) = \sigma_a^2 \quad (3.18)$$

Ved beskrivelsen af FE.5 blev det nævnt at det, ved anvendelse af heteroskedasticitetsrobuste standardfejl, er muligt at beregne gyldige teststatistikker. I analysen vil der som udgangspunkt anvendes *cluster*-robuste standardfejl, vedmindre andet er angivet, og dermed korrigeres standardfejlene for både heteroskedasticitet, klyngeeffekter og seriekorrelation (FE.6). Selve beregningen af disse cluster-robuste standardfejlene er kompliceret og her henvises til Woolridge (2010), dog beregnes robuste standardfejlene nemt i Stata, med en simpel tilføjelse til *xtreg*-kommandoen: *vce(cluster var)* hvor *var* angiver variablen for de unikke sel-skaber: (*Id*).

Hvis alle random effects forudsætninger (FE.1, FE.2, FE.4, FE.5, FE.6 og RE.1-RE.3) er opfyldt, så anses random effects estimatoren både som konsistent, unbiased og efficient ift fixed effects modellen, og bør derfor træde i stedet for fixed effects modellen.

At random effects estimatoren er konsistent, indebærer at dens fejlled er asymptotisk normalfordelt når  $N \rightarrow \infty$ , med et fast  $T$  antal perioder, og F- og t-teststatistikker følger dermed nøjagtige F- og t-fordelinger, hvilket muliggør korrekt statistisk inferens.

Som nævnt ovenfor, så er random effects estimatoren desuden mere efficient end fixed effects, når dens forudsætninger er opfyldt, men fixed effects estimatorens styrke er at den er konsistent og unbiased under færre forudsætning, da den er robust overfor korrelation mellem  $a_i$  og  $X_i$ . Men valget mellem fixed effects og random effects findes i balancen mellem robusthed og efficiens, og bør foretages på baggrund af nogle bredere overvejelser, som vil diskuteres i næste afsnit (3.3).

### 3.3. Valg af model

Valget mellem fixed effects og random effects afgøres i høj grad af om forudsætning RE.2 er opfyldt. Hvis der er korrelation mellem de det selskabsspecifikke fejlled og de øvrige variabler, så bør en fixed effects model som udgangspunkt vælges. Hvorvidt RE.2 er opfyldt, kan undersøges med en Hausman-test (Wooldridge, 2015), som sammenligner parametrene i en random effects model og en kendt konsistent model (fixed effects). Hausman-teststatistikken kan findes ved (StataCorp, 2013a):

$$H = (\beta_{RE} - \beta_{FE})'(V_{RE} - V_{FE})^{-1}(\beta_{RE} - \beta_{FE}) \quad (3.19)$$

Hvor:

- $\beta_{RE}$  er en matrix for random effects parametrene
- $\beta_{FE}$  er en matrix for fixed effects parametrene
- $V_{RE}$  er en kovariansmatrix for random effects parametrene
- $V_{FE}$  er en kovariansmatrix for fixed effects parametrene

$H$ -teststatistikken følger en  $\chi^2$ -fordeling og en forkastelse af  $H_0$  medfører at RE.2 forudsætningen ikke er opfyldt og random effects modellen er inkonsistent. Forudsætningerne for Hausman testen er at fixed effects modellen er kendt konsistent (FE.1-FE.4). Hausman-testen vil gennemføres i Stata med kommandoen: *hausman*.

Foruden den statistiske sikkerhed for om RE.2 er opfyldt, så bør der også gøres nogle konkrete overvejelser (Clark & Linzer, 2015) i valget mellem RE og FE. Et praktisk forhold som medfører at en fixed effects model er uegnet er at fixed effects transformasjonen, som bekendt, fjerner alle periodekonstante effekter, hvilket er et problem hvis de enkelte variabler ikke har tilstrækkelig variation over perioden og det netop er disse variabler som har en teoretisk betydning for analysen. Konkret for denne afhandling, kan det frygtes at der ikke er tilstrækkelige ændringer i ejerskabsvariablerne til at effekten herfra kan undersøges, det forhold diskutes i resultatsnippet (Afsnit 5.1.3: FE.3).

### 3.4. Fremgangsmåde for analysen

På baggrund af overvejelserne omkring valg af model, kan analysens overordnede fremgangsmåde udledes:

1. Først opstilles en fixed effects model (FE)
2. Forudsætningerne FE.1-FE.7 vurderes
3. Hausman-test gennemføres på FE-modellen samt en random effects model (RE) med samme parametre. Fastholdes  $H_0$  i Hausman-testen, vurderes RE.3 som opfyldt og RE-modellen er mere efficient end FE-modellen.
4. De øvrige RE-forudsætninger vurderes

Til sidst vurderes de endelige modellers statistiske og praktiske signifikans, og deres gyldighed diskutes på baggrund af deres forudsætninger.

## 4. Data

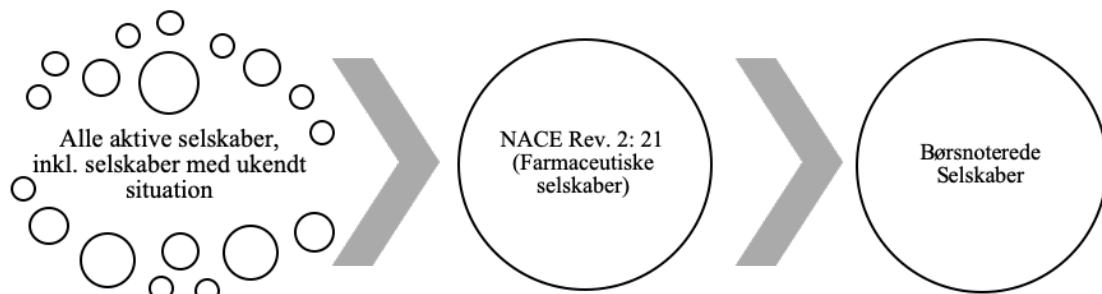
I dette kapitel vil der redegøres for udvælgelsen samt behandlingen af afhandlingens datagrundlag. Denne redegørelse vil ske gennem en forklaring af Orbis-databasen som kilde, en beskrivelse af hvordan datasættet er konstrueret og endelig en vurdering af datasættets begrænsninger og implikationer for analysens kvalitet.

### 4.1. Datagrundlag

Afhandlingens primære datakilde er Orbis Historical – en udvidelse til Orbis – som er en virksomhedsdatabase, fra Bureau van Dijk (BvD), som indeholder finansielle informationer, ejerskabsinformationer, ledelsesoplysninger samt beskrivelser af mere end 300 millioner selskaber (2018). Orbis indeholder regnskabsdata for de seneste 5-10 år på de fleste virksomheder i databasen, dog findes ejerskabsdata kun for seneste regnskabsår, hvorfor Orbis Historical-udvidelsen anvendes: Orbis Historical er et arkiv over Orbis-databasen for hvert år. Dermed kan ejerskabsdata hentes for en længere periode.

Da Orbis-databasen løbende er blevet opdateret med flere selskaber og variabler, så kan der observeres et markant fald i kvaliteten af datasættene, desto ældre de er. Det har også implikationer for det endelige datasæt, hvor antallet af selskabet er steget markant fra 1.498 i år 2008 til 2.273 i år 2018 i det endelige datasæt som beskrives i næste afsnit. Udviklingen, i antal selskaber efter filtreringen, er illustreret i figur 3 nedenfor:

*Figur 3 – Oversigt over søgekriterier*



- 2017: n > 226 mio.
  - 2016: n > 214 mio.
  - 2015: n > 164 mio.
  - 2014: n > 140 mio.
  - 2013: n > 109 mio.
  - 2012: n > 95 mio.
  - 2011: n > 83 mio.
  - 2010: n = ?
  - 2009: n = ?
  - 2008: n = ?
- 2017: n = 132.941
  - 2016: n = 129.556
  - 2015: n = 93.244
  - 2014: n = 85.986
  - 2013: n = 72.887
  - 2012: n = 69.798
  - 2011: n = 47.900
  - 2010: n = ?
  - 2009: n = ?
  - 2008: n = ?
- 2017: n = 2.273
  - 2016: n = 2.080
  - 2015: n = 1.935
  - 2014: n = 1.814
  - 2013: n = 1.606
  - 2012: n = 1.504
  - 2011: n = 1.490
  - 2010: n = 1.476
  - 2009: n = 1.499
  - 2008: n = 1.498

Note: Egen udarbejdelse, på baggrund af data fra Orbis Historical.

## 4.2. Dataudvælgelse

Data- og selskabsudvælgelsen, i Orbis, er gennemført med henblik på at sikre den størst mulige stikprøve, derfor sorteres der alene på to kriterier, fra den samlede Orbis-database:

1. *Farmaceutiske selskaber*: Her er sorteret på den europæiske økonomiske industriklassifikation NACE Rev. 2, hvor kode 21: "*Manufacture of basic pharmaceutical products and pharmaceutical preparations*" er anvendt som filter (eurostat, 2008).
2. *Børsnoterede selskaber*: Her sorteres, som nævnt i afgrænsningen (1.3), på om de farmaceutiske selskaber er børsnoteret eller ej.

Efter sorteringen, så fremkommer en liste med over 1.400 selskaber for alle årene (2008-2018). Som illustreret i figur 3 på forrige side, så øges omfanget, af den samlede database såvel som afgrænsningen, desto nyere databasen er – Heldigvis påvirkes den endelige stikprøve, efter afgrænsningen, ikke i samme udstrækning som den totale database.

Det forhold at databasen udvides markant over perioden kan have negative implikationer ift. datasættets repræsentativitet – en mulig fejlkilde som diskuteres i afsnittet om forudsætningen for stikprøvens tilfældige udvælgelse (5.1.2: FE.2). En yderligere konsekvens af Orbis Historicals periodeafhængige databasestørrelse er ift. antallet og kvaliteten af databasens variabler. Særligt omkring år 2011, hvor Orbis har en større opdatering, sker der en stor stigning i antallet af mulige variabler. De anvendte variabler for perioden før og efter år 2011 oplister i hhv. bilag 2.2 og 2.3. Databruddet har blandt andet resulteret i at enkelte informationer har været umulige at hente og desuden er enkelte variabler beregnet manuelt, disse forhold er diskuteret i næste afsnit.

### 4.2.1 Særligt for år 2008 - 2010

Foranlediget af Orbis-databasens opdateringen omkring år 2011, er der ikke de samme variabler tilgængelige i perioden 2008 – 2010. Derfor præges de 3 år af at mangle flere vigtige variabler. Bl.a. mangler selskabernes markedsværdi, hvilket er en nødvendig variabel til beregningen af Tobin's q. Desuden mangles oplysningerne omkring ejerskabet blandt direktionen samt de tre regnskabstal: Nettogetal, ændringer i arbejdskapital og anlægsinvesteringer (CapEx). Hvor det har været muligt, er disse variabler beregnet manuelt, det gælder for nettogetal (Formel 4.1) samt ændringer i arbejdskapitalen (Formel 4.2). Anlægsinvesteringer og Tobin's q har ikke kunnet beregnes med kongruens ift. Orbis variabler for senere år<sup>12</sup>.

$$Net\ Debt = Loans\ (Short\ Term) + Long\ Term\ Debt - Cash \quad (4.1)$$

$$Changes\ in\ Working\ Capital = Working\ Capital_t - Working\ Capital_{t-1} \quad (4.2)$$

---

<sup>12</sup> Her refereres til bilag 2.2 og 2.3 som redegør for variablerne i hhv. datasættene før og efter 2010.

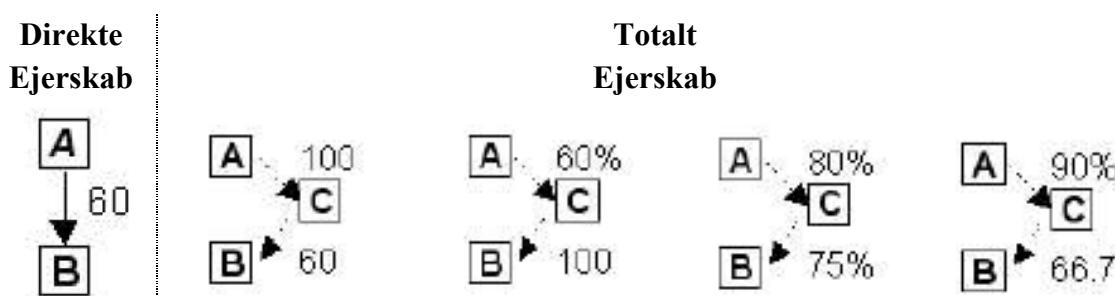
Et yderligere, mindre, forhold afledt af opdateringen af Orbis-databasen omkring år 2011, er at filerne som hentes fra de tidlige år, er et andet format (.xls) og med en anden opsætning end de øvrige datasæt. For at kunne anvende filerne i det samlede datasæt (se næste afsnit 4.3) har det været nødvendigt manuelt at ændre opsætningen samt konvertere dem til et nyere format (.xlsx) så de matcher de øvrige datasæt. Netto gæld og ændringer i arbejdskapital er også beregnet manuelt i denne proces<sup>13</sup>.

#### 4.3. Behandling af data

Når der hentes data fra Orbis Historical, så arkiveres ét datasæt per år, og den oprindelige data er derfor fordelt på 10 datasæt (ét for hvert år: 2008 – 2018). Af denne grund sker databehandlingen af to omgange: Først ensrettes og samles datasættene i Excel med en VBA-makro<sup>14</sup> og dernæst importeres det konsoliderede Excel-datasæt til Stata.

Efter at det endelige datasæt er importeret til Stata<sup>15</sup>, konstateres det hurtigt at én meget interessant variabel er udeladt: Selskabernes land. Variabel var oprindeligt udeladt, ud fra en tese om at store globale virksomheder, ikke påvirkes direkte af hvilket land de har hjemsted i. Det må dog sandes at det argument ikke holder, da ejerskabsstrukturen i høj grad må forventes at være påvirket af hjemlandets love og regler – hvilket underbygges af kompleksiteten af europæiske virksomheders selskabsstrukturer (Faccio & Lang, 2002). For at afbøde for denne væsentlige udeladelse, så måtte der eksporteres en liste over alle selskaber i datasættet fra Stata, som dernæst blev genindlæst hos Orbis, for at hente en variabel som angiver selskabernes ISO-landekode (Se beskrivelse af denne variabel i bilag 3.2)

*Figur 4 – Orbis' klassifikation af direkte og totalt ejerskab*



Kilde: Bureau van Dijk "Type of Percentage: Direct and Total Ownership" 2019), Se elektronisk bilag: <*OrbisOwnership.pdf*>.

Noter: Illustrationen til venstre viser hvordan et 60 % direkte ejerskab, vil se ud. Virksomhed A ejer 60% af virksomhed B. Det totale ejerskab har et utal af mulige variationer, de fire højresidige illustrationer er alle eksempler på at A har et totalt ejerskab på 60 % af virksomhed B.

<sup>13</sup> Nævnte manuelle rettelser: Konvertering, formatering og beregninger er foretaget i de elektroniske bilag <2008.xlsx>, <2009.xlsx> samt <2010.xlsx>, 1. fane i hver Excel-arbejdsmappe stemmer dog overens med det oprindelige datasæt fra Orbis.

<sup>14</sup> VBA (Visual Basic for Applications) er en indbygget og gratis udvidelse til Microsoft Excel.

<sup>15</sup> Se det elektroniske bilag <Makro.xls> for VBA-makroen og <Main.do> for Stata-koden.

#### 4.3.1 Konsolidering af datasættet med VBA

For at konsolidere de 10 datasæt, anvendes VBA-makroerne, som er gemt i det elektroniske bilag <*Makro.xslm*>. De fire makro-knapper, samler datasættene og gemmer dem som ét konsolideret datasæt: <*ConsolidatedDataset.csv*>, som har i alt 61 variabler (Se bilag 3.1).

Majoriteten af de 61 variabler er taget direkte fra Orbis-datasættene: De første 6 variabler er virksomhedsinformationer, 7-19 er regnskabsdata, som er gemt så det altid stemmer overens med senest tilgængelige år. De to dummy-variabler 20 og 21 angiver om hhv. CEO eller CFO ejer aktier i virksomheden. Det var desværre ikke muligt at finde nøjagtige informationer omkring størrelsen af dette ejerskab gennem Orbis. Foruden CEO- og CFO-dummy-variablerne, så er der i datasættet fire dummyvariabler per ejer med mere end 10 % ejerskab: 1) Ejerskabet i procent, 2) Om ejerskabet er direkte eller indirekte (se figur 4), 3) Hvilken type ejeren er: f.eks. bank, familieselskab eller andet industrielt selskab etc. og 4) BvD id-nummer for ejeren. (Se variabelliste i bilag 3.1)

#### 4.3.2 Import og sortering i Stata

Efter konsolideringen af de 10 oprindelige datasæt til én samlet datasæt (<*ConsolidatedDataset.csv*>), så importeres det konsoliderede datasæt til Stata, hvor det behandles yderligere. Først og fremmest, anvendes <*VarList.do*> til at navngive variablerne, så de, i Stata, stemmer overens med navnene fra variabellisten (bilag 3.1). Dernæst tilføjes ISO-landekoderne, så der tilføjes to nominale variabler. *Land* som angiver landekoden for alle selskaber og *Land2* som alene angiver landekoden for selskaberne, hvis landene har minimum 100 observationer (Se desuden linje 41-50 i <*Main.do*> samt bilag 3.2)

Desuden formateres samtlige variabler, så de kan anvendes i analysen: Komma ændres til punktum og *string*-variabler ændres til numeriske værdier.

Den største ændring af datasættet er formateringen af *EjerYType*-variablerne, som grupperes, så den kun har seks udfald (Se tabel 2 på næste side). Denne ændring er nødvendig, da der i de oprindelige datasæt er forskellige definitioner i de 10 oprindelige datasæt. Grupperingen kan ses i det elektroniske bilag <*VarList.do*> linje 100-140 (Se bilag 5.1).

Inden analysen påbegyndes, beregnes desuden performancemålene, som beskrevet i afsnit 2.4: ROA og Tobin's q (ROE var en tilgængelig variabel i Orbis-databasen og er hentet direkte derfra), samt ejerskabsvariablerne, som beskrives i næste afsnit (4.3.3)

Tabel 3 – Ejertypefordelingen for de tre største ejere.

	Ejer1Type	Ejer2Type	Ejer3Type
Financial company	1.624	971	518
Industrial company	3.345	1.660	687
Mutual & Pension Fund/Trust	1.047	866	446
N.A.	3.471	7.648	10.566
One or more named individuals or families	2.948	1.521	583
Other	435	204	70
Total	12.870	12.870	12.870

Note: Fordelingen af de tre første (og største) ejeres selskabstype, ”N.A.” indikerer at der *ikke* er en ejer. Udfald med N.A. i Ejer1Type indikerer derfor at selskabet *ikke* har nogen ejer med over 10 % ejerskab

#### 4.3.3 Ejerskabsvariabler

Før påvirkningen fra ejerskabsstrukturen kan undersøges, er det nødvendigt at klassificere de forskellige ejerskabstyper. Her beregnes først tre dummy-variabler, som angiver om:

- *MajEjer* ( $n = 886$ ): Selskabet har én majoritetsejer med  $> 50\%$
- *IngenEjer* ( $n = 1.039$ ): Selskabet ingen ejere over 10 % har
- *FlereEjer* ( $n = 938$ ): Selskabets 10 største ejere tilsammen ejer  $> 50\%$ .

Sidstnævnte variabel inkluderer *ikke* selskaber med majoritetsejere, selvom disse per definition har en ejerandel over 50 %. Inspirationen til klassificeringen af de tre variabler stammer fra Laeven & Levine (2007), som foruden ovennævnte dummy-variabler også inkluderer to variabler som angiver hhv. størrelsesforskellen mellem de to største ejere og spredningsforholdet mellem de to (Se bilag 3.3 for beregningen af disse):

- *EIME2*: Forskellen i ejerskab mellem den største og næststørste ejer.
- *disp*: Spredningsforholdet mellem de to største ejere.

Da denne afhandling, i modsætning til Laeven & Levine (2007) indeholder data for alle ejere med mere end 10 % ejerskab og ikke kun de to største, så beregnes desuden et koncentrationsmål for alle af ejere ( $> 10\%$ ). I manglen på bedre anvendes et Herfindahl–Hirschman (HH) indeks, som egentlig er et mål markedskoncentrationen indenfor industriøkonomi.

HH-indekset besidder dog nogle interessante karakteristika, som også gør det anvendeligt som koncentrationsmål for ejerskab. En udfordring ved beregningen af HH-indekset, er at det konsoliderede datasæt kun har data for ejere med en andel over 10 %, og mindre ejere derfor ikke inkluderes i beregningen, hvilket kan det påvirke koncentrationsmålet marginalt. HH-indekset beregnes som:

$$HH = \sum_{y=1}^{10} Ejerandel_y^2 \quad (4.3)$$

Hvor  $Ejerandel_y$  angiver den procentuelle andel for ejer  $y$  (Hayes, 2019).

**Tabel 4 – Sortering af datasæt**

	<b>Filter</b>	<b>Observationer</b>
	-	17.145
1	Alle observationer før år 2008 fjernes	16.493
2	Alle observationer fra år 2018 fjernes	16.350
3	Alle duplikerede observationer fjernes	14.506
4	Mangler landekode	12.727
5	Observationer med manglende data fjernes:	
6	Mangler omsætning	11.169
7	Mangler EBITDA	10.867
8	Mangler egenkapital	10.841
9	Mangler ROE	10.012
10	Mangler ROA	9.778
11	Mangler stiftelses år	8.351
12	Mangler nettogæld	8.348
13	Observationer med negativ omsætning fjernes	8.342
14	Selskaber yngre end 10 år fjernes	7.615
15	Selskaber med en markedsværdi under 100 mio. fjernes	5.247
	Endeligt datasæt: 1.567 selskaber & 5.247 observationer	

Note: Liste over de, i datasættet, anvendte filtre og udviklingen i stikprøvestørrelsen.

#### 4.3.4 Filtrering af datasættet

Filtreringen af det endelige datasæt sker ad flere omgange. I første omgang frasorteres redundante observationer, hvilket er alle observationer før år 2008 ( $n = 652$ ) samt observationer fra år 2018 ( $n = 143$ ), desuden fjernes observationer som indgår mere end én gang ( $n = 1.844$ ). Datasættet indeholder adskillige duplikerede observationer, da Orbis-datasættene inkluderer selskaber som ikke er helt opdaterede, f.eks. er der selskaber i 2015-datasættet med data fra 2013, som derfor også indgår i 2014-datasættet.

Foruden de første frasorteringer, fjernes observationer som mangler kritiske variabler ( $n = 4.379$ ), se linje 6-12 i tabel 4 ovenfor. Denne frasortering foretages ud fra formodningen om at selskaber som mangler basale regnskabstal, på grund af underliggende årsager, ikke er relevante for analysen, dermed bør det øge kvaliteten af datasættet at ekskludere disse. Ydermere fjernes seks selskaber fra datasættet, da de er registreret med en negativ omsætning, hvilket tyder på registreringsfejl i Orbis.

Til sidst frasorteres flere selskaber grundet nogle subjektive kriterier: Da problemformuleringen omhandler langsigtet værdiskabelse, frasorteres selskaber yngre end 10 år ( $n = 727$ ) – og altså er stiftet i analysen perioden 2008-2017. Desuden fjernes selskaber med en markedsværdi på mindre end USD 100 mio., jf. afgrænsningen beskrevet i afsnit 1.3.

Det endelige datasæt, efter frasorteringerne, er på 5.247 observationer, fordelt på 1.567 unikke selskaber (Id). Der er stadig selskaber som ikke har data for alle variabler.

#### 4.4. Begrænsninger

Selvom at det endelige konsoliderede datasæt er stort, så har det flere ufordelagtige karakteristika. Én af de store udfordringer gælder opdateringen af Orbis-databasen omkring år - 2011, som gør at enkelte variabler i de ældre datasæt enten er helt udeladt eller beregnet manuelt. Bl.a. har det ikke været muligt at hente data for selskabernes markedsværdi i årene før 2011, hvilket umuliggør beregning af Tobin's q for årene 2008-2010. Desuden mangler der i samme periode informationer omkring ledelsens ejerskab samt anlægsinvesteringer. Konsekvensen af dette forhold, er at modellen på Tobin's q alene kan opstilles p 7-års basis, mens modellerne for ROA og ROE kan opstilles på hele 10-års perioden.

En yderligere problemstilling er at der er store variationer i hvilke selskaber som er med fra år-til-år. Der har været en generel stigning i antallet af selskaber over hele perioden, men mange selskaber er kun med i en kort periode og 27,8 % af alle selskaber er med i under 3 år<sup>16</sup>. I sig selv er det ikke noget problem at datasættet er ubalanceret da både fixed- og random effects modellerne er robuste herfor. Men hvis der f.eks. kun er én observation for et selskab, så kan de ikke inkluderes i modellerne, da selskabet per definition er konstant over tid. Der er hele 645 selskaber (12,3 % af det samlede datasæt) som kun har én observation og de vil derfor ikke inkluderes i modellerne.

Et relateret, men endnu mere kritisk, forhold er hvis de bagvedliggende kriterier for hvilke selskaber som indgår i Orbis-databasen, ændrer sig over perioden. Hvis der er sammenhæng mellem selskaberne og perioden, så vil det være et sikkert brud på forudsætningen om simpel tilfældig udvælgelse (FE.2) og datasættet kan derfor ikke være repræsentativt for hele den globale farmaceutiske industri. I bilag 3.5, er der redegjort for fordelingen af de 5.247 observationer på år og land. Her ses bl.a. at Indien og Indonesien først indgår efter Orbis opdateringen omkring år 2011, og Japan først inkluderes fra år 2016. Det vil være et usandsynligt tilfælde, hvis det ikke er skyldes en underliggende selektionsbias fra Orbis af.

For at teste om den generelle udvidelse af Orbis-databasen har indflydelse på resultaterne, og undersøge konsekvenserne heraf, så vil der i diskussionsafsnittet opstilles modeller som kun inkluderer udsnit af den samlede periode.

---

<sup>16</sup> Specifikt: 27,7 % af alle selskaber, se linje 175-190 i det elektroniske bilag <Main.do> samt grafen i bilag 4.2 for fordelingen af antal år per selskab.

**Tabel 5 – Beskrivende statistikker for inkluderede variabler**

	Observationer	Middelværdi	Std. Afv.	Min	Max
ROA	4864	2,595	23,760	-300,612	136,822
ROE	4864	-,183	57,495	-999,054	300,33
q	3703	2,624	3,573	-,403	100,901
Alder	4864	28,360	25,358	10	339
Aktiv	4864	$2,7 \times 10^5$	$1,19 \times 10^7$	23,537	$1,57 \times 10^8$
LnAktiv	4864	12,632	1,833	3,159	18,874
LnAlder	4864	3,114	,614	2,303	5,826
FGEAR	4864	-,0408	1,024	-15,505	25,918
RDPct	3896	9,758	220,820	-3,693	12.991
disp	4864	,0888	,177	-1	,776
E1ME2	4864	5,215	11,32	-42	73,62
hh	4864	,157	,209	0	1,959

Note: Beskrivende statistikker for alle relevante *ikke*-dumme variabler

## 5. Resultater

I dette kapitel vil der blive redegjort for resultaterne af afhandlingens statistiske analyser, som desuden vil vurderes ift. forudsætningerne for de enkelte modeller. Kapitlet vil tage udgangspunkt i den analyse struktur som er beskrevet i metodeafsnittet (3.4), hvor der først opstilles en fixed effects model, hvorefter dennes forudsætninger vurderes. Dernæst gen nemføres en Hausman-test for at undersøge om den kritiske random effects forudsætning (RE.2) er opfyldt og dernæst opstilles en random effects model.

### 5.1. Fixed effects modellen

Inden den første fixed effects model opstilles, er det nødvendigt at afgøre hvilke variabler som skal inkluderes i modellen. Inspireret af Anderson & Reeb (2003) og Laeven & Levine (2007), så kontrolleres først og fremmest for to vigtige selskabsspecifikke karakteristika: størrelsen, målt på selskabernes totale aktiver og alder, målt på år siden stiftelse. I overensstemmelse med de to artikler, så anvendes en log-transformation af begge variabler, hvilket påvirker variablerne, så store observationer gøres relativt mindre. De to variabler har ingen større fortolkningsmæssig værdi, men deres koefficienter kan fortolkes ved at tage eksponentalfunktionen af log-værdierne:  $e^{\ln(a)} = a$ .

Foruden de to nævnte variabler, så inkluderes også to nøgletal: hhv. selskabets finansielle gearingen, målt ved nettogæld ift. selskabets totale aktiver, samt selskabets udviklingsomkostninger i procent af omsætningen. En vigtig pointe er at udviklingsomkostninger næppe kan materialisere sig i øget indtjening i år 0, så her kunne det være relevant at medtage *lags* fra tidligere år, men grundet den store udskiftning og de mange selskaber som ikke har observationer for hele perioden, har det ikke kunne lade sig gøre i praksis. Udviklingsomkostninger (%) inkluderes derfor som proxy for selskabernes generelle udviklingsniveau – der antages altså at udviklingsniveauet er stabilt over tid, og et højt udviklingsniveau i observationsåret er symptomatisk på et generelt højt udviklingsniveau i de foregående år og vice versa.

Foruden disse fire variabler, så tilføjes en række variabler, for at undersøge påvirkningen fra ejerskabet af virksomheden. Der inkluderes dummy-variabler for om hhv. selskabernes CEO eller CFO er aktionær samt om selskabet er majoritetsejet, *blockholder*<sup>17</sup>-ejere eller har fuldstændig spredt ejerskab. Desuden tilføjes et interaktionsled mellem majoritetsejer-dummy'en og familieejér-dummy'en, som skal indikere hvis majoritetsejeren samtidig er familie. Desuden tilføjes de tre variabler for spredningen og koncentrationen af ejerskabet også: *EIME2*, *disp* og *hh*

Samtlige variabler er forklaret i hhv. afsnit 4.3.3 samt bilag 3.3 og 3.4.

<sup>17</sup> *Blockholders* refererer i afhandlingen til ejere med mellem 10 – 50 % ejerskab.

Som nævnt i dataafsnittet (4.4) er det relevant at undersøge om dataskredet omkring år 2011, samt om selskaber med kun få observationer har indflydelse på resultaterne, derfor opstilles tre modeller, som beregnes i Stata (Se tabel 6 nedenfor), de tre modeller er:

- (1) ROA: hele perioden 2008-2017 medtages
- (2) ROA: kun perioden 2010-2017 medtages
- (3) ROA: hele perioden 2008-2017 medtages, men stikprøven inkluderer kun selskaber med mere end 3 observationer.

**Tabel 6 – Return On Asset, første modeller**

Variabler	(1) ROA	(2) ROA	(3) ROA
ln(Aktiver)	3,540*** (1,100)	5,365*** (1,409)	1,696** (0,775)
ln(Alder)	-4,267*** (0,808)	-10,726*** (3,259)	-3,672*** (0,742)
Finansiel Gearing	-0,622 (0,477)	-0,309 (0,503)	-0,936** (0,471)
Udviklingsomkostninger (%)	-0,004 (0,003)	-0,004 (0,003)	-0,027 (0,017)
Adm. Direktør aktionær?	-0,697 (0,704)	0,055 (0,749)	-0,690 (0,692)
Finansdirektør aktionær?	-3,862* (2,073)	-0,570 (1,647)	-4,080* (2,195)
1 = Hvis ingen ejer > 10%	0,005 (1,290)	-0,599 (1,377)	0,354 (1,224)
1 = Hvis de 10 største ejer samlet er >50%	-1,519 (1,103)	-1,430 (1,153)	-0,527 (0,895)
1 = Hvis største ejer er familie	0,847 (1,417)	0,085 (1,552)	2,160** (1,051)
1 = Hvis majoritetsejer er familie	0,947 (2,384)	-0,038 (2,712)	0,874 (2,340)
Spredningsforhold	-11,969*** (4,175)	-10,305** (4,236)	-8,874*** (2,918)
Forskel på største og næststørste ejer	0,169** (0,066)	0,152** (0,066)	0,149*** (0,055)
Herfindahl-Hirschmann index	-1,379 (2,650)	-1,201 (2,484)	-1,694 (2,745)
Observationer	3.896	3.606	2.798
R-squared	0,028	0,036	0,033
Antal selskaber	1.238	1.200	501
Første år	2008	2010	2008
Antal observationer per selskab	-	-	>3

Robuste standardfejl i parentes

\*\*\* p<0,01, \*\* p<0,05, \* p<0,1

### 5.1.2 Resultater

Resultaterne for de tre modeller er illustreret i tabel 6, på forrige side (Samme modeller beregnet for ROE og Tobin's q findes i bilag 3.1). Overordnet kan det konkluderes at koeficienterne har samme fortegn, men deres størrelse varierer meget fra model til model. Alle tre modeller kan karakteriseres som signifikante, for alle gængse signifikansniveauer:

Model (1):  $F(13,1237) = 4,11; p < 0,000$

Model (2):  $F(13,1199) = 2,56; p < 0,002$

Model (3):  $F(13,500) = 3,95; p < 0,000$

Hverken forskellene i de enkelte modellers variablers signifikans, eller modellernes overordnede signifikansniveau, giver anledning til en yderligere afgrænsning af stikprøven. På baggrund af OLS estimatorens asymptotiske egenskaber, bør flest mulige observationer fastoldes, det medfører at det er model 1, med den største stikprøve, som vil være omdrejningspunktet for den øvrige analyse i dette afsnit.

Det bør bemærkes at selvom model 1, estimeres på baggrund af hele datasættet, så er antallet af observationer lavere (3.896) end det fulde datasæt (5.247). Årsagen hertil er at FE estimatoren ikke inkluderer periodekonstante effekter – derfor vil selskaber med kun én observation automatisk udgå – samt at flere observationer mangler data for nogle variabler, her er særligt udviklingsomkostnings-variablen begrænsende, da 968 observation mangler data herfor, der kan på den baggrund argumenteres for at denne variabel bør udgå af modellen – et forhold som vil diskuteres i afsnit 6.1.1.

I næste afsnit vil fixed effects forudsætningerne for model 1 vurderes:

### 5.1.3 Forudsætninger

#### FE.1 – LINEARITET I PARAMETRENE

Udgangspunktet for modellens validitet er først og fremmest linearitet i parametrene, således at modellen kan stilles op som:

$$ROA_{it} = X_{it} \cdot \beta_i + \alpha_i + u_{it}; \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (5.1)$$

Hvor  $X_{it}$  er en  $1 \times 13$  matrix for de forklarende variabler, mens  $\beta$  er en  $13 \times 1$  matrix for variablene parametre.  $t$  går fra 1 til 10, og modellen estimeres på  $i = 1,238$  selskaber.

Da alle parametrene tager en lineær form i modellen (5.1) må det konkluderes at FE.1 er opfyldt.

#### FE.2 – SIMPEL TILFÆLDIG UDVÆLGELSE

Udvælgelsen af selskaberne i datasættet er allerede nævnt som en mulig fejkilde i dataafsnittet (4.4). Det er muligt at den løbende udvidelse af Orbis databasen, og sædeleshed opdateringen omkring år 2011, ikke er sket vilkårligt men på baggrund af underliggende kriterier, som derved bryder med forudsætningen om tilfældig udvælgelse.

Et yderligere kritikpunkt, vedrører filtreringen af datasættet, som er sket på baggrund af datatilgængelighed, og det kan formodes at der er en reel sammenhæng mellem selskabernes

transparens i deres finansielle rapporter og deres underliggende karakteristika. Derfor er der en risiko for at bestemte selskaber er underrepræsenteret, pga. at de ikke har samme transparens i deres finansielle rapportering.

Validiteten af datasættet generelt, afhænger derfor af en antagelse om at selskaberne i Orbis databasen er tilfældigt udvalgt for hvert år og at der ingen sammenæng er mellem selskabernes specifikke karakteristika og der inddragelse i datasættet.

**FE.3 – PERIODEAFHÆNGIGE VARIABLER OG INGEN PERFEKT MULTIKOLLINEARITET**  
 FE.3 dækker over to forudsætninger for de forklarende variabler i modellen: Først og fremmest skal de være tidsvariante og må derfor *ikke* være konstante over perioden. Dernæst må ingen af variablene forklare de øvrige variabler perfekt.

Det første forhold, at variablene ikke må være konstante over tid, er per definition opfyldt, da periodekonstante variabler fjernes af fixed effects transformationen. Specifikt for denne analyse, kan en udfordring være at meget få selskaber oplever ændringer i de primære variabler og der derfor ikke er tilstrækkelig variation i perioden. Ejerskabsvariablerne er primært defineret som dummy variabler og der skal derfor forholdsvis store ændringer, i den underliggende ejerskabsstruktur, for at variablene ændrer sig. For at undersøge om denne problemstilling er aktuel, så illustrerer tabellen nedenfor, antallet af selskaber som ar ændringer i de vigtigste variabler i perioden.

*Tabel 7 – Ændringer i dummy variable*

Variabler	Antal Ændringer
Adm. Direktør aktionær?	278
Finansdirektør aktionær?	61
1 = Hvis ingen ejer > 10%	91
1 = Hvis de 10 største ejer samlet er >50%	241
1 = Hvis største ejer er familie	233
1 = Hvis majoritetsejer er familie	23
Samlede antal selskaber i modellen	1.483

Note: Tabellen angiver antallet af selskaber som har en ændring for hver ejerskabs dummy variabel.

Af tabellen ses tydeligt at det kun er en begrænset andel af datasættets selskaber, som har ændringer i ejerskabsvariablerne, i løbet af perioden. Dette forhold kan have negative implikationer for fixed effects modellen, og er alt andet lige et argument for at en random effects model kunne være interessant (Se afsnit 5.3 for en diskussion heraf).

Anden del af FE.3 forudsætter at der ingen perfekt multikollinearitet indtræffer i modellen – altså må ingen forklarende variabler, korrelere fuldstændigt med de øvrige variabler. I metodeafsnittet (3.1.1: FE.3) angives hvordan variansinflationsfaktoren (VIF) kan anvendes som indikator for multikollinearitet, i tabel x på næste side er VIF beregnet for alle modellens variabler.

**Tabel 8 – VIF for variablerne i fixed effects modellen**

Variabler	VIF
ln(Aktiver)	27,34
ln(Alder)	26,98
Finansiel Gearing	10,95
Udviklingsomkostninger (%)	10,54
Adm. Direktør aktionær?	2,39
Finansdirektør aktionær?	1,68
1 = Hvis ingen ejer > 10%	1,68
1 = Hvis de 10 største ejer samlet er >50%	1,49
1 = Hvis største ejer er familie	1,49
1 = Hvis majoritetsejer er familie	1,39
Spredningsforhold	1,11
Forskel på største og næststørste ejer	1,01
Herfindahl-Hirschmann index	1,00
<b>Gns. VIF</b>	<b>6,85</b>

Note: Tabellen angiver variansinflationsfaktoren for alle variabler i modellen.

Af tabellen ses at flere af modellens variabler overskrider tommelfingerreglen om at VIF skal være under 10. Her er det vigtigt at huske at VIF-tærskelværdien *kun* er en tommelfingerregel og grænseværdien på 10 er arbitraert valgt. O'brien (2007) argumenteres der for at en VIF på over 10, ikke må være den eneste grund for at udelukke variabler fra en model og at nogle af de gængse værktøjer for at afhjælpe høj VIF virker modsat hensigten. Årsagen til at der iagttages høje VIF-værdier, kan skyldes de relative forskelle i variablene forklaeringsgrader og ikke nødvendigvis en faktisk korrelation imellem dem.

Konklusionen for multikollinearitet, må derfor være at selvom at der iagttages høje variansinflationsfaktorer, så er der ikke nogen umiddelbare argumenter for at fjerne enkelte variabler fra modellen, da årsagen bag høje VIF'er sandsynligvis skyldes de relative forskelle i variablene forklaeringskraft.

#### FE.4: NULFORVENTNING FOR FEJLLEDDET

Forudsætningen om streng eksogenitet, kan som nævnt i metodeafsnittet (3.1.1:FE.4) undersøges med en specifikationstest. Udgangspunktet er Ramsey's RESET test som en F-test for simultan signifikans<sup>18</sup>, hvor der er vil testes for signifikansen af eventuelle kvadratiske og kubiske led i modellen. Nulhypotesen er  $\delta_1 = \delta_2 = 0$ , hvor  $\delta_1$  er koefficienten for de kvadrede residualer, og  $\delta_2$  er koefficient for de kubiske residualer.

På baggrund af RESET-testen kan det konkluderes at fixed effects modellen ikke er specificeret korrekt, da inklusionen af kvadratiske og kubiske effekter, vil alt andet lige få den

<sup>18</sup> Engelsk: *joint significance*

justerede  $R^2$  til at stige. Det kan konkluderes på baggrund af RESET-testen som påviser simultant signifikante 2. og 3. ordens effekter for modellen ( $F(2,1237) = 30,69; p = 0,000$ ).

#### FE.5: HOMOSKEDASTICITET

Homoskedasticitetsforudsætningen er opfyldt, hvis FE modellens fejlled udviser en konstant varians, betinget af de forklarende variabler, over hele tidsperioden. Som diskuteret i metodeafsnittet (3.1.1: FE.5) kan der testes for heteroscedasticitet med en Breush-Pagan-test, men da der allerede er anvendt *cluster*-robuste standardfejl i analysen, så kontrolleres der på forhånd for eventuel heteroskedasticitet i beregningen af F- og t-teststatistikker (Wooldridge, 2010). Det kan retfærdiggøres at anvende robuste standardfejl på baggrund af datasættets størrelse, beregningen af *cluster*-robuste standardfejl kan volde problemer i mindre datasæt. Specifikt når antallet af klynger er højt relativt til antallet af observationer. Foranlediget af størrelsen på datasættet opvejer fordelene ved at anvende *Cluster*-robuste robuste klart ulemporne, da de foruden at korrigere for heteroskedasticitet også afhjælper problemet med seriekorrelation.

#### FE.6: INGEN SERIEKORRELATION

Som nævnt i metodeafsnittet (3.1.1: FE.6) kan en Durbin-Watson test anvendes for at undersøge om der er autokorrelation af første grad, altså autokorrelation én tidsperiode tilbage. Eftersom at der allerede er anvendt robuste standardfejl og Durbin-Watson testen gyldighed afhænger af om modellen er strengt eksogen, hvilket på grund af brud på FE.4 ikke er tilfældet, så gennemføres testen ikke.

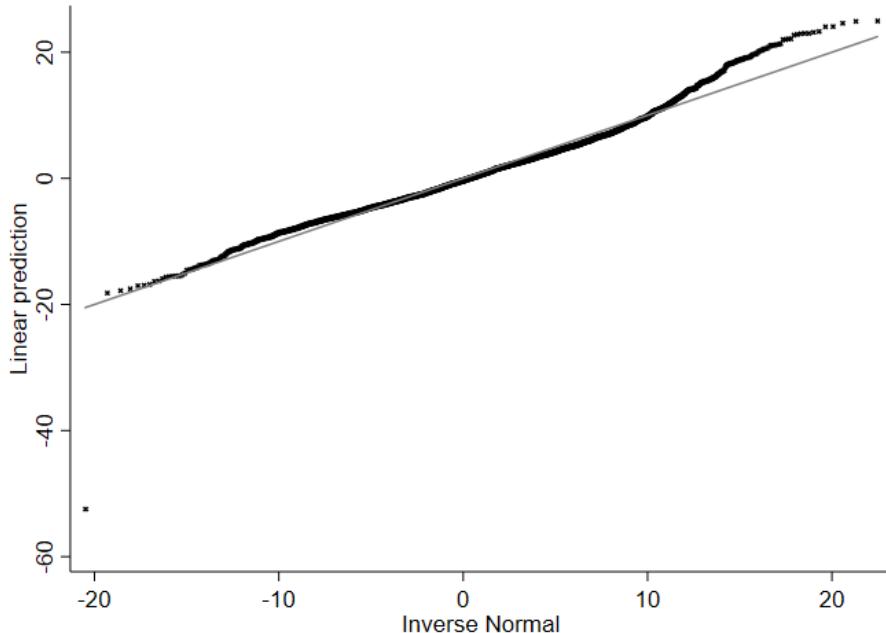
#### FE.7: NORMALITET AF FEJLLEDDET

Den sidste fixed effects forudsætning er noget indskrænkende, da den ved at være opfyldt medfører at FE.4-FE.6 ligeledes er opfyldt.

For at undersøge om fejlledet er normalfordelt, kan flere *goodness-of-fit*-tests for normalitet anvendes. Men da stikprøven er tilpas stor, vil selv små afvigelser fra normalitet, resultere i en signifikant teststørrelse og dermed insinuere at fejlleddet ikke er normalfordelt, derfor er en grafisk illustration mere informativ. Residualerne fra fixed effects modellen er derfor illustreret i et QQ-plot på næste side.

I QQ-plottet (figur 5) observeres tydeligt at det er i "halerne" at den største afvigelse fra normalitet er, hvor særligt én enkelt residual vægter meget tungt i den negative ende. Dog indebærer OLS-estimatorens asymptotiske egenskaber, at residualerne med god grund kan antages at tilnærmelsesvis være normalfordelte (Se afsnit 3.1.1: FE.7). Denne konklusion kan udledes af det store antal observationer (samtidig med en relativ kort periode) og medfører at der bør kunne foretages gyldig statistisk inferens, af både F- og t-teststatistikker, uagtet fordelingen af modellens residualer.

Figur 5 – QQ-plot over fixed effects residualer (model 1)



Kilde: Egen udarbejdelse i Stata, se linje: 305-309 i <Main.do>

#### 5.1.4 Opsummering

I forrige afsnit (5.1.3) blev der redegjort for fixed effects forudsætningerne FE.1-FE.7, med udgangspunkt i metodeafsnittets (5.1.1) beskrivelser heraf. For at opsummere er fixed effects modellens største udfordring bruddet på FE.4, som indikerer at modellen *ikke* er strengt eksogen og derfor hverken konsistent eller unbiased. RESET-testen indikerede at modellen lider af et misspecifikationsbias pga. manglen af 2. eller 3. ordens effekter, men det vil være naivt at tro at det kun er disse effekter som er ekskluderet. Meget få variabler i modellen er signifikante og på grund af den ekstremt lave forklaringsgrad (2,8%) medfører det at modellens praktiske værdi er ret begrænset.

I næste afsnit vil der gennemføres en Hausman-test, som kan anvendes til at vurdere konsistensens af en random effects model ift. den allerede kendte fixed effects model. Som der er beskrevet mere uddybende i afsnittet, så kræver Hausman-testen, for at være gyldig, at forudsætningerne for en fixed effects model er opfyldt. En gyldig Hausman-test kan påvise om den kritiske random effects forudsætning (RE.2) er opfyldt og dermed om random effects modellen skal foretrækkes frem for fixed effects modellen (Wooldridge, 2015).

**Tabel 9 – Fixed- & Random effects koefficienter**

VARIABLES	(4) FE, ROA	(5) RE, ROA	(4) - (5) Forskel
ln(Aktiver)	3,540*** (0,586)	4,223*** (0,350)	-0,689
ln(Alder)	-4,267*** (0,878)	-2,437*** (0,756)	-1,829
Finansiel Gearing	-0,622 (0,442)	0,964** (0,395)	-1,587
Udviklingsomkostninger (%)	-0,004*** (0,001)	-0,005*** (0,001)	0,001
Adm. Direktør aktionær?	-0,697 (0,819)	-1,538** (0,761)	0,841
Finansdirektør aktionær?	-3,862* (2,028)	-0,887 (1,763)	-2,975
1 = Hvis ingen ejer > 10%	0,005 (0,990)	1,196 (0,922)	-1,191
1 = Hvis de 10 største ejer samlet er >50%	-1,519 (0,958)	-2,326*** (0,890)	0,807
1 = Hvis største ejer er familie	0,847 (1,243)	3,179*** (1,095)	-2,332
1 = Hvis majoritetsejer er familie	0,947 (2,957)	2,771 (2,439)	-1,825
Spredningsforhold	-11,969** (4,967)	-13,217*** (4,705)	1,248
Forskell på største og næststørste ejer	0,169* (0,089)	0,252*** (0,081)	-0,082
Herfindahl-Hirschmann index	-1,379 (3,213)	8,534*** (2,677)	-9,913
Observationer	3.896	3.896	
R-squared	0,028	0,016	
Antal selskaber	1.238	1.238	
Standardfejl	Alm.	Alm.	

Note: Tabellen angiver koefficienter, signifikansniveauer samt forskellige for hhv. fixed effects og random effects modellerne, beregnet med almindelige (*ikke* robuste) standfejl.

## 5.2. Hausman Test

Som beskrevet i metodeafsnittet (3.3) så kan en Hausman-test anvendes for at vurdere konsistensen af en model (random effects) i forhold til en, allerede kendt, konsistent model (fixed effects). Validiteten af Hausman-testen afhænger derfor af at fixed effects modellen er konsistent, hvilket er tilfældet hvis fixed effects modellen opfylder forudsætningerne herfor.

Som nævnt i forrige afsnit (5.1) lider fixed effects modellen af misspecificationsbias og bryder således med FE.4, hvilket indebærer at modellen ikke er konsistent når  $n \rightarrow \infty$ . FE.4 er en afgørende og særdeles kritisk forudsætning og bruddet herpå indikerer at der er *omitted variable bias*, som ydermere gør at koeficienterne er upålidelige. Derfor kan det udledes at resultaterne af en Hausman test, ligeledes er upålidelige. I det elektroniske bilag <main.do> (linje 311 - 338) er der dog både udarbejdet en generel Hausman-test for fixed effects og random effects *uden* robuste standardfejl<sup>19</sup>, samt en Sargan-Hansen test<sup>20</sup> som følger fremgangsmåden foreslægt af Arellano (1993) og Wooldridge (2010), da der i Sargan-Hansen testen kan anvendes robuste standardfejl (Schhaffer & Stillman, 2019). Begge test forkaster nulhypotesen: Hausman:  $\chi^2(13) = 149,92$ ;  $p = 0,000$ , Sargan-Hansen:  $\chi^2(13) = 149,92$ ;  $p = 0,000$ ), og havde forudsætningerne for testen været opfyldt, ville konklusionen være at RE.2 ikke er opfyldt og random effects modellen derfor ikke er konsistent.

Eftersom forudsætningerne for Hausman-testen ikke opfyldt, er dens konklusionen ugyldig. Desuden bør værdien af en random effects model, ikke alene vurderes på baggrund af modellens statistiske signifikans, men også for dens praktiske egenskaber. Random effects modeller har den fordel at de tillader inklusion af periodekonstante variabler, hvilket er interessant hvis der er ingen eller meget lav variation i en models vigtige variabler. Det forhold at der er relativt lidt variation i ejerskabsvariablerne gør at random effects modellen kan være interessant at undersøge. Desuden kan der inddrages landeffekter, hvilket ligeledes er interessant for analysen. Af de to grunde, så vil der i næste afsnit beregnes seks forskellige random effects modeller, to for hvert performancemål: én som inkluderer samme variabler som fixed effects modellen og én som inkluderer dummy variabler for landene i datasættet:

- (6) ROA: Uden dummyvariabler for lande
- (7) ROA: Med dummyvariabler for lande
- (8) ROE: Uden dummyvariabler for lande
- (9) ROE: Med dummyvariabler for lande
- (10) Tobin's q: Uden dummyvariabler for lande, kun perioden år 2011-2017
- (11) Tobin's q: Med dummyvariabler for lande, kun perioden år 2011-2017

Det bør bemærkes at de to Tobin's q-modeller, pga. Orbis-opdateringen heromkring, kun har data fra år 2011 og derfor alene er estimeret for syvårsperioden fra år 2011-2017. Resultaterne for modellerne kan ses i tabellen på næste side:

---

<sup>19</sup> Hausman-testen følger ikke en  $\chi^2$ -fordeling, hvis den beregnes på baggrund af robuste standardfejl.

<sup>20</sup> Sargan-Hansen-testen er beregnet i Stata med den brugerprogrammerede kommando: *xtoverid*.

**Tabel 10 – Random effects modeller**

Variabler	(6) ROA	(7) ROA	(8) ROE	(9) ROE	(10) Tobin's q	(11) Tobin's q
ln(Aktiver)	4,223*** (0,568)	5,073*** (0,557)	8,626*** (1,352)	10,866*** (1,422)	-0,77*** (0,152)	-0,75*** (0,164)
ln(Alder)	-2,437*** (0,710)	-2,345*** (0,741)	-1,939 (1,598)	-1,417 (1,670)	0,105 (0,141)	0,057 (0,180)
Finansiel Gearing	0,964** (0,488)	0,008 (0,420)	5,428 (5,139)	2,569 (5,094)	0,186 (0,173)	0,224 (0,183)
Udviklingsom- kostninger (%)	-0,005 (0,003)	-0,004 (0,003)	-0,004 (0,004)	-0,003 (0,004)	-0,000 (0,000)	-0,000 (0,000)
Adm. Direktør ak- tionær?	-1,538** (0,724)	-1,852*** (0,713)	-1,399 (2,139)	-2,141 (2,118)	-0,165 (0,126)	-0,143 (0,127)
Finansdirektør ak- tionær?	-0,887 (1,694)	-1,886 (1,575)	-5,344 (5,955)	-8,188 (5,587)	-0,097 (0,243)	-0,130 (0,245)
1 = Hvis ingen ejer > 10%	1,196 (1,199)	1,029 (1,178)	2,379 (2,930)	2,315 (2,915)	-0,347** (0,155)	-0,356** (0,156)
1 = Hvis de 10 største ejer samlet er >50%	-2,326** (1,000)	-1,795* (0,966)	-5,759* (3,209)	-4,596 (3,119)	0,343** (0,145)	0,322** (0,143)
1 = Hvis største ejer er familie	3,179** (1,239)	1,808 (1,216)	5,290* (3,011)	1,751 (2,852)	-0,305* (0,179)	-0,198 (0,167)
1 = Hvis majoritet- sejer er familie	2,771 (2,291)	2,595 (2,050)	11,668 (7,099)	10,660* (6,243)	0,633 (0,579)	0,589 (0,592)
Spredningsforhold	-13,22*** (4,433)	-12,34*** (4,238)	-24,198** (9,764)	-22,602** (9,462)	-0,472 (0,562)	-0,543 (0,570)
Forskel på største og næststørste ejer	0,252*** (0,067)	0,202*** (0,062)	0,519*** (0,142)	0,410*** (0,134)	-0,001 (0,009)	0,002 (0,009)
Herfindahl- Hirschmann index	8,534*** (2,436)	4,672** (2,303)	20,715*** (6,183)	12,882** (5,932)	-0,768* (0,404)	-0,639 (0,442)
Observationer	3.896	3.896	3.896	3.896	3.116	3.116
R-squared	0,0161	0,0208	0,0133	0,0162	0,0718	0,0717
Antal selskaber	1.238	1.238	1.238	1.238	851	851
Første år	2008	2008	2008	2008	2010	2010
Lande Dummies	-	Ja	-	Ja	-	Ja

Robuste standardfejl i parentes  
 \*\*\* p<0,01, \*\* p<0,05, \* p<0,1

### 5.3. Random effects modellen

Udgangspunktet for random effects modellen er en model bestående af samme variabler som i fixed effects modellen (1). I tabel 9 er koefficienterne for de to modeller illustreret og det kan tydeligt observeres at der er væsentlige forskelle, mellem de to modeller. F.eks. ændrer koefficienten for finansiel gearing sig fra -0,622 (FE) til 0,964 (RE). Dog ændrer de store diskrepanser sig, når der i tabel 10 kontrolleres for lande (model 7) og bliver derved mindre markante.

Det er netop inklusionen af de periodekonstante landeffekter som gør det interessant at opstille random effects modellen, derfor vil udgangspunktet for dette afsnit være model 7 som grundlæggende er samme model som FE model 1, bare med dummy variabler for selskaberne lande, samt naturligvis estimeret med random effects.

I næste afsnit vil random effects forudsætningerne for model 7 vurderes:

#### 5.3.2 Forudsætninger

##### **FE.1 & FE.2: LINEARITET OG TILFÆLDIG UDVALGT STIKPRØVE**

I overensstemmelse med FE.1, kræves at model (7) kan opstilles som:

$$ROA_{it} = X_{it} \cdot \beta_i + \alpha_i + u_{it}: \quad t = 1, \dots, T; \quad i = 1, \dots, N \quad (5.2)$$

Hvor  $X_{it}$  er en  $1 \times 32$  matrix for de forklarende variabler, mens  $\beta$  er en  $32 \times 1$  matrix for variablene antal parametre.  $t$  går fra 1 til 10, og modellen estimeres på  $i = 1,238$  selskaber. Ligesom for fixed effects modellen, så tager alle parametrene en lineær form og FE.1 må vurderes som værende opfyldt.

I forhold til FE.2, så beregnes random effects modellen på samme datasæt som fixed effects modellen og derfor gør de samme overvejelser omkring tilfældig udvælgelse sig gældende (Se afsnit 5.1.3: FE.2).

##### **RE.1: INGEN PERFEKT MULTIKOLLINEARITET**

Da random effects modellen tillader inklusion af periodekonstante variabler, erstatter RE.1, FE.3, så der kun forudsættes at modellen ingen perfekt multikollinearitet har. Ligesom i afsnittet for FE forudsætningerne (5.1.3:FE.3), så anvendes variansinflationsfaktorer til at bedømme graden af multikollinearitet. I tabel 11 nedenfor, illustreres VIF for udvalgte variabler<sup>21</sup>. Ligesom for FE-modellen iagttages store VIF-værdier for særlig størrelse og alder variablerne, men ligesom tidligere diskuteret, så skyldes de høje VIF sandsynligvis de store forskelle i variablene forklaringskraft og de høje variansinflationsfaktorer giver ikke alene noget umiddelbart argument for at fjerne enkelte variabler.

---

<sup>21</sup> Der refereres til linje 381-421 i det elektroniske bilag <*main.do*> for liste over VIF for alle variabler.

**Tabel 11 – VIF for variablerne i random effects modellen**

Variabler	VIF
ln(Aktiver)	40,16
ln(Alder)	35,49
Finansiel Gearing	1,15
Udviklingsomkostninger (%)	1,01
Spredningsforhold	10,79
Forskel på største og næststørste ejer	11,19

Note: Tabellen angiver variansinflationsfaktoren for udvalgte variabler i modellen, se linje 381-421 i det elektroniske bilag <Main.do> for fuld liste over VIF.

#### **FE.4: NULFORVENTNING FOR FEJLEDDDET**

Forudsætningen om streng eksogenitet gælder også for at random effects modellen er *unbiased*. Som forventet indikerer Ramsey's RESET test at der er simultant signifikante 2. og 3. ordens effekter i random effects modellen ( $\chi^2(2) = 2727,24; p = 0,000$ ) og random effects modellen er ligesom FE model 1 *ikke* strengt eksogen.

#### **RE.3 & FE.6: HOMOSKEDASTICITET OG INGEN SERIEKORRELATION**

Da de uforklarede selskabsspecifikke effekter  $a_i$ , ikke fjernes i en random effects model, så udvides homoskedasticitetsforudsætningen til også at gælde disse effekter, betinget af alle øvrige variabler. Ligesom for FE modellen, så korrigeres der både for heteroskedasticitet og seriekorrelation, gennem anvendelsen af *cluster*-robuste standardfejl. Implikationerne af at anvende disse robuste standardfejl er diskuteret i afsnit 5.1.3 (FE.5).

#### **RE.2: INGEN KORRELATION MELLEM $a_i$ OG $X_i$**

Den afgørende random effects forudsætning, RE.2, er uafhængighed mellem de selskabs-specifikke effekter  $a_i$  og de øvrige forklarende variabler. Som diskuteret i afsnit 5.2, så er der konsensus om Hausman testen som indikator for om RE.2 er opfyldt. Det betinges dog af om FE modellens konsistens og derfor implicit at FE-forudsætningerne er opfyldte. Hausman-testen indikerede, som diskuteret i afsnit 5.2, at fixed effects modellen bør vælges, men resultatet er dog begrænset af at FE forudsætningerne ikke er opfyldt.

#### **5.3.3 Opsummering**

Konklusion for random effects model 7, er ligesom for fixed effects modellen, at den pålidelighed er begrænset af dens mangler opfyldelse af forudsætningerne. Modellen lider ligesom FE model 1 af misspecifikationsbias og yderligere så indikerer Hausman testen at den kritiske forudsætning RE.2, heller ikke er opfyldt. Hvorvidt FE eller RE er den bedste model, afgøres dog ikke kun af Hausman testresultatet, men i lige så høj grad på baggrund af nogle praktiske argumenter. I næste kapitel vil begge modellers overordnede resultater, og praktiske relevans, vurderes.

## 6. Diskussion

Indværende kapitel har til hensigt at vurdere afhandlingens og analysens væsentligste kritikpunkter, samt sammenligne analysens resultater med lignende studier. I forrige kapitel blev de forskellige statistiske modellerne vurderet på baggrund af forudsætningerne nævnt i metodekapitlet (3). Dette kapitel vil begynde, hvor forrige sluttede, med at vurdere resultaternes signifikans, både statistisk og praktisk.

### 6.1. Resultaternes validitet

I dette afsnit vil de resultaterne fra forrige afsnit vurderes i et bredere perspektiv. Både fixed effects- og random effects modellen, lider af en meget lav forklaringsgrad på mellem 1,5-3 % for ROA. Den mest forklarende model er Tobin's q, random effects modellen inkl. dummy variabler for lande (model 11), som med en  $R^2$  forklaringsgrad på 7,2 % stadig ligger noget under resultaterne fra tidligere studier (Se afsnit 6.1.2).

Foruden den lave forklaringskraft, så lider modellerne grundlæggende af at de interessante ejerskabsvariabler, kun sporadisk er statistisk signifikante. De to variabler som oftest beregnes som statistisk signifikante, er:  $E1ME2$  og  $disp$ , som begge er mål for fordelingen af ejerskabet mellem de to største ejere:

$$E1ME2 = Andel_1 - Andel_2 \quad (6.1)$$

$$disp = \frac{E1ME2}{Andel_1 + Andel_2} \quad (6.2)$$

Hvor  $Andel_1$  er ejerandelen for den største ejer og  $Andel_2$  er andelen for den næststørste ejer og  $E1ME2$  er f.eks. 10 %, hvis den største og næststørste ejer, ejer hhv. 30 og 20 % af virksomheden og  $disp$  vil i dette tilfælde være 20%.

I betragtning af at modellerne er *biased*, bør der udvises forsigtighed, når resultaterne fortolkes, men det forhold at netop de to variabler, mere eller mindre konsekvent er signifikante, tyder på at type-2 agentproblemer (Majoritet- og minoritetskonflikter, se teoriansnit 2.2.1) kan have betydning for selskabernes performance.

Som nævnt i både resultatkapitlet og i afsnittet ovenfor, så er gyldigheden af resultaterne væsentligt begrænset af forudsætningerne, for modellerne. I resultatafsnittet påvises at fixed effects modellen *ikke* er strengt eksogen, hvilket indebærer at FE modellen er misspecified. Ramsey's RESET test påviste at modellen mangler 2. eller 3. ordens effekter, af de allerede inkluderede variabler. Dette forhold kan give et præg om hvordan modellen kan forbedres, men det er usandsynligt at det er udeladelsen af disse effekter, som alene forårsager forudsætningsbruddet på FE.4.

Random effects modellen lider, foruden de samme problemer som fixed effects modellen, af at det selskabsspecifikke fejlled er korreleret med modellens øvrige variable. Dette brud, på den kritiske random effects forudsætning, blev påvist med en Hausman-test, hvis gyldighed dog begrænses af at FE modellens forudsætninger ikke er opfyldt.

Uanset om den kritiske random effects forudsætning kan anses som opfyldt, eller ej, så lider random effects modellen – ligesom fixed effects modellen – af ikke at være strengt eksogen. Begge statistiske modeller lider af *bias*, og hverken OLS- eller GLS-estimatorerne er konsistente, da forudsætningerne ikke er opfyldt. Forudsætningen for at OLS er konsistent, er svag eksogenitet, men som diskuteret i resultatafsnittet (5.1.3), så kan svag eksogenitet ikke udledes af testen.

Selvom resultaternes gyldighed ikke kan retfærdiggøres, så er det stadig interessant at undersøge hvor sensitive koefficienterne er, og på den måde vurdere resultaternes robusthed. Der er allerede observeret store koefficient-forskelle imellem random effects og fixed effects modellerne, men i næste afsnit (6.1.1) vil yderligere modeller opstilles, med henblik på at undersøge kritikpunkterne, fra de foregående afsnit. Foruden gentagelsen af FE model 1, så er følgende modeller er opstillet i tabel 12 (På side 44):

- (13) Fixed effects model, uden udviklingsomkostnings-variablen.
- (14) Fixed effects model, kun amerikanske selskaber
- (15) Fixed effects model, kun selskaber med data for alle 10 år.

Alle modellerne er med ROA som afhængig variabel.

#### 6.1.1 Sensitivitetsanalyser

De ovennævnte modeller (13, 14 og 15) kan anvendes som sammenligningsgrundlag for den oprindelige fixed effects model (Se tabel 12). I resultatafsnittet 5.1.2, nævnes at det bør undersøges hvilken påvirkning udeladelsen af udviklingsomkostningerne, vil have. Variablen mangler 968 observationer, som udelades hvis den medtages. Desuden er den både statistisk insignifikant samt har en meget lav koefficient uden praktisk betydning. Hvis variablen udelades, og der dermed inkluderes yderligere 968 observationer, medfører det at flere koefficienter ændrer sig markant, bl.a. mere end halveres koefficienten for *ln(Aktiver)*.

De to øvrige modeller er mere restriktive, model 13 inkluderer kun amerikanske selskaber (228 selskaber) og model 14 inkluderer kun de 28 selskaber, som har observationer for alle årene. Koefficienterne i USA-modellen er særligt interessante, da de for ejerskabsvariablerne er meget forskellige fra den overordnede model (1). F.eks. ses at koefficienten for, familie/majoritetsejer-dummyvariablen, er ret stor (om end statistisk insignifikant) og indikerer at hvis selskabet er majoritetsejet af en familie, så bør ROA være 12,83 % højere.

Validiteten af disse sensitivitetsanalyser afhænger – lige som de øvrige modeller i afhandlingen – af datasættets kvalitet. Datasættets størrelse, som burde være dens fordel, er også dets største akilleshæl, da netop størrelsen og kompleksiteten af datasættet kan være et problem. Datasættet er konstrueret på grundlag af mange års data fra Orbis, og hvis de løbende opdateringer af databasen har medført at variablerne ikke er udregnet tilstrækkelig ensartet, så er der risiko for at variablernes variationen er unødig høj og deraf har negative implikationer for variablernes signifikansniveau og modellens forklaringskraft.

*Tabel 12 – Sensitivitetsanalyser*

Variable	(1)	(12) Uden RDPct	(13) Kun USD	(14) Balanceret
ln(Aktiver)	3,540*** (1,100)	1,409* (0,838)	5,860*** (1,919)	-1,938 (1,644)
ln(Alder)	-4,27*** (0,808)	-2,783*** (0,636)	-2,268 (3,331)	-3,868*** (0,745)
Finansiel Gearing	-0,622 (0,477)	-0,673** (0,320)	-1,432 (1,256)	-1,458 (0,949)
Udviklingsomkostninger (%)	-0,004 (0,003)	- -	-0,004** (0,002)	-27,59*** (7,483)
Adm. Direktør aktionær?	-0,697 (0,704)	-1,267** (0,587)	-0,205 (3,960)	-3,012** (1,466)
Finansdirektør aktionær?	-3,862* (2,073)	-3,002* (1,672)	-3,177 (11,326)	-7,040* (3,704)
1 = Hvis ingen ejer > 10%	0,005 (1,290)	0,376 (1,113)	-2,424 (2,894)	-2,147 (2,344)
1 = Hvis de 10 største ejer samlet er >50%	-1,519 (1,103)	-1,652* (0,878)	-3,780 (3,819)	2,754* (1,613)
1 = Hvis største ejer er familie	0,847 (1,417)	0,508 (1,237)	2,149 (4,122)	-0,621 (1,997)
1 = Hvis majoritetsejer er familie	0,947 (2,384)	1,938 (2,310)	12,827 (18,873)	8,691** (3,822)
Spredningsforhold	-11,97*** (4,175)	-11,02*** (3,593)	-14,62 (24,731)	10,08 (11,419)
Forskel på største og næststørste ejer	0,169** (0,066)	0,170*** (0,053)	0,129 (0,593)	-0,228 (0,209)
Herfindahl-Hirschmann index	-1,379 (2,650)	0,350 (1,999)	-14,625 (25,564)	-16,42*** (4,449)
Observationer	3.896	4.864	754	201
Antal selskaber	1.238	1.483	228	23
R-squared	0,028	0,015	0,034	0,381
Standardfejl	Robuste	Robuste	Ukorrigerede	Ukorrigerede

Standardfejl i parentes, robuste eller ukorrigerede.

\*\*\* p&lt;0,01, \*\* p&lt;0,05, \* p&lt;0,1

## 6.2. Sammenligning med lignende studier

De foregående resultat- og diskussionsafsnit har haft fokus på at forklare resultaternes begrænsninger, statistiske signifikans samt robustheden indenfor datasættet. I dette afsnit vil resultaterne sammenlignes med lignende studier.

Igenom afhandlingen er Leaven & Levine samt Anderson & Reeb nævnt som inspirationskilder til metode- og variabelvalg. De to studier kendtegnes ved at undersøge sammenhængen mellem ejerskab og virksomhedsperformance: Leaven & Levine (2007) tager udgangspunkt i effekten af ejerskabskoncentrationen, mens Anderson & Reeb (2003) fokuserer på effekten fra familieejerskab. De primære resultater, fra de to studier, er inkluderet som bilag (Leaven & Levine: bilag 6.1 og Anderson & Reeb: bilag 6.2).

Generelt for begge studier er at de, på tværs af modeller, har betydeligt højere forklaringsgrader og flere signifikante variabler, end indeværende analyse, og det på trods af at flere variabler er defineret ens.

Derudover tager koefficientestimaterne, i begge studier, nogle mere afdæmpede værdier. Eksempelvis observeres en effekt, i model 1 (FE, ROA), på hhv. 3,5 % og -4,27 % for de to kontrolvariablerne: *ln(Aktiver)* og *ln(Alder)*. I Anderson & Reeb (2003) er effekterne herfra kun hhv. 0,5 % og -2,7 % og det på trods af at Anderson & Reeb, ligesom indeværende afhandling, anvender *ln*-transformerede variabler for totale aktiver samt alder.

Hvis der laves en overordnet sammenligning med resultaterne fra de to nævnte studier, så er udfaldet ikke positivt for gyldigheden af resultaterne, i indeværende afhandling. Udfordringerne, er som tidligere nævnt bl.a.: ukongruent data og forudsætningsbrud, som må vurderes som groft invaliderende for analysens resultater. Derfor er det nødvendigt med yderligere analyse, hvis at der skal påvises en kausal sammenhæng mellem virksomheders ejerskabsstruktur og deres rentabilitet.

### 6.2.1 Oplæg til videre analyse

Efter at have slået fast at resultaterne i indeværende afhandling, i bedste fald, skal tolkes med stor forsigtighed, så vil dette afsnit fokusere på hvordan fremtidige studier kan forbedres. Det vigtigste forbedringspunkt er at sikre bedre konsistens i datasættet. Variablerne bør, så vidt muligt, beregnes ensartet, for hele perioden. I Anderson & Reeb (2003) er denne problemstilling håndteret med antagelsen om, at ejerskabet er uændret, over perioden. Det kan virke som en stor antagelse, men som det blev illustreret i afsnit 5.1.3 (FE.3), så er det reelt kun få selskaber som oplever større ændringer i ejerskabet, over 10-års perioden.

En yderligere forbedring, som også vedrører datasættet, er at der bør inddrages flere variabler, for derigennem at afhjælpe modellernes misspecifikationsproblemer. Det vil være særlig interessant at inddrage *lagged effects* for udviklingsomkostningerne, da deres langsigtede påvirkning af rentabiliteten, i så fald, bedre vil kunne inkluderes. Desuden bør variablerne for direktørernes ejerskab kunne uddybes, f.eks. ved at inkludere variabler for optimizationsbaseret aflønning samt størrelsen af ejerskabet i procent.

## 7. Konklusion

---

Afhandlingens primære omdrejningspunkt har været den empiriske analyse af sammenhængen mellem ejerskabsstruktur og rentabilitet, indenfor den globale farmaceutiske industri. Den overordnede konklusion, er at der *ikke* har kunnet påvises nogen kausal sammenhæng mellem ejerskabsstruktur og rentabilitet. Denne konklusion udledes på baggrund af de statistiske modellers lave forklaringsgrader og sandsynlige *bias*.

Selvom at det ikke været muligt at påvise nogen kausal sammenhæng mellem børsnoterede farmaceutiske selskaber og deres performance, og derved besvare opgavens centrale forskningsspørgsmål, så der undervejs redegjort for problemformuleringens delspørgsmål. Disse svar, vil kort opsummeres nedenfor (Der refereres til delspørgsmål 1-3 i afsnit 1.2.1):

1. Med udgangspunkt i agent- og kontraktteori bør ejerskab og ejerskabsstrukturer karakteriseres ud fra deres evne til at begrænse potentielle omkostninger, som opstår af interessekonflikter mellem ejere og ledere. Begge teorier afstedkommer at høj ejerskabskoncentration, og i særdeleshed lederejerskab er et ønskværdig ideal og bør pga. afhjælplingen af type-1 agentproblemer, manifestere sig i bedre performance.
2. I metode- og resultatafsnittet er der redegjort for de generelle overvejelser, som bør gøres om valget mellem fixed effects og random effects. Random effects modellen er betinget af én meget restriktiv forudsætning, om at det selskabsspecifikke fejlled skal være ukorreleret med de øvrige variabler (RE.2). Denne forudsætning kan testes med en Hausman test, hvis validitet dog afhænger af om de generelle FE forudsætninger er opfyldt og desuden bør valget også foretages på baggrund af nogle praktiske overvejelser, da random effects har den fordel, at der kan inddrages tidsinvariante variabler i modellen.
3. Som diskuteret i diskussionsafsnittet (6.2), så er der væsentlige diskrepanser mellem indeverende afhandling og resultaterne i lignende studier. Dette forhold skal dog ses i lyset af udfordringerne for afhandlingens analyse. Disse udfordringer vil sandsynligvis kunne udbedres, hvis der dataproblemerne kunne afhjælpes, f.eks. hvis der kan findes adgang til en bedre datakilde.

Den overordnede konklusion for analysen er derfor at der *ikke* har kunnet påvises nogen praktisk eller statistisk signifikant kausal sammenhæng mellem virksomhedernes ejerskabsstruktur og deres rentabilitet. En sammenhæng som afhandlingens teoretiske udgangspunkt ellers indikerer at der bør være. Konklusionen er dog betinget af problemer med det underliggende datasæt, som formentlig ikke har været tilstrækkelig kongruent på tværs af virksomhederne og år. Den manglende ensartethed har formentligt resulteret i en del uhensigtsmæssig og uforklaret variation i både fixed effects og random effects modellerne.

## 8. Bibliografi

---

- Anderson, R. C., & Reeb, D. M. (2003). Founding-Family Ownership and Firm Performance: Evidence from the S&P 500. *58*(3), 1301-1328. doi:doi:10.1111/1540-6261.00567
- Andres, C. (2008). Large shareholders and firm performance—An empirical examination of founding-family ownership. *Journal of corporate finance*, *14*(4), 431-445.
- Arellano, M. (1993). On the testing of correlated effects with panel data. *Journal of econometrics*, *59*(1-2), 87-97.
- Bharadwaj, A. S., Bharadwaj, S. G., & Konsynski, B. R. (1999). Information technology effects on firm performance as measured by Tobin's q. *Management science*, *45*(7), 1008-1024.
- Bodie, Z., Kane, A., & Marcus, A. J. (2014). *Investments* (10. ed.). Maidenhead: McGraw-Hill Education.
- Cadbury, A. (1992). *Report of the Committee on the Financial Aspects of Corporate Governance* (C. o. t. F. A. o. C. Governance Ed.): Gee.
- Clark, T. S., & Linzer, D. A. (2015). Should I use fixed or random effects? *Political Science Research and Methods*, *3*(2), 399-408.
- DiMasi, J. A. (1995). Trends in drug development costs, times, and risks. *Drug Information Journal*, *29*(2), 375-384.
- NACE Rev. 2: Statistical classification of economic activities in the European Community, (2008).
- Faccio, M., & Lang, L. H. (2002). The ultimate ownership of Western European corporations. *Journal of financial economics*, *65*(3), 365-395.
- Gould, W. (2019). Interpreting the intercept in the fixed-effects model.
- Grossman, S. J., & Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration. *Journal of political economy*, *94*(4), 691-719.
- Guba, E. G. (1990). The Paradigm Dialog. 17-27.
- Gyldendal. (2017). Corporate governance. Retrieved from [http://denstoredanske.dk/Samfund,\\_jura\\_og\\_politik/Jura/Selskabsret/corporate\\_governance](http://denstoredanske.dk/Samfund,_jura_og_politik/Jura/Selskabsret/corporate_governance)
- Hart, O. (2016). *Incomplete Contracts and Control*. Nobel Lecture.
- Hayes, A. (2019). Herfindahl-Hirschman Index (HHI). Retrieved from <https://www.investopedia.com/terms/h/hhi.asp>
- Holderness, C. G. (2003). A survey of blockholders and corporate control. *Economic policy review*, *9*(1).
- Holderness, C. G., & Sheehan, D. P. (1988). The role of majority shareholders in publicly held corporations: An exploratory analysis. *Journal of financial economics*, *20*, 317-346.
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of financial economics*, *3*(4), 305-360.
- Laeven, L., & Levine, R. (2007). Complex ownership structures and corporate valuations. *The Review of Financial Studies*, *21*(2), 579-604.
- O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & quantity*, *41*(5), 673-690.
- Schhaffer, M. E., & Stillman, S. (2019). help for xtoverid. Retrieved from <http://fmwww.bc.edu/repec/bocode/x/xtoverid.html>

- Shleifer, A., & Vishny, R. W. (1997). A survey of corporate governance. *The journal of finance*, 52(2), 737-783.
- Short, H. (1994). Ownership, control, financial structure and the performance of firms. *Journal of economic surveys*, 8(3), 203-249.
- Smith, A. (1776). *An Inquiry Into the Nature and Causes of the Wealth of Nations*.
- StataCorp. (2013a). Stata Manual 13, xt. rhausman. Retrieved from <https://www.stata.com/manuals13/xt.pdf>
- StataCorp. (2013b). Stata Manual 13, xt. Retrieved from <https://www.stata.com/manuals13/xt.pdf>
- Thomsen, S., & Conyon, M. (2012). *Corporate Governance: Mechanisms and Systems*: McGraw-Hill.
- Thomsen, S., & Pedersen, T. (2000). Ownership structure and economic performance in the largest European companies. *Strategic Management Journal*, 21(6), 689-705.
- Tobin, J., & Brainard, W. C. (1976). *Asset markets and the cost of capital*. Retrieved from Type of Percentage: Direct and Total Ownership. (2019).
- Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*: MIT Press.
- Wooldridge, J. M. (2015). *Introductory Econometrics: A Modern Approach*: Cengage Learning.

Forsidebilledet er stillet til rådighed af Pixabay gennem pexels.com og anvendes i overensstemmelse med betingelserne opelistet i Creative Commons Zero (CC0).

# Bilag

---

<b>1.</b>	<b><u>OVERSIGT OVER ELEKTRONISKE BILAG.....</u></b>	<b>49</b>
<b>2.</b>	<b><u>BESKRIVELSE AF ORBIS DATASÆTTENE .....</u></b>	<b>50</b>
2.1.	ORBIS HISTORICAL: SØGEKRITERIER .....	50
2.2.	ORBIS HISTORICAL: VARIABLE (2011 – 2017).....	50
2.3.	ORBIS HISTORICAL: VARIABLE (2008 – 2010).....	51
<b>3.</b>	<b><u>BESKRIVELSE AF DET SAMLEDE DATASÆT.....</u></b>	<b>52</b>
3.1.	KONSOLIDERET DATASÆT .....	52
3.2.	EKSTRA ORBIS-VARIABLER .....	53
3.3.	BEREGNEDE VARIABLER.....	53
3.4.	BEREGNEDE DUMMY-VARIABLER.....	53
3.5.	FORDELINGEN AF OBSERVATIONER PÅ LANDE OG ÅR. ....	54
<b>4.</b>	<b><u>BILAG TIL ANALYSEN.....</u></b>	<b>55</b>
4.1.	FIXED EFFECTS MODEL: ROE OG $Q$ .....	55
4.2.	FORDELING AF SELSKABER PER ANTAL OBSERVATIONSÅR.....	56
4.3.	K-DIAGRAM, NORMALITET AF FIXED EFFECTS FEJLLED .....	56
4.4.	KORRELATIONSMATRIX OVER MODELLENS VARIABLER .....	57
<b>5.</b>	<b><u>UDKLIP FRA ELEKTRONISKE BILAG.....</u></b>	<b>58</b>
5.1.	LINJE 100-140 FRA < <i>VARLIST.DO</i> >.....	58
5.2.	LINJE 41-50 FRA < <i>MAIN.DO</i> > .....	58
5.3.	LINJE 175-190 FRA < <i>MAIN.DO</i> > .....	59
5.4.	LINJE 305-309 FRA < <i>MAIN.DO</i> > .....	59
5.5.	LINJE 311-338 FRA < <i>MAIN.DO</i> > .....	59
5.6.	LINJE 381-421 FRA < <i>MAIN.DO</i> > .....	60
<b>6.</b>	<b><u>RESULTATER FRA TIDLIGERE STUDIER.....</u></b>	<b>61</b>
6.1.	LEAVEN & LEVINE (2007): TABLE 5 .....	61
6.1.	ANDERSON & REEB (2003): TABLE V .....	62

## 1. Oversigt over elektroniske bilag

---

Vedhæftet i *BscEjerskab.zip* findes følgende bilag:

Navn	Format	Beskrivelse
OrbisOwnership	.pdf	Sekundær kilde: Orbis beskrivelser er ejerskabsvariablerne: <i>Direct- og Total Ownership</i> .
VBA	.pdf	Visual Basic Application kode fra <i>Makro.xls</i>
Main	.do	Stata Do-fil som indeholder alle kommandoer anvendt til databehandling, analyse og eksport af tabellen. Beskrivelser er angivet i filen.
Main	.pdf	Tekstfil over koden i <i>Main.do</i> .
Varlist	.do	Stata Do-fil indeholdende navne og labels til variablerne fra <i>ConsolidatedDataset.csv</i> . Do-filen anvendes i <i>Main.do</i> . Beskrivelser er angivet i filen.
Varlist	.pdf	Tekstfil over koden i <i>Varlist.do</i> .
ExVar	.dta	Stata datasæt indeholdende BvD Id-kode (BvD) og tilhørende landekode (countryisocode), er hentet fra Orbis som csv-fil og konverteret i Stata.
LIST	.csv	Kommasepareret fil, til import i Orbis. Filen er generet i Stata, for at kunne hente manglende lande ISO-koder.
ConsolidatedDataset	.csv	Kommasepareret fil, til import i Stata. Filen er generet af makroen i <i>Makro.xls</i> arket.
Makro	.xls	Excel-arbejdsmappe med Visual Basic for Application makro-knapper til konsolidering af datasæt fra Orbis.
2008	.xlsx	Konverteret datasæt, Orbis Historical 2008
2009	.xlsx	Konverteret datasæt, Orbis Historical 2009
2010	.xlsx	Konverteret datasæt, Orbis Historical 2010
2011	.xlsx	Datasæt, Orbis Historical 2011
2012	.xlsx	Datasæt, Orbis Historical 2012
2013	.xlsx	Datasæt, Orbis Historical 2013
2014	.xlsx	Datasæt, Orbis Historical 2014
2015	.xlsx	Datasæt, Orbis Historical 2015
2016	.xlsx	Datasæt, Orbis Historical 2016
2017	.xlsx	Datasæt, Orbis Historical 2017

## 2. Beskrivelse af Orbis datasættene

---

### 2.1. Orbis Historical: Søgekriterier

1. *Alle aktive selskaber i databasen*
2. *Farmaceutiske Selskaber (NACE Rev. 2: 21)*
3. *Børsnoterede selskaber*

### 2.2. Orbis Historical: Variable (2011 – 2017)

1. *Company Name*
2. *BvD ID number*
3. *ISIN Number*
4. *Date of Incorporation*
5. *Last available year*
6. *Market Capitalisation\**
7. *Operating Revenue\**
8. *Net Income\**
9. *Depreciation & Amortization\**
10. *EBITDA\**
11. *Total Shareholders Equity\**
12. *Total Assets\**
13. *Research & Development expenses\**
14. *Shareholder – Name\*\**
15. *Shareholder – BvD ID number\*\**
16. *Shareholder – Country ISO code\*\**
17. *Shareholder – Type\*\**
18. *Shareholder Direct % \*\**
19. *Shareholder Total % \*\**
20. *DM Full Name\*\*\**
21. *DM Individual or Company\*\*\**
22. *DM Type of role\*\*\**
23. *DM Also a shareholder\*\*\**
24. *Total Liabilities and Debt\**
25. *Net Debt\**
26. *ROE using P/L before tax\**
27. *Changes in Working Capital\**
28. *Capital Expenditures\**

\* Seneste regnskabsår ( $t$ ) og to år tilbage ( $t - 2$ )

\*\* Ny kolonne for hver ejer med ejerskab > 5 %.

\*\*\* Ny kolonne for hver CEO/CFO med ejerskab i virksomheden.

### 2.3. Orbis Historical: Variable (2008 – 2010)

1. *Company Name*
2. *BvD ID number*
3. *ISIN Number*
4. *Date of Incorporation*
5. *Last available year*
- 6.
7. *Operating Revenue\**
8. *Net Income\**
9. *Depreciation & Amortization\**
10. *EBITDA\**
11. *Total Shareholders Equity\**
12. *Total Assets\**
13. *Research and Development Expenses\**
14. *Shareholder – Name\*\**
15. *Shareholder – BvD ID number\*\**
16. *Shareholder – Country ISO code\*\**
17. *Shareholder – Type\*\**
18. *Shareholder Direct % \*\**
19. *Shareholder Total % \*\**
- 20.
- 21.
- 22.
- 23.
- 24.
25. *Net Debt\**
26. *ROE using P/L before tax\**
27. *Changes in Working Capital\**
- 28.

\* Seneste regnskabsår ( $t$ ) og to år tilbage ( $t - 2$ )

\*\* Ny kolonne for hver ejer med ejerskab > 5 %.

Da Net Debt og Change in Working Capital ikke er variabler i Orbis i perioden før 2012, så er disse beregnet som hhv.:

$$\text{Net Debt} = \text{Loans (Short Term)} + \text{Long Term Debt} - \text{Cash and Cash Equivalents}$$

$$\text{Changes in Working Capital} = \text{Working Capital}_t - \text{Working Capital}_{t-1}$$

### 3. Beskrivelse af det samlede datasæt

---

#### 3.1. Konsolideret datasæt

No.	Fork.	Beskrivelse
1	Id	Selskabets Id-nummer i datasættet
2	Navn	Selskabsnavn
3	BvD	Bureau van Dijk Id-nummer
4	ISIN	Selskabets ISIN-nummer
5	Stift	År for selskabsstiftelse.
6	År	Anvendt regnskabsår
7	MV	Markeds værdi
8	Oms	Omsætning
9	Net	Nettoindkomst
10	Afs	Afskrivninger (Depreciation & Amortization)
11	EBITDA	EBITDA, Orbis-beregnet
12	Egen	Bogført egenkapital
13	Aktiv	Totale Aktiver
14	RD	Udviklingsomkostninger (Research & Development)
15	Passiv	Totale Passiver (Inkl. Gæld)
16	Gæld	Nettogaeld
17	ROE	ROE, Orbis-beregnet
18	ArbKap	Ændringer i arbejdskapitalen år-til-år
19	CapEx	Anlægsinvesteringer (Capital Expenditure)
20	CEO	Dummy: 1 = Adm. direktør er aktionær, 0 = hvis ikke
21	CFO	Dummy: 1 = Finansdirektør er aktionær, 0 = hvis ikke
22 -31	EjerYPct	Angiver ejerskabsprocenten af ejer Y
32 -41	EjerYD	Dummy: Angiver om ejerskabet for ejer Y, er direkte (1) eller total (0)
42 -51	EjerYType	Angiver typen af ejer Y (Fra Orbis)
52 -61	EjerYId	Angiver BvD Id-nummer for ejer Y

\* EjerYPct, EjerYD og EjerYId, er kun inkluderet for ejere med enten et direkte eller indirekte ejerskab over 10 %, og kun de 10 største.

### 3.2. Ekstra Orbis-variabler

No.	Fork.	Beskrivelse
62	Land	Angiver selskabets landekode (ISO)
63	Land2	Angiver selskabets landekode (ISO) for alle lande hvor: $n > 100$

### 3.3. Beregnede Variabler

No.	Fork.	Beregnet som
64	FGEAR	$= Gæld / Egen$
65	ROA	$= 100 \times (EBITDA - Afs) / Aktiv$
66	q	$= (MV \times 1000 + Gæld) / Aktiv$
67	LnAktiv	$= \ln(Aktiv)$
68	Alder	$= År - Stift$
69	LnAlder	$= \ln(Alder)$
70	RDPct	$= RD/Oms$
71	CapExPct	$= CapEx/Oms$
72	ArbKapPct	$= ArbKap/Oms$
...		
77	E1ME2	$= Ejer2Pct - Ejer1Pct$ (hvis $Ejer2Pct > 0$ )
78	disp	$= E1ME2 / (Ejer2Pct + Ejer1Pct)$
79	hh	Summen af EjerYPct^2, for alle ejere.

Beregninger foretaget på variablerne beskrevet i bilag 3.1.

### 3.4. Beregnede dummy-variabler

No.	Fork.	Beskrivelse
73	AlderD	Angiver om selskabet er ændre (1) eller yngre end (0) 50 år.
74	MajEjer	Angiver om selskabet har én majoritetsejer (1) eller ej (0). $> 50\%$ .
75	IngenEjer	Angiver om selskabet har én ejer med $> 10\%$ (1) eller ej (0).
76	FlereEjer	Angiver om selskabets 10 største ejere samlet har $> 50\%$ ejerskab. (1) hvis ja eller (0) hvis nej.
77	FamSel	Angiver om selskabets største ejer er familie (1) eller ej (0).
78	MajFam	Interaktionsled som angiver om majoritetsejeren er familie (1) eller ej (0): MajEjer $\times$ FamSel

### 3.5. Fordelingen af observationer på lande og år.

	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Australien	1	3	3	6	8	8	7	8	9	10
Canada	3	3	5	4	7	9	10	10	9	13
Kina	5	52	111	152	154	167	181	252	305	364
Tyskland	4	5	6	7	8	12	9	11	15	17
Frankrig	7	0	8	17	12	13	18	18	18	16
Storbritanien	3	4	6	9	12	13	16	20	19	20
Indonesien	0	0	0	0	32	44	42	42	55	56
Israel	1	2	3	4	5	5	7	6	7	6
Indien	0	0	0	0	32	44	42	42	55	56
Iran	0	0	0	1	2	7	9	10	7	13
Japan	0	0	0	0	0	0	0	0	41	45
Sydkorea	13	16	39	37	49	47	57	68	75	82
Caymanøerne	5	4	14	15	15	16	18	18	15	15
Malaysia	1	1	3	3	4	5	5	4	6	5
Rusland	2	2	3	3	4	4	3	3	2	7
Sverige	3	4	3	5	5	8	9	12	12	8
Taiwan	0	0	7	15	19	29	32	33	30	29
USA	31	39	52	74	90	97	126	121	117	116
Vietnam	0	1	2	4	4	2	4	6	10	16
Øvrige	25	32	45	49	52	71	80	95	102	109
Total	107	171	316	410	488	562	638	742	860	953

Note: Fordelingen af observationer på lande (Land2) og regnskabsår (År).

## 4. Bilag til analysen

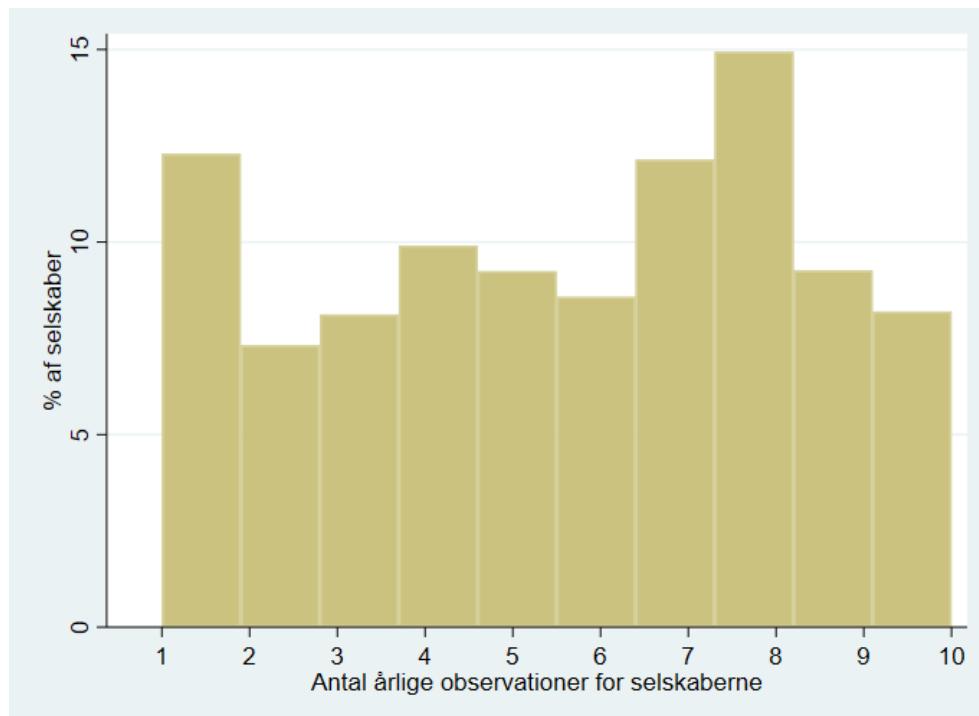
### 4.1. Fixed effects model: ROE og $q$

Variabler	(1) ROE	(5) ROE	(6) ROE	(7) Tobin's $q$	(8) Tobin's $q$
ln(Aktiver)	9,113*** (3,238)	12,927*** (4,034)	5,653* (3,021)	-1,50*** (0,45)	-1,47*** (0,49)
ln(Alder)	-5,24*** (2,015)	-26,81*** (9,023)	-3,443* (1,950)	1,94** (0,96)	1,97* (1,00)
Finansiel Gearing	2,763 (6,065)	3,875 (6,755)	1,384 (6,208)	0,30 (0,22)	0,38 (0,25)
Udviklingsomk. (%)	-0,003 (0,003)	-0,002 (0,003)	-0,036 (0,022)	-0,00 (0,00)	-0,00 (0,00)
Adm. Direktør aktionær?	0,316 (2,362)	0,298 (2,582)	1,313 (2,253)	-0,25* (0,14)	-0,30* (0,16)
Finansdirektør aktionær?	-15,059* (7,814)	-15,641* (8,982)	-17,19** (8,737)	0,08 (0,27)	-0,03 (0,30)
1 = Hvis ingen ejer > 10%	-0,087 (3,149)	0,008 (3,337)	1,535 (3,016)	-0,37** (0,18)	-0,34 (0,21)
1 = Hvis de 10 største ejer samlet er >50%	-4,275 (3,875)	-4,700 (4,126)	-2,718 (2,821)	0,40*** (0,15)	0,34** (0,16)
1 = Hvis største ejer er familie	-2,101 (3,591)	-2,578 (3,846)	0,205 (3,028)	-0,07 (0,20)	-0,20 (0,18)
1 = Hvis majoritetsejer er familie	13,346 (11,125)	16,274 (14,128)	2,073 (5,334)	0,67 (0,66)	0,92 (0,71)
Spredningsforhold	-24,8** (10,343)	-23,811* (12,312)	-17,72** (8,866)	-0,23 (0,53)	-0,19 (0,53)
Forskel på største og næststørste ejer	0,394** (0,169)	0,405** (0,202)	0,322** (0,151)	-0,00 (0,01)	-0,00 (0,01)
Herfindahl-Hirschmann index	-0,854 (7,770)	-1,429 (8,652)	-0,829 (7,632)	-0,41 (0,62)	-0,74 (0,64)
Observationer	3.896	3.606	2.798	3.116	2.484
R-squared	0,019	0,028	0,015	0,09	0,09
Antal selskaber	1.238	1.200	501	851	492
Første år	2008	2010	2008	2010	2010
Antal observationer per selskab	-	-	>3	-	>3

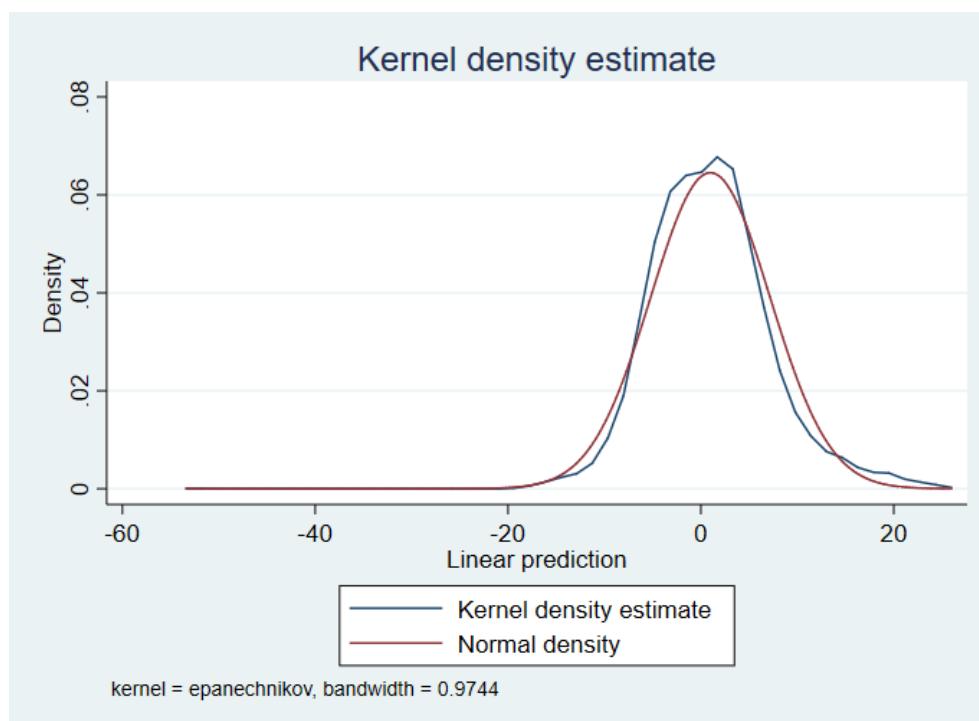
Robuste standardfejl i parantes

\*\*\* p<0,01, \*\* p<0,05, \* p<0,1

#### 4.2. Fordeling af selskaber per antal observationsår



#### 4.3. K-diagram, normalitet af fixed effects fejlled



#### 4.4. Korrelationsmatrix over modellens variabler

	ROA	ROE	q	Alder	Aktiv	LnAktiv	LnAlder	FGEAR	RDPct	CEO	CFO	IngenEjer	FlerEjer	FamSel	FamSel	MajFam	disp	E1ME2	hh
ROA	1																		
ROE	.780	1																	
q	-.249	-.212	1																
Alder	.184	.160	-.133	1															
Aktiv	.0899	.1	-.0764	.265	1														
LnAktiv	.346	.305	-.291	.364	.626	1													
LnAlder	.227	.189	-.139	.918	.22	.334	1												
FGEAR	.279	.296	-.118	.147	.155	.331	.191	1											
RDPct	-.095	-.0522	.012	-.024	-.0110	-.041	-.0318	-.043	1										
CEO	-.008	-.0244	.019	-.097	-.050	-.102	-.108	.0223	-.007	1									
CFO	.0118	-.0139	-.041	-.023	.091	.09	-.047	.079	-.004	.244	1								
IngenEjer	-.0492	-.022	-.024	.062	.213	.099	.07	-.0362	.014	-.003	.046	1							
FlerEjer	-.0476	-.0476	.045	-.037	-.053	-.045	-.062	-.0131	-.008	.028	.01	-.198	1						
FamSel	.0869	.0669	.005	-.031	-.092	-.157	-.011	.077	-.014	.238	-.008	-.220	.103	1					
MajFam	.0589	.046	-.011	.051	.008	.041	.042	.027	-.006	.036	.016	-.061	-.062	.278	1				
disp	.0596	.052	-.049	-.018	-.099	-.028	-.01	.023	-.007	.027	-.017	-.26	.289	.09	.106	1			
E1ME2	.082	.07	-.054	.0219	-.065	.002	.031	.019	-.012	.003	-.021	-.218	.272	.054	.159	.925	1		
hh	.162	.125	-.067	.131	-.0353	.097	.123	.061	-.023	-.076	-.019	-.332	.127	-.003	.239	.199	.325	1	

## 5. Udklip fra elektroniske bilag

### 5.1. Linje 100-140 fra <Varlist.do>

```

97 *** CONSOLIDATE VARIABLE: EjerYType ***
98 foreach var of varlist v42-v51 {
99 * Replace duplicates
100    replace `var' = "Employees/Managers/Directors" if (`var' == "Employees, managers, directors")
101    replace `var' = "NA" if (`var' == "Public (publicly listed companies)")
102    replace `var' = "Foundation/Research Institute" if (`var' == "Foundation, research Institute")
103    replace `var' = "Hedge fund" if (`var' == "Hedge funds")
104    replace `var' = "One or more named individuals or families" if (`var' == "I")
105    replace `var' = "Mutual & Pension Fund/Nominee/Trust/Trustee" if (`var' == "Mutual and pension fund, nominee, trust, trustee")
106    replace `var' = "Mutual & Pension Fund/Nominee/Trust/Trustee" if (`var' == "E")
107    replace `var' = "Private equity firm" if (`var' == "Private Equity firms")
108    replace `var' = "Public authority, State, Government" if (`var' == "Public authority, state, government")
109    replace `var' = "Other unnamed shareholders, aggregated" if (`var' == "Unnamed private shareholders, aggregated")
110 * Sort Hedge fund, PE, venture capital, banks & Insurance into "Financial Company"
111    replace `var' = "Financial company" if (`var' == "F")
112    replace `var' = "Financial company" if (`var' == "Hedge fund")
113    replace `var' = "Financial company" if (`var' == "Y")
114    replace `var' = "Financial company" if (`var' == "Private equity firm")
115    replace `var' = "Financial company" if (`var' == "P")
116    replace `var' = "Financial company" if (`var' == "Venture capital")
117    replace `var' = "Financial company" if (`var' == "V")
118    replace `var' = "Financial company" if (`var' == "Insurance company")
119    replace `var' = "Financial company" if (`var' == "A")
120    replace `var' = "Financial company" if (`var' == "Bank")
121    replace `var' = "Financial company" if (`var' == "B")
122 *Sort Corporate into "Industrial Company"
123    replace `var' = "Industrial company" if (`var' == "Corporate")
124    replace `var' = "Industrial company" if (`var' == "C")
125 *Sort Self, State, foundation & employee ownership into "Other"
126    replace `var' = "Other" if (`var' == "D")
127    replace `var' = "Other" if (`var' == "Self ownership")
128    replace `var' = "Other" if (`var' == "H")
129    replace `var' = "Other" if (`var' == "Public authority, State, Government")
130    replace `var' = "Other" if (`var' == "S")
131    replace `var' = "Other" if (`var' == "Foundation/Research Institute")
132    replace `var' = "Other" if (`var' == "J")
133    replace `var' = "Other" if (`var' == "L")
134    replace `var' = "Other" if (`var' == "M")
135    replace `var' = "Other" if (`var' == "Employees/Managers/Directors")
136    replace `var' = "Other" if (`var' == "Other unnamed shareholders, aggregated")
137 *If no owner > 10% then = "NA"
138    replace `var' = "NA" if (`var' == "Z")
139    replace `var' = "NA" if (`var' == "")
140 }

```

### 5.2. Linje 41-50 fra <Main.do>

```

41 encode countryisocode, gen(Land) label(Country) // Encode as numeric variable named 'Land'
42 bysort Land: egen temp = count(Land) // Create temporary-variable counting number of observations
43 gen LandTemp = countryisocode if (temp > 100) // Create new variable if temp > 100
44 encode LandTemp, gen(Land2) // Encode variable as numeric.
45 lab var Land "Landekode, ISO" // Label new variables
46 lab var Land2 "Landekode, ISO, hvis Obs > 100" // Label new variables
47 drop if _merge == 1 //Drop if missing match from master
48 drop if _merge == 2 //Drop if missing match from Orbis country-list
49 drop vl-_merge //Drop variable-duplicates.
50 drop temp-LandTemp //Drop temporary variables

```

### 5.3. Linje 175-190 fra <Main.do>

```

175  bysort Id: gen nyear=[_N]
176  tab nyear // See number of observations per company in table below
177  /*      nyear |      Freq.    Percent     Cum.
178  -----
179      1 |      645     12.29    12.29
180      2 |      384      7.32    19.61
181      3 |      426      8.12    27.73
182      4 |      520      9.91    37.64
183      5 |      485      9.24    46.88
184      6 |      450      8.58    55.46
185      7 |      637     12.14    67.60
186      8 |      784     14.94    82.54
187      9 |      486      9.26    91.80
188     10 |      430      8.20    100.00
189  -----
190      Total |      5,247    100.00    */

```

### 5.4. Linje 305-309 fra <Main.do>

```

305  * Normality
306  xtreg ROA $covariates $ownership , fe vce(cluster Id)
307  predict rs
308  kdensity rs, normal
309  qnorm rs

```

### 5.5. Linje 311-338 fra <Main.do>

```

311  * Generalized Hausman Test (Non-robust S.E.)
312  xtreg ROA $covariates $ownership , fe
313  estimates store FE
314      outreg2 using "table3", ///
315          replace word nocons dec(3) label ///
316          keep($covariates $ownership) ///
317          title("Table 3") ///
318          addtext(Std.Afv., Alm.)
319  xtreg ROA $covariates $ownership , re
320  estimates store RE
321      outreg2 using "table3", ///
322          append word nocons dec(3) label ///
323          keep($covariates $ownership) ///
324          title("Table 3") ///
325          addtext(Std.Afv., Alm.)
326  hausman FE RE
327  /* Test: Ho: difference in coefficients not systematic
328
329          chi2(13) = (b-B)'[(V_b-V_B)^(-1)](b-B)
330          =        149.92
331          Prob>chi2 =      0.0000 */
332  * Sargan-Hansen test using XTOVERID (Cluster-robust S.E.)
333  xtreg ROA $covariates $ownership , re vce(cluster Id)
334  xtoverid
335  /* Test of overidentifying restrictions: fixed vs random effects
336  Cross-section time-series model: xtreg re robust cluster(Id)
337  Sargan-Hansen statistic 137.507 Chi-sq(13) P-value = 0.0000 */

```

## 5.6. Linje 381-421 fra <Main.do>

```

381 *** RANDOM EFFECTS ASSUMPTIONS ***
382 * Multicollinearity
383 xtreg ROA $covariates $ownership $REvar, re vce(cluster Id)
384 vif, uncentered
385 /* Variable |      VIF      1/VIF
386 -----
387     LnAktiv |    40.16    0.024901
388     LnAlder |    35.49    0.028175
389     FGEAR |     1.15    0.870082
390     RDPct |     1.01    0.991847
391     CEO |     1.53    0.652733
392     CFO |     1.17    0.852309
393     IngenEjer |    1.71    0.584739
394     FlereEjer |    1.43    0.697102
395     FamSel |     1.92    0.520616
396     MajFam |     1.54    0.647741
397     disp |    10.79    0.092696
398     ElME2 |    11.19    0.089335
399     hh |     2.74    0.364415
400     Land2 |
401     1 |     1.10    0.909030
402     2 |     1.11    0.904805
403     3 |     3.58    0.279189
404     4 |     1.21    0.826489
405     5 |     1.25    0.798998
406     6 |     1.28    0.781665
407     7 |     1.11    0.899694
408     8 |     1.12    0.891346
409     9 |     1.44    0.694703
410     10 |     1.02    0.975839
411     11 |     1.28    0.783254
412     12 |     2.25    0.443582
413     13 |     1.21    0.826447
414     14 |     1.06    0.942868
415     15 |     1.03    0.967343
416     16 |     1.08    0.929614
417     17 |     1.37    0.730987
418     18 |     2.85    0.350419
419     19 |     1.01    0.994760
420 -----
421     Mean VIF |    4.35      */

```

## 6. Resultater fra tidligere studier

### 6.1. Leaven & Levine (2007): Table 5

**Table 5**  
Valuation and complex ownership structures

	(1)	(2) Majority and widely held Basic	(3) Other firm factors	(4) Exclude majority	(5) Exclude majority and widely held	(6) Multiple blockholders	(7) Multiple blockholders and other firm factors	(8) UK
Cash-flow-1	0.015*** (0.005)	0.019*** (0.007)	0.015* (0.008)	0.017*** (0.005)	0.023*** (0.006)	0.037*** (0.017)	0.031 (0.022)	0.031 (0.022)
Control-1 minus Cash-flow-1	-0.009*** (0.003)	-0.008 (0.004)	-0.008* (0.004)	-0.014* (0.007)	-0.009 (0.007)	-0.001 (0.008)	-0.018 (0.014)	-0.017 (0.014)
Cash-flow-1 minus Cash-flow-2	-0.009*** (0.003)	-0.012*** (0.003)	-0.011* (0.005)	-0.020** (0.008)	-0.022** (0.008)	-0.040*** (0.018)	-0.040** (0.017)	-0.037** (0.017)
Sales growth	0.381 (0.308)	0.384 (0.308)	0.449*** (0.202)	0.619*** (0.240)	0.685*** (0.201)	0.927* (0.491)	1.683*** (0.286)	1.663*** (0.299)
Majority	0.105 (0.158)	0.105 (0.158)	0.091 (0.118)	0.091 (0.118)	0.091 (0.118)	0.091 (0.118)	0.091 (0.118)	0.091 (0.118)
Widely held			0.273** (0.118)	0.439** (0.169)				
Log of total assets			-0.150 (0.092)	-0.150 (0.092)				
Asset tangibility				-1.160*** (0.360)				-0.228 (0.143)
Investment ratio				2.126*** (0.366)				-0.231 (0.143)
Leverage				-0.900 (0.533)				-1.715* (0.843)
Founder				0.949 (0.692)				-1.726* (0.843)
Cash-flow-1 minus Cash-flow-2 * UK								-0.880 (0.728)
Observations	1638	1638	1195	1233	975	464	330	330
R <sup>2</sup>	0.12	0.12	0.16	0.14	0.16	0.18	0.26	0.26

This table presents OLS regressions that control for the presence and incentives of the two largest shareholders. The dependent variable is Tobin's  $Q$ . The regressions are of the form:  $Tobin's Q = \alpha X + \beta(Cash-flow-1; Control-1 minus Cash-flow-1; Cash-flow-1 minus Cash-flow-2) + \gamma(Other firm characteristics) + u$ , where  $u$  is the error term. All regressions include the cash-flow rights of the largest shareholder (Cash-flow-1), the difference between control rights and cash-flow rights of the largest shareholder (Control-1 minus Cash-flow-1), and real sales growth over the preceding three years (Sales growth). Regression (2) controls for whether the controlling owner holds more than 50% of the voting rights (Majority) and whether no shareholder owns more than 10% of control rights (Widely held). Regression (3) controls for the ratio of fixed assets to total assets (Asset tangibility), the ratio of capital expenditure to assets (Investment ratio), the ratio of total debt to assets (Leverage), and a dummy variable that takes a value of one if the founder of the company is still on the board of the company (Founder). Regression (4) excludes majority-owned and widely held firms. Regression (6) includes only firms with multiple blockholders. Regression (7) adds the firm-level control variables to regression (6). Regression (8) adds an interaction between the Cash-flow-1 minus Cash-flow-2 variable and a dummy variable that takes a value of one if the firm is located in the United Kingdom (UK) to regression (7). All regressions control for country- and industry-fixed effects with clustering of standard errors at the country level. Standard errors in parentheses. \*Significant at 10%; \*\*sign significant at 5%; \*\*\*sign significant at 1%.

## 6.1. Anderson &amp; Reeb (2003): Table V

	Return on Assets (Using EBITDA)	Return on Assets (Using Net Income)	Tobin's <i>q</i>
Intercept	0.272 (9.02)	0.205 (14.89)	3.760 (17.92)
Family ownership	0.099 (3.41)	0.042 (2.27)	0.663 (2.11)
(Family ownership) <sup>2</sup>	-0.161 (3.82)	-0.076 (2.65)	-1.070 (2.55)
Officer/directors own (less family)	0.048 (0.60)	0.023 (0.52)	1.447 (1.68)
Unaffiliated blockholders	-0.010 (2.74)	-0.013 (4.73)	-0.356 (10.86)
Outside directors	-0.021 (1.84)	-0.005 (0.75)	-0.068 (0.73)
CEO equity-based pay	0.005 (0.74)	0.005 (1.17)	0.192 (3.09)
R&D/sales	0.199 (2.48)	0.013 (0.24)	4.702 (7.20)
LT debt/total assets	0.052 (2.59)	-0.135 (12.56)	-1.026 (7.92)
Return volatility	-0.190 (7.01)	-0.161 (7.60)	-1.897 (9.82)
Ln(total assets)	0.005 (2.11)	-0.004 (3.57)	-0.098 (5.94)
Ln(firm age)	-0.027 (5.26)	-0.010 (3.83)	-0.197 (5.72)
Adjusted <i>R</i> square	0.367	0.278	0.408
Inflection point (%)	30.8	27.6	31.0