

# ET STUDIE AF LATENTE FAKTORER I ESG-SCOREN

EFFEKTEN AF ESG SCORER PÅ VIRSOMHEDERS FINANSIELLE PERFORMANCE<sup>\*</sup>

Martin Larsen<sup>†</sup>, Frederik Munck Bigom<sup>‡</sup>

Kandidatspeciale: 181,115 tegn / 80 normalsider

Vejleder: Lisbeth La Cour<sup>§</sup>

15. maj 2020

## Abstract

This paper investigates the effects of the different levels of ESG scores, and their underlying data, on the financial performance of S&P500 companies. The study is based on Thomson Reuters ESG data between 2005 and 2018. Initially a thorough sampling process is executed, including data imputation, to take care of the missing observations. The final sample includes 45 of the initial 500 companies and their respective ESG-score, pillar-scores, sub-scores and 21 of their fundamental environmental- and raw data variables.

The analysis, based on univariate time series analysis and multidimensional panel data regression analysis, aim to investigate the magnitude of ESG determinants on log-return. By conducting a thorough modelling process, where log-returns are fitted with each and every variable as external covariate, 345 models are identified. These models show that the magnitude of ESG determinants are indifferent. This is confirmed in extension by the results of the multidimensional regression analysis, which furthermore shows that the initially omitted inter-individual differences have no impact on the results.

The robustness tests of both the univariate time series analysis and the multidimensional regression analysis is checked by conducting multiple tests, which include, tests of informationcriteria, imputation, influential observations and timeintervals. These measures are taken in order to validate the methodological choices made in this paper. In short, by validating the robustness of the sampling and modelling process, this paper establishes that both ESG-scores and their underlying determinants have no significant impact on returns of S&P500 companies.

---

\*Tak til vores vejleder, Lisbeth La Cour.

<sup>†</sup>Cand.merc.(mat) - erhvervsøkonomi og matematik, Copenhagen Business School. Studienummer: 93800

<sup>‡</sup>Cand.merc.(mat) - erhvervsøkonomi og matematik, Copenhagen Business School. Studienummer: 92941

<sup>§</sup>Økonomisk institut, Copenhagen Business School

# **Indholdsfortegnelse**

<b>1</b>	<b>Introduktion og motivation</b>	<b>5</b>
<b>2</b>	<b>Problemformulering og opgavestruktur</b>	<b>6</b>
<b>3</b>	<b>Litteratur</b>	<b>7</b>
3.1	Generelle bæredygtige investeringer og virksomheders finansielle performance . . . . .	7
3.2	Individuelle effekter af E, S og G . . . . .	8
3.3	ESG som scoringsmål . . . . .	10
3.4	ESG-scoren . . . . .	11
<b>4</b>	<b>Afgrænsning</b>	<b>14</b>
<b>5</b>	<b>Videnskabsteoretisk tilgang</b>	<b>14</b>
<b>I</b>	<b>Metode</b>	<b>15</b>
<b>6</b>	<b>Data imputering</b>	<b>15</b>
<b>7</b>	<b>Tidsserieanalyse</b>	<b>16</b>
7.1	Stationaritet . . . . .	16
7.2	Autokorrelation . . . . .	17
7.3	Autoregression, AR . . . . .	18
7.4	Moving average, MA . . . . .	19
7.5	ARMA . . . . .	19
7.6	ARMA-X . . . . .	19
7.7	Heteroskedasticitet . . . . .	20
7.8	ARCH . . . . .	20
7.9	GARCH . . . . .	21
7.10	Estimation . . . . .	21
7.11	Normalitet . . . . .	22
7.12	Model selektion . . . . .	23
7.13	Mean Average Error . . . . .	24
<b>8</b>	<b>Regressionsanalyse</b>	<b>25</b>
8.1	Transformation . . . . .	25
8.2	Indflydelsesrige observationer . . . . .	25
8.3	Paneldata regression . . . . .	26
8.4	Pooled regressionsmodel . . . . .	26
8.5	Fixed effekts regressionsmodel . . . . .	27
8.6	Random effekts regressionsmodel . . . . .	27
8.7	Model valg . . . . .	28

8.8	Gauss-Markov antagelser for OLS . . . . .	28
8.8.1	Robuste standardfejl . . . . .	31
<b>9</b>	<b>Data</b>	<b>32</b>
9.1	Dataindsamling . . . . .	32
9.1.1	Datafrekvens . . . . .	32
9.1.2	Finansiel data . . . . .	33
9.1.3	ESG data . . . . .	33
9.1.4	Vedrørende årlige estimerer . . . . .	34
9.1.5	Kontrolvariable . . . . .	34
<b>II</b>	<b>Resultater</b>	<b>36</b>
<b>10</b>	<b>Data</b>	<b>36</b>
10.1	Sampling . . . . .	39
10.2	Imputering af manglende data . . . . .	41
<b>11</b>	<b>Tidsserie analyse</b>	<b>43</b>
11.1	Fordelingsantagelser . . . . .	43
11.2	Stationaritet i finansiel data . . . . .	45
11.3	Stationaritet i de eksterne kovariate . . . . .	45
11.4	Autokorrelation i finansiel data . . . . .	47
11.5	Heteroskedasticitet . . . . .	48
11.6	ARMA-X/GARCH modellerne . . . . .	49
11.7	Modelresultater . . . . .	55
<b>12</b>	<b>Paneldata regressionsanalyse</b>	<b>62</b>
12.1	Transformering af variable . . . . .	62
12.2	Tjek af indflydelsesrige observationer . . . . .	63
12.3	Model-setup . . . . .	64
12.4	Model valg . . . . .	66
12.5	Gauss-Markov antagelser for OLS . . . . .	69
12.6	Modelresultater . . . . .	73
<b>III</b>	<b>Robusthed i modellerne</b>	<b>77</b>
<b>13</b>	<b>Robusthed i univariate modeller</b>	<b>77</b>
13.1	GARCH(1,1) antagelse og valg af informationskriterie . . . . .	77
13.2	Datafrekvens . . . . .	80

<b>14 Robusthed i multidimensionelle modeller</b>	<b>82</b>
14.1 Indflydelsesrige observationer . . . . .	82
14.2 Modellering med ændret tidsinterval . . . . .	84
14.3 Test af dataimputering . . . . .	84
14.4 Afsluttende bemærkning . . . . .	85
<b>IV Diskussion</b>	<b>87</b>
<b>15 Manglende data</b>	<b>87</b>
<b>16 ESG-scorer leverandør</b>	<b>88</b>
<b>17 Afsluttende om bias og kausalitet</b>	<b>89</b>
<b>18 Metodevalg</b>	<b>89</b>
<b>V</b>	<b>92</b>
<b>19 Konklusion</b>	<b>92</b>
<b>20 Fremtidig forskning</b>	<b>93</b>
<b>VI Appendix</b>	<b>100</b>
<b>Appendix A Dataarbejde</b>	<b>100</b>
<b>Appendix B Tidsserieanalyse</b>	<b>102</b>
<b>Appendix C Regressionsanalyse</b>	<b>134</b>
<b>Appendix D Robusthed</b>	<b>144</b>

## 1 Introduktion og motivation

I de seneste år har interessen for bæredygtighed oplevet en markant forøgelse. Den øgende efterspørgsel blandt forbrugere verden over, har motiveret virksomheder til at sætte bæredygtighed på dagsordenen. Alene siden 2014 er salget af bæredygtige produkter i USA steget næsten 20% hen imod slutningen af 2018, hvilket svarer til en fire gange større årlig vækstrate, sammenlignet med konventionelle produkter (Nielsen, 2018). Flere undersøgelser viser, at især de yngre generationer har mest fokus på bæredygtighed, hvorfor tendensen blot forventes forstærket henover de kommende år (Gallup, 2019).

I december 2015, nåede 195 af verdens lande til enighed om den første globale klima-aftale, ved COP21 i Paris. Aftalen indebærer et samlet ønske om at reducere verdens samlede CO<sub>2</sub>-udledning, for at mindske den globale opvarmning (EU, 2015).

Dette har naturligvis haft en stor effekt på verdensøkonomien og i høj grad virksomheder, som har skulle tilpasse sig de nye krav. Ligeledes har det haft stor indflydelse hos investeringsfonde, hvor bæredygtige investeringer har tiltrukket nye investorer. Det er enhver virksomheds opgave at bidrage til samfundet, foruden at levere god finansiell performance, ifølge Larry Fink, CEO for verdens største kapitalforvalter (Blackrock, 2020).

Harward Business Review beskriver den nye bæredygtige agenda blandt investorerne, i deres Maj-Juni 2019 udgave under overskriften *"The Investor Revolution"*. Her beskrives hvordan de store investeringsfirmaer er blevet *"too big to let the planet fail"*. Således "tvinges" investeringsfonde og især pensionsselskaber til at tænke langsigtet (HBR, 2019).

I 2019 blev mere end \$20 milliarder af nytilkomne aktiver investeret i bæredygtige investeringsforeninger og ETF'er alene i USA. Dette svarer til næsten en firdobling sammenlignet med 2018 (Morningstar, 2020). Fondene er baseret på nænsomt at integrerer ESG, 'environmental', 'social' og 'governance' i deres investeringsproces, og videre at opnå målbare bæredygtige effekter, foruden finanzielle afkast. Her investeres i høj grad efter virksomheders ESG score på baggrund af overordnede parametre.

Med henblik på ovenstående, findes det interessant at undersøge hvordan ESG scoreren og dens determinanter har indflydelse på virksomheders finansielle performance. ESG scorene er blevet et populært mål for bæredygtighed, da de overordnede niveauer er lette af forholde sig til, og sammenligne mellem virksomheder. ESG scorene har hidtil været brugt meget overfladisk og få akademiske analyser har dykket ned i determinanterne og selve opbygning af ESG scorene. Problemer som manglende data punkter, fejlagtigt vægtning og "matchmaking" er alle noget der ønskes større fokus på i forbindelse med denne undersøgelse.

## 2 Problemformulering og opgavestruktur

Med henblik på den forøgede interesse for ESG investeringer beskrevet ovenfor, ønskes det at undersøge følgende spørgsmål:

- Hvilken grad af indflydelse har de forskellige niveauer af ESG scorer på virksomheders finansielle performance, i form af afkast på det underliggende aktiv?
- Hvilke underliggende ESG faktorer ligger til grund for indflydelsen?

Problemformuleringen vil besvares med udgangspunkt i en todelt metodisk tilgang, hvor henholdsvis tidsserieanalyse og panel data regressionsanalyse vil belyse ESG scorenes effekter og aspekter. Indledningsvis vil en grundig datasortering og efterfølgende imputering, tage hånd om manglende observationer.

Efterfølgende vil den todelte tilgang muliggøre forskellige modeleringsmetoder, der tillader at undersøge virksomheders finansielle performance både univariat og multidimensionelt. Først vil der gennemgås en univariat tidsserieanalyse for at undersøge eventuelle signifikante sammenhæng mellem ESG-variablene og aktiekursen. Hernæst vil Virksomhedernes aktieafkast agere responsvariabel i multidimensionelle regressionsmodeller, som også har til formål at undersøge indflydelsen af ESG-variablene. Resultaterne vil derfor ligge til grund for besvarelsen af problemformuleringen, da indflydelsen og retningen af statistisk signifikante forklarende variable undersøges. Afslutningsvis vil modellernes robusthed testes, og resultaterne diskuteres.

For at imødekomme analysens kompleksitet, og bidrage til en grundig besvarelse af problemformulering, er følgende hypoteser formuleret:

### **Overordnet hypotese: Effekten af ESG scorernes underliggende niveauer er større end i det overordnede niveau**

Den overordnede hypotese refererer til projektets første problemstilling, og foreslår, at der i det bagvedliggende ESG data ligger en større forklarende effekt, end i de konstruerede ESG scorer. De øvre niveauer af ESG-scorerne er alle baseret på det rå underliggende data, og percentilt rangeret herefter. De konstruerede scorer skabes på baggrund af en standardiseringprocess ved vægtning af underliggende faktorer indenfor forskellige industrier. Derfor teoretiseres der om hvorvidt essentiel information går tabt i konstruktionsprocessen. Det forventes da, at der lettere kan findes sammenhænge og effekter for rådataen end for de konstruerede ratings.

### **Underhypotese 1: Høje ESG scorer har en positiv indflydelsesgrad på virksomheders finansielle performance**

Den første underhypotese refererer til den generelle forventning om sammenhængen mellem ESG scorer og virksomheders finansielle performance, og foreslår en positiv sammenhæng hertil, med udgangspunkt i den generelle konsensus på området. Tankegangen er her at virksomheder med høje ESG ratings i højere grad vil undgå dårlig omtale og potentielle skandaler, som kan medføre finansielle tab. I en metaundersøgelse, af mere end 2000 studier, vedrørende bæredygtige investeringer,

konkluderes at langsigtede ansvarlige/bæredygtige investeringer har en positiv effekt for virksomheders finansielle performance. Dette kommer især til udtryk i Nordamerika og fremspirende markeder (Friede et al., 2015). Det forventes da, at høje ESG scorer afspejles positivt i virksomhederne finansielle performance.

### **Underhypotese 2: Miljø variable har en signifikant indflydelse på virksomheders finansielle performance**

I almen omtale kategoriseres miljø som den klassiske bæredygtighedsfaktor. På baggrund af øgede politiske tiltag, som skærper virksomhederne miljømæssige krav, forventes det at virksomhedernes miljømæssige initiativer har en direkte indflydelse på deres finansielle performance. Yderligere tænkes det at den øgede interesse for bæredygtighed blandt forbrugere, har medført at virksomheder i højere grad har været nødsaget til at inkorporere miljømæssige initiativer i deres forretningsgrundlag. Det forventes derfor at miljømæssige variable har en signifikant indflydelse på virksomhedernes finansielle performance.

## **3 Litteratur**

Der findes en lang række undersøgelser der forsøger at beskrive sammenhængen mellem bæredygtige investeringer og virksomheders finansielle performance og dertilhørende finansielle afkast. Følgende afsnit har til formål at gennemgå fremgangsmåden samt resultaterne fra relevante studier og litteratur på området. Først og fremmest vil den generelle tilgang til bæredygtige investeringer og virksomheders finansielle performance gennemgås. Hernæst vil studier af de individuelle effekter blive gennemgået, hvorefter ESG-scoringsprincippet og studier heraf vil blive redegjort for. Yderligere bringes en gennemgang af ESG-scoren og dens konstruktionsprocess.

### **3.1 Generelle bæredygtige investeringer og virksomheders finansielle performance**

Bæredygtige investeringer har vundet stort indpas i den finansielle investeringsverden igennem de sidste årtier. Bæredygtighed er et vidt begreb, hvorfor der findes mange forskellige metoder til at investere bæredygtigt. ESG investeringer kendes tilbage fra starten af 1980'erne, hvor der refereres til investering på baggrund af virksomheders sociale ansvar (CSR), investeringer i socialt ansvarlige fonde (SRI), eller ekskludering af sortlistede aktier (Ribando and Bonne, 2010). Bæredygtige investeringer kan ses som et alternativ til konventionelle investeringsmuligheder, og bør derfor også performancemåles relativt til disse. I forhold til ESG initiativer er det væsentlige altså om disse skaber værdi for virksomhederne, som dermed afspejles i aktiekursen.

Der argumenteres i litteraturen for ESG investeringer som en outperformende strategi, da værdien af ESG initiativer undervurderes af markedet. Dette bygger på grundlæggende arbejde om markedsefficiens, hvor (Ball and Brown, 1968) og (Bernard and Thomas, 1989) beskriver aktiers tendens til post-earnings-announcement drift (PEAD). PEAD beskriver aktiers tendens til at bevæge sig i samme retning i en længere periode, som følge af en overraskende finansiel nyhed. Yderligere beskriver (Jegadeesh and Titman, 1993) aktiers tendens til at følge momentum. Disse teorier anses

for værende robust dokumentation imod det efficiente marked, og er derfor konsistent med markedernes underreaktion på ESG nyheder. Yderligere kan virksomheders immaterielle ESG initiativer have indflydelse på markedets underreaktion i overensstemmelse med (Edmans, 2011). På den anden side argumenteres der for at den eneste sociale ansvarlighed en virksomhed forpligter sig til, er at maksimere aktionærernes værdi (Friedman, 1970). Dette er i overensstemmelse med den neoklassiske økonomiske tankegang, at investeringer i bæredygtighed blot medfører øgede omkostninger og dermed mindsker rentabiliteten (Palmer et al., 1995); Baumol and Blackman (1991).

Yderligere argumenteres for at investeringer i ESG kan vise sig at underperforme relativt til aktier uden hensynstagning til ESG. Hvis en stor gruppe investorer fravælger en gruppe aktier på baggrund af deres ringere grad af ESG initiativer, vil det medføre initiale lavere afkast for disse aktier. På sigt vil de dog opnå større relative afkast grundet deres undervurderede prissætning (Merton, 1987). I et studie af (Hong and Kacperczyk, 2009) viser det sig at sociale normer leder til frasortering af sortlistede aktier. Studiet viser dog at på sigt vil de sortlistede aktier opnå relativt højere afkast. Helt grundlæggende har ekskludering af bestemte aktier en klar ulempe, da diversifikation reducerer risiko uden at reducere forventede afkast (Markowitz, 1959). Men med udgangspunkt i Markowitz' moderne porteføljeteori viser studier af data fra 1992 til 2007, af blandt andet (Kempf and Osthoff, 2007) og (Statman and Glushkov, 2009), at der opnås et positivt merafkast ved at investere i bæredygtige porteføjler.

Et studie af (Borgers et al., 2013), viser dog at dette merafkast forsvinder i årene op til 2009. Et følgende studie af (Halbritter and Dorfleitner, 2015) viser videre at effekten heller eksisterer i årene op til 2012. Selve kausaliteten er et omdiskuteret emne indenfor ESG og bæredygtige investeringer. Spørgsmålet er her, om det er virksomhedernes rentabilitet der muliggør bæredygtige investeringer, i højere grad end at ESG initiativerne leder til højere rentabilitet (Roberts and Whited, 2013).

### 3.2 Individuelle effekter af E, S og G

#### Environmental: specifikke effekter

Den nok mest populære kategori indenfor bæredygtige investeringer er den miljømæssige del. I et studie af effekten mellem virksomheders miljømæssig ledelse og deres finansielle performance, påvises der et stærkt link mellem miljømæssige initiativer og fremtidige finansielle forbedringer (Klassen and McLaughlin, 1996). Undersøgelsen bygger på event-studier i USA, og belyser virksomhedernes kumulative abnormal afkast som følge af positive miljømæssige begivenheder. Ydermere bringer deres tværsnitsundersøgelse argumenter fra grundlæggende teorier om efficiente markeder, der postulerer at aktiekurser er en proxy for virksomheders finansielle performance.

En alternativ fremgangsmåde til at undersøge samme effekt bringes af (Cohen et al., 1995). Her tages udgangspunkt i en porteføljeteoretisk tilgang, hvor der konstrueres forskellige porteføljer fra det amerikanske S&P 500 indeks, på baggrund af virksomhedernes forskellige miljømæssige tiltag. Igen viser undersøgelsen en positiv sammenhæng mellem de miljømæssige tiltag og virksomhedernes finansielle performance. Der argumenteres yderligere for en stigende effekt over tid, som følge af emnets popularitet. Dette er i nyere tid blevet bekræftet i et event studie af (Yadav et al., 2016), hvor der påvises at amerikanske virksomheders miljømæssige performance har en positiv indflydelse

på deres værdi. Fælles for disse studier er deres tilgang til at måle finansiel performance, nemlig på baggrund af aktieafkast.

I et studie af (Guenster et al., 2011) tilføjes en ekstra dimension til finansiel performance ved at bruge Tobin's q: ratioen mellem fysiske aktivers markedsværdi og deres udskiftningsværdi. Ved brug af regressionanalyse på månedligt data af virksomheders miljøefficiente faktorer, finder de at miljømæssige ledere på markedet ikke nødvendigvis initialt sælges til en premium i forhold til virksomheder uden miljøefficiens. Her fremgår det at effekten øges over tid. Yderligere argumenteres der for, at virksomheders miljøbevidsthed ikke nødvendigvis afspejles i kapitalmarkederne, men realiseres gennem alternative kanaler, i overensstemmelse med (Alexander and Buchholz, 1978) og (Porter and Van der Linde, 1995). Sidstnævnte stiller sig yderligere kritisk overfor tidligere økonometriske studier som viser at miljømæssige reguleringer blot øger omkostninger og skader virksomheders konkurrenceygtighed, som blandt andre (Hazilla and Kopp, 1990).

I nyere tids litteratur viser et studie af (Daszyńska-Zygadło et al., 2016) da også en negativ effekt mellem virksomheders finansielle performance og miljøbevidsthed på baggrund af en sektorbaseret regressionsanalyse. Studiet bygger på nyere tids ESG-scorings datagrundlag og finder at miljø-scoren har signifikant negativ indflydelse på virksomheders performance.

### **Social: specifikke effekter**

Effekten af virksomheders sociale performance i forhold til deres finansielle performance er også et velstuderet område. Grundlæggende studier om corporate social performance (CSP), viser at god ledelse og sociale forhold associeres med bedre finansiel performance (Ullmann, 1985); (Waddock and Graves, 1997) og (McGuire et al., 1988).

Yderligere viser et metastudie af 52 undersøgelse at den positive effekt kan bekræftes, og netop at dette er den generelle konsensus (Orlitzky et al., 2003). I forbindelse med det tidligere nævnte metastudie af (Friede et al., 2015) viste det sig dog også, at det sociale aspekt var den underkategori med den mindst positiv effekt. Det vises endda yderligere i et studie af britiske virksomheders sociale performance og deres aktieafkast at sammenhængen er negativ (Brammer et al., 2006). Her estimeres en multifaktormodel over virksomhederens variation i afkast. Denne er baseret på porteføjler med forskellige sammensætninger af social performance. Herefter bakkedes resultaterne op ved brug af regressionsanalyse med fokus på aktieafkast i forhold til de tre indikatorer: 'miljø', 'medarbejdere' og 'samfund'. Deres hovedresultat viser at virksomheder med høje standarder af CSP opnår lavere afkast, mens de virksomheder med lavest CSP scorer klarer sig bedre end markedet generelt. Endnu et studie baseret på regressionsanalyse finder dog at effekten er neutral (McWilliams and Siegel, 2000), og effekten må siges at være ufuldstændig undersøgt.

### **Governance: specifikke effekter**

Den sidste kategori, er som de andre også blevet undersøgt i høj grad. Grundlæggende arbejde af (Gompers et al., 2003) skabte det velkendte G-indeks, som på baggrund af 24 ligevægtige elementer kan prædictere virksomheders værdi og performance. Indekset er en proxy for styrkeforholdet mellem ledere og aktionærer, og er blevet en velkendt metode til at måle virksomheders beskyttelse mod

overtagelse. G-indekset er konstrueret på 1500 amerikanske virksomheder i perioden 1990-1999, og studiet viser at virksomheder med god governance har en tendens til at have højere afkast, end virksomheder med dårlig governance.

(Bebchuk et al., 2013) kritisere dog G-indekses og udvider undersøgelsen til 2008, som viser at de abnormale afkast er insignifikante i perioden 2000-2008, og indikerer dermed at effekten er forsvundet. I deres kritik argumenterer de for at blot 6 af de 24 elementer er relevante, mens resten er støj. Efterfølgende har et studie af (Gu and Hackbarth, 2013) undersøgt relationen mellem governance og aktieafkast fra G-indekssets, og viser at forklaringen er koncentreret omkring virksomhedernes transparens. Argumentet stemmer overens med at virksomheder med stor transparens er mere værdifulde overtagelseskandidater, da potentielle opkøbere kan byde mere præcist og er i bedre stand til identificere potentielle synergieffekter. Som proxy for transparens benyttes gennemsnitlige tidsserier af estimeret værdier for spredning, fejl og volatilitet.

I et nyere studie af (Auer and Schuhmacher, 2016) undersøges ESG-ratings indflydelse på porteføljers Sharpe ratio, baseret på Sustainalytics data over 600 virksomheder fra både Asian, Europa og USA i tidsperioden 2004-2012. Sharpe ratioen, som indikerer forholdet mellem risiko og afkast ved en investering, viser sig at stige ved at ekscludere virksomheder med dårlig 'governance' rating, mens dette ikke er tilfældet for hverken 'environmental' eller 'social'. Det skal dog tages i betragtning at perioden er relativt kort, hvilket kan have indflydelse på insignifikansen af 'E' og 'S'. Ydermere skal det tages i betragtning at tidsperioden strækker sig henover finanskrisen, hvorfor volatiliteten påvirkes kraftigt, og dermed også sharp ration.

I forhold til virksomheders 'governance', spiller aktionærerne en væsentlig rolle. Hidtil har den beskrevne litteratur bygget på aktionærerne som aktive i forhold til at udvælge aktier, men passive i forhold til ejerskab. I et studie af (Dimson et al., 2015) undersøges indflydelsen af aktivt ejerskab på ESG med data fra over 2000 investorer fordelt på over 600 amerikanske virksomheder i perioden 1999-2009. Resultatet viser at succesfulde aktive ejerskabstiltag kan resultere i værditilførsel til virksomhederne. Undersøgelsen viser samme resultat ved kategorisering af ejerskabstiltag opdelt på 'ES' og 'G'.

### 3.3 ESG som scoringsmål

Den store interesse for bæredygtige investeringer har ledt til en stor efterspørgsel på ESG information. Typisk ESG data er meget komplekst og svært at sammenligne. Dette har medført indførelsen af et mere standardiseret sammenligningsgrundlag. I starten af årtusindeskiftet begyndte en større konsolideringen indenfor ESG-rating industrien (Avetisyan and Hockerts, 2017). De store analysehuse begyndte at publicere disse mere standardiserede ESG ratings, som senere har afført en lang række studier heraf. Dette har også afledt et sammenligningsgrundlag mellem tidligere studier af bæredygtighed og nyere studier baseret på mere transparent data. Fælles for disse nye studier er at de undersøger selve ESG-scoren, samt 'E', 'S' og 'G' hver for sig. Disse undersøgelser begrænses altså til de overordnede niveauer, uden videre at gå i dybden med det underliggende data.

Hovedargumentet for denne begrænsning er, at underliggende information samles og opsummeres i hovedkategorierne, som afspejler virksomhedernes mere overordnede og generelle tiltag indenfor de

forskellige kategorier af bæredygtighed. Der findes dog også studier som undersøger nedre niveauer af ESG scorernes opbygning, men her er den generelle tendens at én specifik variabel undersøges i henhold til et givet emne. Eksempelvis undersøges sammenhængen mellem kønsfordelingen i virksomheders bestyrelser i forhold til deres ESG performance, i et studie af (Velte, 2016). Ved multivariat regressionsanalyse af ESG data fra over 200 virksomheder fra Tyskland og Østrig, findes at kvindelige medlemmer i bestyrelser, har en positiv indflydelse på virksomhedernes ESG performance.

### **Opfølgende bemærkning**

På baggrund af overstående gennemgang af litteratur, er det tydeligt at bæredygtighed i den finansielle verden er et velstuderet område. Nye metodiske tilgange til ESG data har muliggjort det at sammenligne klassiske ESG studier med nyere studier på baggrund af et mere standardiseret datagrundlag. På baggrund af den undersøgte litteratur viser det sig at der en general tendens til ikke at forholde sig til den empiriske dybde som ESG-scorene indeholder. Derfor finder dette studie det relevant at undersøge ESG scorernes bagvedliggende data for at bidrage til litteraturen. For at kunne nærmereudlede bagvedliggende data bringes en dybdegående forklaring af ESG scorernes konstruktion.

### **3.4 ESG-scoren**

ESG er en betegnelse der dækker over virksomheders evne til at operere bæredygtigt indenfor forskellige parametre, som udtrykkes ved en scoring. Thomson Reuters Asset4 database indeholder en bred vifte af forskellig ESG data. Databasen er udarbejdet for at skabe en transparent og objektiv måling af virksomheders relative ESG performance, på baggrund af rapporter og virksomhedsanalyser. Databasen indeholder ESG-scorer på over 7.000 børsnoterede virksomheder verden over, med data for omkring 1.000 af dem tilbage fra 2002. Databasen udvides løbende og opdateres i forhold til virksomhedernes regnskabsår, foruden kontinuerligt ved ESG påvirkende hændelser. Scoringen foregår ved en percentil ranking metode, hvorfor scorerne ligger mellem 0 og 100. Således kan virksomhedernes ESG-scorer let sammenlignes relativt til andres. Dataprocesseen starter ved indhentning af de mere end 400 ESG målinger fra regnskaber, ngo-hjemmesider, CSR rapporter, selskabshjemmesider, nyhedskilder og børsrelaterede dokumenter. Herefter gennemgår dataen en lang række kvalitetstestjek for fejl og mangler, og standardiseres for sammenlignelighedens skyld.

### **Opbygning af ESG-scoren**

ESG-scoringen er opbygget igennem fire niveauer. Første niveau er den overordnede ESG-score som består af de tre hovedkategorier: 'Environmental', 'Social' og 'Governance', som altså repræsentere andet niveau. Disse består yderligere af ti underkategorier, som samlet udgør det tredje niveau. Sluteligt dækker disse ti underkategorier tilsammen over 178 essentielle ESG målinger i det sidste niveau.

Den miljømæssige kategori 'E' defineres ud fra de tre bæredygtige aspekter: 'ressourceforbrug', 'emissioner' og 'innovation'. Resourceforbrugs-scoren reflekterer den enkelte virksomheds bedrifter indenfor reduktion af materialeforbrug, energi- og vandforbrug, foruden miljøvenlig effektivisering af virksomhedens værdikæde. Emissions-scoren inkluderer målbar udledning af CO<sub>2</sub> og lignende, men

også virksomhedens engagement og effektivitet i forhold til at reducere emissioner. Innovations-scoren reflektere virksomhedens evne til innovativt at skabe bæredygtige teknologier og miljøvenlige produkter for at reducere kunders miljømæssige byrde.

Det socialt ansvarlige aspekt 'S' defineres ud fra de fire underkategorier: 'arbejdsstyrke', 'menneskerettigheder', 'fællesskab' og 'produktansvarlighed'. Arbejdsstyrke-scoren reflekterer virksomhedens evne til at efterleve arbejdstilfredshed gennem en sund og sikker arbejdsplads, samt at udvise diversitet og lige rettigheder. Menneskerettigheds-scoren er baseret på virksomhedens evne til at efterleve det fundamentale bag menneskerettighedskonvention. Samfunds-scoren måler virksomhedens engagement i at beskytte offentlig sundhed, at udvise godt borgerskab og respektere etiske virksomhedsrelaterede retningslinjer. Produktansvarligheds-scoren reflektere i hvilken grad en virksomhed er i stand til at integrere forbrugerens sundhed, sikkerhed og integritet i deres produkter.

Den sidste kategori 'G' dækker over virksomhedens styring og defineres ud fra de 3 underkategorier: 'ledelse', 'aktionærer' og 'CSR strategi'. Ledelses-scoren afspejler i hvor høj grad virksomheden følger og praktiserer ledelsesmæssige principper. Aktionær-scoren dækker over virksomhedens arbejde for at yde lige vilkår for aktionærerne, og undgå for eksempel et uønsket opkøb. CSR strategi-scoren afspejler implementering af virksomhedens samfundsansvar, heraf CSR: "Corporate Social Responsibility".

De 10 underkategorier består af 178 udvalgte faktorer, som både er målbare enheder, men også virksomhedens politik eller holdning til et emne. Det er netop disse som er grundlaget for hele ESG-scoringen. Det er også faktorerne som styrer vægtningen af af de forskellige kategorier. Eksempelvis består 'emissioner' af 22 faktorer, som derfor har en vægt på 12% ud af de 178. Et samlet overblik over vægtningen ses i følgende tabel:

Kategori	Ledende indikator	Faktorer	Vægt	Samlet
Environmental	Ressourceforbrug	19	11%	34%
	Emissioner	22	12%	
	Innovation	20	11%	
Social	Arbejdsstyrke	29	16%	35.5%
	Menneskerettigheder	8	4.5%	
	Samfund	14	8%	
	Produktansvarlighed	12	7%	
Governance	Ledelse	34	19%	30.5%
	Aktionærer	12	7%	
	CSR strategi	8	4.5%	

Table 1: ESG vægtning

Den faktiske vægtning varierer dog mellem industrier, da alle faktorer ikke har samme relevans for forskellige industrier. Således opnås en mere fair scoring, hvor eksempelvis faktoren 'ansvarlig kapitalforvaltning' kun har indflydelse på virksomheder i den finansielle sektor. I følgende tabel ses et udvalg af emissioners underliggende faktorer:

Faktor	Beskrivelse
Emission reduktions politik	Har virksomheden en politik vedrørende reducering af emissioner?
Total CO <sub>2</sub> ækvivalente emissioner	Total CO <sub>2</sub> og CO <sub>2</sub> ækvivalente emissioner i ton.
Direkte CO <sub>2</sub> ækvivalente emissioner	Direkte CO <sub>2</sub> og CO <sub>2</sub> ækvivalente emissioner i ton.
Indirekte CO <sub>2</sub> ækvivalente emissioner	Indirekte CO <sub>2</sub> og CO <sub>2</sub> ækvivalente emissioner i ton.
Affald total	Total mængde affald produceret emissioner i ton.
Affald genbrugeligt	Total mængde affald genbrugt over total mængde affald produceret produceret i ton.
Affalts reducerende initiativer	Forsøger virksomheden aktivt at reducere eller genbruge affald?
Kommerciel risici og muligheder ved klimaforandringer	Er virksomheden klar over de kommersielle risici og/eller muligheder ved klimaforandringer?
Total NOx og SOx emissioner	Total NOx og SOx emissioner i ton.
Bæredygtig transport	Har virksomheden taget initiativ til miljøvenlig transportation af produkter eller medarbejdere?

Table 2: Emissions faktorer

Det bemærkes her at direkte målbare enheder angives som værdier, mens de binære kvalitative spørgsmål antager værdien 0.5 eller 1. Desuden haves en polaritetsindikator, som angiver om hvorvidt en høj værdi er positiv eller negativ. Scoreren for en enkelt virksomhed beregnes relativt til andre virksomheder indenfor samme industri, på baggrund af de udvalgte relevante faktorer. Dette sker ved en percentil rangering, som udregnes på baggrund af antallet af virksomheder med en værre værdi end den pågældende, antallet af virksomheder med den samme værdi inklusiv den pågældende og antallet af virksomheder med en værdi:

$$\text{score} = \frac{\text{antal virksomheder med en værre værdi} + (\text{antal virksomheder med samme værdi})/2}{\text{antal virksomheder med samme værdi}}$$

På den måde rangeres virksomhederne indenfor hver relevant faktor og et gennemsnit af disse afgør scoreren indenfor den omhandlende kategori. Den virksomhed med det bedste gennemsnit indenfor en kategori tildes altså den bedste vurdering indefor den indikator. Det betyder altså at selvom en virksomheden har den bedste vurdering indefor sin industri, kan den i realiteten performe dårligere sammenlignet med virksomheder fra andre industrier med lavere scorere (Refinitiv, 2019).

## 4 Afgrænsning

Denne undersøgelse begrænser sig til at undersøge indflydelsen af de forskellige ESG-score mål, ved brug af Thompson Reuters ESG-scoring. Undersøgelsen er endvidere begrænset til at determinere indflydelsesgraden af det underliggende miljø-data. Her er det underliggende 'social' og 'governance' data udeladt, bortset fra fem "kontrolvariable".

Samtidig er denne undersøgelse begrænset til at undersøge indflydelsen af ovenstående data på virksomheder i indekset S&P500. Her er den primære analyse begrænset til de månedlige lukkekurser, hvor afkast bruges som mål for virksomhedernes finansielle performance.

Undersøgelsen er begrænset til at tjekke sammenhængen mellem ESG-data i et tidsserie- og panel data regression setting. Undersøgelsen er derfor afgrænset fra porteføljetekniske metoder og klassiske finansielle modeller, så som CAPM.

Ydermere er out-of-sample prædiktion udeladt fra denne opgave. Undersøgelsens fokus er altså ikke at prædiktere fremtidige værdier, men derimod at determinere indflydelsen af de enkelte variable.

## 5 Videnskabsteoretisk tilgang

De tre videnskabsteoretiske begreber ontologi, epistemologi og metodologi danner grundlag for projektets overvejelser om at besvare problemformuleringen bedst muligt.

Videnskabeligt kan der tages udgangspunkt i en enten subjektiv eller objektiv ontologi, hvilket netop reflekterer en fortolkning af, hvornår og hvordan noget eksisterer. Det ontologiske standpunkt afspejler derfor individets tilgang til problemstillingen, hvoraf denne undersøgelse tager udgangspunkt i en objektiv ontologi. Dette er ensbetydende med en overbevisning om, at den virkelighed projektets problemformulering søger at undersøge, har en objektiv eksistensform som er uafhængig af subjektive fortolkninger. For eksempel afspejler de finansielle data, der er benyttet i dette projekt, en objektiv virkelighed, fordi de ikke er påvirket af individuelle og subjektive holdninger.

Epistemologi betyder erkendelsesteori, og refererer til leren om den opnåede viden, samt reflekterer hvorledes denne viden er opnået. I denne sammenhænge er den essentielle videnskabsteoretiske overvejelse, hvordan man kan opnå gyldig viden til belysning af den problemstilling, der ønskes besvaret. Projektets undersøgelser bygger på historisk data udarbejdet af professionelle analytikere, og vurderes derfor som værende repræsentative og hensigtsmæssige til besvarelse af problemformuleringen. Efterfølgende analyser og konklusioner bygger på logiske overvejelser og et grundigt litteratarbejde.

Metodologi refererer til leren om forskningsmetoder, deres forudsætninger og kritikken af disse. Projektet arbejder ud fra en hypotetisk deduktiv metode, hvor hypoteser danner grundlag for en efterfølgende undersøgelse.

## Part I

# Metode

I følgende afsnit vil de grundlæggende principper og essentielle teorier og metoder blive introduceret. Først og fremmest vil opgavens data-metodiske tilgang blive gennemgået, som inkluderer både univariat tidsserieanalyse og multidimensionel regressionanalyse, også kendt som paneldata regressionsanalyse. Indledningsvis bringes en gennemgang af populære termer brugt i univariat tidsserieanalyse, såsom stationaritet, autokorrelation og normalitet. Her vil blandt andet Autoregressioner, Moving Average termer, ARMA-X, ARCH og GARCH modeller bliver introduceret.

Hernæst introduceres de klassiske paneldata regressionsmodeller 'Pooled regressions model', 'Fixed effekts model' og 'Random effekts model'. Efterfølgende introduceres essentielle test af blandt andet enhedsspecifikke- og tids-effekter, samt korrelation mellem disse. Herudover introduceres principperne bag Gauss-Markovs antagelser for OLS, som inkluderer test af multikollinearitet, homoskedacitet og autokorrelation.

Slutteligt opsummeres, gennemgås og forklares, hvordan og hvorfor de enkelte modeller er benyttet, og herunder præcis hvad de kan tilføje for at besvare hoved- og underhypoteser.

## 6 Data imputering

Imputering er en statistisk proces, hvor manglende data punkter bliver erstattet med nye genererede datapunkter. Ifølge (Barnard and Meng, 1999)'s research, er der 3 hovedproblemer med manglende datapunkter: ekstra bias i analysen, besværliggøre analyse af data og reducere brugbarheden af resultaterne. Til at overkomme flere af de problemer, findes flere forskellige metoder til at imputere datapunkter. Fordelen ved et imputeret datasæt er at det kan indgå med medmindre bias til både univariat og multidimensionel statistisk analyse (Gelman and Hill, 2006). Da der i denne undersøgelse er benyttet R pakken 'AMELIA II', hvilken benytter Expectation-Maximation med bootstrapping til at imputere datapunker, gennemgås antagelser, fordele og ulemper metodisk. I følgende vil der blive skelnet mellem det observerbare data  $D^{obs}$ , det uobserverbare data  $D^{mis}$  og det fulde datasæt  $D$ .

### Antagelse: Multivariat Normalfordelt data

Imputerings algoritmen i 'AMELIA II' pakken antager, at data approksimativt følges en fælles multivariat normalfordeling:

$$D \sim \mathcal{N}(\mu, \sigma)$$

Den multivariate normalfordeling kan vise sig at være en grov approximation af den "sande" fordeling af data. (Schafer and Olsen, 1998) viser på baggrund af empiriske undersøgelser, at der er evidens for, at antagelsen ikke mindske effektiviteten af imputeringerne, selv ved brug af kategoriske og/eller mixed data.

### **Antagelse: 'Missing at Random' (MAR)**

Et af hovedproblemerne, når der arbejdes med manglende data, er man kun har observerbare data for  $D^{obs}$ , og ikke det fulde data sæt  $D$ , hvorfor det er en nødvendig antagelse, at  $D^{mis}$  er MAR. Betydning af dette er, at mønstret for det manglende data kun er betinget af  $D^{obs}$ , og ikke  $D^{mis}$ . Med  $M$  defineret som indikator matrix for observationerne, altså  $m_{ij} = 1$  hvis  $d_{ij} \in D^{mis}$  og  $m_{ij} = 0$  ellers. Kan MAR defineres som:

$$p(M|D) = p(M|D^{obs})$$

En svagere betingelse end MAR, er defineret som 'Missing Completely At Random' (MCAR), hvor  $D^{mis}$  også er uafhængig af  $D^{obs}$ .

### **Multipel Imputering**

Hvor normal data imputering beskæftiger sig med at imputere værdier til ét enkelt datasæt, beskæftiger multipel imputering sig med generere  $m$  datasæt ud fra de manglende værdier i ét. Herefter kan man 'bagge' værdierne fra datasættende direkte, for at lave et enkelt datasæt til fremtidig analyse. Mere populært er dog separat analyse er alle  $m$  imputerede datasæt, hvor de separate resultater sidst kan kombineres for at få det endelige resultat. (Honaker and King, 2010) argumenterer for, at den anvendte multiple imputationsmetode er en direkte forbedring med henblik på glatte tidstrends, ændringer på tværs af cross-sectionale enheder, og korrelation over tid og rum. Det fremfører da, at kompleksitet-sniveauet af modelleringen, vil øge i kraft af mængden af datasæt, der skal analyser stiger.

Expectation Maximization algoritmen (Dempster et al., 1977) er praktisk egnet til data, der skal modelleres ved brug af multivariate regressioner og/eller PCA analyse. Dette skyldes, at algoritmen bibrænder relationerne blandt variablene. EMB algoritmen kombinerer herefter EM algoritmen med bootstrapping, se (Efron, 1992), for at generere  $m$  datasæt til fremtidig analyse, eller kombination. En teknisk gennemgang af selve EM og EMB algoritmen er ude for opgavens omfang, men er gennemgået af (Honaker and King, 2010).

## **7 Tidsserieanalyse**

Tidsserieanalyse benyttes grundlæggende til at forecaste, fortolke og teste hypoteser om økonomisk data. For finansielt data anses afkast som en tilfældig (random) variable over tid, hvilket for en given periode er en tidsrække. Tidsrækkeanalyse er essentiel til at modellere tidsafhængigt data. Dette inkluderer effekter som stationaritet, afhængighed og autokorrelation.

### **7.1 Stationaritet**

Stationaritet er fundamentalt indenfor tidsserieanalyse. Aktieafkast anses typisk som stokastiske fra en periode til en anden, men dette er ikke nødvendigvis gældende for middelværdien og standardafvigelsen. En tidsrække er strengt stationær, hvis fordelingen er invariant uagtet en tidsændring, hvilket kræver, at for ethvert  $m$  og  $n$  gælder, at fordelingen af  $(Y_1, \dots, Y_n)$  er lig  $(Y_{1+m}, \dots, Y_{n+m})$ . Stærk stationaritet er i praksis en svær betingelse at overholde, da det kræver, at det stokastiske aspekt er konstant over tid. Derfor antages ofte svag stationaritet. En tidsserie siges at være svagt stationær

hvis middelværdien, variansen og kovariansen er tidsinvariante. Det gælder altså, at  $(Y_1, Y_2, \dots)$  er en svagt stationær process hvis:

$$\begin{aligned} E(Y_t) &= \mu(\text{en konstant}) \text{ for alle } t \\ Var(Y_t) &= \sigma^2(\text{en positiv konstant}) \text{ for alle } t \\ Cov(Y_t, Y_s) &= \gamma(|t - s|) \text{ for alle } t \text{ og } s, \text{ som funktion af } \gamma(h) \end{aligned}$$

Funktionen  $\gamma(h)$  er autokovariansfunktionen, som udtrykker kovariansen af den laggede tidsperiode mellem  $t$  og  $s$  (Ruppert, 2010).

Ikke alle tidsserie er stationære, men i praksis kan stationaritet ofte opnås ved brug af simple transformationer, eksempelvis for finansielle data viser det sig, at tidsserier for prisen ofte ikke er stationære. For at imødekommme dette kan tidsserien transformeres til det logaritmiske transformerede aktie afkast, som i praksis overholder den svage stationaritetsbetingelse. Det skal dog nævnes, at en stationær historisk tidsserie ikke nødvendigvis er stationær i fremtidige tidsperioder. Test af stationaritet i en tidsserie kan udføres ved brug af en Augmented Dickey-Fuller Test. En fundamental betingelse for stationaritet i en tidsserie er, at alle rødder i nedenstående polynomie har en absolut værdi større end ét:

$$1 - \phi_1 x - \cdots - \phi_p x^p \quad (1)$$

Augmented Dickey-Fuller tester nulhypotesen om, at der findes en enhedsrod (unit-root), der har en absolut værdi lig med én i ligning (1).

## 7.2 Autokorrelation

Autokorrelation beskriver afhængigheden mellem en tidsserie og dens tidligere værdier. Autokovariansfunktionen benyttes til at beskrive kovariansen mellem  $Y_t$  og sit eget lag,  $h$ :

$$\gamma(h) = Cov(Y_t, Y_{t+h}) \quad (2)$$

Autokovariansfunktionen estimeres ved brug af sample mean  $\bar{Y}$  og sample variansen  $s^2$ . Da kovariansen mellem  $Y_t$  og  $Y_{t+h}$  er uafhængig af  $t$ , benyttes datapunkterne defineret med  $h$  antal lags imellem, til at estimere sample autokovariansfunktionen:

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-h} (Y_{t+h} - \bar{Y})(Y_t - \bar{Y}) \quad (3)$$

Autokorrelationsfunktionen (ACF) udtrykkes ved at normere kovariansfunktionen med variansen:

$$\rho(h) = Corr(Y_t, Y_{t+h}) = \frac{\gamma(h)}{\gamma(0)} \quad (4)$$

ACF kan testes individuelt for et givet antal lags ved at teste nulhypotesen, at den givne autokorrelationskoefficient er lig 0. Nulhypotesen afvises ved et signifikansniveau på 5%. Alternativt kan en simultan test benyttes til at teste om, hvorvidt en gruppe af nulhypoteser alle er sande eller ej. En sådan test kaldes en Ljung-Box test, som tester nulhypotesen:

$$H_0 : \rho(1) = \rho(2) = \dots = \rho(K) = 0, \text{ for } K \text{ lags} \quad (5)$$

Ved en simultan test af autokorrelation reduceres således også risikoen for fejlagtigt at konkludere, at en eller flere af disse er ikke-negative (Ruppert, 2010). I overensstemmelse med finansiell litteratur om CAPM teori, bør afkastet af et aktiv ikke være forudsigtigt, hvorfor der ikke eksisterer autokorrelationer. Derfor kan en test af 0 autokorrelation benyttes til at validere antagelserne ved det efficiente marked (Tsay, 2010).

Ønsker man at teste for autokorrelation i en model med afhængige lags, kan en Breusch-Godfrey test med fordel benyttes i stedet (Maddala and Lahiri, 1992). Nulhypotesen er her, at der ikke eksisterer nogen autokorrelation i op til p lags, altså igen:

$$H_0 : \rho(1) = \rho(2) = \dots = \rho(p) = 0, \text{ for } p \text{ lags} \quad (6)$$

Breusch-Godfrey testen kan derfor vise sig at være et bedre alternativ end Ljung-Box testen, når der arbejdes med paneldata eller ARMA/GARCH modeller.

### 7.3 Autoregression, AR

En autoregressiv proces med et lag, AR(1) proces, er i sin simpleste form givet som:

$$Y_t - \mu = \phi(Y_{t-1} - \mu) + \epsilon_t \quad (7)$$

Hvor  $\epsilon_t$  er hvid støj,  $WN(0, \sigma^2)$ , og  $\mu$  og  $\sigma$  er parametre i modellen.  $\epsilon_t$ . Her er  $\mu$  middelværdis parametret for modellen og det ses derfor, at  $Y_t - \mu$  har middelværdi på 0 for alle t. Det andet parametre  $\sigma$  vægter tidligere værdier af  $Y'_t$ 's effekt på  $Y_t$ , og udtrykket på højre side  $\phi(Y_{t-1} - \mu)$  kan derfor beskrives som "hukommelsen" i modellen (Ruppert, 2010). AR(1) processen kan udvides til en AR(p) process, hvor p'et indikerer, at den autoregressiv proces er afhængig af flere tidsperioder:

$$Y_t - \mu = \phi_1(Y_{t-1} - \mu) + \phi_2(Y_{t-2} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + \epsilon_t \quad (8)$$

Autoregression er i bund og grund en regression, hvor der bruges beregnede vægte af forrige værdier, af en given variabel, til at beregne værdien af variablen til tiden t.

Det vil sige, at et stokastisk stød kan have en uendelig lang indflydelse på processen via støjleddet, hvor netop størrelsen på  $\phi_i'$ erne er mængden af det stokastiske stød, som bringes med videre. Stødet i sig selv er en indirekte effekt, hvor tiden t-p har en direkte effekt på  $Y_{t-p}$ , som så har en direkte effekt på  $Y_{t-(p-1)}$ , hvilket har en direkte effekt på  $Y_{t-(p-2)}$  osv (Ruppert, 2010).

## 7.4 Moving average, MA

Et alternativ til en AR(p) model er en moving-average model. Denne defineres som følgende af (Ruppert, 2010):

$$Y_t = \mu + \theta_1 \epsilon_t + \theta_2 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (9)$$

En moving average proces vægter de forrige stokastiske led direkte, hvorimod der kun arbejdes med en indirekte effekt i AR(p) modellen. Parameteren  $\theta_i$  er vægten af det stokastiske stød til tidspunkt t-i og angiver effekten af netop dette stød på vores forklarende variabel. Da den forklarende variabel kun er bundet op i q perioder, er effekten af stødende ikke vedvarende, hvilket er en væsentlig forskel i forhold til AR(p) processen.

## 7.5 ARMA

Ofte kan AR(p) og MA(q) modeller være besværlige, da de tit kræver en høj orden af lags for at være signifikante. Dette giver anledning til en kombinationsmodel, ARMA, hvor aspekterne fra AR og MA modellerne samles. Princippet bygger på en kompakt og påholdende model, hvor en ARMA(p,q) model ser således ud (Ruppert, 2010):

$$(Y_t - \mu) = \phi_1(Y_{t-1} - \mu) + \dots + \phi_p(Y_{t-p} - \mu) + \epsilon_t + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} \quad (10)$$

Den stokastiske variabel er altså bundet i dens eget lags, samt de laggede værdier af hvid-støjs processen. Hvor:

$$\epsilon_t \sim WN(0, \sigma^2) \quad (11)$$

Introducere backshift operatoren, kan det vises, at er en hvid-støjs proces også er en ARMA(0,0):

$$B^h * Y_t = Y_{t-h} \quad (12)$$

ARMA(p,q) processen kan nu skrives op vha. backshift operatoren:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(Y_t - \mu) = (1 + \theta_1 B + \dots + \theta_q B^q)\epsilon_t \quad (13)$$

Og en ARMA(0,0) process kan reduceres til den simple form:

$$(Y_t - \mu) = \epsilon_t \quad (14)$$

ARMA modellen har altså et kompleksitets niveau, som både er nemt og overskueligt at arbejde med, og ARMA modeller med få lags kan altså både benyttes praktisk såvel som teoretisk.

## 7.6 ARMA-X

ARMA modellen kan med fordel udvides til at inkluderer eksterne kovariate, populært kaldet en ARMAX model. Her indikerer x'et inklusion af de eksterne kovariate. ARMAX modellen kombinerer

den linære regression med ARMA fejledet. (Ruppert, 2010) definerer modellen som:

$$\mathbf{Y}_t = \beta_0 + \beta_1 \mathbf{X}_{t,1} + \dots + \beta_p \mathbf{X}_{t,p} + \epsilon_t \quad (15)$$

Hvor det for fejledet er defineret som i (indsæt ref til ovenstående formel):

$$(1 - \phi_1 B - \dots - \phi_p B^p) \epsilon_t = (1 + \theta_1 B + \dots + \theta_q B^q)$$

Her er  $u_1, \dots, u_n$  hvid støj. ARMA-X modellen benyttes ofte i forbindelse med modellering af økonomisk data, hvor der ønskes at teste indflydelse af forskellige økonomiske faktorer på fx. afkast eller lignende. Her kræves det, at de eksterne kovariate også er stationære for at estimaterne af de estimerede koeficienter er konsistente (Hyndman and Athanasopoulos, 2018).

## 7.7 Heteroskedasticitet

I forrige afsnit er ARMA(p,q) modellen beskrevet. Netop denne model forbeholder sig, at den betingede varians er konstant. Dette er ikke en holdbar betingelse, når der modelleres med empirisk data. Det er både vist og bredt accepteret, at den betingede varians ofte udviser klyngedannelse, også kaldet heteroskedasticitet. For at imødekomme dette, introduceres i følgende afsnit modeller, der imødekommer heteroskedasticitet.

## 7.8 ARCH

Ønsker man i stedet at inkluderer et autokorreleret støjled i, hvor vi hidtil har brugt betinget kostant varians, kan ARCH modellen benyttes. Der tages udgangspunkt i en autokorreleret regressionmodel, hvor den betingede varians modelleres ved brug af en ARCH model. Nedenstående ses modelopsætningen:

$$Y_t = f(X_{1,t}, \dots, X_{p,t}) + a_t \quad (17)$$

$$a_t = \sigma_t \epsilon_t \quad (18)$$

Hvor  $f(\cdot)$  er  $E(Y_t | X_{1,t}, \dots)$  og  $\epsilon_t$  iid  $N(0, 1)$ . Her angiver  $a_t$  støjprocessen, der modelleres i form af en ARCH(1) proces:

$$a_t = \epsilon_t \sigma = \epsilon_t \sqrt{\omega + \alpha a_{t-1}^2} \Leftrightarrow a_t^2 = \epsilon_t^2 (\omega + \alpha a_{t-1}^2) \quad (19)$$

For at sikre positiv varians i modellen, skal der gælde et par betingelser. Først og fremmest skal  $\omega > 0$ , hvor  $\omega$  intuitivt kan tolkes som langtidsniveauet for variansen. Dernæst skal  $\alpha \geq 0$  og slutteligt er stationariteten af modellen betinget ved at  $\alpha < 0$ , som vi kender fra AR-processen. Ud fra (19) er den betingede heteroskedasticitet givet ved:

$$Var(a_t | a_{t-1}, \dots) = (\omega + \alpha a_{t-1}^2) \quad (20)$$

Hvilket for den generelle ARCH(p) er givet ved:

$$\sigma_t^2 = \omega + \alpha_1 a_{t-1}^2 + \alpha_2 a_{t-2}^2 + \dots + \alpha_p a_{t-p}^2 \quad (21)$$

I ARCH(p), modellen sikrer  $\sum_{t=1}^p \alpha_t < 1$  stationaritet, da modellen konvergerer mod den ubetingede varians. Ligesom i ARCH(1) er det gældende, at  $\omega > 0$  og  $\alpha_t \geq 0$  for  $t = 1, \dots, p$ . ARCH(p)-modellen modellerer altså autokorrelationen i støjprocessen autoregressivt, således at forrige stød i volatiliteten påvirker volatiliteten i fremtidige perioder. Dette gør, at model typen egner sig til at modellere heteroskedasticitet.

## 7.9 GARCH

Til tider kan det vise sig nødvendigt at udvide modellen for støjprocessen i ARCH(p) modellen. Dette kan gøres ved brug af en GARCH(p,q) model. Ved at inkluderer autokorrelation i selve støjprocessen, ser GARCH(p,q) modellen således ud(Ruppert, 2010):

$$a_t = \sigma_t \epsilon_t \quad (22)$$

Hvor:

$$\sigma_t = \sqrt{\omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2} \quad (23)$$

Det første led er fra ARCH(p) modellen. Det andet led sikrer, at tidligere værdier af volatiliteten bliver lagt ind i volatiliteten til tid  $t$ . Derved har stød i volatiliteten en mere vedvarende effekt i GARCH(p,q) modellen. Antagelserne, der sikrer positiv varians i GARCH(p,q) modellen, ligner det kendte fra ARCH(p) modellen. Det kræves, at  $\omega > 0$ ,  $\alpha_t \geq 0$  og  $\beta_t \geq 0$  for  $t = 1, \dots, p$ .

Dernæst sikrer  $\sum_{t=1}^p \alpha_t + \sum_{t=1}^q \beta_t < 1$  stationaritet. Den fremtidige varians vil konvergerer imod det langsigtede gennemsnit for den ubetingede varians med en rate på  $1 - \alpha - \beta$ .

## 7.10 Estimation

I undersøgelsen benyttes flere forskellige R funktioner til estimation af modellerne, hovedsageligt er der dog benyttet 'auto.arima()' og 'ugarchfit()' til at estimere ARMA-X/GARCH modellerne. Da 'auto.arima()' og 'ugarchfit()' funktionerne i R benytter sig af Maximumlikelihood estimation, vil dette, meget kort, blive redegjort for i følgende afsnit. Der introducereres log-likelihood funktionen for en AR(1) proces, som er givet ved:

$$\mathcal{L}(\boldsymbol{\theta}) = \left( \frac{1}{\sqrt{2\pi\sigma_\eta}} \right)^n \sqrt{1-\phi^2} \exp\left\{-\frac{(Y_1 - \mu)^2}{2\sigma_\eta^2/(1-\phi^2)}\right\} \prod_{i=2}^n \exp\left\{-\frac{(Y_t - c + \phi Y_{t-1})^2}{2 * \sigma_\eta^2}\right\}$$

Her er  $Y_1, \dots, Y_d$  variable i tidsrækken, hvor  $Y_t$  er normalfordelt som  $(Y_t | Y_{t-1} = y_{t-1}) \sim N(c + \phi y_{t-1}, \sigma_\eta^2)$ . Log-likelihood funktionen maksimeres over  $\boldsymbol{\theta} = \mu, \phi, \theta$ , for at estimerer de enkelte parametre

tre. Da der er antaget normalitet for at maksimere log-likelihoodfunktionen, kan det vise sig praktisk at tjekke om disse er overholdt, inden model estimation. Der er et empirisk grundlag for, at finansielle data ikke altid i praksis overholder normalfordelingsantagelserne. I følgende afsnit vil der blive introduceret normalitet.

## 7.11 Normalitet

I statistisk analyse er det vigtigt at teste fordelingsantagelserne, der står forud for modelvalget. Oftest antages data at være normalfordelt, hvor, ønsker der at gå udover dette, skal en anden fordelings benyttes til at ”tweake” estimationsprocessen. Normalfordeling er defineret ved 2 parametre, middelværdig og varians. Det svarer i normalfordelingens tilfælde til lokation og skala parametre, som populært bruges til at sammenligne forskellige fordelinger. Det er dog ikke altid nok at definere standardiserede fordelinger ud fra lokation og skala parametre, hvorfor højere ordens momenter med fordel kan benyttes i stedet. Standardnormalfordelingen (Tsay, 2010):

$$\begin{aligned} \mathbf{Z} &= \frac{\mathbf{Y} - E(\mathbf{Y})}{sd(\mathbf{Y})} \\ E(\mathbf{Z}) &= 0, \quad Var(\mathbf{Z}) = 1 \end{aligned} \tag{24}$$

Skævhed er mål for graden af asymmetri og er defineret som det tredje moment af den standardiserede fordeling.

$$Sk(\mathbf{Z}) = E \left\{ \frac{\mathbf{Y} - E(\mathbf{Y})}{sd(\mathbf{Y})} \right\}^3 \tag{25}$$

Negativ skævhed vil være karakteriseret ved en fordeling der ”læner” sig mod højre og vice versa.

Det fjerde moment af den standardiserede fordeling er kendt som Kurtosis. Kurtosis er mål for, hvor høj grad sandsynligheden er koncentreret i midten eller (især) i halen af fordelingen:

$$Kur(\mathbf{Z}) = E \left\{ \frac{\mathbf{Y} - E(\mathbf{Y})}{sd(\mathbf{Y})} \right\}^4 \tag{26}$$

En fordeling med tunge haler vil være præget af høj kurtosis og vice versa. Skævhed og Kurtosis kan, ligesom middelværdi og varians, for en standardiseret fordeling, sammenlignes med de kendte mål for en standardnormalfordeling for at teste om data er normalfordelt. En normalfordeling er karakteriseret ved en skævhed på 0 og en kurtosis på 3. Dette kan udføres ved brug af en Jarque Bera test og/eller i QQ-plots. Jarque Bera tester under nulhypotesen, om at datas skævhed er 0 og kurtosis er 3, hvor teststørrelsen er givet ved:

$$JB = n \left\{ \frac{\widehat{Sk}(\mathbf{Z})^2}{6} + \frac{(\widehat{Kur}(\mathbf{Z}) - 3)^2}{24} \right\} \tag{27}$$

QQ-Plots plotter empiriske mod teoretiske kvantiler og en lineær sammenhæng vil derfor være en indikator på, at normalfordelingen er et godt fit til data.

## 7.12 Model selektion

Når der modelleres ARMA-X/GARCH modeller skal den optimale orden af modellerne bestemmes. Selve ordenen i ARMA modellen  $(p_m, q_m)$  skal udvælges, hvor ordenerne  $(p_v, q_v)$  af GARCH modellen også skal bestemmes. Da vi i denne opgave både benytter rene ARMA modeller, men også introducerer ARMA/GARCH modeller, vil to metodiske fremgangsmåder blive beskrevet.

I denne undersøgelse benyttes både en teoretisk såvel som en empirisk fremgangsmåde. For en specifik tidsserie kan performance målet Aikake Information Criterion benyttes (AIC). AIC benyttes i en bred vifte af statistiske områder, og er givet ved:

$$AIC = -2\log(\mathcal{L}(\hat{\theta}_M L)) + 2p$$

Her er  $\mathcal{L}(\hat{\theta}_M L)$  likelihoodfunktionen, hvor parametrene er estimeret ved brug af maximum likelihood estimation.  $p$  er antal parametre i den specificerede model. AIC minimere den negative log-likelihood for modellen, hvor  $p$  indgår i termet som en straf for forøgelse af parametre. Det fremgår derfor at AIC er defineret med en trade-off balance i mellem likelihood og modelkompleksitet. Termet  $2p$  er oftes kaldet ”kompleksitets termet”.

AIC har et par ulemper. I små samples er der en substancial sandsynlighed for modellerne, valgt på baggrund af AIC, vil indeholder for mange parametre, altså overfitting kan forekomme (McQuarrie and Tsai, 1998). Derfor kan AIC kan yderlige udvides, hvilket netop R funktionen ’auto.arima’ gør brug af. En udvidelse af AIC kaldet AIC corrected (AICc) er defineret som (Cavanaugh, 1997):

$$AICc = AIC + \frac{2p^2 + 2p}{n - p - 1}$$

Lader man  $n \rightarrow \infty$ , konvergerer det tilføjede term i AICc mod 0, og AICc konvergerer i praksis mod AIC. Sammenlignes AICc med AIC, ses det at AICc primært er AIC en ”stregere” kompleksitets term, der netop straffer modelkompleksiteten yderligere på baggrund af sample size.

Slutteligt introducerer Bayesian Information Criterion (BIC), som første gang var publiseret og introduceret i (Schwarz et al., 1978). BIC er tæt relateret til AIC, og ligesom AICc, er BIC mere følsom over for antal parametre i modellen. Det er derfor anerkendt at BIC generelt søger en ”simplere” model end AIC. BIC beregenes som:

$$BIC = -2\log(\mathcal{L}(\hat{\theta}_M L)) + \log(n)p$$

Igen søges der er minimere informations kriteriet.

### 7.13 Mean Average Error

Inde for statistisk analyse er der flere forskellige former for ”performance mål”. To af de mest populære er ’Mean Absolute Error’ og ’Root Mean Squared Error’. MAE er givet ved:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} = \frac{\sum_{i=1}^n |e_i|}{n}$$

Hvor  $y_i$  er de fittede eller prædikterede værdier, og  $x_i$  er de sande værdier. MAE er derfor kan den gennemsnitlige absolute fejl for en givent model. RMSE er givet ved:

$$RMSE = \sqrt{MSE(\hat{\theta})}$$

Og er altså en udvidelse af MAE. På baggrund af (Willmott and Matsuura, 2005)’s undersøgelse af de to performance mål, er der valgt i denne opgave udelukkende at benytte MAE. Willmott, viser blandt andet at tvetydigheden i RMSE kan give problemer i forbindelse med modelsammenligning.

## 8 Regressionsanalyse

Til at undersøge sammenhænge mellem virksomheders finansielle performance og deres ESG-scorer benyttes regressions analyse. Til dette tages udgangspunkt i mindste kvadraters metode, også kaldet ordinary least squares (OLS), som er en standard metode til regressions analyse. For en lineær model vælges parameter estimater for hver af de forklarende variable, således at summen af de kvadrerede residualer minimeres. Her vil virksomhedernes logaritmetransformerede aktie-afkast repræsentere responsvariablen, mens de forklarende variable vil bestå af de enkelte ESG-scorer fra de forskellige niveauer. Her undersøges hvilke forklarende variable, der har signifikant forklaringsgrads af aktieafkastet, foruden vigtigheden af dem hver især. Ideen er at optimere modellen sekventielt ved at reducere den en variabel af gangen. Ved en række af F-test fjernes den respektive variabel med den laveste p-værdi, således modellens fit forbedres uden tab af forklaringsevne. Processen fortsættes, indtil alle de tilbageværende variable har signifikant forklarende effekt. Denne process er kendt som backwards elimination (Fox, 2015). Indledningsvis bør en række overvejelser dog gennemgås, som herunder vil blive specificeret.

### 8.1 Transformation

Transformationer af data kan oftest simplificere udførelsen og lette undersøgelsen ved statistisk modellering. Hvis der eksisterer et ikke-linært forhold mellem responsvariablen og de enkelte forklarende variable, kan en passende transformation gøre forholdet lineært uden at ændre observationernes orden. Dette er især relevant ved arbejde med lineære regressioner, hvor lineariteten forsøger at beskrive forholdet således, at uanset udgangspunktet vil en specifik ændring af  $x$  føre til en tilsvarende ændring af  $y$ . (Fox, 2015) Transformationen kan ske af både responsvariablen og de forklarende variable. Projektets responsvariabel Afkast er allerede logaritmetransformert, i relation til statistisk teori omkring stationaritet af afkast (Ruppert, 2010), hvorfor relevante transformationer er rettet mod de forklarende variable. Hertil vil logaritme transformationer benyttes, da disse især er velegnet til at håndtere finansiel data, da transformationen ”trækker” store værdier ind til en mere koncentreret fordeling. Ydermere kan normalisering være behjælpelig til at transformere data med vidt forskellige proportioner og skala, for så vidt muligt at undgå at ekstreme observationer ”trækker” for meget.

### 8.2 Indflydelsesrige observationer

OLS er baseret på mindste kvadraters metode, som minimerer de kvadrerede residualer, hvorfor indflydelsesrige observationer og outliers kan påvirke estimatorer og resultere i bias. En outlier er en enkelt observation, som ligger langt fra regressionslinjen og dermed har et stort residual. Outliers kan øge variansen af parametre, men vil oftest ikke ændre resultaterne markant, medmindre de også er indflydelsesrige observationer. Til gengæld kan de lede til type II fejl, hvor nulhypotesen fejlagtigt ikke afvises, således potentielt relevante variable fejlagtigt fjernes grundet manglende signifikans (Stock and Watson, 2010). En indflydelsesrig observation har selvstændig indflydelse på regressionslinjen, hvorfor det er vigtigt at have disse in mente. Til at undersøge om eventuelle selvstændige observationer påvirker til overdrevent bias på paramter estimatorne, kan DFBETA beregnes. DFBETA er

et mål for forskellen mellem et parameter estimat med og uden hver observation. Der er altså et DFBETA tilhørende hver observation og parameter estimat. Generelt vil en stor DFBETA værdi antyde, at en observation har overdreven indflydelse ved estimation af en given parameter. Til at bedømme grænsen for indflydelsesrigne observationer anbefales den standardiserede grænse på  $2/\sqrt{n}$ , som er justeret for størrelse (Belsley et al., 1980).

### 8.3 Panedata regression

Longitudinal data, også kendt som Panel data, består af målinger i to dimensioner. Her er gentagende observationer målt over en tidsperiode for et antal af enheder. Enhedsdimensionen, også kaldt 'cross-section', repræsenteres i projektet af virksomheder. Panel data har flere fordelagtige egenskaber. Datastrukturen tillader at arbejde med individuelle virksomheder med flere forskellige gentagende observationer over tid (Ruppert, 2010). Her er især fleksibiliteten ved modellering af forskelle mellem virksomheder relevant. Den grundliggende panel data regressions model kan opskrives som ses i ligning 28:

$$y_{it} = \sum_{t=1}^n \beta_i x_{it} + \sum_{t=1}^n z_i \alpha + \epsilon_{it} \quad (28)$$

hvor  $x_{it}$  er de forklarende variable for virksomhed  $i$  til tid  $t$ ,  $\beta_i$  er dets tilhørende koefficient,  $z_i \alpha$  repræsentere heterogeniteten, altså den individuelle effekt for virksomhederne og  $\epsilon_{it}$  repræsenterer det idiosynkratiske fejlled. Hvis  $z_i$  er observeret for alle virksomheder, er det en ordinær lineær regressions model, som kan løses ved mindste kvadraters metode. Dette er dog sjældent tilfældet, hvorfor der også findes flere udgaver med forskellige modelstrukturer til at tage højde for forskellige scenarier (Greene, 2018). Disse vil herunder blive gennemgået.

### 8.4 Pooled regressionsmodel

I en pooled regressionsmodel er data "pooled" i forhold til tid og virksomhed. Det betyder derfor, at 'cross-section' dimensionen af virksomheder antages værende ikke unikke, og videre at der ikke er en effekt over tid. Derfor kan modellen også ses som tilfældige træk fra en population i en lineær model. Det er derfor også essentielt ved estimationen om, hvorvidt modellen kan benyttes i hver tidsperiode. Hvis  $z_i$  blot indeholder et konstantled for hvert individ, kan OLS benyttes som efficient estimator for  $\alpha$  og  $\beta$ . Den 'pooled' regressions model kan skrives som (29):

$$y_{it} = \alpha + \sum_{t=1}^n \beta_i x_{it} + \epsilon_{it} \quad (29)$$

hvor det bemærkes, at interceptet,  $\alpha$ , er konstant på tværs af alle virksomheder. Når panel data 'pooles' kan det ofte resulterer i heterogen bias, hvis de individuelle effekter har forklaringsgrad. Disse effekter vil derfor blive opsamlet i fejlledet, og på den måde danne grundlag for biased og inkonsistente koefficenter (Greene, 2018).

## 8.5 Fixed effekts regressionsmodel

Den 'fixed' effekts regression bruges til at kontrollere for 'omitted' eller 'udeladte' variable i paneldata, når de udeladte variable varierer mellem virksomheder, men er tidsinvariante. På den måde har hver virksomhed også sit eget intercept,  $\alpha_i$ . Disse repræsenteres af et sæt af koefficienter, som afspejler indflydelsen af de udeladte variable, som er forskellige fra individerne, men er tidsinvariante. De kan altså ses som effekten af at være "indenfor" en virksomhed. Den fixed effect regressions model kan skrives som følger i ligning 30:

$$y_{it} = \alpha_i + \sum_{t=1}^n \beta_i x_{it} + \epsilon_{it} \quad (30)$$

hvor det ses, at hældningskoefficienten for de forklarende variable  $\beta_i$  er ens på tværs af individerne. Modellen tillader desuden, at de uobserverede individuelle effekter er korrelerede med de forklarende variable,  $x_{it}$ . Ydermere ses fejlleddet,  $\epsilon_{it}$ , som repræsenterer den 'fixed effekt' fra udeladte variable, som både er afhængig af tid og individ. Fejlleddet er enten en- eller to-vejs, afhængigt af de uobserverede individuelle effekter  $\mu_i$  og den uobserverede tidseffekt,  $\lambda_t$ . En virksomheds uobserverede individuelle effekter kunne eksempelvis være deres forretningsstrategi eller ledelsesegenskaber, som er tidsinvariante. Ligeledes kunne uobserverede tidseffekter være politiske reguleringer, som er virksomhedsinvariant. Et to-vejs fejltid opskrives som følger i ligning 31:

$$\epsilon_{it} = \mu_i + \lambda_t + v_{it} \quad (31)$$

hvor  $\mu_i$  repræsenterer de uobserverede individuelle effekter, og  $v_{it}$  er det idiosynkratiske fejlled. Hvis det antages at  $\mu_i$  og  $\lambda_t$  er 'fixed' parametre til estimation og den resterende støj,  $v_{it}$  er  $N(0, \sigma_v^2)$ , haves en to-vejs fixed effects fejlleds-komponent model repræsenteret ved ligning (30) og (31) (Baltagi, 2015). Termen fixed effekts estimator kaldes også for "within" estimatoren, da den refererer til estimation af koefficienterne i regressionsmodellen inklusiv de 'fixed' effekter, som i praksis er 'dummy'-variable.

## 8.6 Random effekts regressionsmodel

Den afgørende forskel mellem den 'fixed' effekts model og den 'random' effekts model er, hvorvidt de uobserverede individuelle effekter har indflydelse på de forklarende variable eller ej. Modsat den 'fixed' effekts model, hvor de udeladte variable indholdes i interceptet og fejlleddet, har disse altså indflydelse på de forklarende variable, hvorfor effekten nu er 'random'. Det antages for den 'random' effekts model, at de individuelle effekter er ukorreleerde med de forklarende variable, hvorfor de tidsinvariante koefficienter kan estimeres. Fordelen er nu, at antallet af parametre som skal estimeres bliver reduceret med forbehold for potentiel inkonsistent estimatorer, hvis antagelsen viser sig ikke at være passende. Den 'random' effekt regressions model kan skrives, som følger af ligning 32:

$$y_{it} = \sum_{t=1}^n \beta_i x_{it} + (\alpha + \mu_i) + \epsilon_{it} \quad (32)$$

hvor der bemærkes, at interceptet  $\alpha$  er konstant på tværs af individerne, da effekten af uobserverede varibale er indeholdt i fejlledet, med enheds- og tidsspecifikke effekter (Greene, 2018).

## 8.7 Model valg

Til at teste valget mellem de forskellige modeller tages udgangspunkt i en F-test, som sammenligner 'fixed' effekts modellen med den 'pooled' OLS. Her opstilles nulhypotesen, at alle de observerede og uobserverede 'fixed' effekter er lig nul for samtlige virksomheder. Hvis nulhypotesen afvises, må den alternative hypotese accepteres, at der eksisterer mindst en tids- eller enhedsspecifik effekt, hvilket udelukker den 'pooled' regressions model. Dette kan opskrives som følger af ligning 33:

$$F = \frac{(SKR_{pool} - SKR_{FE}/\nu_{pool} - \nu_{FE})}{SKR_{FE}/\nu_{FE}} \quad (33)$$

hvor SKR er summen af kvadraterne på residualerne, og  $\nu$  repræsenterer modellernes frihedsgrader. Alternativt kan Breusch-Pagan's Lagrange multiplier (LM) test benyttes til at teste nulhypotesen, at variansen af tids- eller enhedsspecifikke effekter er lig nul. Hvis nulhypotesen afvises, er der en signifikant 'random' effekt, og da vil en 'random' effekts model således være i stand til at håndtere heterogenitet bedre end en 'pooled' OLS regression. LM testen følger en  $\chi^2$ -fordeling med en frihedsgrad og kan skrives som følger af ligning 34:

$$LM = \frac{NT}{2(T-1)} \left[ \frac{\sum_{i=1}^N (\sum_{t=1}^T \hat{\epsilon}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T \hat{\epsilon}_{it}^2} - 1 \right]^2 \quad (34)$$

hvor N er antallet af virksomheder, og T er antal tidsperioder. Således kan eksistensen af enhedsspecifikke og tids effekter undersøges, og en Hausman test kan benyttes til at vælge mellem en fixed effekts og en random effekts model. I ligning 35 ses Hausmans test, som sammenligner de to modeller under nulhypotesen, at de individuelle effekter er ukorreleerde med de forklarende variable:

$$HT = \frac{\hat{\beta}_{FE,\nu} - \hat{\beta}_{RE,\nu}}{(\sigma^2 \hat{\beta}_{FE,\nu} - \sigma^2 \hat{\beta}_{RE,\nu})^{0.5}} \quad (35)$$

hvor  $\hat{\beta}$  repræsenterer modellernes koefficient estimater med deres tilsvarende varians  $\sigma^2$ . Sammenligningen tester altså, om random effekts estimatet er insignifikant forskellig fra 'fixed' effekts estimatet. De to estimater sammenlignes altså, hvor den 'fixed' effekt er konsistent under både nulhypotesen og alternativhypotesen, mens 'random' effekts estimatet kun vil være konsistent under nulhypotesen. Derfor vil en signifikant forskel mellem de to estimator betyde, at en 'fixed' effekts model er at foretrække. Omvendt hvis nulhypotesen ikke kan afvises, er de to modeller konsistente, hvor en 'random' effekts model i praksis ofte vil foretrækkes.

## 8.8 Gauss-Markov antagelser for OLS

Brugen af OLS motiveres af Gauss-Markov teorem. Hvis OLS regressionen overholder en række af antagelser, siges den ifølge Gauss-Markovs teorem at være den bedste lineære unbiased estimator (den siges, at være BLUE) af parametre. Her refererer "bedste" til den laveste varians af estimatorer i forhold til andre unbiased lineære estimatorer. Unbiased estimator refererer til, at gennemsnittet af dens stikprøvemålsfordeling er lig den sande værdi for de estimerede parametre, altså gennemsnittet

af populationen. I sin grundform antager Gauss-Markov's teorem tre betingelser (udover antagelsen om lineære parametre i regressionsmodellen), for at en OLS estimator er BLUE. Antagelserne er baseret på regressionens fejlled og danner grundlag for sande tilfældige fejlled. Antagelserne betinger, at fejlenes population har middelværdi nul, at fejlene har konstant varians, og at fejlene er ukorrelerede for forskellige observationer (Stock and Watson, 2010). Gauss-Markovs teorem er et klassisk udgangspunkt for de forudsættende antagelser bag OLS regressionen. I praksis splittes antagelserne typisk op i en længere række af mere specifikke antagelser. Disse vil herunder vil blive gennemgået med udgangspunkt i Egerod (2016), med supplement fra (Stock and Watson, 2010) og (Wooldridge, 2010).

### **Antagelse 1: Linearitet**

Den første antagelse forudsætter linære parametre i regressions modellen. Dette implicerer, at den afhængige variabel kan skrives som en linearkombination af de forklarende variable plus et fejlled. Dette betyder dog ikke endvidere, at der nødvendigvis er et lineært forhold mellem den afhængige variabel og de enkelte forklarende variable. De forklarende variable kan nemlig tage ikke-lineære form, så længe de kan transformeres således, at parametrene tager passende lineær form. Et brud på antagelsen vil resulterer i biased estimerater, som i gennemsnit ikke vil ramme den sande værdi. Den lineære sammenhæng kan undersøges ved blandt andet at plotte de forklarende variable mod den afhængige, hvor også potentielle data transformationer eksplickeres. Desuden kan quantile-quantile (QQ) plots og histogrammer tydeliggører om de forklarende variable fordelagtigt kan normaliseres, eller om de helt kan fjernes fra regressions modellen.

### **Antagelse2: Tilfældig udvælgelse**

Det skal gælde, at fejledets population har middelværdi lig nul for at overholde antagelsen om tilfældig udvælgelse og dermed opnå repræsentativt data. Hvis ikke dette er tilfældet, har man biased estimerater, som i gennemsnit vil ramme en forkert værdi.

### **Antagelse 3: Ingen perfekt multikollinearitet**

Den tredje antagelse omhandler forholdet mellem de forklarende variable, som ikke må være perfekt korreleret med andre forklarende variable. Sagt på en anden måde, må ingen af de forklarende variable være en perfekt linearkombination af de andre forklarende variable. Hvis dette er tilfældet kan modellen slet ikke estimeres, da det matematisk vil kræve en division med nul. Problemet kan varetages ved enten at transformere eller helt fjerne de pågældende variable. I tilfældet uden en fuldstændig korrelation, men stadig med en høj multikollinearitet, vil OLS stadig fungere, men en forøgelse af standardfejlene vil forekomme, hvorfor estimererne bliver mindre efficiente. Problemet med perfekt multikollinearitet er let genkendeligt, da statistiske programmer (i vores tilfælde R) helt vil undlade at estimere koefficienter og standardfejl for de relevante variable. Her må en variabel fjernes, hvor der tages højde for variationen, den er med til at forklare i den afhængige variable. Typisk vil der ikke forekomme perfekt multikollinearitet, men der kan forekomme situationer med høj multikollinearitet, hvorfor det skal tages højde for. Her kan en beregning af varians inflations faktoren (VIF) for de fork-

larende variable belyse, hvor meget variansen af regressionskoefficienten øges på grund af kollinearitet. Her vil kvadratroden af VIF indikerer hvor mange gange større standardfejlen for de estimerede koeficienter er, i forhold til ikke at være korreleret med andre forklarende variable. Derfor betyder en VIF på 1, at den pågældende variable er ukorreleret med de andre, mens en høj VIF naturligvis indikerer en høj grad af korrelation. Alternativt kan en korrelationsmatrice ligeledes belyse den potentielle eksistens af multikollinearitet.

#### **Antagelse 4: Eksogenitet**

Den fjerde antagelse kræver, at de forklarende variable er ukorrelerede med fejlleddet. OLS modellens fejledd repræsenterer eksogeniteten, da denne indeholder alle tænkelige variable, som ikke er inkluderet i regressionsmodellen, men påvirker den afhængige variabel. Derfor er fejlleddet også uobserverbart, og et brud på antagelsen vil indikere ikke-tilfældige fejl, og dermed repræsenterer endogent bias. Endogent bias forekommer typisk på tre måder, enten fra (relevante) udeladte variable, omvendt kausalitet eller ved målefejl. Bias fra udeladte variable kommer fra de variable, som er korrelerede med den afhængige, men er udeladt fra regressionsmodellen. Disse vil være repræsenteret i fejlleddet og påvirke variationen. Udeladte variable kan medføre, at korrelationen mellem to variable forsvinder eller omvendt, at korrelation mellem to variable i modellen opstår, hvis der tages højde for en udeladt variabel. Endogent bias kan også forekomme ved omvendt kausalitet mellem den afhængige variabel og de forklarende variable. Dette vil resultere i, at OLS-estimatoren udover at opfange de ønskede effekter af de forklarende variable også indeholder effekten af den afhængige variabel. Afslutningsvis kan også målefejl af både den afhængige variabel og de forklarende variabel resultere i endogent bias. Det er værd at bemærke, at antagelsen om eksogenitet ikke direkte kan testes, men i stedet bør tages højde for gennem en nøje udvælgelse af forklarende variable. Dette indebærer overvejeser om både benyttede variable, men også udeladte variable.

#### **Antagelse 5: Homoskedacitet & autokorrelation**

Det antages, at modellens fejledd har konstant varians, heraf homoskedacitet, og yderligere, at der ingen korrelation forekommer i fejlleddet komponenter heraf autokorrelation som specificeres i afsnit 7.2. Homoskedaciteten indebærer, at uanset værdien af den uafhængige variable kræves, at OLS-estimaterne har samme præcision. Hvis antagelsen ikke er opfyldt, vil det resultere i inefficiens og forkert estimerede standardfejl, hvorfor OLS ikke længere vil være den bedste estimator. Autokorrelation er et typisk fænomen indenfor tidsserie analyse, hvor tidsserien kan opleve inertি. Finansielle tidsserier udviser ofte *hukommelse* omkring, at variationen fra en periode ikke er uafhængige af den forrige, hvorfor fejlene for eksempel kan have sæsonprægede effekter (Greene, 2018). Residualerne benyttes ofte til at undersøge potentiel heteroskedacitet ved at plotte dem mod modellens fit. Hvis residualerne har samme afstand til regressionslinjen, uagtet værdien af den uafhængige variable, indikeres en ens varians. Omvendt vil forskellig afstande føre til uens varians, hvorfor modellen for forskellige værdier af den uafhængige variabel, vil have forskellig grad af præcision. Her kan Whites standardfejl indføres for at opnå robuste standardfejl (White, 1980). Til formelt at teste for heteroskedacitet i en lineær regressionsmodel kan en Breusch-Pagan test udføres (Breusch and Pagan, 1980).

Videre kan der testet for autokorrelation i op til  $p$  lags ved at benytte en Breusch-Godfrey test, som beskrevet i afsnit 7.2.

### **Antagelse 6: Normalfordelte fejlled**

Afslutningsvis siges modellen at have normalfordelte fejlled, som vil medføre, at stikprøvemåls-fordelingen også vil være det. Antagelse er dog *ikke* en nødvendighed for at sikre, at en OLS estimator er BLUE, hvorfor den heller ikke er en Gauss-Markov antagelse. Den centrale grænseværdidisætning siger for OLS, at *nok* observationer i et datasæt vil aflede en normalfordelt stikprøvemåls-fordeling uagtet fejlledets fordeling. ”Nok” observationer er et vidt begreb, og det kan være nyttigt at undersøge residualernes fordeling ved brug af QQ-plots.

#### **8.8.1 Robuste standardfejl**

Gauss-Markov antagelserne udgør de ideelle betingelser for OLS, men vil i typisk realiteten aldrig være opfyldt perfekt. I en situation med antagelser der ikke er overholdt, vil det dog stadig være muligt at håndtere disse, og stadig have den bedste estimator. Problemet ved overtrædelse af antagelserne er altså ikke forkerte koefficient estimater, men derimod lede til fejlfortolkninger i modellen. Til at tage højde på dette indføres robuste standardfejl i overensstemmelse med (Millo, 2017) og (Manuel, 1987). Mere specifikt indføres det ved robust estimator af kovariansmatricen af koefficienter. Den robuste kovariansmatrice er netop konsistent ved eksistensen af både autokorrelation og heteroskedacitet. Derfor kan brugen af en robust kovariansmatrice indføres til at håndtere overtrædelse af antagelserne om autokorrelation og heteroskedacitet uden at ændre værdien af selve koefficientestimaterne.

## 9 Data

I det følgende afsnit vil processen for indsamlingen af datasættet, som ligger til grund for undersøgelsen blive gennemgået. Herunder følger en beskrivelse af datasættet, foruden en gennemgang af dataspecifikke beslutninger med henblik på fremtidig sortering, filtrering og udvælgelse.

### 9.1 Dataindsamling

Standard & Poor's 500 indeks (SPX) er valgt som datagrundlag for projektet af flere forskellige årsager. Aktieindekset spænder bredt omkring amerikanske industrier og anses for at være en af de mest repræsentative indeks af det amerikanske aktiemarked, foruden at være en af de mest alment fulgte aktieindeks. Indeksets popularitetsniveau ligger derfor også til grund for et bredt dataudvalg. Dette er især relevant for data med ringere grad af tilgængelighed som ESG-scorer. I modsætningen til traditionelle finansielle nøgletal, som offentliggøres i virksomheders regnskab, bliver virksomhederne scoret af uafhængige tredjeparts selskaber. Investorernes efterspørgsel på disse ratings er stigende, hvorfor også virksomhedernes ønske om transparens er vokset. Derfor har et populært indeks som S&P500 naturligvis også en høj prioritet hos analytikerne.

#### 9.1.1 Datafrekvens

Da ESG data bliver udgivet årligt, skulle der i beslutningsprocessen, tages et valg om hvorvidt der skulle hentes data på daglig, ugentligt eller månedlig frekvens. Et af hovedargumenterne for at benytte sig af daglig frekvens er antallet af datapunkter der kan inkluderes i modellerne. Herudover er der også meget variation, der går tabt, hvis der benyttes en højere datafrekvens end dagligt. Dog kan der argumenteres for, at meget af den variation, der er indeholdt i de enkelte måneder, er støj, som ikke tilfører nogen ekstra værdi for potentielle tidsseriemodeller eller multidimensionelle regressioner. Det primære fokus er at teste effekten af ESG variablene og derfor skabe gode rammer for datakvaliteten af disse.

En daglig frekvens af aktiekurser medføre, at variationen potentiel stagner, da det samme datapunkt skal "carries forward" 365 gange. Derfor er der endeligt valgt at føre en primær analyse, hvor en månedlig frekvens af aktiekurserne er benyttet. Samtidig er der dog hentet daglig data, for netop at tjekke robustheden af de potentielle modeller.

### **9.1.2 Finansiel data**

Finansiel data er hentet ved hjælp af Thomson Reuters Eikon database. Her er følgende variable hentet på månedlig basis i perioden fra 1. januar 2005 til d. 31. december 2018, hvilket giver 168 datapunkter for:

- Aktiekurs
- Net indkomst før skat
- Samlede rapporterede aktiver
- Totale forpligtelser

Disse er hentet for de 500 unikke aktier i S&P 500, som fordeler sig på branche-klassifikationsstandarden ”Global Industry Classification Standard” (GICS) sektorer:

- Finans (67)
- Forsyning (28)
- Sundhedspleje (61)
- Telekommunikation (23)
- IT (68)
- Materialer (28)
- Industri (69)
- Energi (28)
- Ejendomme (32)
- Konsumentvarer (33)
- Forbrugsgoder (63)

### **9.1.3 ESG data**

Thomson Reuters Asset4 database er benyttet til at indhente ESG data for de fornævnte selskaber. Al data er indhentet for de første 3 niveauer af decidederede scorer som beskrevet i afsnit 3.4. For det bagvedliggende ESG scoringsdata er der 31 kvantitative og 7 kvalitative variable udvalgt, i henhold til projektets afgrænsning. Et overblik over disse kan ses i tabel 3 og 4. Den nedre del af ’Environmental’ kategorien består af i alt 61 faktorer, hvoraf 33 er udvalgt. Først og fremmest er alle 26 kvantitative faktorer indhentet, mens de 7 kvalitative er udvalgt på baggrund af deres datagrundlag. Fælles for mange af de underliggende faktorer er den ringe kvalitet af data og mangel på observationer. Netop derfor er de få variable med data for hele tidsperioden indsamlet, for at opnå konsistens. De resterende

5 variable: 'bestyrelsens gennemsnitlige siddetid', 'bestyrelsens kønsfordeling', 'bestyrelsens kompenstation', 'senior executives kompenstation' og 'executives kønsfordeling' er indhentet, som en slags kontrol. Dette gøres med udgangspunkt i (Core et al., 1999) og (Velte, 2016), som beskriver henholdsvis effekten af virksomhedsstyring, executives lønkompenstation og virksomheds kønsfordeling i forhold til virksomhedens performance. Fordelen ved netop disse variable er yderligere, at datakvaliteten er høj i forhold til de resterende.

#### **9.1.4 Vedrørende årlige estimer**

Da ESG data kun er udgivet årligt, er der i denne undersøgelse benyttet en kendt samplings strategi kaldet "carry forward". Carry forward metoden replikerer den forrige værdi, til det næste datapunkt i rækken, hvis og kun hvis, netop denne observation er tom. I afsnit 6 er begreberne MAR og MCAR gennemgået. Udover det finder der også en tredje kategori 'Missing Not at Random (MNAR). I denne undersøgelse er der taget udgangspunkt i at ESG data, inden for et bestemt år, tilhører en katogorien MNAR. Dette udgangspunkt skyldes, at netop de årlige mål for de enkelte variable er konsistente over hele den årlige periode. Det vil sige, at denne undersøgelse ikke tager højde for potentielle ændringer inden for de enkelte år. Der skal dog skelnes imellem de månedlige datapunkter, som ikke er tilgængelige i Thomson Reuters database og de manglende årlige estimer af ESG variablene. De manglende årlige estimer antages, alt andet lige, at følge en MAR struktur, mere om dette i afsnit 10.1.

Carry forward metoden er derfor brugt på de månedlige datapunkter fra januar, for at udfylde de manglende datapunkter fra februar til december.

#### **9.1.5 Kontrolvariable**

Tre finansielle kontrolvariable benyttes til at øge robustheden af projektets resultater. Dette er afkastningsgrad, størrelse og finansiel gearing. Disse er udvalgt på baggrund af grundlæggende arbejde indenfor finansiel teori af (Fama and French, 1992) og (Fama and MacBeth, 1973). Disse beskriver faktorer, der har signifikant indflydelse på det forventede afkast af aktier, som senere har dannet grundlag for en lang række af projekter; blandt andre (Chua et al., 2010). Afkastningsgrad, også kaldet totalkapitalforrentning, givet ved net indkomst før skat / samlede rapporterede aktiver, beskrives som forklarende for virksomheders finansielle performance jævnfør (Derwall, 2007). Bevæggrundlaget kommer fra idéen, at virksomheder med højere afkastningsgrad potentielt oplever en højere grad af positivt merafkast, baseret på deres stærkere finansielle performance. Den anden kontrolvariable størrelse beskrives i litteraturen med signifikant forklaringsgrad af virksomheders finansielle performance. Til at kontrollere effekten af størrelsesforskellen mellem virksomheder benyttes logaritmen af virksomheders samlede aktiver. Yderligere beskrives virksomheders økonomiske størrelse med evident indflydelse på deres sociale ansvarlighed (Ullmann, 1985). Større og mere etablerede virksomheder har i højere grad en tendens til at efterleve åbentlyse socialt ansvarlige projekter, da de med generelt større omtale, omdømme og opmærksomhed i højeregrad må efterleve investorernes bæredygtige efterspørgsel (Waddock and Graves, 1997). Endnu en grund til at inddrage størrelse er, at der for større virksomheder eksisterer mere tilgængelig information, og derfor vil effekten af finansielle nyheder være forstørret hos

mindre virksomheder med mindre tilgængelig information og vice versa (Freeman, 1987). Den sidste kontrolvariable, finansiel gearing, defineres som totale forpligtelser / samlede rapporterede aktiver. Ideen er her, at finansiel gearing afspejler virksomheders niveau af risiko, hvor mere risikable virksomheder vil reagere kraftigere ved udsving i aktiemarkederne. Derudover viser et studie, at mindre finansielt gearede virksomheder investerer mere i CSR (Kim et al., 2012). Sluteligt skal det nævnes at der benyttet den samme 'carry forward' metode til at udfylde de manglende månedlige datapunkter for kontrolvariablene.

## Part II

# Resultater

I følgende afsnit vil både del- og hovedresultaterne blive præsenteret. Dette inkluderer en metodisk gennemgang af selve processen og overvejelser der ligger bag disse, inden hovedresultaterne vil blive præsenteret. Først og fremmest vil projektets dataresultater blive præsenteret. Her vil en grundlæggende beskrivelse af samplingsprocessen bringes for at belyse det endelige datagrundlag. Hernæst vil resultaterne af tidsserieanalyse blive præsenteret metodisk. Slutteligt følger resultater af panel data regressionsanalysen. Hertil hører også en metodisk gennemgang.

## 10 Data

I følgende afsnit vil behandling af data blive gennemgået, herunder: sortering, rensning og transformationer. Samtidig vil der blive præsenteret en dataopsummering af de forklarende variable og kontrolvariable. Som udgangspunkt er der hentet 7 kvalitative variable og 49 kvantitative. Dette er inden en metodisk sampling og tjek af eventuelle manglende datapunkter. I modsætning til tidligere ESG analyser, som typisk er lavet på niveau et og to af ESG-scorene, er der her inkluderet en stor del af rådata, altså det data som de overordnede ESG-scorer er konstrueret af. Dette åbner blandt andet op for et større sammenligningsgrundlag og en reel diskussion om validiteten af konstruktionen af ESG scorene. Som benævnt i afsnit 3, har denne undersøgelse sin plads i litteraturen, da der ikke førhen er gennemgået dybdegående analyse af det bagvedliggende data for ESG-scorene. For samtlige ESG variable kan de kategoriske deskriptive statistikker ses i tabel 3 herunder:

Variable	n	m	% total
<b>Grønne bygninger</b>	346	-	-
Sandt	-	2508	40.82%
Falsk	-	3636	59.18%
<b>Miljø distributionskæde politik</b>	339	-	-
Sandt	-	2543	41.65%
Falsk	-	3601	58.35%
<b>Miljø kontroverser</b>	277	-	-
Sandt	-	108	1.76%
Falsk	-	6036	98.24%
<b>Affaldsreducerende initiativer</b>	361	-	-
Sandt	-	3551	57.80%
Falsk	-	2593	32.20%
<b>Vedvarende energiforbrug</b>	114	-	-
Sandt	-	2482	40.40%
Falsk	-	3662	59.60%
<b>Vedvarende og grøn energiforbrug</b>	275	-	-
Sandt	-	773	12.58%
Falsk	-	5371	97.42%
<b>Reduktion af giftige kemikalier</b>	127	-	-
Sandt	-	1084	17.64%
Falsk	-	5060	82.36%

Table 3: Beskrivelse af kategoriske variable

De deskriptive statistikker for de kontinuerlige variable fremgår i tabel 4, hvor variablene er inddelt efter, hvilken niveau de tilhører længst til venstre, hvor der til sidst yderligere er de finansielle variable.

Variable	n	Mean	Median	SD	Min	Max
<b>ESG Niveau 1</b>						
ESG Score	500	58.5475	59.646	17.4945	8.6601	97.8982
<b>ESG Niveau 2</b>						
Enviromental	500	57.3727	59.1285	22.7582	5.8582	99.0872
Social	500	60.5043	61.6034	19.3762	4.8623	99.0307
Governance	500	57.6881	59.5764	20.929	2.8165	99.0561
<b>ESG Niveau 3</b>						
Emissioner	500	58.7325	62.9032	28.2024	0.1953	99.8674
Innovation	500	53.7385	48.8636	25.116	0.2092	99.8
Ressourceforbrug	500	59.5231	64.726	28.869	0.1587	99.8674
CSR Strategi	500	61.4987	66.5335	27.8445	0.1518	99.9817
Ledelse	500	58.8671	62.308	27.2606	0.052	99.9817
Aktionære	500	52.0384	52.894	27.936	0.0743	99.9488
Samfund	500	73.9343	78.9634	20.6764	1.3158	99.8498
Menneskerettigheder	500	55.6047	45.9503	24.5247	8.5938	99.8092
Produktansvarlighed	500	58.5944	57.7708	24.5432	0.6173	99.6774
Arbejdsstyrke	500	56.0048	57.9266	27.3045	0.1684	99.875
<b>ESG Niveau 4</b>						
CO2 Uddeling - Type 1	335	6363736.1418	220432	18058123.6516	0	156300000
CO2 Uddeling - Type 2	328	1157449.0366	336815	2431894.4438	0	28500000
CO2 Uddeling - Type 3	266	12067398.5893	114853.5	58662204.9793	0	1005311900
CO2 Uddeling - Total	350	7868667.7564	815457.5	20289797.1636	0	166000000
CO2 Uddeling - Estimeret	496	4511315.0616	263759	14820978.1137	77	166000000
CO2 Uddeling - / USD	350	0.0006	0	0.0016	0	0.0155
Electricitet - Produceret	103	121389022.3411	3400000	206364210.0803	0	939852000
Electricitet - Købt	264	14869256.429	2504726.7585	65112571.9999	0	953395200
Energi - Direkte Produceret	116	16902774020.3694	5500000	272432416254.564	0	5261041607760
Energi - Direkte Købt	296	66620191.8994	6203292.436	214806333.2283	158	2325653517
Energi - Totalt Forbrug	296	50192577.6702	6300000	159022687.2405	27.97	2270173320
Affald - Farligt	166	1839083.6943	1801.5	19511263.4888	0	293463600
Affald - Ikke Farligt	156	6110405.4537	46712.826	44675294.8793	0	459433340
Affald - Farligt / USD	151	0.0003	0	0.0027	0	0.0533
Affald - Genbrug / Total	225	0.0008	0	0.0077	0	0.1341
Affald - Total Genbrug	232	189712.5231	25631.378	686432.5928	3.791	9775537.432
Affald - Total	225	5974307.6958	58722	52891038.2594	67.803	741824000
Affald - Total / USD	213	0.5876	0.6263	0.2464	0	1
Ozon Nedbrydende Stoffer	46	10368.3545	1.2	55430.6947	0	385710
Vedvarende Energi - Produceret	122	9215554.2702	178200	22678076.9619	0	167409954
Vedvarende Energi - Købt	155	7953003.9235	559521	70683657.7854	0	1778400000
Vedvarende Energi - Total	154	5533992.7114	414684	16283161.1131	0	167409954
Vedvarende Energi - / Forbrug	132	0.1816	0.0952	0.2169	0	1
Vandforbrug - / USD	274	0.0448	0.0004	0.1857	0	2.6251
Vandforbrug - / Genbrug	211	58.5781	62.98	24.4381	0.004	99.97
Vand fjernet- Total	274	582100782.1058	6197596.5565	2893091106.5325	627.5	39184542662
Bestyrrelsen - Gns. Siddertid	497	9.0718	8.7826	3.4509	0	28.8611
Bestyrrelsen - Kønsfordeling	497	17.3423	16.6667	9.5754	0	62.5
Bestyrrelsen - Kompensation	497	2946998.4489	2331850	25825655.6821	34000	1929824018
Senior Executive - Kompensation	497	30811241.464	24461581.5	27526331.4008	0	900144285
Executive - Kønsfordeling	497	13.4306	12.5	12.0965	0	100
<b>Finansielle Variable</b>						
Aktiekurs	500	63.6812	43.6705	101.2803	0.8521	3508.22
Net Indkomst Før Skat	500	567167214.825	21000000	1765919447.1379	-60556000000	126603000000
Samlede Rapporterede Aktiver	500	55885696194.6754	13046391500	195389024606.528	1000	2615183000000
Totale Forpligtigelser	500	44355360283.1236	7862127000	175787048781.576	4570000	2356227000000

Table 4: Beskrivelse af kontinuerlige variable

## 10.1 Sampling

For at gøre datasættet ensartet og optimere kvaliteten af datapunkterne, laves en dybdegående analyse af kvaliteten af de enkelte ESG variable. På baggrund af analyse af forskellige opsummeringstabeller i R, kan det ses at datakvaliteten af ESG-scorer variablene generelt er høj, hvorimod kvaliteten generelt er lav for miljø- og rådata. For at optimere data i forhold til at kunne besvare både hovedhypotesen og de enkelte underhypoteser, er det derfor nødvendigt at lave en metodisk gennemgang af kvaliteten af de enkelte variable. Der ønskes, som nævnt, i denne undersøgelse at finde effekten af de enkelte determinanter og ikke blot benytte de overordnede scorer. Det ses ydermere at kvaliteten af data er lav, er der lavet en model til at kontrollere datakvaliteten, for at inkludere det bedst mulige miljø- og rådata.

Samplingsprocessen er udarbejdet gennem mange iterationer og derfor er nedenstående fremgangsmåde inkluderet for at lette forståelsen. Der er forsøgt at visualisere denne samplingsprocessen i tabel 5.

I software programmet 'R' er det muligt at summere manglende datapunkter, hvorfor der er itereret igennem samtlig miljø- og rådata, så antallet af valide datapunkter kan findes, givet et predefineret cut-off niveau. Modellen er her bygget til at tage højde for et trade-off imellem mængden af kvalitative variable overfor mængden af virksomheder, som kan inkluderes i analysen.

Ud fra en "trial and error" forsøgs metode er der besluttet at det maksimale antal manglende datapunkter i en given variable ikke må overskride 50%. Dette skyldes flere ting. For det første er kan det volde estimations problemer når de manglende datapunkter udgør mere end 50%. Samtidig ses det at når tolerancen bliver sat ned mod 25% falder antallet af brugbare virksomheder til seks. Videre tages der høje for målet med samplingsprocessen, nemlig at bibringe det bedst mulige sammenligning grundlag til fremtidig analyse. Derfor ønskes der hverken 200 virksomheder, som har variable med op til 95% mangelnde datapunkter, mens der heller ikke ønskes en smal sample på seks virksomheder med komplet data.

Slutteligt er der fjernet 10 af de 37 kvalitative variable (reelt skete dette via en trial and error proces i samhørighed med valg af cut-off niveau), da disse var af svingende kvalitet. Her er der taget højde for, om der er nok datapunkter til at lave et paneldatasæt med en rimelig population.

I tabel 5 er der forsøgt at skabe et overblik over de udvalgte selektions muligheder, hvor det er muligt at se, hvordan trade-off balancen mellem antal virksomheder og antal kvalitative variable er. Tabellen viser, hvordan antallet af variable gradvist er nødsaget til at blive frasorteret, jo strengere tolerance der er på procentvis manglende datapunkter. I kolonnen længst til venstre er de forskellige variable listede, hvor der ud fra disse ses det tilhørende antal virksomheder givet et specificeret cut-off niveau. Eksempelvis har variablen 'vedvarende energi / forbug' henholdsvis nul og to virksomheder, som indeholder op til 5% og 10% datapunkter.

Cut-off niveau	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
Vedvarende Energi - Total	0	0	0	0	0	0	0	0	0	0
Ozon Nedbrydende Stoffer	0	0	3	3	6	8	8	11	12	12
Vedvarende Energi - / Forbrug	0	2	6	6	9	14	14	18	22	22
Vedvarende Energi - Købt	0	3	7	7	12	16	16	23	27	27
Vedvarende Energi - Produceret	3	5	7	7	12	15	15	18	23	23
Electricitet - Produceret	5	8	10	10	13	16	16	19	23	23
Energi - Direkte Produceret	4	7	11	11	19	22	22	24	29	29
CO2 Udledning - Type 3	0	4	14	14	37	64	64	91	107	107
Vandforbrug - / Genbrug	8	12	22	22	44	60	60	76	91	91
Affald - Ikke Farligt	11	16	24	24	37	45	45	55	66	66
Affald - Farligt / USD	15	21	30	30	48	61	61	80	93	93
Affald - Total Genbrug	14	20	30	30	49	60	60	84	97	97
Affald - Farligt	17	24	37	37	48	58	58	66	78	78
Affald - Total / USD	17	24	37	37	46	56	56	65	75	75
Electricitet - Købt	18	30	41	41	54	71	71	84	102	102
Affald - Genbrug / Total	27	38	47	47	64	78	78	98	107	107
Affald - Total	27	38	47	47	65	79	79	98	107	107
Energi - Direkte Købt	24	38	50	50	80	101	101	119	144	144
Energi - Totalt Forbrug	29	38	58	58	91	109	109	132	150	150
CO2 Udledning - Type 2	30	53	65	65	89	124	124	157	180	180
Vand - Total Tilbagetrækning	38	50	66	66	85	99	99	118	132	132
Vandforbrug - / USD	38	51	67	67	85	99	99	118	132	132
CO2 Udledning - Type 1	38	60	73	73	102	133	133	167	188	188
CO2 Udledning - Total	62	89	108	108	136	169	169	202	227	227
CO2 Udledning - / USD	62	89	108	108	135	168	168	202	227	227
CO2 Udledning - Estimeret	0	0	275	275	396	429	429	442	451	451
Bestyrrelsen - Kønsfordeling	80	220	328	328	374	410	410	428	439	439
Bestyrrelsen - Gns. Siddetid	321	351	368	368	397	427	427	439	448	448
Bestyrrelsen - Kompensation	323	351	369	369	399	429	429	441	450	450
Senior Executive - Kompensation	275	342	369	369	398	427	427	441	451	451
Executive - Kønsfordeling	325	352	370	370	399	429	429	441	450	450

Table 5: Test af antal manglende datapunkter

### Afsluttende bemærkninger om samplingsprocessen

Udover sampling af variable, og virksomheder, er der også tjekket, hvorvidt yderligere reduktion af tidsdimensionen kunne forbedre datakvaliteten. Dette forekommer, hvis de manglende observationer er grupperet i starten eller slutningen af tidsserierne. Det viser sig at det ikke er muligt at reducere antallet af observationer til fordel for at inkludere flere forklarende variable eller virksomheder. Dette skyldes, at de manglende datapunkter er pænt fordelt over hele tidsperioden fra 2005 til 2018.

Den endelig sample indeholder 45 af de 500 virksomheder, med tilhørende ESG-score, 3 søjlescorer, 10 under-scorer og 21 af de 39 miljø- og rådata variable, hvor 6 af dem er kvalitative, mens de resterende 15 er kvantitative. Denne endelig sample er, som sagt, baseret på frasortering med et 50% cut-off niveau. Den opfattende samplingsprocess tyder på at der er en basal mangel på datakvalitet i ESG-scorene.

## 10.2 Imputering af manglende data

Grundet store mængder manglende data, er det nødvendigt at imputere manglende datapunkter. Et problem ved datasæt med mange manglende datapunkter og flere variable, er at der er kan være meget få ”overlap” i datapunkterne. Herved kan en proces, hvor der slettes rækker med manglende datapunkter medføre at man slutteligt er nødsaget til at frasortere mange af de ”gode” datapunkter.

Derfor er det en fordel at imputere manglende datapunkter for at undgå forøget bias, der kan forekomme ved at fjerne de oprindelige datapunkter. Dette bias forekommer da de fjernede datapunkter kunne have haft en signifikant indflydelse på eventuelle modeller. I denne analyse udgør data imputeringen derfor en stor del af det endelige data for mange af miljø- og rådata variablene.

For at imputere datapunkter er pakken ’AMELIA II’ benyttet i R. Fordelen ved at benytte netop denne pakke er, at det er muligt at specificere grænser for de imputerede datapunkter. Ved at benytte disse grænser er det muligt at sørge for, at de imputerede datapunkter ligger i mellem minimum og maksimum værdierne af de enkelte variable for hver virksomheder. Dermed undgås der at middelværdien bliver ”skubbet” for meget i den ene eller den anden retning. Yderligere pointer om fordele ved data imputering til tidsserie analyse er bragt i afsnit 6. Der antages ligeledes, at MCAR og normalfordelingsantagelserne er overholdt for data.

Da der er benyttet multipel imputering, er der imputeret henholdsvis 3 datasæt, som fremover vil blive refereret til som imputering(1), og yderligere 3 datasæt til fremtidig analyse, benævnt i denne undersøgelse som imputering(2). Imputering(1) er lavet på baggrund af 3 bootstrappede datasæt, hvor de 3 datasæt er blevet lagt sammen, og de gennemsnitlige estimerer af de manglende datapunkter fundet. Imputering(1) er lavet med formål til både univariat tidsserie analyse og multidimensionel regressionsanalyse. Imputering(2) er lavet med formål til robusthedstjek, hvor ideén er, at imputering(2) kan bruges til at tjekke robustheden af imputering(1). Imputering(2) er konstrueret på samme facon som imputering(1). I tabel 6 ses en opsummeringstabell af undervariablene før data imputering. For kortfattedhed er der én enkelt tabel af data opsummering for de imputerede datasæt, nemlig for imputering(1), som ses i tabel 7.

	Middelværdi	Median	Standard Afvigelse	Minimum	Maksimum
CO2 Udledning Total	3165434.5792	994185	6919737.8141	45260	68284041
Energi - Totalt Forbrug	32596610.5567	10363120.2	83033396.5424	9707	868311006
CO2 Udledning / USD	1e-04	0	2e-04	0	0.0011
Vandforbrug / USD	0.0035	4e-04	0.0137	0	0.1172
Affald - Genbrug / Total	0.0019	0	0.0128	0	0.1341
Affald - Total Genbrug	118098.9897	39261.13	243122.0768	178	3152000
Affald - Total	13428236.7977	76248.97	86893072.5238	578	741824000
Vand - / Genbrug	58.0249	63	25.1556	0.004	99.829
Vand - Total Tilbagetrækning	44860274.0558	8517000	164330799.4448	4107	1.738e+09
Affald - Total / USD	0.5805	0.629	0.2524	0	1
Bestyrrelsen - Gns. Siddetid	8.5359	8.7273	2.296	1.9286	14.8636
Bestyrrelsen - Kønsfordeling	20.5755	20	8.2384	0	45.4545
Bestyrrelsen - Kompensation	3019546.969	2893020	1255791.734	297941	14685739
Senior Executive - Kompensation	38763995.0461	33804404	21952296.5341	1564816	189004246
Executive - Kønsfordeling	15.7039	15.3846	10.9109	0	50
n	45	45	45	45	45

Table 6: Data opsummering

	Middelværdi	Median	Standard Afvigelse	Minimum	Maksimum
CO2 Udledning - Total	3308351.7699	1062500	6836610.7621	45260	68284041
Energi - Totalt Forbrug	36260741.893	10693298.558	90629072.7501	9707	868311006
CO2 Udledning / USD	1e-04	1e-04	2e-04	0	0.0011
Vandforbrug / USD	0.0041	6e-04	0.0133	0	0.1172
Affald - Genbrug / Total	0.0028	0	0.0136	0	0.1341
Affald - Total Genbrug	146950.2882	54303.5	243201.2379	178	3152000
Affald - Total	18858092.3459	99434	89817084.8575	578	741824000
Vand - / Genbrug	57.6322	61.5593	24.6027	0.004	99.829
Vand - Total Tilbagetrækning	49520659.9607	10249088.91	158761716.3718	4107	1.738e+09
Affald - Total / USD	0.5777	0.6174	0.2469	0	1
Bestyrrelsen - Gns. Siddetid	8.5327	8.7273	2.3013	1.9286	14.8636
Bestyrrelsen - Kønsfordeling	20.5475	20	8.2617	0	45.4545
Bestyrrelsen - Kompensation	2969788.4013	2847564.5	1275386.3047	297941	14685739
Senior Executive - Kompensation	38787765.086	33804404	22061890.7822	1564816	189004246
Executive - Kønsfordeling	15.7039	15.3846	10.9109	0	50
n	45	45	45	45	45

Table 7: Data opsummering af imputering (1)

Som det fremgår af ovenstående tabeller er der minimal forskel på middelværdi, standardafvigelse og median, før og efter imputering(1). Dette tyder på robuste imputeringer af datapunkter. Imputering(1) kan derfor sagligt benyttes til fremtidige analyser. Ydermere ses et opsummeringstabelf af imputering(2) i appendix på side 100.

## 11 Tidsserie analyse

I dette afsnit vil der først blive introduceret forskellige antagelser, hvilke skal overholdes eller tages in mente når der skal modelleres tidsserier. Hernæst vil der blive behandlet data, for at se om stationaritetsbetingelsen er overholdt blandt de eksterne kovariate. Dernæst vil der blive gennemgået selve estimeringsprocessen, og de forskellige overvejelser der ligger bag de valg der er truffet i denne. Herefter præsenteres modellernes og resultaterne vil blive gennemgået i lineær facon, for at tjekke hvorvidt hovedhypotesen og underhypoteserne kan blive tildels be- eller afkraeftet. Hernæst tolkes der hvordan datagrundlaget har præget selve analysen. En modelsammenligning ved brug af MAE vil blive gennemgået for at teste validiteten af de endelige modeller, samt yderligere at fremme sammenligningsgrundlaget. Slutteligt, vil der blive samlet op på hovedpointerne fra analysen, hvorefter der med fordel kan pointeres til den videre proces.

### Note omkring masseestimering af ARMA/GARCH modeller

Det skal pointeres at der i denne del af undersøgelsen arbejdes med et førhen uset antal forskellige ARMA/GARCH modeller i ESG litteraturen. Dette har flere implikationer. For det første vil der i dele af næste afsnit blive estimeret og tjekket mere  $15 * 45 * 16$  modeller udelukkeende ved brug af miljø- og rådata som eksterne kovariate. Det betyder endvidere, at det er undersøgt mere end 10,080 modeller i dele af analysen, hvorfor det i praktis ikke er muligt at lave ekstrem finpudsning af de enkelte modeller. Dog er der i forhold til bedst mulig praksis udført modelkontrol og tjek over en bred kam, netop for at estimere påholdende modeller. Der tages derfor forbehold for at enkelte eller flere modeller, kan vise sig at være sub-optimale.

### 11.1 Fordelingsantagelser

Finansiel data bliver ofte og traditionelt antaget at være normalfordelt. Det er også blevet vist at afkast teoretisk følger en log-normalfordeling (Black and Scholes, 1973). I praksis viser det sig tit at data har tungere haler end man ser i en log-normalfordeling. I figur 1 ses fire histogrammer for månedlige log-afkast på tilfældigt udvalgte S&P500 virksomheder. Ud over histogrammet ses to kurver. Den røde er densiteten af den specifikke virksomheds log-afkast, mens den sorte er en normalfordelt tæthed med middelværdi og varians svarende til virksomhedens. Det ses at log-afkastet ikke ser normalfordelt ud i alle tilfælde, da flere plots har bredere haler og større sandsynlighedskoncentration i midten end den tilsvarende normalfordeling.

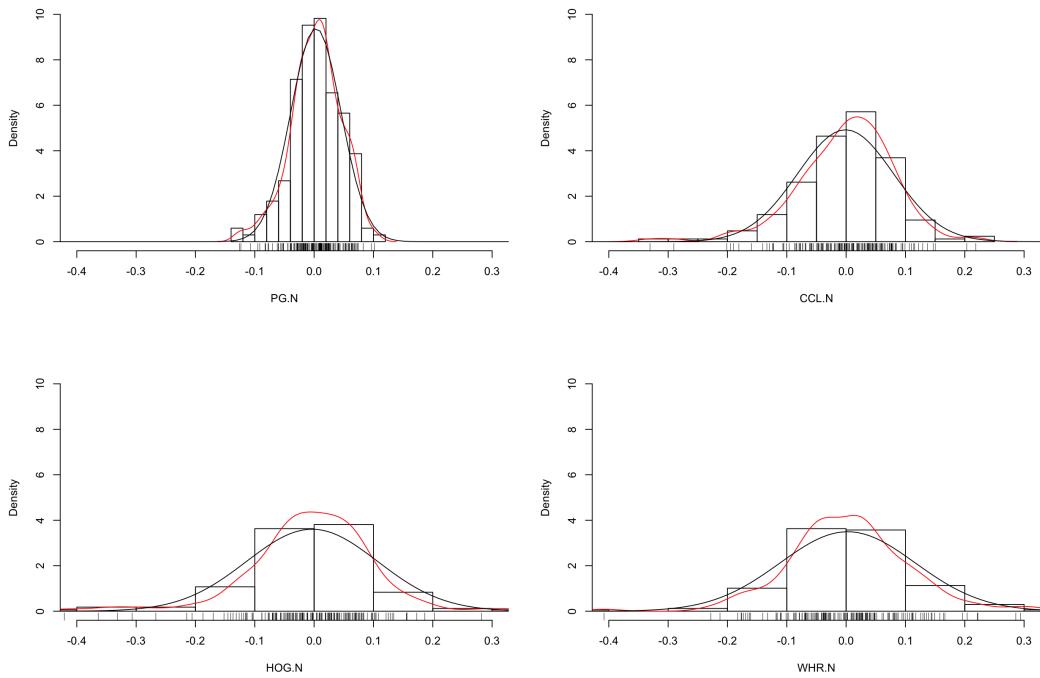


Figure 1: Tæthed af månedlige log afkast

Dette ses også svagt i QQ-normal-plots på figur 2 på side 44, hvor halepunkterne ligger skævt af den teoretiske fordeling. Lignende tendenser tegner sig i figurene for de resterende virksomheder. Fordelings- og QQ-plots for de resterende virksomheder kan findes i appendix på side 102. Tabel 8, på side 45, viser momenterne for de samme fire virksomheder.

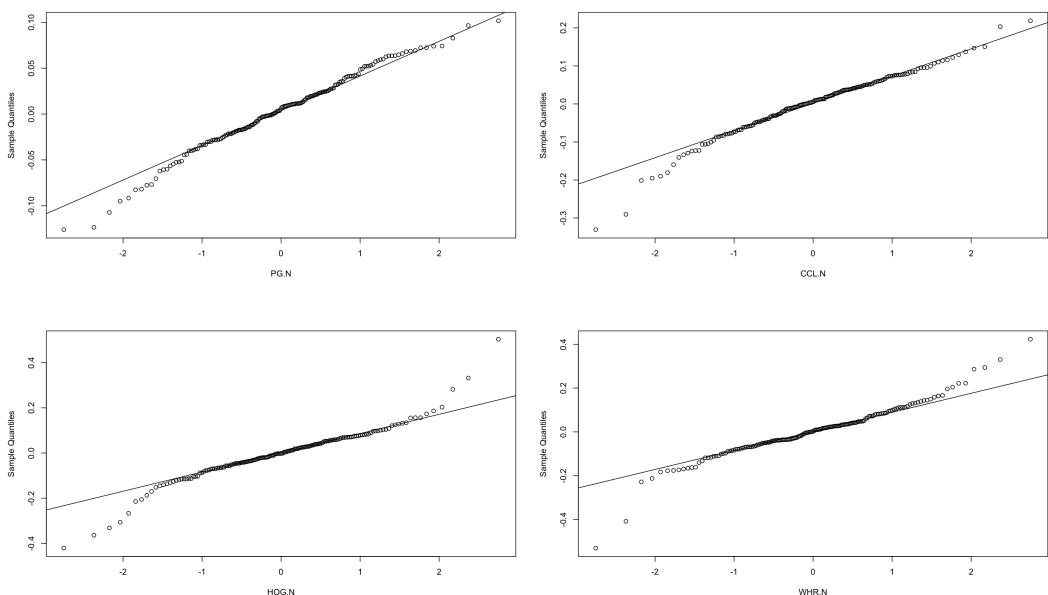


Figure 2: Normal Q-Q Plot

Det bekræftes ligeledes i momenter for de 45 virksomheder, at flere af fordelingerne er venstre skævvredet og sandsynlighedkoncentrationen, kurtosis, er større end tre, som kendes fra normalfordelingen. Det fremgår præcist at af de 45 virksomheder der arbejdes med, har 44 af dem log-afkast der følger en venstre skævvredet fordeling. Tester man momenterne ved brug af en Jarque-Bera test, denne ses i tabel 37 på side 112 i appendix, fremgår det også at momenterne for skævhed og kurtosis generelt ikke ligner en normalfordeling for 34 ud af de 45 virksomheder. De tunge haler og den generelle skævhed kunne tyde på at det er fordelagtigt at benytte en t-fordeling, som også er vist at fitte finansielle data godt. Dette skyldes at skævheden og formen kan estimeres i denne fordeling (Theodossiou, 1998). Ligeledes findes der andre fordelinger der kunne passe data godt, så som generaliseret fejl-fordeling.

Ticker	Middelværdi	Varians	Skewness	Kurtosis
PG	0.00305	0.00180	-0.37310	3.28413
CCL	-0.00093	0.00657	-0.73202	4.93194
HOG	-0.00343	0.01225	-0.03670	7.09316
WHR	0.00259	0.01303	-0.28263	6.91379
Median <sub>45</sub>	0.00377	-0.51333	-0.51333	4.59668

Table 8: Momenter

### Estimater af ikke-normalfordelt data

Da Quasi-maximum likelihood estimatoren (QMLE), som defineret i (White, 1982) er vist både at være konsistent og asymptotisk normal (Yao and Brockwell, 2006) er der valgt at arbejde videre med en normalfordelings specifikation, trods af ovenstående resultater. Dette valg understøttes ligeledes af at Jarque-Bera testen ikke er konsekvent for de 45 tidsserie.

## 11.2 Stationaritet i finansiel data

For at tjekke stationaritet, benyttes en Augmented Dickey-Fuller test (ADF), som beskrevet i afsnit 7.1. I tabel 38 på side 113 ses resultaterne af ADF testen. Det fremgår at nulhypotesen kan afvises for alle 45 virksomheder, og derfor egner log-afkast sig til at blive modelleret som en stationær tidsserie i alle 45 tilfælde.

## 11.3 Stationaritet i de eksterne kovariate

For at sikre stationaritet i de eksterne kovariate og opnå konsistente parameterestimater, tjekkes først plots af udvalgte kovariate:

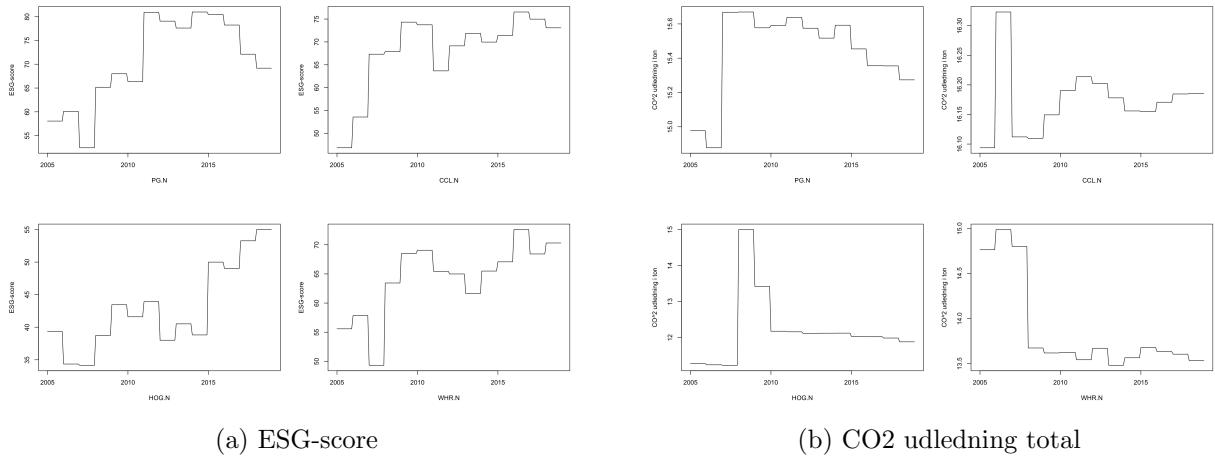


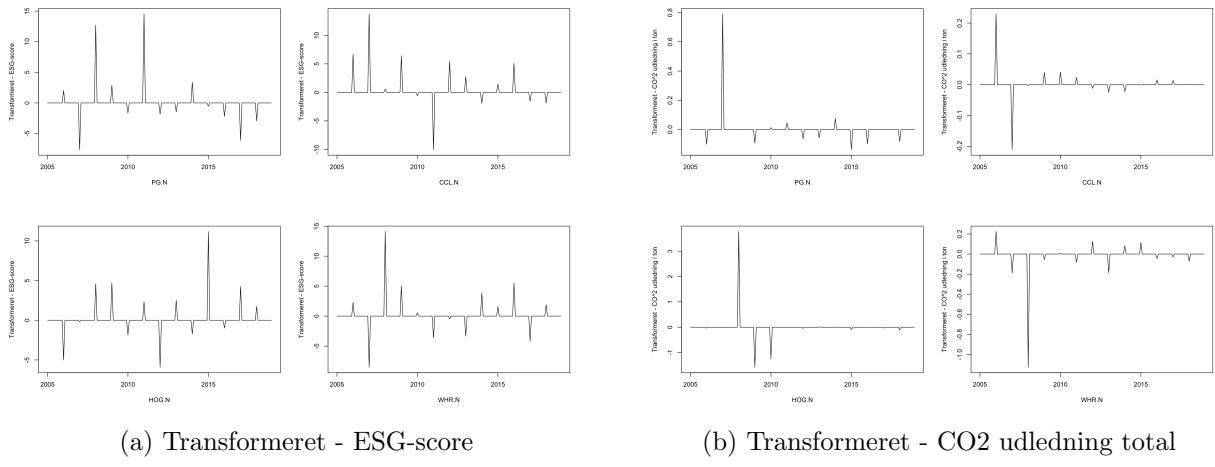
Figure 3: Eksterne kovariate

Det fremgår tydeligt at de to plottede variable ikke er stationære. Det er altså nødvendigt at teste om alle variablene overholder stationaritetsbetingelserne. De transformerede variable for WHR, kan ses i appendix på side 115. Der tjekkes igen for stationaritet ved brug af en ADF test. For kortfattethed er denne kun bragt for PG, CCL, HOG og WHR:

	PG	CCL	HOG	WHR
Test størrelse	-2.23081	-3.75155	-3.04024	-1.60484
P-værdi	0.47994	0.02307*	0.14208	0.74123

Table 9: Augmented Dickey-Fuller test af ESG-scoren

Det fremgår at 'CO<sub>2</sub> emission' ikke kan modellers stationært. Det kræves altså enten en differentiering eller en transformation af denne. Der er her valgt at transformere alle de forskellige scorer på samme måde som aktiekursen. Den nye variable til tiden  $t$  kommer til derfor til at være  $ESG_{transformeret} = ESG_t - ESG_{t-1}$ . For de kvantitative variable i niveau 4 er den logaritmiske ændring brugt på præcis samme måde som for aktiekursen,  $CO_2_{(transformeret)} = \log(CO_2(t)/CO_2(t-1))$ .



(a) Transformeret - ESG-score

(b) Transformeret - CO2 udledning total

Figure 4: Transformerede Eksterne kovariate

Det bekræftes ligeledes af en ADF test, se tabel 15, at de transformerede variable er stationære. Nedenstående fremgår ADF-testen for de fire virksomheder for 'CO<sub>2</sub> emission'. Nulhypotesen afvises på baggrund af testen på et 1%’s signifikans niveau. Desuden har transformationerne af de eksterne kovariate indflydelse på foltolkningsevnen af fremtidige resultater. Dette er et nødvendigt trade-off, for at få konsistente parameterestimater. Alle ADF test resultater af de transformerede variable kan findes på side 118 i appendix.

	PG	CCL	HOG	WHR
Test størrelse	-5.29807	-5.0921	-5.09916	-5.2783
P-værdi	0.01**	0.01**	0.01**	0.01**

Table 10: Augmented Dickey-Fuller test af den transformerede ESG-score

I følgende afsnit, vil de transformerede variable blive omtalt på normal vis. Det vil altså sige at refereres der til 'ESG' er det den transformerede ESG-score der refereres til.

### Kategoriske variable i modellerne

Da de kategoriske variable udviser meget lidt variation gør manglen på datakvaliteten at de ikke er særlig spændende at modellere univariat. Derfor er de kategoriske variable udeladt i følgende analyse, mere omkring dem følger senere i paneldata regressionsanalyse afsnittet.

#### 11.4 Autokorrelation i finansiell data

Autokorrelation i data kan blandt andet tjekkes ved at plotte ACF’erne. Disse fremgår for de samme fire tidsserier af henholdsvis PG, CCL, HOG og WHR, i figur 5, på side 48. Her er de stiplede blå linjer den kritiske værdi for et 5% signifikansniveau i en Ljung-Box test. Det fremgår at flere af tidsserierne, faktisk ikke viser store tegn på autokorrelation for få lags. Her er ACF plots lavet med 30 lags, for at

illustrere både kortsigtet og langsigtet autokorrelation. Ljung-Box testen kan ikke afvise at tidsserierne følger en hvidsstøjsproces for 36 af de 45 virksomheder. Dette tyder på at ikke alle tidsserier vil give lige gode ARMA modeller. For at teste hypotesen omkring ESG variablenes indflydelse på tidsserierne, modellers der dog ARMA modeller for samtlige virksomheder. Dette skyldes to primære argumenter. For det første ønskes der at bideholde et så stort sammenligningsgrundlag som muligt. For det andet er det vigtigt for en saglig analyse, at der netop bliver modelleret autokorrelation i de tidsserier der viser tegn på det. Ljung-Box testen ses på 41 på side 119 i appendix.

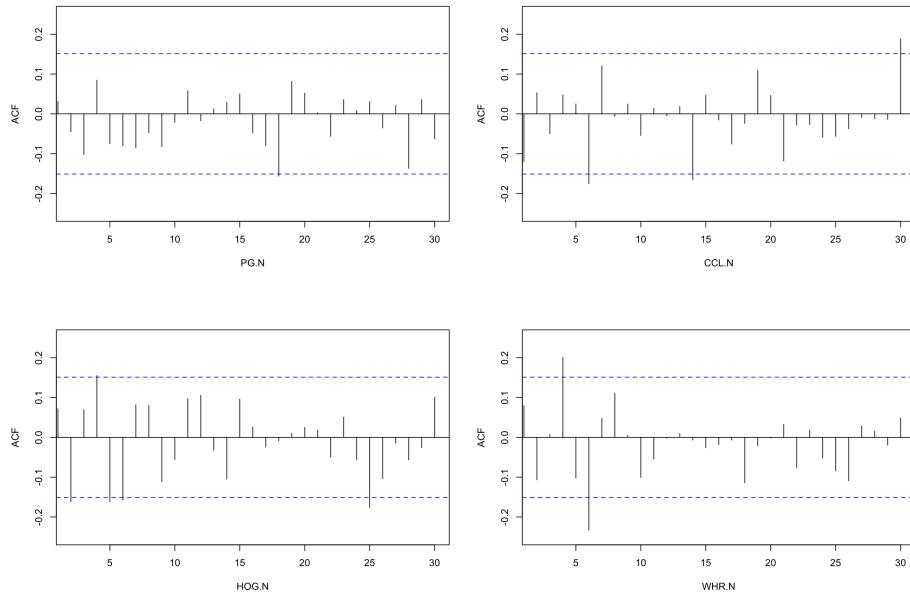


Figure 5: ACF Plot

## 11.5 Heteroskedasticitet

Som belyst i teoriafsnittet, er det typisk for finansielle data at demonstrere volatilitetsklyngedannelse. Nedenstående, i figur 6, er der plottet log-afkast for flere af virksomhederne. Det ses at både, WHR, CCL og HOG viser tegn på kortvarig volatilitetsklyngedannelse omkring 2015. Netop heteroskedasticitet ønskes modelleret, hvorfor ARMA/GARCH modeller egner sig godt til data. Ingen refereres der til appendix på side 120 for, hvor de resterende log-afkast kan ses.

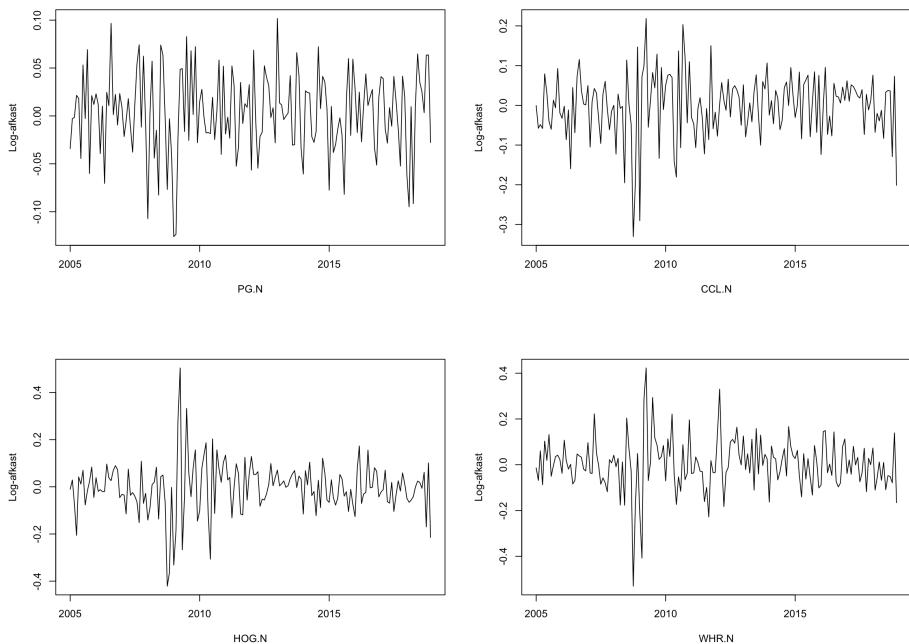


Figure 6: Månedlige log afkast

Det ses her at størstedelen af tidsserierne sandsynligvis ikke udviser tegn på kompleks heteroskedasititet. Ud fra plots er det ikke muligt at konkludere om største delen af tidsserierne egner sig til komplekse ARMA/GARCH modeller. For de tidsserier der viser tegn på klyngedannelserne er det dog praktisk at undgå at benytte rene ARMA-modeller. Dette begrundes på baggrund af teorien i afsnit . (Hansen and Lunde, 2005) lavede en dybdegående empirisk benchmarkanalyse med 330 modeller, hvori de ikke fandt nogen tegn på at GARCH(1,1) performer værre end mere komplekse volatilitetsmodeller for finansielle data. På baggrund af deres omfattende undersøgelse og et ønske om at estimere påholdende modeller, er der i følgende analyse benyttet udelukkende GARCH(1,1) modeller. Ingen refereres der til at effekten af de eksterne kovariate er i fokus, hvorfor netop påholdende modeller foretrækkes.

### Afsluttende kommentar

Det fremgår altså at det transformerede finansielle data egner sig til modellering ved brug af ARMA-X/GARCH modeller. For at besvare de opstillede hypoteser vil disse model derfor blive modelleret med henblik på at tjekke effekten af ESG-data på netop log-afkast. Estimationsprocessen gennemgås i følgende afsnit.

### 11.6 ARMA-X/GARCH modellerne

ARMA-X/GARCH estimation kan udføres både simultant og i en to-delt estimationsmetode. I en to-delt fremgangsmåde, estimeres ARMA parametre først, hvorefter GARCH parametre estimeres ved brug af de valgte ARMA parametre. En af de større ulempes ved en to-delt estimationsproces er dog

at de fundne parametre kan vise sig at afvige fra den ”rigtige model”. Det er dog vist at estimationen ved brug af en to-delt metode, vil give parameterestimater der er approksimativt tæt på de rigtige parametre. Dette gælder specielt for moderate til store samples (Tsay, 2010). I følgende er der benyttet en ”simultan” estimationsmetode da der på forhånd kendes parametrene for GARCH modellen.

### Modelvalg og valg af informationskriterie

Der er estimeret ARMA-X/GARCH modeller med forskellige modelspecifikationer, for at udvælge modellen med bedst fit. Her viste initiale undersøgelser, lavet med BIC og AICc, at flere af tidsserierne ville blive reduceret til ARMA(0,0). Dette skyldes med stor sikkerhed at både AICc og BIC straffer på baggrund af antal parametre, se appendix for AICc af WHR 42. For at undgå underfitting er der sluteligt valgt at fremtidige fits vurderes ud fra AIC alene. Der fremgår et eksempel i tabel 11 for virksomhed HOG, med ESG-scoren som ekstern kovariat. Selektionen på baggrund af AIC er begrænset til maksimalt 3 AR-lags og 3-MA lags. Denne restriktion er indført for at undgå overfitting, samt ønsket om påholdende modeller. Til modellering af ARMAX(p,q)/GARCH(1,1) modellen er funktionen ’ugarchfit’ brugt i R.

Lags	q = 0	q = 1	q = 2	q = 3
p = 0	-1.789449	-1.773176	-1.746623	-1.725326
p = 1	-1.774108	-1.743608	-1.717862	-1.717725
p = 2	-1.744814	-1.716146	-1.752162	-1.673851
p = 3	-1.732853	-1.729792	-1.675770	<b>-1.797841</b>

Table 11: AIC for ARMA(p,q)/GARCH(1,1) modeller

Lags	q = 0	q = 1	q = 2	q = 3
p = 0	<b>-1.899424</b>	-1.884746	-1.876789	-1.874086
p = 1	-1.885678	-1.873773	-1.866622	-1.885080
p = 2	-1.874979	-1.864906	-1.879517	-1.859801
p = 3	-1.881613	-1.897148	-1.861720	-1.882387

Table 12: BIC for ARMA(p,q)/GARCH(1,1) modeller

For virksomhed HOG ses det at en ARMA(3,3) ville være optimal på baggrund af AIC. Derimod fremgår det at ARMA(0,0) er den optimale modelspecifikation ved brug af BIC som informationskriterie. Som beskrevet i afsnit 7.12 har det altså også i praksis en signifikant effekt på modelspecifikationer, om man benytter det ene eller det andet informationskriterie. Der kunne argumenteres for at begge informationskriterier kunne bruges i en samhørighed. Omfanget af denne analyse tillader ikke enkeltstående vurderinger på baggrund af antallet af modeller der skal estimeres. I tabel 13 ser vi at ARMA(2,3) er den mest favorable model på tværs af virksomhederne, når der modelleres med ESG som ekstern kovariat. Det ses også at de komplekse modeller generelt har lavest AIC, hvilket tyder på at tilførelse yderligere parametre kunne forbedre AIC. Da der ikke ønskes yderlige kompleksitet er valget, af informationskriterie og maximalt antal lags, indtil videre tilfredsstillende.

Lags	$q = 0$	$q = 1$	$q = 2$	$q = 3$
$p = 0$	1	0	0	0
$p = 1$	0	2	0	0
$p = 2$	0	4	8	12
$p = 3$	0	0	9	9

Table 13: Antal virksomheder med optimal modelspecifikation

### Modelselektion i ARMA/GARCH modellerne

Da formålet med analysen er at fortolke på effekten af de eksterne kovariate og ikke prædiktion, modelleres ARMA(p,q)-X/GARCH(1,1) modellerne således at de tre forskellige søjle-scorer er inkluderet enkeltvis som eksterne kovariate. Det samme gør sig gældende for både under-scorene og miljø- og rådata variablene. Det betyder altså at der først modelleres 3 modeller for hver virksomhed, en med hver søjle, yderligere 10 modeller for hver virksomhed ved brug af under-scorene og slutteligt 15 modeller for hver virksomhed ved brug af miljø- og rådata. Dette udføres således at effekten af hver ekstern kovariat kan isoleres. Det skal nævnes at dette valg potentiel har signifikant indflydelse på de kommende resultater. Dette skyldes at der potentielt kunne opnås modeller med bedre fit, hvis hver enkelt virksomheds log-afkast blev fittet med de optimale eksterne kovariate. Her kan det dog diskuteres, hvorvidt det er vigtigt at kigge på selve fittet, i forhold til AIC, eller om det er vigtigt at kigge på selve effekten af de enkelte kovariate. P-værdierne fra en t-test benyttes til at vurdere, hvorvidt de eksterne kovariate har nogen basis for at blive inkluderet i tidsseriemodellerne. Fordelen ved t-testen frem for AIC er at der bliver valgt modeller på baggrund af de eksterne kovariates signifikans.

I tabel 14 ses et udsnit af ARMA-X/GARCH(1,1) modellerne, for tidsserie HOG, med miljø- og rådata.

Ekstern kovariat	$\mu$	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	xreg	$\omega$	$\alpha_1$	$\beta_1$
CO <sub>2</sub> udledning total	-0.00078	-0.75588	-0.91356		0.80364	0.99597	-0.04325	0.00073	0.36673	0.60314	
Energi forbrug total	-0.00229	-0.5605	0.05164	-0.16727	0.69052		0.02365	0.00095	0.34574	0.588	
CO <sub>2</sub>	-0.00206	-0.58339	0.05239	-0.14204	0.69758		-0.02136	0.00098	0.34745	0.58656	
Vandforbrug total/USD	-0.00198	-0.57143	0.05581	-0.15307	0.6937		0.00762	0.00093	0.33782	0.59775	
Affald genbrug/total	-0.00251	-0.11704	-0.90865	0.08518	0.24398	0.91156	0.00602	0.00079	0.25047	0.69058	
Affald genbrug total	-0.00253	-0.11788	-0.90675	0.08271	0.24322	0.90767	-0.00537	8e-04	0.25257	0.68705	
Affald total	-0.00227	-0.56619	0.05409	-0.15967	0.69483		0.01872	0.00092	0.34001	0.59716	
Vand forbrug/genbrugt	-0.00255	-0.56652	0.04862	-0.1532	0.69042		0.07525	0.00095	0.3539	0.58457	
Vand total tilbagetrækning	-0.00138	-0.73574	-0.90493		0.79692	0.98966	-0.04053	0.00077	0.38169	0.58785	
Affald total/USD	-0.0026	-0.56568	0.049	-0.1535	0.69019		0.06455	0.00095	0.35397	0.58423	
Gennemsnitlig siddetid i bestyrelsen	-2e-04	-0.72872	-0.92016		0.79463	0.98312	0.20697	0.00086	0.40468	0.56539	
Bestyrelsens kønsfordeling	-0.00062	0.59408	-0.98636	0.1487	-0.41719	1.00525	0.16412*	0.00116	0.51214	0.45658	
Bestyrelsens kompensation	-0.00203	-0.5713	0.05461	-0.15287	0.69533		-0.02961	0.00095	0.34426	0.5913	

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 14: ARMA/GARCH modeller med miljø- og rådata for tidsserien HOG

Ovenstående tabel viser at flere af tidsserierne kan modelleres med eksterne kovariate. Det fremgår også at der er flere tidsserie, hvor ingen af de eksterne kovariate viser sig at være signifikante. Der tegner sig derfor et foreløbigt billede af inkonsistent signifikans blandt de 45 tidsserier. For kortfattedhed

er de resterende modeller for både ARMA(p,q)/GARCH(1,1) udeladt. Slutteligt er alle de endelige modeller samlet, hvor signifikante eksterne kovariate for hver enkelt tidsserie og ESG niveau er samlet. Dette udgør i alt tre omfattende tabeller, som fremgår i appendix på side 127.

## Residual diagnostik

Slutteligt kan modellernes fit vurderes ud fra ACF plot af residualerne. Her er der igen for kortfattedhed kun bragt fire af ARMA(p,q)-X/GARCH(1,1) for ARMA modellerne med ESG scorer.

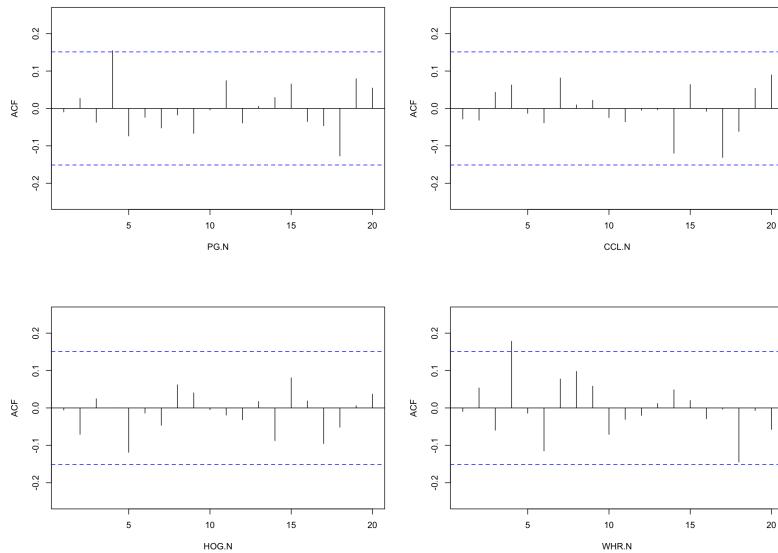


Figure 7: ACF - ARMA/GARCH residualer

Ljung-Box båndende kan vises at give sub-optimale resultater, når der testes data der indeholder afhængige lags. Derfor benyttes en Breusch-Godfrey test til at understøtte ovenstående plot. Breusch-Godfrey testen er beskrevet i afsnit 7.2, og har samme nulhypotese som Ljung-Box testen. Breusch-Godfrey tests er udført med op til 20 lags på ESG modellerne, hvor de forventes at kunne generaliseres til de resterende modeller. Resultaterne for Breusch-Godfrey testene viser at, der ikke kan afvises at autokorrelation i residualerne er 0 for samtlige tidsseriemodeller. Dette tyder på at modellerne opfanger autokorrelation i data, hvilket også kan ses i de standardiserede residualer i figur 8, da disse ikke viser tegn på autokorrelation. Nedenstående i tabel 51 ses Breusch-Godfrey test for modeller med ESG-score af fire gennemgående virksomheder:

	PG	CCL	HOG	WHR
Test størrelse	14.7608	14.83362	21.35298	25.10277
P-værdi	0.78993	0.78585	0.37662	0.19753

Table 15: Breusch-Godfrey test af standardiserede residualer for ESG-scorer modeller

Breusch-Godfrey test resultaterne for de resterende modeller fremgår i tabel 51 på side 126 i appendix. Nedenstående ses QQ-plots for de standardiserede residualer af ESG modellerne:

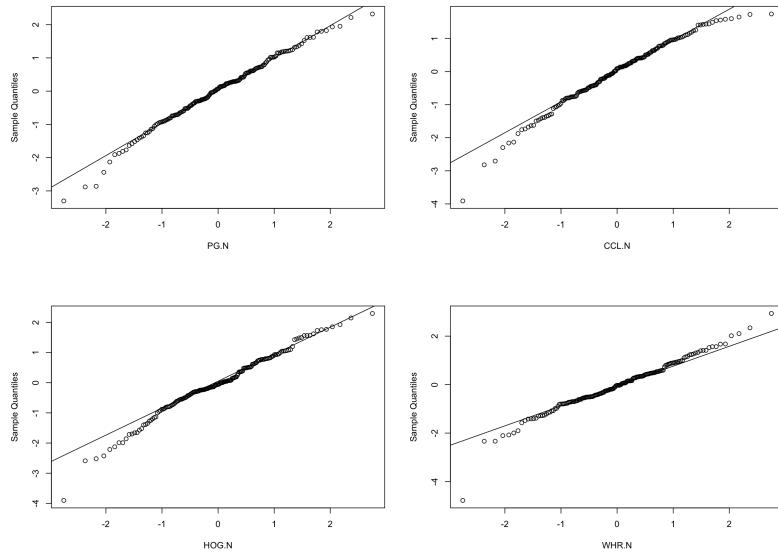


Figure 8: QQ Plot - ARMA/GARCH std. residualer

Ud fra QQ-plottene kunne det tyde på at normalfordelingsantagelsen er overholdt, da residualerne ligger relativt glat op af linjen. Det bekraeftes ligeledes af en ADF test, hvor der ikke kan afvises at de standardiserede residualer følger en normalfordeling ved et 5% signifikans niveau. Beslutningen om at beholde normalfordelingsantagelsen er altså valideret.

	PG	CCL	HOG	WHR
Test størrelse	-5.05850	-5.23757	-5.08559	-5.71891
P-værdi	0.01**	0.01**	0.01**	0.01**

Table 16: Augmented Dickey-Fuller test af standardiserede residualer

De endelige modeller med ESG-scoren er vist på side 54 i tabel 17.

Table 17: ARMA-X/GARCH(1,1) med ESG score

Virksamhed	$\mu$	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	ESG-score	$\omega$	$\alpha_1$	$\beta_1$
STT	0.0075019969	0.7211878232	-0.5834024828	-0.9147117386	-0.7540582325	0.5762154125	0.1712797027	-0.0076284849*	0.000908708	0.2411285785	0.6489015647
C	0.00812577851	-1.0113336057	-0.4676396466	0.2982747564	1.1034582875	0.9919557104	-0.0068297252***	-0.0068297252***	0.0006988526	0.4382828406	0.5607168411
BAC	-0.0370406922	-1.1661815745	-0.476523213	-0.476756744	-0.6133858498	1.1125506891	0.0033937319***	0.0006785619	0.290537852	0.6879876259	
ABT	0.0023305942	0.1894092912	0.726523213	-0.890487434	-0.0981303733	-1.19532119411	1.0066050343	0.004418258***	3.672e-06	1.241e-07	0.998972685
AMGN	0.0080520527	1.1130130507	-0.0140897677	-0.7859642292	-0.9999999691	-0.9999999691	-0.009159046*	0	7.38e-07	0.9983233666	
BAX	0.0098584455	0.9530947666	-0.11301305135	-0.2737905467	1.177667795	0.4442383562	0.001031539	-0.2426029569	0.003576623	4.2868e-06	0.99899218
BDX	0.005889756	1.3422739898	-1.273792719	0.3634111088	-1.4047005467	1.5133932088	-0.4399245849	-0.0084403124***	5.8059e-06	1.1691e-06	0.998910298
BMY	0.0108164855	-0.3971273959	-0.0610062314	0.2746722059	-0.4145751572	0.0016109451	0	6.69e-08	0.9980544236		
LLY	0.0078219449	-1.11631313347	-0.4666518394	-0.1226203515	1.177667795	0.0006778012	0.0013184022	0.0012104357	0.3410683441	0.2017636321	
MDT	0.0037311479	0.9614040775	-0.9999999896	-0.0061009806	-1.1110372965	0.005506459*	0.0012104357	0.4617203144	0.269059603		
DGX	0.0056890744	-0.12293398946	0.9013550074	-0.9685112358	-0.96859740281	-0.96859740281	-0.0054097327***	-0.0054097327***	5.8008e-06	3.4777e-06	0.997372626
A	0.0102263132	0.8685112358	-0.5714795355	-0.6810150418	0.0410607984	0.7449392393	0.6914211207	0.0001402661	0.0004611197	0.2266178333	0.7170202337
MRK	0.0064728407	-0.2189598897	-0.224756328	0.4997155921	-0.2958159199	-0.5310870422	0.1440575907	-0.00234185012	0.0001231172	0.064308453	0.8934210568
T	0.0022298453	0.224756328	0.7333290981	-0.7346658503	0.7106770516	-0.0328706465	0.8628750712	-0.561296805	0.0002208012	0.0757206645	0.83965285605
AMD	0.00699565597	0.1853230981	-0.0316276487	0.245870616	-0.0328706468	0.1936897668	0.1273641358	0.00153992	0.1330283196	0.7943669667	
AMAT	0.00931323652	-0.5666930911	-0.5666930911	1.4011032519	-0.9889702624	0.1931190093	-0.1898410603	0.005910634***	0.0057323353	0.16958161707	0.7596607847
CSCO	0.0073972958	-1.4978892668	-1.0239883297	-1.0422481229	0.1275453104	0.9998177871	0.00595965131	0.050759675	0.9322765377		
HPQ	0.0079192359	0.4567121124	-0.9287479894	0.4567121124	1.017329754	0.0021321562	-0.0111917069***	0	0.0063737812	0.7090263432	
INTC	0.0052169909	-0.4186320211	-1.1884526968	1.7147228214	1.1884526968	-0.0032993886	6.45582e-05	3.821e-06	0.998686695		
IBM	0.0019562996	-1.7147228214	-0.8578275179	-0.247067559	0.996030013	0.008825168	0	0.9767816056			
MSFT	0.00861199635	0.2688842	-0.4223580369	-0.1306113642	0.9220451179	-0.25413086743	0.0439029006	-0.0102723119*	0	2.964e-07	0.998986949
TXN	0.011230582	-0.118779638	-0.13124769231	-0.1111762132	-0.9935956474	0.0576891353	-0.0028966275***	-0.0002067727	0.0376992652	0.8793785908	
BLL	2.023e-05	-0.118779638	-0.13124769231	-0.16471070474	0.5102676958	0.0582447772	-6.96685e-05***	0.0003232863	0.1422650907	0.6462885181	
FCX	0.0082881776	-0.59394644744	-0.1540357124	0.1210830497	-1.4224459233	1.0477963524	-0.005226933***	0.000433563	0.1491538359	1.68e-08	
NEM	-0.00099111637	-0.2844582785	-0.9400064584	0.2213019274	1.0143833155	-0.0054460494*	0.00114736622	0.1710138295	0.7195336333		
CMI	0.0117757877	0.1959276626	-0.7357622846	0.1872399272	-0.166829202064	0.7562302718	-0.0066081664	0.0003197162	0.1866718979	0.7965790058	
ITW	0.0106794435	-0.8863857459	-0.7703334315	1.1220382787	1.1590137061	0.0576891353	0.0003397148	0.1422650907	0.7550475328		
IR	0.017190063	1.4625200055	-1.1540357124	0.1210830497	-1.4224459233	1.0477963524	0.0582447772	0.1914072623	0.7661138196		
MMM	0.0047382942	-0.6540689773	-0.8119856388	0.7243223311	1.0246316336	-0.005226933***	0.001016880159	0.7385329451			
UPS	0.0017013989	-0.2352040783	-0.680919146	0.3928140367	0.1303416243	0.4678215548	-0.582391106	-0.0054460494*	0.0004877708	0.0372207084	0.7994834091
RTN	0.0060761295	-1.131252428	0.1764156042	-1.0201255831	1.1438462377	0.004421516***	0.0004474216***	0.0002233572	0.4726789316	0.8687564164	
HES	0.0258141752	0.5120348129	0.4574195006	-0.574924448	-0.394936435	-0.1177581496	-0.0055975149***	0.0031089164	0.47258e-06	0.685586853	
BKR	-0.009401141	COP	0.0103788981	-0.8495438065	0.6411543664	0.8131620404	0.8892371921	-0.6750152373	-0.9397209961	0.0002684457	0.3404876114
CPB	0.0021612645	-1.9472046446	-0.109688664	1.9401368541	1.0204358419	-0.110547978	-0.001060373799	0.0002518443***	0.175126473	0.7461747465	
CAG	-0.010461355	-0.081430984	1.0340035149	0.1688577463	-1.1100582858	-0.171054978	0.002518443***	0.0004904936	0.3494204499	0.5775666922	
TAP	0.003124218	-0.0320129209	0.8641099764	-0.0980511796	-0.9155647533	0.2185816673	0.000333354	0.0018923406	0.2762789316	0.3206455428	
GIS	0.0032075262	0.0928766489	-0.9637335173	0.0847918185	1.0260993147	0.1868596729	-0.00101105615	3.0264e-06	0	0.9989999897	
KMB	-0.0018799502	-0.3626275079	0.4453742159	0.8419950791	0.3432240881	-0.6274959505	-0.8806466174	-2.8073e-05	7.13334e-05	0.1216798316	0.8440471034
MO	0.0080795436	-1.5143814092	-0.5535579346	-0.0032501486	0.5145513453	0.0042688695***	0.0001833348	0.0531717895			
PG	0.0034625709	0.9740579809	-0.0986184811	-0.9999999866	-0.004188291*	1.077e-06	5e-10	0.998994355			
CCL	0.0022616374	-1.1269915016	-0.8533511892	-0.1072078615	1.063144646	0.8739065145	-0.0041981393	0.000431726	0.1603218106	0.77711684413	
HOG	-0.0110200831	0.224210591791	-0.2066997914	0.8628238216	-0.2455653034	0.1331492333	-1.0353541065	-0.0090297602***	0.0006119698	0.3649283607	0.608694757
WHR	0.0059516573	0.9833185367	-0.0554767409	-0.999999871	-0.999999871	-0.00660523151***	0.00111187198	0.1625776002	0.7566003626		

Note:

\* p&lt;0.1; \*\* p&lt;0.05; \*\*\* p&lt;0.01

## 11.7 Modelresultater

Nedenstående ses opsummeringstabel 57, som indeholder alle resultaterne for samtlige ARMA(p,q)-X/GARCH(1,1) modeller der modelleret med enten ESG-, søjle-, underscorer eller miljø- og rådata. Modeller der har nul signifikante kovariate er ekskluderet fra tabellen.

Data grundlag	ESG-score	Pillar-score	Under-score	Miljø- og Rådata
ESG-score	25			
Miljø-score		14		
Governance-score		17		
Social-score		17		
Samfunds-score			23	
CSR-score			22	
Emmisions-score			19	
Miljø innovations-score			21	
Menneskerettigheds-score			23	
Ledelse-score			17	
Produktansvarligheds-score			13	
Ressource-score			15	
Aktionær-score			18	
Arbejdskraft-score			20	
Gennemsnitlige siddetid - Bestyrelsen				20
Senior Executive Kompensation				18
CO2 udledning total				1
CO2 Emission - Total/USD				4
Affald - Total/USD				20
Vandforbrug - Total/USD				18
Affaldsgenbrug - Total				2
Affaldsgenbrug / Affald				5
Affald - Total				13
Vand Total Tilbagetrækning				20

Table 18: Tidsserier med signifikante kovariate

### ESG-scoren som ekstern kovariat

Hovedsageligt er der relativt mange modeller der har signifikante eksterne kovariate. Ud af de 45 univariate tidsserier, ses der flere tegn på at ESG-scorene har en signifikant indflydelse på log-afkast. Dette skyldes at  $\sim 55.555\%$  af virksomhederne har en signifikant ekstern kovariat på baggrund af t-testen. Det er yderst overraskende at der opnås et så stort antal modeller med signifikante eksterne kovariate, da flere af de forgående steps foretaget i denne undersøgelse, potentieligt kunne have forværet de endelige resultater. Der kunne argumenteres for at, på trods af flere foranstaltninger imod det, at overfitting er skyld i mængden af antal signifikante modeller med eksterne kovariate. Blandt de signifikante modeller med ESG-score har 11 positivt og 14 negativt fortegn på estimatet. Det er altså ikke muligt at finde en konsistent retning af ESG-scorens indflydelse på log-afkast i de univariate modeller. Effekten svinger fra  $-0.0028$  til  $0.0059$ , hvorfor modellernes ikke antages at kunne generaliseres. Med henblik på underhypotese 1 kan der foreløbigt ikke konkluderes at ESG-scorene, på baggrund af de 45 virksomheder i S&P500 med højst data kvalitet, har en signifikant positiv indflydelse på det månedlige afkast.

### **Søjle-scorer som eksterne kovariate**

Af analysen ses det at søjle-scorerne, som eksterne kovariate, ikke umiddelbart forbedrer de konstruerede ARMA(p,q)-X/GARCH(1,1) modeller. Mængden af signifikante modeller i de enkelte søjle scorer er nogenlunde ens med mængden af signifikante modeller med ESG-scoren. På baggrund af dette kan det derfor ikke afvises at sammenhængen er konsistent. Det fremgår derfor foreløbigt at niveau et og to i ESG data er tilnærmelsesvis ens. Ingen ses det er fortegnet på parametret er inkonsistent for alle 3 søjle-scorer i tabel 19. Der ses foreløbig ikke en bemærkelsesværdig forskel på E'et, S'et og G'ets indflydelse, hvilket der kan argumenteres for primært skyldes den tætte relation mellem de to niveauer. Her refereres der til selve konstruktionsprocessen af ESG-scoren. Der kan foreløbigt ikke konkluderes noget yderlige med henblik på hovedhypotesen og underhypotese 2.

Table 19: ARMA-X/GARCH(1,1) med Søjle-score

Virksomhed	$\mu$	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	ESG-score	$\omega$	$\alpha_1$	$\beta_1$
Governance - AMD	-0.10434	-0.31031	-1.01791	0.4484	1.11782	0.11846	0.00133***	0.00951	0.14787	0.48312	0.01408
Governance - BDX	-0.03416	0.38774	0.40134	-0.64412	-0.4239	0.06555	0.00046***	0	0.05	0.9	
Governance - BMY	0.00857	1.19483	-0.50136	-1.58321	0.90796	-0.27795	-6e-05*	1e-05	0	0.999	0.11391
Governance - CAG	-0.02343	-0.65122	0.0619	0.83262	-0.05549	-0.29348	0.00038***	0.00058	0.43312	0.56328	0.18687
Governance - GIS	-0.01697	0.36792	0.9495	-0.24872	-1.06713	-0.42745	0.00028***	1e-05	0.02	0.97834	0.27888
Governance - HES	0.06267	0.6804	0.45285	-0.69555	-0.35627	-0.20459	-0.00117***	3e-05	0.00048	0.99725	
Governance - HOG	0.04432	1.02445	-0.18349	-0.98103	0.76676	0.05422	-0.00095***	0.00048	0.34619	0.65281	
Governance - HPQ	0.05889	1.16795	-0.98704	-0.93909	0.76676	0.11729	-0.00094***	0.00127	0.10255	0.68963	0.27964
Governance - IBM	0.0483	0.38511	0.5954	-0.49309	-0.56499	-0.0956	-0.00068***	0	0.05006	0.90001	
Governance - INTC	0.02427	-0.50166	-0.97279	0.65024	1.09792	0.16014	-0.00033***	0	0.01258	0.98353	
Governance - ITW	0.00823	0.79238	-1.15608	-0.62965	1.08802	0.30634	-3e-05***	0.00085	0.34941	0.42075	
Governance - MDT	0.03137	0.90557	-1				-0.00038***	0.00122	0.40289	0.29611	
Governance - MMM	-0.0762	-0.65367	-0.81092	0.72274	1.02399	0.00095**	0.00042	0.106	0.74186		
Governance - MO	0.09134	0.07486	0.95338	-0.06438	-0.98444	-0.00062***	1e-05	0.01086	0.98787	0.19299	
Governance - MRK	0.0127	-0.18204	-0.60454	-0.08492	0.86111	-0.46806	-0.00012***	0.00013	0.07593	0.87995	0.50867
Governance - RTN	0.12375	1.11571	-1.10771	-1.01446	1.07316	0.20697	-0.00141***	0.00039	0.06345	0.80015	
Governance - T	0.02659	0.71618	0.46085	-1.08225	-0.34275	-0.71791	-0.00043***	0.00038	0.1289	0.69127	0.87997
Miljø - AMAT	-0.03498	1.55995	-1.04728	-1.68117	1.24675	0.09188	0.00051***	0.00053	0.11006	0.80014	-0.09863
Miljø - AMGN	0.04692	-0.45404	-0.60239	0.31058	0.52597	0.41676	-0.00043***	0	0	0.99811	-0.65779
Miljø - BDX	0.03501	-0.15652	-0.94108	-0.05784	1.06767	-0.09637	-0.00045***	0.00082	0.13944	0.49527	-0.23558
Miljø - CPB	0.02833	0.23369	0.73838	-0.38746	-0.68674	-0.00025	-3e-04***	0.00023	0.11118	0.82487	
Miljø - DGX	-0.00386	0.86117	-1				0.00015***	0	0	0.999	
Miljø - FCX	-0.01455	0.2795	0.57136	-0.27618	-0.48694	0.00036***	2e-05	0.05	0.9	-0.31595	
Miljø - GIS	-0.06129	1.07139	-0.322	-0.91871	1.10672	0.02637	0.00086*	0	0	0.999	
Miljø - HES	0.13369	0.64633	0.28863	-0.70558	-0.08097	-0.005637	-0.00159***	1e-05	0.05617	0.93507	-0.27093
Miljø - IBM	0.08478	1.03434	0.48781	-1.23656	-0.28919	-0.76461	-0.00095*	0.00022	0	0.9223	0.74151
Miljø - ILY	-0.04166						0.00068*	0.00126	0.32324	0.26251	
Miljø - MO	0.01451	-0.16195	0.67872	0.26226	-0.84645	-9e-05***	1e-05	0	0.99898		
Miljø - MRK	0.04421	-0.14289	-0.66704	-0.05438	0.88412	-0.44405	-0.00051***	0.00011	0.07078	0.89407	0.49686
Miljø - STT	0.13214	0.10588	0.87915	-0.08084	-1.00577	-0.00186***	0.0011	0.30871	0.60142		
Miljø - UPS	-0.07862	-0.12407	0.55969	-0.49616	-0.71085	0.00094***	0.00041	0.34343	0.58292	0.16635	
Social - A	0.13719	-0.312323	0.18394	0.33944	0.23228	-0.85812	-0.00147***	6e-05	0.11869	0.86102	0.16318
Social - AMD	0.16377	0.07606	-0.80715	0.14552	0.95731	0.56376	-0.00118***	0.00171	0.04171	0.88952	-0.37633
Social - BAX	0.02167	-0.73696	0.56236	0.64824	-0.74796	0.87498	-0.00017***	0	0	0.999	-1.04257
Social - BDX	0.23019	-0.49585	-0.82978	0.49064	1.03476	-0.00232***	0.00041	0.09231	0.74272		
Social - BMY	0.02226	1.17772	-0.46102	-1.53502	0.83331	-0.31119	-3e-04***	1e-05	0.09875	0.16385	
Social - CAG	0.04034	-1.61175	-0.4624	1.77501	0.58557	0.21642	-0.00044***	3e-04	0.3308	0.66791	-0.26494
Social - COP	0.07421	0.95479	-1				-0.00082***	0.00048	0.27645	0.66523	
Social - CSCO	-0.05245	-1.37573	-0.47286	1.26605	0.24115	0.00079***	0.00012	0.05103	0.92339		
Social - DGX	-0.01622	0.50207	0.19943	-0.7124	-0.25787	4e-04***	0	0.05	0.9	-0.08273	
Social - HES	0.02502	0.40161	0.59399	-0.38252	-0.53774	0*	0.00248	0.00073	0.75025	-0.16679	
Social - IBM	0.08478	-1.72843	-1.45598	1.82	1.71696	-0.74122	-0.00088***	0	0	0.999	0.94724
Social - MDT	0.17054	-0.73824	-0.4321	0.72638	0.29801	-0.57592	-0.00174***	0.00128	0.33194	0.29171	0.68829
Social - MMM	-0.06756	0.23749	-0.23249	-0.29798	0.2771	0.68717	0.00085***	1e-04	0.09301	0.87448	-1.03733
Social - MO	0.08255	0.04308	0.9398	0.00532	-0.98653	-0.00069***	2e-05	0.03991	0.95867	0.13466	
Social - MRK	0.06493	-0.09068	-0.65571	-0.13152	0.88102	-0.43099	-0.00065***	0.00014	0.07786	0.8696	0.46162
Social - PG	0.03212	0.85069	-0.15585	-0.86357	-0.15585	-0.00036***	0	0.07582	0.91551		
Social - TAP	0.03213	1.11846	-0.34661	-1.6094	0.79832	-0.2052	-0.00044***	0.00174	0.52663	0.14211	0.18574

Note: \*p&lt;0.1; \*\*p&lt;0.05; \*\*\*p&lt;0.01

### **Under-scorer som eksterne kovariate**

I modsætning til ESG- og søjle-scorene, er der for kortfattethed valgt at udelade de samlede resultater for alle under-scorene. Disse kan ses i appendix fra side 127. Det ses at ud af de 450 modeller er der 194 modeller med en signifikants i den eksterne kovariat. Det svarer til 43.111% af modellerne, hvor 'Samfunds-scoren' og 'CSR-scoren' de eksterne kovariate der er signifikans i flest modeller (23 og 22).

Det viser sig at de ti under-scorer alle har flere modeller med signifikante kovariate. Her er der ikke én underscore som har konsistent fortegn i blandt de tilhørende signifikante modeller. De univariate ARMA-X/GARCH modeller tyder derfor på at effekten af en ændring i under-scorene hverken er udelukkende positiv eller negativ. Dette bekræftes ligeledes af de gennemgåede resultater i litteraturen. Ydermere er mange af modellerne med signifikans potentiel overparametreret. Det ses blandt andet i mængden af ARMA(3,3) modeller der er signifikante. Det fremgår også at der er en trend, nede fra og op i tabellen, hvor jo færre parametre der indgår i modellerne, jo færre modeller er der med signifikant ekstern kovariat. Dette tyder på at der ikke er opnået meget mer-information ved at dykke ned i niveau tre.

Slutteligt peger mængden af signifikante modeller på at der ikke er et statistisk grundlag for at konkludere konsistent signifikans for selv de bedste modeltyper. Her tegner der sig et lignende billede for de resterende modeller med signifikante kovariate, hvor 'produktansvarligheds-scoren' er den mindst repræsenterede, med "kun" 13 signifikante tidsseriemodeller.

Det fremgår derfor at hovedhypotesen indtil videre ikke kan bekræftes. Yderlige undersøgelser kræves ligeledes for at afkræfte den, da der ses tendenser til signifikans i underscorene. Med henblik på underhypoteserne ses der ikke nogle mønstre i resultaterne der peger mod at miljø-scorene skulle have en større sammenhæng med afkast end de resterende scorer.

### **Miljø- og rådata som eksterne kovariate**

Bevæger man sig ud af de Thompson-Reuters konstruede scorer og benytter de bagvedliggende data, som i denne opgave benævnes som miljø- og rådata, findes der signifikans i 10 af de 15 benyttede variable i mindst én model, se appendix 133. Dog er resultaterne meget spredte. Det fremgår at 'bestyrelsen gennemsnitlig siddetid', 'affald - total/USD' og 'vand total tilbagetrækning' er de tre eksterne kovariate med flest signifikante modeller. Igen ses det at fortegnet er inkonsekvent, hvorfor mange af de samme pointer bragt omkring under-scorene gør sig gældende igen. Dette inkluderer, potentiel overparametrerede modeller hvor sammenhængen med stor sikkerhed er mere tilfældighed end signifikans på trods af t-testen.

Med henblik på hovedhypotesen, ses der ikke nogen mer-information i miljø- og rådataen på baggrund af de univariate modeller. De informationer de estimerede modeller kan tilføje undersøgelsen understøtter derfor ikke hovedhypotesen. Der ses dog heller ikke nogen "top-down" informationsforøgning, altså, hvor de konstruerede scorer giver en direkte forøgelse i fit af data.

### **Omkring konsistente modeller**

Det fremgår i enkelte tidsserier, at der er blevet udvalgt modeller, hvilke ikke overholder de klassiske stationaritets betingelser for parametrene. Dette tyder igen på potentiel overfitting, da de modeller

der ikke overholder de givne betingelser strider imod finansiell teori. Det skal sluteligt nævnes, at på trods af enkelte tidsserier, med tilhørende eksplorative modeller, er helhedsbilledet prioriteret i dette afsnit. Dog er det fordelagtigt at teste en mere påholdende og restriktiv estimationsprocess, hvilken er gennemgået i part III.

### **Fravalg af out-of-sample prædiktion**

Da denne undersøgelse ikke er interesseret i selve prædiktionsevnen, er der fravalgt at teste denne på et nyt datasæt. Som følge af samplingsprocessen er der valgt at benytte alt data til at konstruere modellerne. Følgende analyse er derfor restrikeret til in-sample fits. Det skal nævnes at in-sample fits kan være prægede af bias ved brug af overfittede modeller, der er dog taget visse forhold i forrige afsnit for bedst muligt at undgå resultater præget af dette.

### **Test af modellerne**

I følgende afsnit er der udvalgt modeller lavet med fire eksterne kovariate. Her vurderes det at der kan drages paralleller mellem de udvalgte modeller og de resterende modeller. Dette er gjort for at overskueliggøre hovedpointerne af denne analyse. Der gennemgås samlet fem modeller, som sammenlignes i forhold til MAE som performance mål. Her refereres der til standard-modellerne som er ARMA/GARCH(1,1) modellerne uden eksterne kovariate, hvor de indeholder ESG-scoren, miljø-scoren, ressource-scoren og variablen 'Affald total/USD'.

For at sammenholde de forskellige modeller, testes modellernes fit. In-sample fits er test af modellens fit på samme datagrundlag. Dette er med til at danne et yderlige billede af hvorvidt de eksterne kovariate er værd at inkluderer i de enkelte modeller. Testdatasættet er valgt i hele perioden fra 2005 til 2018 for at danne et så bredt billede som muligt. For kortfattedhed testes der kun MAE af enkelte tidsserier, nemlig CCL, HON, PG og WHR. Netop disse er valgt på baggrund af et repræsentativt grundlag i et flertal af de signifikante modeller, der indeholder henholdsvis under-scorer eller miljø- og rådata som eksterne kovariate. Nedenstående ses in-sample fittet for den rene ARMA/GARCH(1,1)-model, ESG modellen, 'miljø-score' modellen og sluteligt 'ressource-score modellen'. Disse er modeller er udelukkende af virksomheden HOG.

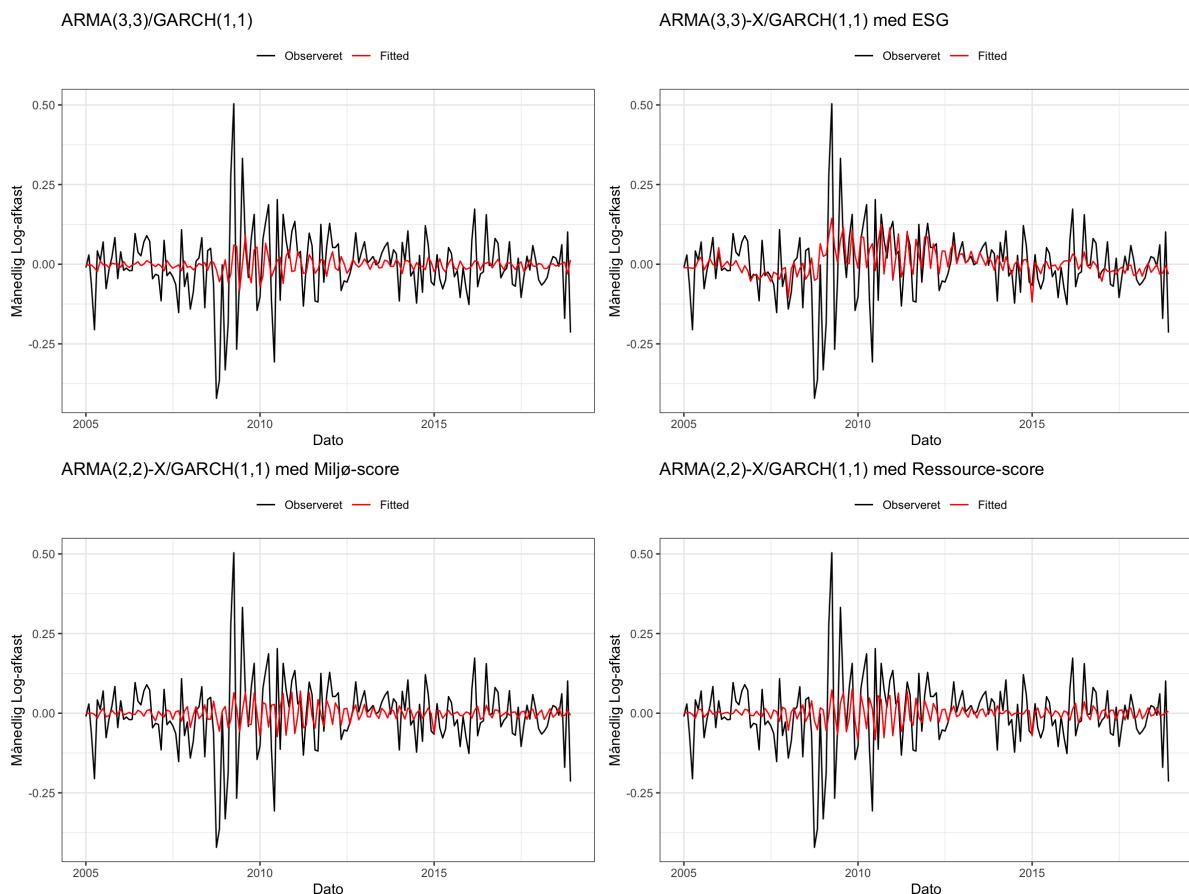


Figure 9: In-Sample fits af HOG

Af figur 9 fremgår det at, de fittede værdier er tilnærmelsesvis ens for de fire modeller. Der ses at både ARMA(3,1)/GARCH(1,1) modellen og ARMA(3,1)-X/GARCH(1,1) modellerne har meget forsigtige fit for log-afkast, hvor store udsving i log-afkastet ikke bliver opfanget af modellerne. I øvre højre hjørne ses modellen med ESG-scoren som ekstern kovariate. Dette fit er en smule mere markant end de andre, hvilket tyder på at flere af dynamikkerne i log-afkast opfanges af ESG-scoren. Trods det og trods de signifikante eksterne kovariate, er fittet af data dog stadig relativt dårligt. For yderligere at teste denne påstand beregnes MAE for de fire udvalgte tidsserier. Her er MAE beregnet for de forskellige niveauer, hvor modellerne med 'ressource-scoren' og 'affald total/USD' variablene er udvalgt. Valget er truffet med udgangspunkt i mængden af modellerne der viste signifikans. Det skal nævnes at MAE er beregnet udelukkende på de signifikante modeller.

Virksomhed	Standard	ESG	Miljø	Ressource	Affald total/USD
PG	0.03087763	0.03157617	0.03167617	0.03211453	0.03208724
CCL	0.06077352	0.05909165	0.05610895	0.05558270	0.05986785
HOG	0.07769570	0.07012558	0.07477859	0.07738171	0.07577897
WHR	0.08188402	0.08062951	0.08071887	0.08024244	0.08056013
Gennemsnit	0.05587269	0.05541629	0.05585491	0.05518465	0.05534357

Table 20: MAE af udvalgte modeller

Det fremgår at MAE er en grad lavere for flere af modellerne med eksterne kovariate. Det ses også at forbedringen af fit ved brug af ESG-scoren, kun er lokaliseret i modellen for HOG. Pointerne bragt i tidligere afsnit, så som dårligt fit og insignifikante forbedringer på trods af inklusion af de eksterne kovariate, styrkes yderligere. Dermed er der bredt grundlag for at underhypotese 2 ikke ser gyldig ud. Det kan igen bekræftes at samtlige modeller er mindre gode fits til data grundet den relativt høje MAE for in-sample fits, hvilket indikerer at empirisk forecasting, med stor sandsynlighed, ikke ville være muligt. Tidsseriemodellerne har derfor vist sig ikke at være anvendelig i praksis. Dette bekræftes ligeledes af den gennemsnitlige MAE, hvor forskellen på modellerne er insignifikant.

### Afsluttende bemærkninger

Som det har fremgået i denne analyse, har datakvaliteten generelt præget diverse estimerater. For det første har det vist sig at modellering med månedlige log-afkast, på et datagrundlag med 168 observationer, ikke har vist sig fyldestgørende nok til at modellere en ”sand” model. Det er umiddelbart ikke til at vide om modellering af log-afkast i daglig frekvens ville forbedre det generelle fit, men det kunne tænkes. For det andet har omfanget af analysen lidt under indskærpninger, hvor miljø- og rådata er den primære faktor. Her kan det vise sig at selve imputeringsprocessen har været med til at tilføje ekstra tilfældighed i de eksterne kovariate, selvom tidligere dataarbejde ikke tyder på det.

Tages resultaterne som en helhed, fremstår der faktisk flere gode pointer. For det første ses det på baggrund af den univariate tidsserieanalyse, at det er inkonklusivt at de lavere niveaur er ESG data har den direkte mer-information. Dette fremstår i både MAE for modellerne og en generel indifferent mængde af signifikante modeller på tværs af alle de eksterne kovariate. Hovedhypotesen ser derfor ikke ud til at holde stik, hvorfor yderligere analyser kan vise sig nyttige. For det andet ses der ikke nogen sammenhæng mellem ESG-data og log-afkast. Dette inkluderer både, ESG-scoren, søjle-scorene, under-scorene og det udvalgte miljø- og rådata. Der er altså foreløbigt ikke nogle resultater der tyder på at effekten af ESG-score på virksomheder finansielle performance er positiv, hvorfor underhypotese 1 hverken kan be- eller afkræftes. Sidst ses det også at miljø variablene ikke viser sig at give et markant bedre fit end de resterende variable. Her kræves ligeledes yderligere analyse for at undersøge hvilken indflydelse de eksterne kovariate har på log-afkast.

Hernæst kunne det være interessent at modellere data multidimensionelt. En multidimensionel analyse åbner, for det første, op for et større datagrundlag, hvorved performance af eventuelle multidimensionelle modeller kan vise sig at være bedre end de univariate. Udo over det åbner en multivariatanalyse muligheden for at benytte de kategoriske variables potentialle til fulde, netop da variation på tværs af virksomheder kan modellers.

## 12 Panedata regressionsanalyse

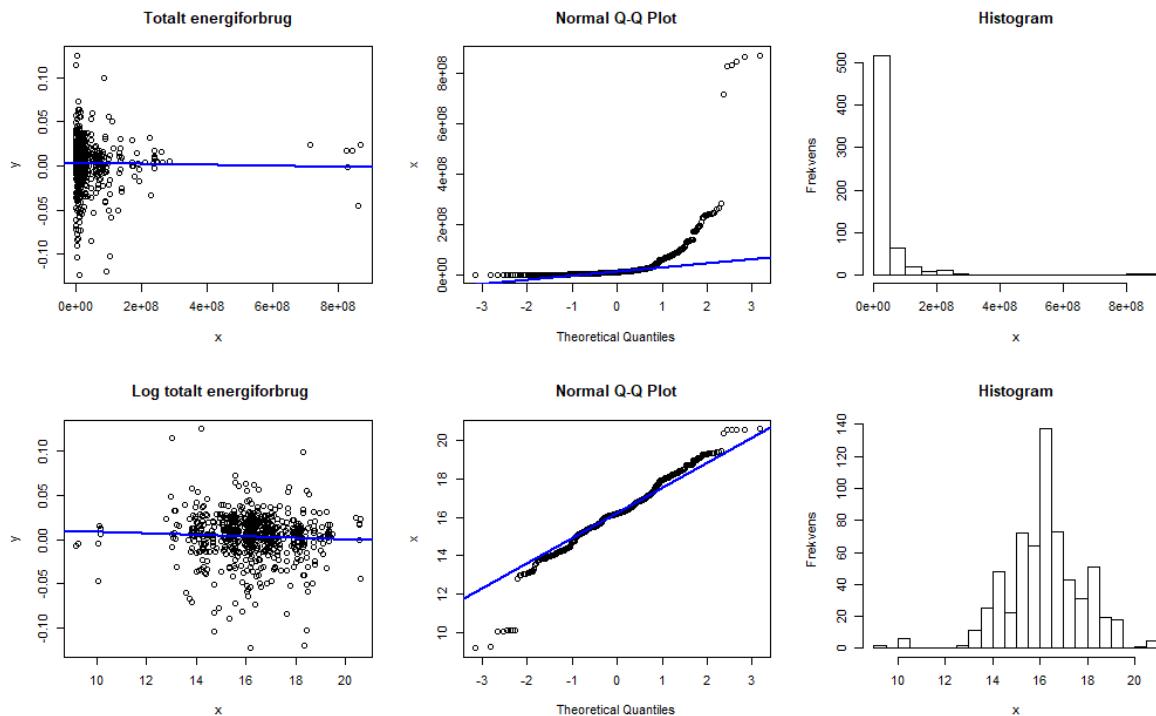
ESG-scorernes indflydelse undersøges videre ved panedata regressionsanalyse. Således modelleres effekterne for virksomhederne i et multidimensionelt setup. Dette indebærer en transformering af datasættet til panedata format. Resultaterne vil i det følgende afsnit blive præsenteret. Først og fremmest vil transformering af variable gennemgås. Efterfølgende vil den firdelte regressionsanalyses model-setup blive gennemgået, hvor regressionernes modelspecifikation introduceres. Herefter undersøges valget mellem en 'Pooled', 'Fixed' eller 'Random' effekts model, ved passende F test, Breusch-Pagan Lagrange Multiplier test og Hausman test. Efterfølgende testes den valgte model for Gauss-Markov antagelserne for OLS, herunder multikollinearitet, eksogenitet, homoskedacitet og autokorrelation. Afslutningsvis beskrives de reducerede modeller, med henblik på projektets hypoteser.

### Indledende bemærkninger

Panedata tillader modellering af mere komplekse handlingsmønstre end i dataset bestående af en enkelt "cross-section" eller tidsserie. Dette kan naturligvis have sine fordele, men kan samtidig skabe ekstra udfordringer og kræve mere præcision. En fordel ved modellering af panedata er egenskaben der tillader at kontrollere for individuel heterogenitet. Det er ikke altid muligt at inkludere alle tænkelige variable, som kan have en potentiel effekt på undersøgelsen. Som tidligere beskrevet kan udeladte variable lede til biased estimerater, som panedata kan håndtere i modsætning til klassisk tidsserieanalyse.

#### 12.1 Transformering af variable

Indledningsvis undersøges potentielle transformationer af de forklarende variable med udgangspunkt i henholdsvis et scatterplot, et QQ-plot og et histogram. Disse er konstrueret for kontrolvariablene samt for ESG-scorenes underliggende data i niveau 4. Da ESG-scorerne i niveau 1, 2 og 3 i forvejen bygger på en percentil ranking, ønskes de ikke transformert. Plottene er konstrueret på baggrund af årlige gennemsnit for de 45 enkelte virksomheder, og kan ses på side 134 til 139 appendix. Den generelle tendens viser at især de få store virksomheder påvirker til en højre skævhed. Dette resulterer naturligvis også i, at de eksempelvis bruger mere energi og udleder mere CO<sub>2</sub>. For at imødekomme de ekstreme værdier, logaritme transformeres variablene: 'Størrelse', 'CO<sub>2</sub> udledning total', 'affald total genbrug', 'affald total', 'vand total tilbagetrækning', og 'energiforbrug total' som eksemplificeret herunder:



Således ”trækkes” de ekstreme værdier ind mod en mere centreret fordeling, for at de så vidt muligt ikke har en overrepræsenteret effekt som kan lede til biased estimatorer. De to forklarende variable ’bestyrelsen kompensation’ og ’senior executives kompensation’ indeholder ligeledes ekstremt store værdier, men ser ved første øjekast ikke lige så skæve ud som de øvrige tilfælde. Værdierne variere dog i et stort spænd, hvorfor disse også logaritmetransformereres. Dette gøres for at minimere afstanden fra den største til den mindste observation. Variable som blandt andet ’vandforbrug / genbrug’ og ’vandforbrug total / USD’ er angivet i ratioer og ønskes derfor ikke yderligere transformert. Det samme gør sig gældende for variablene ’bestyrelsen kønsfordeling’ og ’executives kønsfordeling’ som allerede er angivet i procent. Dette gælder naturligvis også for de binære ”sandt” eller ”falsk” variable: ’miljøkontroverser’, ’affaldsreducerende initiativer’, ’vedvarende og grøn energiforbrug’, ’miljø distributionskæde politik’, ’grønne bygninger’ og ’reduktion af giftige kemikalier’, som alle antager værdien 0 eller 1.

## 12.2 Tjek af indflydelsesrige observationer

Indledningsvis foretages en undersøgelse af potentielle indflydelsesrige observationer ved brugen af DFBETA. Grundet projektets komplekse datastruktur, foretogs undersøgelsen som udgangspunkt univariat for hver virksomhed for hver variabel. Ideen er som udgangspunkt at få et indblik i eksistensen af potentielle outliers, mere end at fjerne dem. Dette skyldes at der en risiko for at udelukke reelle effekter, ved at fjerne outliers. Da datagrundlaget i forvejen bygger på imputering er der yderligere specifiseret ydre grænser for observationerne. Desuden ønskes der i projektet at undersøge virkelighedsnært data, hvorfor der ikke ønskes at fjernes flere virksomheder, i det i forvejen strengt reducerede datasæt. Indflydelsesrige observationer holdes in mente, og undersøges nærmere i afsnit 14.1.

### 12.3 Model-setup

På samme måde som i den univariate tidsserianalyse undersøges ESG-variablene i de fire enkelte niveauer. Hermed kan de fire regressioner opstilles på deres generelle form, med nummerering til senere praktisk brug.

Regressionen for det første niveau, altså for den overordnede ESG-score ser ud som følger:

$$1 : y_{it} = z_i\alpha + \beta_1 \text{ESG-score}_{it} + k_{it} + \epsilon_{it} \quad (36)$$

hvor  $k_{it}$  repræsenterer de tre kontrolvariable: 'totalkapitalforrenting', 'størrelse' og 'finansiel gearing'. På samme vis kan de resterende regressioner opskrives, hvorfor regressionen for det andet niveau, med de tre søjlescorer ser ud som følger:

$$2 : y_{it} = z_i\alpha + \beta_1 \text{Environmental}_{it} + \beta_2 \text{Social}_{it} + \beta_3 \text{Governance}_{it} + k_{it} + \epsilon_{it} \quad (37)$$

Hernæst kan regressionen for det tredje niveau opstilles:

$$\begin{aligned} 3 : y_{it} = z_i\alpha &+ \beta_1 \text{Emissioner}_{it} + \beta_2 \text{Innovation}_{it} + \beta_3 \text{Ressourceforbrug}_{it} + \beta_4 \text{Arbejdsstyrke}_{it} \\ &+ \beta_5 \text{Menneskerettigheder}_{it} + \beta_6 \text{Samfund}_{it} + \beta_7 \text{Produktansvarlighed}_{it} + \beta_8 \text{Ledelse}_{it} \\ &+ \beta_9 \text{Aktionærer}_{it} + \beta_{10} \text{CRSstrategi}_{it} + k_{it} + \epsilon_{it} \end{aligned} \quad (38)$$

Og slutteligt opstilles den sidste regression med de tilbageværende variable efter datasortering, hvorfor regressionen for det fjerde niveau ser ud som følger:

$$\begin{aligned}
4 : y_{it} = & z_i \alpha + \beta_1 \text{AffaldGenbrug}/\text{Total}_{it} + \beta_2 \text{AffaldTotalGenbrug}_{it} + \beta_3 \text{VandforbrugGenbrug}_{it} \\
& + \beta_4 \text{CO}_2 \text{UdledningTotal}_{it} + \beta_5 \text{EnergiTotaltForbrug}_{it} + \beta_6 \text{VandTotalTilbagetrækning}_{it} \\
& + \beta_7 \text{CO}_2 \text{Udledning/USD}_{it} + \beta_8 \text{BestyrelseKønsfordeling}_{it} + \beta_9 \text{VandforbrugTotal/USD}_{it} \\
& + \beta_{10} \text{SeniorExecKompensation}_{it} + \beta_{11} \text{ExecKønsfordeling}_{it} + \beta_{12} \text{BestyrelsenGnsSidetid}_{it} \\
& + \beta_{13} \text{MiljøKontroverser}_{it} + \beta_{14} \text{VedvarendeGrønEnergiforbrug}_{it} + \beta_{15} \text{GrønneBygninger}_{it} \\
& + \beta_{16} \text{BestyrelseKompensation}_{it} + \beta_{17} \text{KemikalieReduktion}_{it} + \beta_{18} \text{VandforbrugTotal/USD}_{it} \\
& + \beta_{19} \text{AffaldsreducerendeInitiativer}_{it} + \beta_{20} \text{MiljøDistributionskæde}_{it} + \beta_{21} \text{AffaldTotal}_{it} \\
& + k_{it} + \epsilon_{it}
\end{aligned} \tag{39}$$

## Model specifikationer

De tre generelle modeller, som blev gennemgået i afsnit 8, besidder hver deres fordele og ulemper, hvorfor de alle vil blive taget i betragtning til undersøgelsen.

Fordelen ved en 'pooled' model er, at hvis virksomhedernes effekter individuelt viser sig at være ens, betinget af visse variable, tillader den 'pooled' model at undersøge enkelte virksomheder på baggrund af observationer af andre. Dette kræver dog udeblivelsen af både tids- og enhedsspecifikke effekter.

Fordelene ved specifikationerne af en 'fixed' effekts model er, at den tillader de tids- og enhedsspecifikke effekter at være korrelerede med de forklarende variable. Ulempen er til gengæld, at antallet af ukendte parametre stiger med antallet af observationer. Den 'fixed' effekts model kræver parameterestimation af hver enheds- eller tidseffekt, altså estimation af koefficienterne af de respektive 'dummy'-variable. Dette kan substansialt reducere modellens forklaringskraft, og øge estimaternes standardfejl. Således vil de 'fixed' effekter allerede forklare en stor del af variationen i den afhængige variabel. De mange parametre i den 'fixed' effekts model medfører også et stort tab af frihedsgrader. Yderligere tillader 'fixed' effekts estimatoren ikke estimation af koefficienter der er tidsinvariante. Modsat gælder der for en 'random' effekts model at antallet af parametre forbliver konstant når størrelsen af samplet stiger, og videre tillades estimation af effekten af tidsinvariante variable. En fordel ved den 'random' effekts model er at tabel af frihedsgrader i den 'fixed' effekts model kan undgås, hvis de ubserverede individuelle effekter kan antages tilfældige. Den største forskel mellem den 'fixed' og den 'random' effekts er effekten af 'dummy'-variable. Et parameter estimat af en given 'dummy'-variabel er en del af interceptet i en 'fixed' effekts model, og fejleddet for en 'random' effekts model (Baltagi, 2015).

Som udgangspunkt danner projektets data grundlag for, at en fixed effekts model er velegnet. Dette skyldes at 'within' estimatoreren, tager højde for de ubserverede individuelle forskelle mellem virksomheders karakteristika, som påvirker aktiekursen, eksempelvis virksomhedernes forretningsmodel. Hermed ses der isoleret set for virksomhederne på, hvorledes de forklarende variable påvirker responsvariablen. Modellerne indeholder enten en tids- eller enhedsspecifik effekt, eller begge dele. I forhold til tidsspecifikke effekter er der tale om effekter, som for forskellige virksomheder er korrellerede indenfor samme tidsperiode, mens enhedsspecifikke effekter omhandler effekter der variere mellem virksomheder, men er uafhængige af tiden. Det forventes at visse tidsafhængige effekter eksisterer, da

eksempelvis generelle tendenser i aktiemarkederne vil afspejles i alle aktiekurser, uden nødvendigvis at relatere sig til specifikke virksomheder. Ydermere er forskellige virksomhedsspecifikke variable udeladt, hvorfor en to-vejs effekt vil blive inkluderet. En Breusch-Pagan Lagrange multiplier test for to-vejs effekter, bekræfter senere eksistensen af disse. En to-vejs effekt kan lede til komplikationer for ubalancede og store panel dataset med for mange enheder (Baltagi, 2015). Men eftersom projektets data har gennemført en grundig organisering, og manglende datapunkter er imputeret, haves et balanceret panel dataset, som hverken er for kort eller langt, hvorfor der ikke dannes grundlag for eventuelle bekymringer.

## 12.4 Model valg

Til at undersøge hvilken af de tre generelle modeller som er passende for hver af de fire regressioner, er de fulde modeller konstrueret for både en 'Pooled', 'Random' og 'Fixed' effekts model. Konstruktionen er sket på baggrund af det balancede panel datasæt med 168 tidsobservationer for hver af de 45 virksomheder. Resultatetet af regressionsanalysen for det tredje ESG niveau ses i tabel 21 herunder, med parameter estimerater og deres tilhørende standardfejl i parentes. Resultat for de tre resterende regressioner kan ses i tabel 52, 53 og 54, i appendix.

	<i>Respons variabel:</i>		
	Logaritmisk afkast		
	Pooled	Random	Fixed
Emissioner	-0.00004 (0.0001)	-0.00002 (0.0001)	-0.0001 (0.0001)
Innovation	-0.00003 (0.00004)	-0.00004 (0.00003)	-0.00002 (0.00005)
Ressourceforbrug	0.0001 (0.0001)	-0.00000 (0.0001)	0.0001 (0.0001)
CSR strategi	-0.00004 (0.0001)	-0.00003 (0.0001)	-0.00001 (0.0001)
Ledelse	0.0001 (0.00004)	0.00004 (0.00004)	-0.00001 (0.00005)
Aktionære	0.00002 (0.00004)	-0.00000 (0.00003)	-0.00002 (0.00005)
Samfund	-0.0001 (0.0001)	-0.0001 (0.0001)	-0.0001 (0.0001)
Menneskerettigheder	0.00000 (0.0001)	0.00001 (0.0001)	0.00003 (0.0001)
Produktansvarlighed	-0.00003 (0.00005)	-0.00004 (0.00004)	-0.00004 (0.0001)
Arbejdsstyrke	0.0001* (0.0001)	0.0001*** (0.0001)	0.0002*** (0.0001)
Totalkapitalforrentning	-0.024 (0.031)	-0.036 (0.028)	-0.106*** (0.032)
Finansiell gearing	0.005 (0.006)	0.002 (0.005)	0.034*** (0.011)
Størrelse	-0.002*** (0.001)	-0.002*** (0.001)	-0.012*** (0.003)
Intercept	0.054*** (0.019)	0.057*** (0.017)	
Observationer	7,560	7,560	7,560
Virksomheder	45	45	45
R <sup>2</sup>	0.002	0.003	0.006
Adjusted R <sup>2</sup>	0.001	0.001	-0.024
F Statistic	1.354 (df = 13; 7546)	22.874**	3.510*** (df = 13; 7335)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 21: Regressions Resultater

I tabellen kan det ses, at fælles for de tre regressioner er en stor overvægt af insignifikante variable, og yderligere at de har meget lave R<sup>2</sup> værdier. Det er dog endvidere først og fremmest interessant at undersøge hvilken af de tre modeller som er passende for de forskellige regressioner. Til at teste dette, tages udgangspunkt i en F-test. Her sammenlignes den 'fixed' effekts model med en 'pooled' OLS regressions model på baggrund af goodness-of-fit, for at teste tilstedeværelsen af individuelle og tids-effekter. I tabel 22 ses resultatet af F-testen for hver af de fire regressioner:

Regr.	F-værdi	P-værdi
1 : $y_{it}$	13.906	0.00
2 : $y_{it}$	13.686	0.00
3 : $y_{it}$	13.823	0.00
4 : $y_{it}$	13.798	0.00

Table 22: Resultater af F-test

Fra tabellen kan det ses at nulhypotesen kan afvises for alle 4 regressioner. Dette indikerer altså at der eksistere minimum en tids- eller enhedsspecifik effekt, og at den fixed effekts model derfor er foretrukket over en pooled OLS model.

Herefter benyttes en Breusch-Pagan Langrange multiplier test til at undersøge om variansen af tids- eller enhedsspecifikke effekter er lig 0. Hermed undersøges den random effekts model i forhold til en pooled OLS regressionsmodel, på baggrund af en  $\chi^2$  fordeling med en frihedsgrad. Resultatet af Breusch-Pagan Lagrange Multiplier tests for de fire regressioner ses i følgende tabel:

Regr.	$\chi^2$	P-værdi
1 : $y_{it}$	7.2902	0.006933
2 : $y_{it}$	7.8696	0.005027
3 : $y_{it}$	7.5841	0.005889
4 : $y_{it}$	8.2168	0.004150

Table 23: Breusch-Pagan Lagrange Multiplier test

Fra tabel 23 kan de ses at p-værdien for hver af de fire regressioner er markant lavere end et 5% signifikansniveau, og dermed at variansen af tids- eller enhedsspecifikke effekter afviger fra 0. Dermed er den random effekts model bedre i stand til at tage højde for heterogeniteten end en pooled OLS regression.

Som forventet har det altså vist sig at både den 'fixed' effekts model og den 'random' effekts model er bedre egnet end en 'pooled' OLS regressionsmodel. Til at vælge mellem de to modeller benyttes en Hausman test, som sammenligner de to modeller og tester om der er signifikant forskel mellem deres estimerater. Resultatet af Hausman test for de fire regressioner ses i følgende tabel:

Regr.	$\chi^2$	P-værdi
1 : $y_{it}$	33.256	0.000001059
2 : $y_{it}$	34.029	0.000006641
3 : $y_{it}$	37.852	0.000304200
4 : $y_{it}$	39.603	0.023580000

Table 24: Hausman test

Tabellen viser at de fire regressioner igen udviser konsistente resultater. Endnu engang kan nulhypotesen afvises, hvorfor individuelle effekter er signifikant korrelerede med mindst en forklarende variabel og dermed at en fixed effekts model er foretrukket. I overensstemmelse med forventningerne

til resultaterne, viste det sig altså, at en 'fixed' effekts model og 'random' effekts model foretrækkes fremfor en 'pooled' regressionsmodel, og yderligere at en 'fixed' effekts model foretrækkes fremfor en 'random' effekts model. Samtlige regressioner udviste overensstemmende resultater hvilket muliggøre et validt sammenligningelighedsgrundlag. Den endelige model er altså en 'fixed' effekts model med en to-vejs effekt.

## 12.5 Gauss-Markov antagelser for OLS

I det følgende afsnit vil resultaterne af Gauss-Markov antagelserne for OLS blive præsenteret. Disse er udarbejdet for hver af de fire regressioner, for at sikre den bedst mulige kontrol og grundlag for sammenligning. Det skal dog nævnes, at ved statistisk modellering vil antagelserne sjældent være opfyldt perfekt. Resultaterne kan med fordel fortolkes eller behandles, således at de stadig er i stand til at forklare noget om virkeligheden (Egerod, 2016).

### Linearitet og tilfældig udvælgelse

Den første antagelse forudsætter linearitet i parametrene. Modellerne er specificeret med linearitet i parametrene, og yderligere er transformation af forklarende variable foretaget til at håndtere ikke-linære forhold. Derfor er antagelsen overholdt. I forhold til antagelsen om tilfældig udvælgelse undersøges middelværdien af residualerne. Der gælder for hver af de fire regressioner at middelværdien af residualerne er approksimativt nul, hvorfor denne antagelse er overholdt.

### Ingen perfekt multikollinearitet

Til at teste for multikollinearitet beregnes varians inflations faktoren (VIF) for hver af de fire regressioner, ekspliceret i tabel 25 herunder:

Variable	Y1	Y2	Y3	Y4
<b>ESG Niveau 1</b>				
ESG Score	1.156727			
<b>ESG Niveau 2</b>				
Enviromental	1.527233			
Social	1.656733			
Governance	1.206835			
<b>ESG Niveau 3</b>				
Emissioner	1.948481			
Innovation	1.235825			
Ressourceforbrug	2.299410			
CSR Strategi	1.756000			
Ledelse	1.144954			
Aktionære	1.100297			
Samfund	1.344725			
Menneskerettigheder	1.610079			
Produktansvarlighed	1.292279			
Arbejdsstyrke	1.768105			
<b>ESG Niveau 4</b>				
CO2 Udledning - Total	7.014414			
CO2 Udledning / USD	4.188543			
Energi - Totalt Forbrug	3.682295			
Affald - Genbrug / Total	3.555704			
Affald - Total Genbrug	3.484815			
Affald - Total	4.396532			
Affald - Total / USD	<b>112.587</b>			
Vandforbrug - Total / USD	3.151275			
Vandforbrug / Genbrug	<b>113.092</b>			
Vand - Total Tilbagetrækning	3.870198			
Bestyrrelsen - Gns. Siddetid	1.118268			
Bestyrrelsen - Kønsfordeling	1.283068			
Bestyrrelsen - Kompensation	1.558404			
Senior Executive - Kompensation	2.010906			
Executive - Kønsfordeling	1.263305			
Grønne Bygninger	1.619033			
Miljø Distributionskæde Politik	1.824873			
Miljø Kontroverser	1.198694			
Affaldsreducerende Initiativer	1.448686			
Vedvarende og Grøn Energiforbrug	1.216443			
Reduktion af Giftige Kemikalier	1.334508			
<b>Kontrol Variable</b>				
Totalkapitalforrentning	1.064233	1.072913	1.076737	1.171683
Fínansiell gearing	1.127218	1.127400	1.227166	1.311752
Størrelse	1.251827	1.268531	1.297053	2.112489

Table 25: VIF

En god tommelfingerregel siger, at en  $VIF > 10$  indikerer en høj korrelation der bør undersøges (Finlay, 2013). I tabellen er disse markeret med fed. Det kan her ses at ingen af de tre første regressions variable har en VIF der overstiger grænsen, mens der blot er to for den sidste regression. Dette indikerer altså at de to variable 'affald - total/USD' og 'vandforbrug-/genbrug' er overflødige for modellen, som kan prædikteres ud fra de andre forklarende variable. Dette kan potentielt skyldes at de to variable er ratioer, som for alle værdier ligger tæt på 0. Fra tabellen kan det ydermere ses, at halvdelen af variablene for regression 4 har en  $VIF > 2.5$ . Denne mere konservative grænse anses af blandt andre (Johnston et al., 2018), som et betydeligt niveau af kollinearitet. Derfor undersøges

korrelationen yderligere mellem variablene i regression 4, ved brug af dens korrelationsmatrice, som kan ses i tabel 35 i appendix. Korrelationsmatricen er konstrueret uden de seks kategoriske variable 'grønne bygninger', 'miljø distributionskæde politik', 'miljø kontroverser', 'affaldsreducerende initiativer', 'vedvarende grøn energiforbrug' og 'reduktion af giftige kemikalier', som alle da også har en lav VIF. Et nærmere kig på korrelationsmatricen viser først og fremmest (uden den store overraskelse), at de to variable med ekstreme VIF-værdier er 99.54% korrelerede. Desuden viser korrelationsmatricen at 'CO2 udledning - total' og 'energi - totalt forbrug' er 83.64% korrelerede, som det eneste andet par med en korrelation på mere end 80%. Dette afspejles således også i en relativt høj VIF på 7.01 for CO2 Udledning - Total. Da den dog stadig ikke overstiger grænsen på 10, er der ikke grund til større bekymringer om multikollinearitet i nogle af regressionerne, foruden de to variable med meget høj korrelation. Dette vil der senere hen blive taget højde.

For at undgå problemer med multikorrelation kan pågældende variable enten helt udelades eller transformeres på passende vis. Da flere af de forklarende variable i den sidste regression er blevet transformeret, bringes for nysgerrighedens skyld en tabel over VIF og korrelationsmatricen for det sidste niveau uden relevante transformationer. Disse kan findes i tabel 55 og 36 i appendix. Uden videre bemærkninger viser det sig, at 8 variable, før deres respektive transformationer, har en  $VIF > 10$ .

## Eksogentitet

Endvidere undersøges modellerne for potentiel endogenitet, altså om mindst én forklarende variable er korreleret med fejlleddet. Endogent bias forekommer typisk på grund af udeladte relevante variable. Fra det grundlæggende ESG niveau, er adskillige variable fravalgt grundet dårlig eller ikke eksisterende data. Fælles for disse er netop at de alle kunne have været relevante at inddrage, med henblik på en dybdegående analyse af ESG-scorerne. Videre eksistere der potentiel endogent bias, ved manglen på yderligere inddragelse af kontrolvariable, med potentiel indflydelse på virksomheder finansielle performance og andre virksomhedsspecifikke nyheder. I forhold til omvendt kausalitet og målefejl rejses ikke yderligere bekymring. Dog kan de enkelte forklarende variable potentielt have en effekt på virksomhedernes afkast, mens det også kan gælde omvendt. Potentielle målefejl kan forekomme igennem den lange dataprocess hos Thomson Reuters, men gennemgang af adskillige rutinetjeks, foruden projektetsrensning og datasortering, fjerner disse bekymringer.

## Homoskedacitet og autokorrelation

Indledningsvis kan et kig på residualer vs. fittede plots give et indblik i homoskedaciteten. Disse kan ses i figur 13 i appendix. Ved første øjekast ses fire næsten identiske plots, som alle udviser en forholdsvis konstant varians, men med afstikkende værdier. Til at teste for homoskedacitet tages udgangspunkt i en Breusch-Pagan test, med nulhypotesen, at der forekommer homoskedacitet. Resultateterne for de fire regressioner ses i følgende tabel:

Regr.	BP	P-værdi
1 : $y_{it}$	113.32	0.00
2 : $y_{it}$	133.21	0.00
3 : $y_{it}$	180.72	0.00
4 : $y_{it}$	388.24	0.00

Table 26: Breusch-Pagan test

Tabellen viser tydeligt at Breusch-Pagan testen indikerer heteroskedacitet for samtlige regressioner, hvorfor det er nødvendigt at benytte robuste standardfejl.

Herefter benyttes en Breusch-Godfrey test, til at undersøge eksistensen af autokorrelation i panel modeller. Resultaterne af disse ses i følgende tabel, med antal lags som standarden for testen. Testen benytter det mindste antal af observationer henover tidsperioden:

Regr.	$\chi^2$	P-værdi
1 : $y_{it}$	553.68	0.00
2 : $y_{it}$	553.09	0.00
3 : $y_{it}$	554.55	0.00
4 : $y_{it}$	561.76	0.00

Table 27: Breusch-Godfrey test af autokorrelation

Nulhypotesen for testen lyder, at der ikke eksisterer autokorrelation i modellen. Det fremgår tydeligt at tabellen at der eksisterer autokorrelation i fejlene, hvorfor der skal korrigeres for dette.

### Normalfordelte fejller

For at vurdere normalheden af fejllledene for de fire regressioner, undersøges QQ-plots for residualerne:

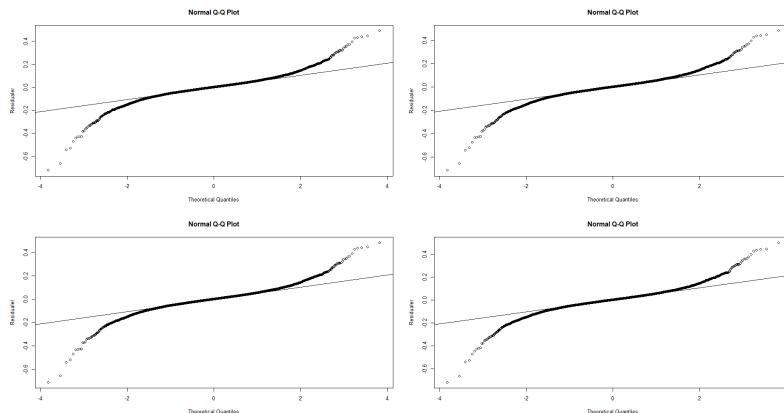


Figure 10: QQ Plot - To-vejs Fixed Effekts modeller

I figur 10 ses at fejllledene viser sig nogenlunde normalfordelt, men med tunge haler. Dette er gældende for hvert af de 4 tilfælde, og er ikke den største overraskelse, givet at der arbejdes med finansielt data. Som beskrevet i afsnit 8.8, er antagelsen om normalfordelte fejller ikke strengt nødvendig,

og en lempelse af normalfordelingsantagelsen vil i praksis ikke påvirke parameter estimaterne. I relation til den centrale grænseværdidisætning haves ingen yderligere bekymringer om de tilnærmelsesvis normalfordelte residualer med tunge haler.

### Afsluttende bemærkninger

Afslutningsvis kan vi altså konkludere at modellerne generelt overholder antagelserne for 'BLUE', ved mindre korrektion. Til at tage hånd om multikollinearitet i regression 4 vil 'affald - total / USD' udelades af modellen, mens indførelsen af en robust kovariansmatrice benyttes til at korrigere for autokorrelation og heteroskedacitet. Således vil koeficienterne i de endelige regressionsmodeller overholde antagelserne, og vil derfor være BLUE.

## 12.6 Modelresultater

På baggrund af de indledende modelspecifikationer præsenteret i de forrige sektioner foretages nu en individuel optimering af regressionerne til at udvælge de enkelte variable således at modellerne fitter data bedst muligt, ved at fjerne insignifikante variable. Hver af de fire regressioner er optimeret ved 'backwards elimination' og kan findes i tabel 28. Med henblik på forhåndværende regressionsresultater for de fulde modeller, forventes det at en større del af de forklarende variable elimineres ved processen som er foretaget proceduredrevet. Fremgangsmåden inkludere alle forklarende variable og fjerner variablen med den højeste p-værdi. Metoden er desuden god til at undgå manglende relevante variable (Weisberg, 2005). Videre er den let anvendelig, i modsætningen til flere alternative metoder. Samtidig fokuseres der på signifikansen af de enkelte variable, i modsætning til eksempelvis AIC. AIC bruges generelt til se på modellernes helhed, fremfor signifikansen af enkelte variabel, med risiko for overfitting.

Til sammenligning kan de fulde to-vejs 'fixed' effekts modeller findes i tabel 21 og videre i tabel 52, 53 og 54 i appendix. Ved nænsom sammenligning kan effekten af at bruge robuste standardfejlses, da nogle variable mister signifikans grundet den øgede varians.

Indledningsvis bringes hovedpointen af et nyligt studie omkring to-vejs 'fixed' effekts modeller af (Kropko and Kubinec, 2020). Her argumenteres for øget kompleksitet i fortolkningen af regressionssresultater, med to-vejs effekter, da det kræver en to-dimensionel fortolkning af henholdsvis individuelle og tids-effekter. Således ses fortolkningen altså som en virksomheds performance i forhold til sig selv over tid, og samtidig relativt til de resterende virksomheder.

### ESG-scoren

Tages et kig på kolonne (1) i tabel 28, bemærkes det først og fremmest, at der ikke er en signifikant effekt mellem den overordnede ESG score og virksomheders finansielle performance. En reducering af modellen er naturligvis udeladt, da elimination af 'ESG-score' blot ville efterlade en regressionsmodel bestående af de tre kontrolvariable. De tre kontrolvariable viser sig til gengæld at være signifikante, hvilket gør sig gældende for hver af de fire regressioner. Disse er dog ikke af speciel interesse, men blot inddraget på baggrund af relevant litteratur, for at øge robustheden mod bias ved udeladte variable. Tages videre et kig på modellens  $R^2$ -værdi, som udtrykker hvor stor del af variationen af den afhængige variabel der kan forklares af de uafhængige, ses det at den er meget lav, nærmest

ubetydelig. Videre ses at den justerede  $R^2$ -værdi er negativ, hvilket yderligere er et udtryk for at forklaringsevnen mod responsvariablen er meget lav eller ligefrem ubetydelig. Til sammenligning ses det at  $R^2$ -værdien for den fulde model, er betydeligt større. Dette skyldes naturligvis at den fulde model inkluderer de 'fixed' effekter, hvorfor den også performer bedre, da de 'fixed' effekter forklarer mere af variation i responsvariablen. Det interessante er dog endvidere ikke at undersøge de 'fixed' effekts 'dummies', hvorfor de også er udeladt i tabellen og  $R^2$ -værdien. I henhold til underhypotese 1, kan der på baggrund af regressionsanalysen ikke konkluderes nogle effekter mellem den overordnede ESG scorer og virksomheders finansielle performance, grundet statistisk insignifikans.

### Søjle-scorerne

Tages et kig på kolonne (2) i tabel 28, kan samme konklusioner drages. Ingen af de tre forklarende variable findes signifikante. Det kan ses at 'social' er den sidste tilbageværende variabel efter backwards elimination. Men på baggrund af statistisk insignifikans haves ikke et grundlag til at drage nogle konklusioner om effekten. I henhold til tidligere omtalte meta-studie konkluderes da også at 'E', 'S' og 'G' ikke udviser selvstændig indflydelse på virksomheders finansielle performance (Friede et al., 2015).

### Under-scorerne

I kolonne (3) i tabel 28, viser det sig at 2 ud af de 10 forklarende variable er signifikante. Fælles for disse er at de begge tilhører den 'sociale' kategori, hvor de står for 68% af vægtningen. Hvis der drages parallelle mellem kolonne (3) og (2), er det interessant at de signifikante variable i det tredje niveau på sin vis afspejles i det øvre niveau, ved at være den mindst insignifikant variabel. Der kan endvidere ikke drages nogle konklusion om regression 2 på baggrund af regression 3, men blot observeres at der er en vis sammenhæng mellem niveauerne. I henhold til underhypotese 2 indikerer regressionsresultatet endvidere at de 'sociale' variable har større effekt end 'environmental' variable, alene baseret på variabeludelukkelsen. For at bakke dette op testes koefficienterne for de respektive variable, ved en F-test, med nulhypotesen at koefficienten for en specifik variabel er lig koefficienten for en anden. Testene er foretaget enkeltvis ved test af en koefficient mod én anden af gangen. Det viser sig at 'arbejdsstyrke' er signifikant forskellig fra både 'emissioner' og 'innovationer' fra 'environmental' kategorien. Det skal yderligere bemærkes, at til trods for at de to variable findes signifikante, har modellen en  $R^2$ -værdi på 0.005, hvorfor man skal være påpasselig med at drage nogle konklusioner. Hvis man dog skulle kigge nærmere på de to forklarende variable, ses at 'samfund' har en negativ effekt, mens 'arbejdsstyrke' har en positiv effekt. En eventuel forklaring på arbejdsstyrkens positive effekt kan relateres til værdien i gode medarbejderforhold. Hvis en virksomhed har gode arbejdsforhold kan den tiltrække dygtigere medarbejdere, holde længere på eksisterende medarbejdere, og øge deres arbejdsomhed og effektivitet, resulterende i bedre afkast, i henhold til relevant litteratur om virksomheders arbejdsstyrke (Becker et al., 1997); (Faleye and Trahan, 2011). Ingen skal det dog nævnes, at der ikke bør lægges for meget i fortolkningen, grundet de generelt lave  $R^2$ -værdier.

	<i>Respons variabel:</i>			
	Logaritmisk afkast			
	(1)	(2)	(3)	(4)
ESG Score	0.00005 (0.0001)			
Social		0.0001 (0.0001)		
Samfund			-0.0001* (0.0001)	
Arbejdsstyrke			0.0002** (0.0001)	
Affaldsreducerende Initiativer				-0.007** (0.004)
Vedvarende og Grøn Energiforbrug				0.007* (0.004)
Vandforbrug - Total / USD				-0.213*** (0.056)
Bestyrelsen - Kønsfordeling				0.0003** (0.0001)
Totalkapitalforrentning	-0.093* (0.048)	-0.094** (0.048)	-0.103** (0.045)	-0.107** (0.046)
Finansiell gearing	0.032** (0.014)	0.033** (0.013)	0.034*** (0.012)	0.029** (0.013)
Størrelse	-0.012*** (0.004)	-0.012*** (0.004)	-0.012*** (0.004)	-0.013*** (0.004)
Observationer	7,560	7,560	7,560	7,560
Virksomheder	45	45	45	45
R <sup>2</sup>	0.005	0.005	0.005	0.007
Adjusted R <sup>2</sup>	-0.025	-0.024	-0.024	-0.023
R <sup>2</sup> (Full)	0.285	0.285	0.286	0.287
F-test (reducet vs. fuld)		0.259	0.401	0.853
F Statistic	8.403*** (df = 4; 7344)	8.81*** (df = 4; 7344)	8.49*** (df = 5; 7343)	6.273*** (df = 8; 7340)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 28: Optimerede to-vejs Fixed Effekts modeller med robuste standardfejl

## Miljø- og rådata

Tages et kig på kolonne (4) i tabel 28, ses de 4 tilbageværende signifikante variable efter modeloptimeringen. I henhold til den overordnede hypotese, er det altså muligt at finde signifikante effekter i ESG-scorernes bagvedliggende data i højere grad end for de øvre niveauer. Der er da også væsentligt flere forklarende variable i den nedre regression, hvorfor der alt andet lige vil være en større sandsynlighed for at enkelte variable har signifikant forklaringseffekt på responsvariablen. R<sup>2</sup>-værdien indikerer at modellen kan forklare 0.7% af variationen i virksomheders aktieafkast, men den negative justerede R<sup>2</sup>-værdi indikerer dog at forklaringen af responsvariablen er ekstremt lav, eller ligefrem ubetydelig. Dette betyder endvidere insignifikans af de forklarende variable, til trods for at F-værdien indikerer at modellen er signifikant (Torres-Reyna, 2010). Med forbehold for modellens insignifikans viser

effekterne dog en hvis grad af overensstemmelse med relevant litteratur. I et meta-studie af (Dixon-Fowler et al., 2013) vises at proaktive miljømæssige initiativer ikke øger virksomheders rentabilitet mere end reaktive initiativer. De affaldsreducerende initiativer afleder altså større omkostninger end fordele, heraf negativt fortegn. I forhold til bestyrelsens kønsfordeling ses også en positiv effekt, i overensstemmelse med tidligere beskrevne studie af (Velte, 2016).

### Afsluttende bemærkninger

Igennem forrige afsnit er ESG-scorene blevet undersøgt i de fire niveauer ved paneldata regressionsanalyse, i forlængelse af den univariate tidsserieanalyse. Til trods for grundigt modelvalg på baggrund af passende test, og indførelse af en robust kovariansmatrice til håndtering af autokorrelation og heteroskedacitet, viste det sig at ingen af de fire regressioner udviste betydelig forklaringsgrad. Dette er i overensstemmelse med de forrige resultater fra den univariate tidsserieanalyse.

I forhold til projektets overordnede hypotese viste det sig altså inkonklusivt om effekterne i det grundlæggende niveau har en større forklaringsgrad end de øvrige niveauer. Endvidere kan underhypotese 1 hverken be- eller afkræftes grundet statistisk insignifikans. I forhold til underhypotese 2, viste det sig at 'social' var den mindst insignifikante kategori, med to signifikante forklarende variable i niveau 3. Der kan ligesom i forrig analyse dog ikke drages nogle videre konklusioner.

Ved udvidelsen fra univariat tidserianalsyse til multidimensionel regressionsanalyse viste det sig at der kunne drages paralleller mellem resultaterne af disse. Paneldata regressionanalise har dermed bidraget til at validere eksisterende konklusioner yderligere. Indenfor projektets rammer har en to-delt analytisk tilgang til undersøgelse af problemformulering kastet lys over den ønskede undersøgelse. Til trods for den omfattende metodiske gennemgang, fandtes dog ingen stærke argumenter for, at kunne bekræfte effekterne til undersøgelse, med henblik på hypoteserne omkring disse. Derfor kan en robusthedsanalyse benyttes, til at validere resultaterne, eller yderligere belyse potentielt indflydende effekter.

## Part III

# Robusthed i modellerne

Følgende afsnit er bragt for at teste validiteten af resultaterne i både den univariate tidsserieanalyse og panaldata regressionsanalysen. I dette afsnit gennemgås de metodiske valg, som potentielt kan have haft en signifikant indflydelse på resultaterne. Dette inkluderer robusthedstest af valget af informationskriterie, dataimputeringen, datafrekvensen, indflydelsesrige observationer og tidsintervallet.

## 13 Robusthed i univariate modeller

### 13.1 GARCH(1,1) antagelse og valg af informationskriterie

I denne undersøgelsen er analysen baseret på flere forskellige valg der har indsnævret specifikationerne af de endelige ARMA-X/GARCH modeller. To af de mest essentielle valg, er valget af volatilitetsmodel og valget af AIC som informationskriterie. Valget af GARCH(1,1) som volatilitetsmodel er i høj grad baseret på det empiriske research af (Hansen and Lunde, 2005). Robustheden i disse valg ønskes derfor testet, for at tjekke hvorvidt disse valg potentielt har forværret eller forbedret de endelige ARMA-X/GARCH modeller.

#### ARMA-X/GARCH modeller i en to-delt estimationsproces med AICc

I nedenstående afsnit gennemgås en to-delt estimeringsmetode. Her modelleres der først rene ARMA modeller, hvorefter der testes for ARCH/GARCH effekter. Modellerne er udvalgt ved brug af AICc som informationskriterie. For kortfattethed er flere af tabellerne bragt i appendix.

I dette afsnit er der først estimeret ARMA-X modeller, hvor fittet vurderes ud fra AICc. Der fremgår i tabel 29 en tabel med de initiale modelvalg. Til modellering af ARMA-X(p,q) modellerne er funktionen 'auto.arima' brugt i R (Hyndman et al., 2007). Det ses umiddelbart at størstedelen af de estimerede ARMA-X modeller er ARMA(0,0). Dette skyldes ændringen i informationskriteriet. Dette bekræfter den initiale test i den univariate hovedanalyse, som pegede på at flere af tidsserierne ville have optimale ARMA(0,0) modeller ved benyttelse af AICc og BIC.

Lags	q = 0	q = 1	q = 2	q = 3	q = 4	q = 5
p = 0	11	1	0	0	0	0
p = 1	2	1	0	0	0	1
p = 2	0	1	5	3	1	1
p = 3	0	0	4	3	2	2
p = 4	0	0	0	1	1	1
p = 5	0	0	1	1	1	1

Table 29: Antal virksomheder med optimal modelspecifikation for ESG-score modellerne

## Modelselektion i ARMA-X modellerne

Hernæst findes de rene ARMA modeller med signifikante eksterne kovariate. I tabel 30 ses et udsnit af ARMA modellerne, for tidsserien WHR modelleret med miljø- og rådata.

Eksterne Kovariat	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	xreg
CO <sub>2</sub> udledning total	0.6553872867	-0.9051493947	0.2299569854	-0.5890597213	0.8368283833	9e-10
Energi forbrug total	0.6556505952	-0.9049768702	0.2301943019	-0.5893124195	0.8368359095	2e-10
CO <sub>2</sub> udledning total/genbrug	0.6543883673	-0.9053391946	0.2290295816	-0.5889746125	0.8375082601	0.0076271426
Vandforbrug total/genbrugt	0.6549054212	-0.905587359	0.2291932968	-0.5893086581	0.8377775974	0.000347445
Affald total/genbrugt	0.6544034444	-0.9051454374	0.2289795987	-0.5892040187	0.8373510212	0.0092735111
Affald genbrugt	0.6525616652	-0.9043910227	0.2283160273	-0.5880070676	0.8363194601	1.58e-08
Affald total	0.6549435055	-0.9047360711	0.2298812142	-0.5889967423	0.836570157	7.8e-09
Vandforbrug total/genbrugt	0.6534313887	-0.9045652745	0.2286390813	-0.5886402376	0.8366075374	6.73648e-05
Vand total tilbagetrækning	0.6553312427	-0.9049624792	0.229979191	-0.5891674317	0.8368096798	3e-10
Affald total/USD	0.6533454942	-0.9045997728	0.22857021	-0.5886409244	0.8366757937	0.0067777504
Bestyrelsens gns. siddetid	0.6558297223	-0.9051613857	0.2301308296	-0.5895038787	0.8370799364	0.0001834334
Bestyrelsens kønsfordeling	0.6563970771	-0.9053027125	0.2310892532	-0.589280012	0.8369719351	9.40765e-05
Bestyrelsens kompensation	0.6559702019	-0.9054473898	0.2302504207	-0.5894906767	0.8373042633	4e-10
Senior Executive kompensation	0.6521486542	-0.9032712893	0.2288239602	-0.5878283033	0.834799817	2e-10
Executive kønsfordeling	0.6572347977	-0.9071898728	0.2288440452	-0.5923477965	0.840475682	0.0007483052

Table 30: Rene ARMA modeller med miljødata for tidsserien WHR

Det fremgår at der er flere tidsserier, hvor ingen af de eksterne kovariate viser sig at være signifikante. Der ses igen inkonsistent signifikans blandt de 45 tidsserier. Slutteligt er alle endelige modeller samlet, hvor signifikante eksterne kovariate for hver enkelt tidsserie er inkluderet, præcis på samme facon som i tidligere analyse. En bemærkelsesværdig detalje er her, hvordan der intet mønster er, med henblik på hvilke modeller der er signifikante. Det ses derfor at resultaterne er meget følsomme overfor valget af informationskriterie.

## ARCH/GARCH effekter

For at tjekke for ARCH & GARCH effekter udføres Breusch-Godfrey tests igen. Disse viser at, der ikke kan afvises at autokorrelation i residualerne er 0 for alle modellerne. Derfor egner data sig som udgangspunkt ikke til at fitte en ARMA/GARCH model i disse tilfælde. Breusch-Godfrey testresultaterne fremgår i tabel 56 på side 146 i appendix for ESG modellerne. Ydermere kan residual plots findes på side 144.

## Resultater af robusthedstest

En lignende opsummeringstabell, som bragt i resultatafnittet, kan findes i appendix på side 147. Det ses at der er markant forskel på antal signifikante modeller. Specielt modellerne med ESG-scoren og søjle-scorene viser sig at være insignifikante ved brug af AICc. Her bekræftes undersøgelsens valg af AIC som informationskriterie, da det tyder på at modeller i robusthedsundersøgelsen er underparametrerede. Et plot med de fittede værdier mod de rigtige værdier af log-afkast for HOG er igen konstrueret:

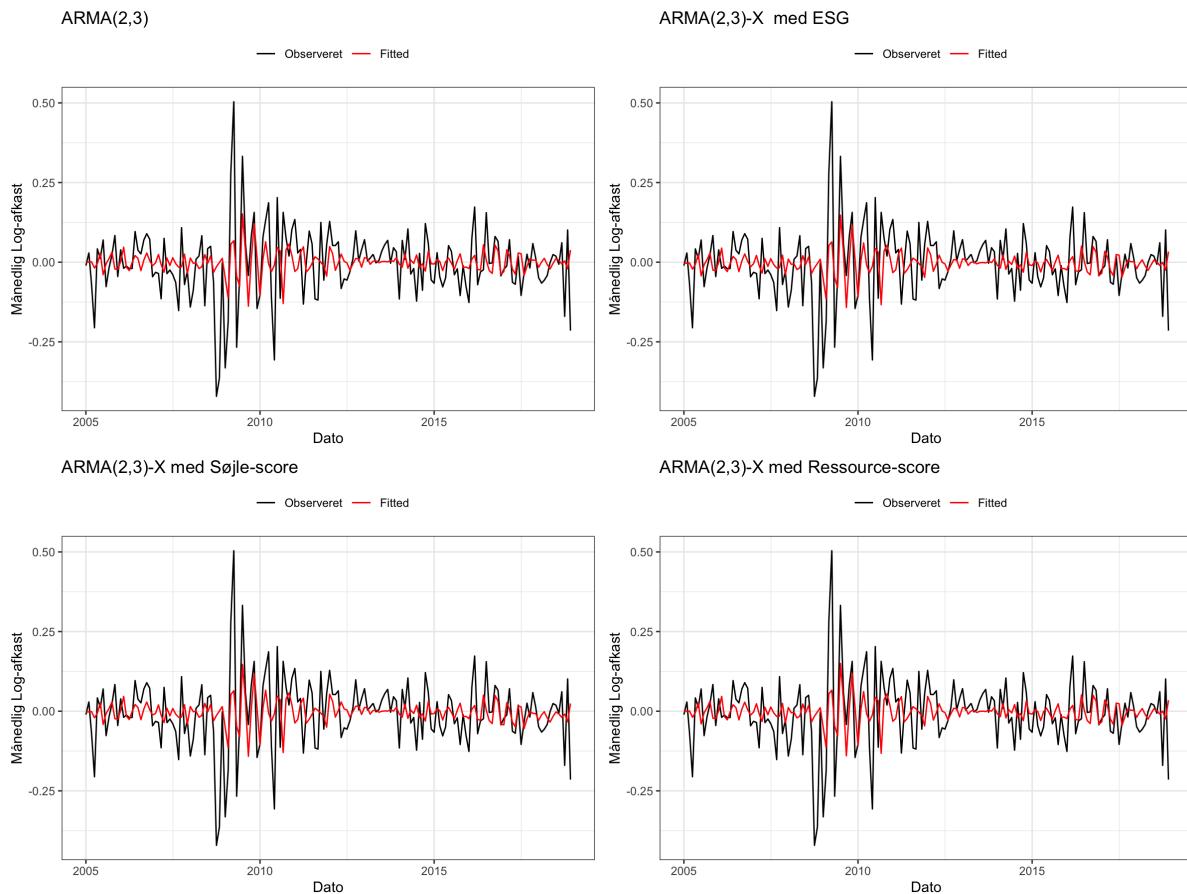


Figure 11: In-Sample fits af LLY

Sammenholdes dette med figur 9, ses det at den rene ARMA-X model har mindre forsigte estimerater. Det tyder også på at ikke alle modellerne konstrueret i dette afsnit er underfittede. Ingen beregnes MAE for de fire udvalgte virksomheder, samt gennemsnittet. Dette gøres igen for standard-modellerne, ESG-modellerne, miljø-modellerne, ressource-modellerne og Affald total/USD-modellerne.

Virksomhed	Standard	ESG	Miljø	Ressource	Affald total/USD
PG	0.03335273	0.03318234	0.03315616	0.03314624	0.03318227
CCL	0.05986411	0.05985975	0.05990524	0.05993000	0.05991717
HOG	0.07388833	0.07404942	0.07404884	0.07401862	0.07474461
WHR	0.07972261	0.07971519	0.07972272	0.07971623	0.07972187
Gennemsnit	0.05717290	0.05697920	0.05702132	0.05683255	0.05696528

Table 31: MAE af udvalgte modeller

Her ses det ikke overraskende at den gennemsnitlige MAE er en grad højere for disse modeller end for modellerne i hovedanalysen.

### Afsluttende bemærkninger

Det fremgår altså at antagelser om at benytte GARCH(1,1) ikke viste sig at være optimal, da modeller baseret på AICc ikke havde signifikant autokorrelation i residualerne. Dog har valget af en GARCH(1,1) model ikke haft signifikant indflydelse på resultaterne.

Ydermere viser det sig at resultaterne er følsomme overfor valget af informationskriterie. Det fremgår også at valget af AIC frem for AICc, har givet mere komplekse modeller. Dette ses i at den gennemsnitlige model i analysen er en ARMA(3,3)-X/GARCH(1,1), hvor denne robusthedtests gennemsnitlige model er en ARMA(0,0)-X.

Der vurderes dog ikke at de anvendte modeller i analysen er overfittede. Der ses derfor ikke nogen grund til at invalidere hovedresultaterne i den univariate tidsserieanalyse på baggrund af valget af informationskriterie. På trods af det, skal nogle af de konklusioner bragt i resultatafsnittet tages med en vis usikkerhed. Dette robusthedstjek viste, hvor forskellige resultater det er muligt at ende ud med, på baggrund af to valide og veldiskuterede fremgangsmåder.

## 13.2 Datafrekvens

Slutteligt ønskes der at teste valget af datafrekvens. Dette gøres for at imødekomme nogle af de problemer der kan have forekommeth ved at modellere med månedlige log-afkast. Fordelen ved at benytte en daglig frekvens er primært at der er betydeligt flere datapunkter. Den primære ulempe er, som tidligere nævnt, at der tilføres mere støj til de eksterne kovariate. Tidligere i denne undersøgelse blev der argumenteret for at en månedlig datafrekvens ville være optimalt til både en univariat tidsserieanalyse og en videregående paneldata analyse. Da selve mængden af de endelige datapunkter er relativt lav i de univariate tidsserier ønskes datafrekvensen primært testet her.

For kortfattedhed, generaliseres robusthedstjekket til én enkelt tidsserie, HOG. Her tjekkes resultaterne af ARMA-X(p,q)/GARCH(1,1) modeller i analysens ånd. Der er modelleret to modeller, en uden eksterne kovariate og en med ESG-score. Modelvalget er baseret på AIC, hvor metodiskfremgang af modelvalg er udeladt. Nedenstående ses log-afkast plottet mod en ARMA(1,3)/GARCH(1,1) og en ARMA(2,1)-X/GARCH(1,1):

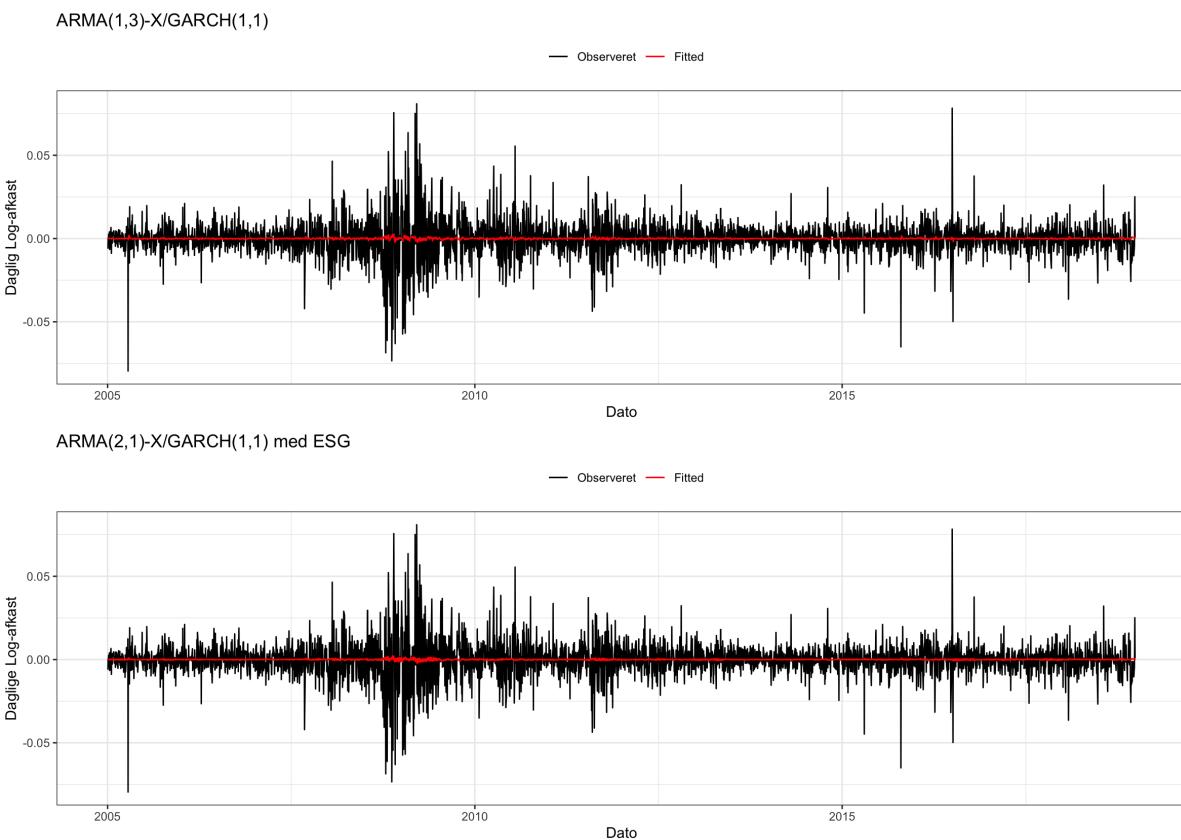


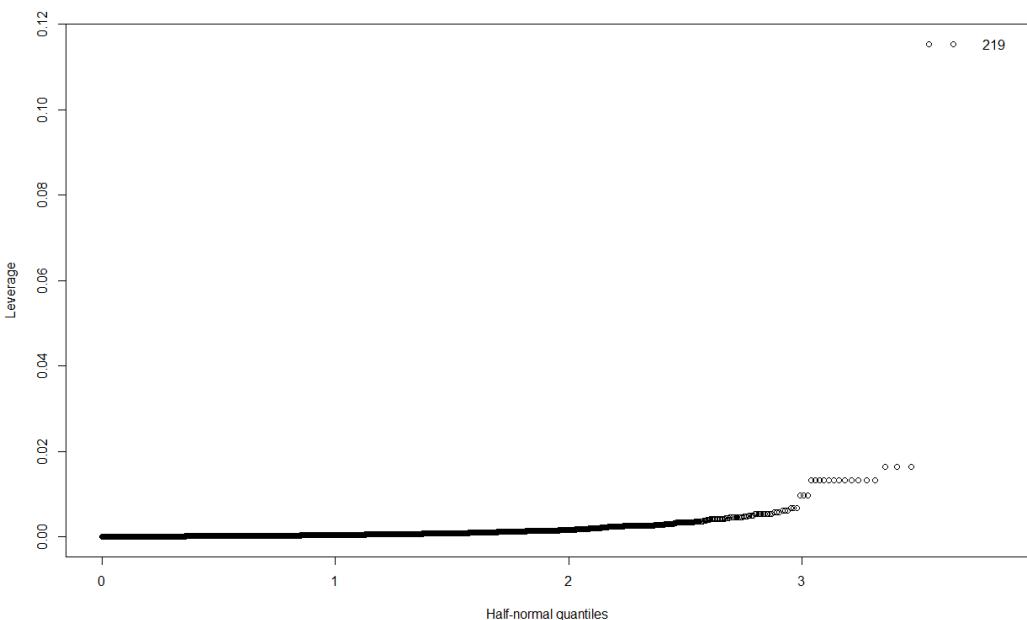
Figure 12: In-Sample fits af HOG

Hvor de førhen estimerede månedlige modeller har vist tendenser til forsigtige fits, fremgår det at modellen for de daglige log-afkast er yderst forsiktig, både med og uden ekstern kovariat. Noget kunne tyde på at der er mindre støj i de månedlige modeller. Samtidig er det uvist, hvorvidt modellerne på månedligt niveau, er komparativt overfittet. Yderligere fremgår det også i figur 12 at ingen yderligere dynamikker er tilført ved at benytte ESG scoren. Dette kan begrundes i den forholdsmaessige manglende variation af denne. Denne præliminære analyse af daglige afkast tyder derfor ikke på at valget af månedlige afkast frem for daglige, har haft en negativ indvirkning på resultaterne. Det kan dog konkluderes at resulaterne er følsomme overfor valget af datafrekvens.

## 14 Robusthed i multidimensionelle modeller

### 14.1 Indflydelsesrige observationer

Regressionsanalyse kan være følsom overfor indflydelsesrige observationer. Det lineære forhold kan være domineret af et mindre antal af indflydelsesrige observationer. Gennem projektets analyse blev ingen outliers fjernet på baggrund af et ønske om reelt data. Til at undersøge modellernes følsomhed overfor indflydelsesrige observationer udregnes det 'leverage', hvor specifik observation for hver virksomhed har af indflydelse på det lineære forhold. Med andre ord udregnes et mål for hvor følsom forholdet er til en specifik observation. Til dette benyttes et 'half-normal' plot, som sorterer observationerne efter deres 'leverage'. Et half-normal plot for fjerde regression kan ses herunder:



Plottet viser at specielt 3 observation har stor effekt, og indflydelse på forholdet. Disse er tre efterfølgende månedlige observationer for virksomheden FCX. Yderligere ses 3 selvstændige grupperinger af observationer, for henholdsvis MO, IR og BKR. De resterende observationer ligger alle sammen på en ret linje. Tager man et grundigt kig på aktierne individuelt, ses potentielle forklaringer for hvorfor netop disse skulle være ekstra indflydelsesrige. FCX oplevede et drastisk fald på næsten 90% fra kurs 62 til 8 på grund af finanskrisen i 2008, hvor det generelle marked oplevede et fald på 50% fra top til bund. Observationen er i den forstand en real observation og på sin vis ikke en outlier, da det ønskes at opnå så virkelighedsnære resultater som muligt. Tilsvarende oplevede BKR også et stort fald som følger af finanskrisen, mens IR i 2017 foretog en corporate action, resulterende i et stort aktiespring. De omtalte virksomheder fjernes nu fra datasættet for at teste resultaternes følsomhed overfor observationerne. Hele processen for udvælgelse af model, og videre test af disse er gennemgået til kontrol for overensstemmelse. Dette er nødvendigt at tjekke for at sikre at modellerne opfylder samme forudsætninger. For kortfattethed er en gennemgang af disse udeladt, men som forventeligt resulterer processen i samme endelige modelvalg for samtlige regressioner, og yderligere konsitens med

Gauss-Markov antagelserne. Resultatet af de optimerede regressioner kan ses herunder i tabel 32:

	<i>Respons variabel:</i>			
	Logaritmisk afkast			
	(1)	(2)	(3)	(4)
ESG Score	0.0001 (0.0001)			
Social		0.0002* (0.0001)		
Arbejdsstyrke			0.0002** (0.0001)	
Affaldsreducerende Initiativer				0.005** (0.003)
Vedvarende og Grøn Energiforbrug				0.007* (0.004)
Affald Genbrug Total				0.001* (0.001)
Affald Total				-0.001*** (0.001)
Bestyrelsen - Kønsfordeling				0.0004*** (0.0001)
Totalkapitalforrentning	-0.018 (0.059)	-0.023 (0.059)	-0.031 (0.058)	-0.035*** (0.047)
Finansiel gearing	0.027* (0.015)	0.028* (0.015)	0.030** (0.015)	0.018 (0.015)
Størrelse	-0.010*** (0.004)	-0.0112*** (0.004)	-0.011*** (0.004)	-0.011*** (0.003)
Observationer	6,888	6,888	6,888	6,888
Virksomheder	41	41	41	41
R <sup>2</sup>	0.003	0.003	0.004	0.005
Adjusted R <sup>2</sup>	-0.029	-0.029	-0.028	-0.027
R <sup>2</sup> (Full)	0.287	0.288	0.288	0.289
F Statistic	4.229*** (df = 4; 6676)	4.693*** (df = 4; 6676)	6.006*** (df = 4; 6676)	3.887*** (df = 8; 6670)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 32: Optimerede to-vejs Fixed Effekts modeller med robuste standardfejl, uden potentielle outliers

Til sammenligning med regressionsresultaterne for samtlige 45 virksomheder tegner sig et rimeligt lignende resultat ved første øjekast. Generelt ses endnu lavere niveauer af R<sup>2</sup>-værdier, for de i forvejen ubetydeligt små værdier. Det bemærkes endvidere at variablen 'social' er blevet signifikant i modsætning til tidligere, men med pludseligt insignifikante kontrolvariable, resulterer dette altså ikke i en forøget R<sup>2</sup>-værdi. For tredje regression er arbejdsstyrke stadig signifikant, mens samfund ikke længere findes signifikant. For den fjerde regression er 3 af de 5 tilbageværende signifikante variable de samme. På baggrund af denne inkonsistens og med henblik på R<sup>2</sup>-værdierne, kunne det altså tyde på at signifikansen er mere et tilfælde end en reel sammenhæng.

På baggrund af den ovenstående robusthedsundersøgelse af potentielle indflydende observationer,

kan det altså konkluderes at de ekstreme tilfælde ikke ledte til markante resultatændringer. Dette bekræfter yderligere tidlige overvejelser omkring reelle data, og ønsket om at modellere så virkelighedsnært data som muligt.

## 14.2 Modellering med ændret tidsinterval

Projektets observationer strækker sig over 168 månedlige tidsobservationer i årene 2005-2018. Eftersom finanskrisen er indeholdt i perioden, findes det interessant at undersøge resultaternes følsomhed i forhold til de særlige omstændigheder på aktiemarkederne i tidsperioden. Med yderligere skærpet fokus fra forrige afsnit, hvor indflydelsesrige observationer udviste ekstreme udsving, i og omkring finanskrisen, ønskes den generelle tendens for perioden undersøgt. Samtidig giver denne undersøgelse mulighed for at tjekke om effekterne kan konkretiseres yderligere for kortere tidsperioder, i tråd med tidlige nævnte studier beskrevet i afsnit 3.1, hvor effekten af virksomheders bæredygtige initiativer i forhold til deres finansielle performance viste sig forskellig for 2 efterfølgende tidsperioder.

Indledningsvis skal det nævnes at robustheden af modellering med et ændret tidsinterval er begrænset til regressionsanalysen. Der gælder for de univariate tidsserier, at datagrundlaget er så smalt at en ændret tidshorisont med stor sandsynlighed vil medføre dårlige estimater. Dette skyldes at der kræves et hvis antal observationer for at opfange sæsonprægede komponenter, og håndtere dem på tilpas vis (Box and Tiao, 1975).

På baggrund af de nævnte overvejelser, inddeltes datasættene i to 7-årige tidsperioder af 84 månedsobservation fra henholdsvis 2005-2011 og 2012-2018. Således isoleres første periodes markedsnedtur, mens anden periode afspejler markedets efterfølgende optur. Ingen er hele modelvalgs processen og de dertilhørende test foretaget, men udeladt for kortfattethed. Endnu engang gælder præcis samme modelspecifikationer, og regressionsresultaterne af de reducerede to-vejs 'fixed' effekts modeller for hver af de fire regressioner fordelt over de 2 tidsperioder, kan ses i tabel 33 herunder:

I tabellen ses resultaterne ved opdeling i de to under-tidsperioder. Til sammenligning refereres til tabel 28 med resultaterne for den fulde periode. Først og fremmest viser kolonne (1) og (2) i tabel 33 at 'ESG score' er signifikant i begge under-tidsperioder, i modsætning til resultatet for hele tidsperioden. Dog findes ingen markant stigning i  $R^2$ -værdien på baggrund af opdelingen af tidsperioder. Med fortsat negative justerede  $R^2$ -værdier, er det irrelevant at drage nogle konklusioner om effekterne af disse. Dette er en tendens der gør sig gældende for samtlige modeller i tabel 33. Tages videre et kig på kolonne (3) og (4), ses det hvorledes 'governance' er signifikant i den sene tidsperiode, i modsætning til den tidlige periode og resultaterne fra den fulde model, hvor ingen variable er signifikante. Noget kunne tyde på at signifikansen er styret af tilfældige sammenhæng, som yderligere bekræfter manglen på forklaringsevne, afspejlet i  $R^2$ -værdierne.

Det viste sig altså at modellerne var robuste overfor ændringer i tidsperioden.

## 14.3 Test af dataimputering

For at teste selve dataimputeringsprocessen, er der tidlige lavet endnu en dataimputering, benævnt som imputering(2). Idéen er at benytte imputering(2) til at validere kvaliteten af den benyttede imputering(1). En data opsummeringstabell af imputering(2) kan, som tidlige benævnt, ses i tabel

	Respons variabel:							
	Logaritmisk afkast							
	(1) (2005-2011)	(2) (2012-2018)	(3) (2005-2011)	(4) (2012-2018)	(5) (2005-2011)	(6) (2012-2018)	(7) (2005-2011)	(8) (2012-2018)
ESG Score	0.0003* (0.0002)	-0.0003* (0.0001)						
Social			0.0002 (0.0001)					
Governance				-0.0002** (0.0001)				
Arbejdsstyrke					0.0002* (0.0001)			
CSR Strategi						-0.0002** (0.0001)		
Produktansvarlighed						-0.0003** (0.0001)		
Energiforbrug Total							-0.001* (0.001)	
Affaldsreducerende initiativer							-0.007** (0.003)	
Bestyrelsen - gns. siddetid							-0.001** (0.001)	
Executives kønsfordeling							0.0005** (0.0001)	
Vandforbrug total / USD							0.431*** (0.093)	
Affald Genbrug / Total							-0.548*** (0.118)	-2.564*** (0.147)
Miljø Distributionskæde								-0.027*** (0.007)
Grønne bygninger								0.005*** (0.002)
CO2 udledning total								-0.012** (0.005)
Bestyrelsen kompensation								-0.012* (0.007)
Totalkapitalforrentning	-0.155*** (0.046)	-0.010 (0.042)	-0.155*** (0.046)	-0.015 (0.040)	-0.154*** (0.046)	-0.004 (0.048)	-0.162*** (0.044)	-0.006 (0.040)
Finansiell gearing	0.053** (0.025)	0.033** (0.014)	0.053** (0.025)	0.034** (0.014)	0.056** (0.026)	0.038** (0.015)	0.039 (0.029)	0.042*** (0.015)
Størrelse	-0.018*** (0.005)	-0.012*** (0.004)	-0.0192*** (0.005)	-0.013*** (0.004)	-0.019*** (0.005)	-0.012*** (0.004)	-0.019** (0.005)	-0.003 (0.004)
Observationer	3,780	3,780	3,780	3,780	3,780	3,780	3,780	3,780
Virksomheder	45	45	45	45	45	45	45	45
R <sup>2</sup>	0.008	0.004	0.008	0.004	0.008	0.006	0.011	0.012
Adjusted R <sup>2</sup>	-0.028	-0.032	-0.028	-0.032	-0.027	-0.030	-0.026	-0.024
F Statistic	7.254*** (df = 4; 3648)	3.505*** (df = 4; 3648)	7.301*** (df = 4; 3648)	3.508*** (df = 4; 3648)	7.577*** (df = 5; 3647)	4.239*** (df = 5; 3647)	4.424*** (df = 9; 3643)	5.672*** (df = 8; 3644)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 33: Optimerede to-vejs Fixed Effekts modeller med robuste standardfejl, for sub tidsperiode 2005-2011 og 2012-2018

34 i appendix. Valideringsprocessen er naturligvis kun foretaget for det underliggende ESG data, og resultatet af den reducerede model kan findes i tabel 60 i appendix. På trods af en uafhængig dataimputteringsproces, vist det sig at der ikke var forskel på de endelige resultater. Herved viste det sig, at hvis de imputerede datapunkter ikke ændrer markant på middelværdi, median, minimum, maksimum og standardafvigelse for de enkelte variable, er analyseresultaterne robuste. Dermed kan der bekræftes at variationen der kan forkomme med at benytte multipel imputering ikke har haft en signifikant indflydelse på analysen.

#### 14.4 Afsluttende bemærkning

Robusthedsanalysen viste at de univariate tidsseriemodeller er yderst følsomme valg af information-skriterie. Dette afspejles dog ikke i resultaterne, men primært i modelvalget. Videre fandtes at de samme modeller også er følsomme overfor datafrekvens. Her menes der dog ikke at valget af månedlig

datafrekvens har haft en negativ indflydelse på analysen, tværtimod. Gennem de tre yderlige robusthedstest af regressionsanalysen, viste det sig at resultaterne generelt var upåvirket af ændringerne. Først og fremmest viste det sig at resultaterne ikke drives af få indflydelsesrige observationer. Ved inddeling af datasættene i to tidsperioder i forhold til markedstrenden skabt af finanskrisen, viste sig heller ikke nogen signifikant forskel i resultaterne. Dette var ligeledes gældende for robustehdstesten af imputeringen.

# Part IV

## Diskussion

### 15 Manglende data

#### Teoretiske principper

I denne undersøgelse er der antaget at de manglende datapunkter følger en 'MAR' struktur, hvorfor der er imputeret datapunkter til at færdiggøre analysen. Hvis de manglende datapunkter ikke følger en 'MAR' struktur og derimod en 'MCAR' struktur er selve dataimputeringsprocessen ikke valid. Dette skyldes at når data følger en 'MCAR' struktur er det ikke muligt at finde sammenhænge på kryds og tværs af data, hvorfor imputering af datapunkterne praktisk talt er umulig (Van Buuren, 2018). Der kan argumenteres for at hvis datapunkterne i denne opgave har en 'MCAR' struktur, skulle en signifikant anden fremgangsmåde være benyttet. Her ville det have været fordelagtigt at beholde det initiale datagrundlag, hvorefter de enkelte modeller kunne estimeret med de manglende datapunkter. Dette ville have fjernet det indførte bias, ved imputering, men dog have estimeret modeller på et meget tyndt datagrundlag. Det er dog i praksis ikke muligt at validere, om manglende data følger en 'MCAR', 'MAR' eller 'MNAR' struktur. Sluteligt vil en 'MNAR' struktur af data også invalidere resultaterne i denne undersøgelse. Dette skyldes at i en 'MNAR' struktur er det essentielt at de manglende datapunkter bliver estimeret på baggrund af både strukturen i det kendte data, men også strukturen i de manglende datapunkter. I praksis ville der derfor skulle kendes den underliggende struktur i de manglende datapunkter, inden imputering.

#### Indflydelse på ESG scorer

På baggrund af undersøgelsen, er det værd et diskutere nærmere, i hvilken setting de manglende datapunkter har en stor indflydelse. I et studie viser (Mooij, 2017) at selve omkostningerne af rapporteringsprocessen for, de mere end 400, ESG mål kan vise sig at have en signifikant indflydelse på, om de enkelte virksomheder vælger at rapportere ESG data eller ej. Det er derfor interessant at teoretisere, hvorvidt det store omfang af manglende datapunkter kunne skyldes selve rapporteringsomkostningerne. Det kan konkluderes at ESG-scorerne er utsat for en stor mængde bias, både hvis og hvis ikke rapporteringsomkostningerne kan vises at være grunden til manglende data. Denne undersøgelse viste at kun 45 ud af 500 virksomheder havde tilstrækkelig data i det underliggende niveau, til modellering af log-afkast. Disse 45 virksomheder havde stadig så dårlig data, at flere af de tiltænkte variable måtte udelades fra analysen. Selv efter at flere af variablene var udeladt, var det stadig nødvendigt at imputere datapunkter, hvor få af variablene havde en inkomplethed på op til 50%. Det vil altså sige at selve datagrundlaget er yderst tvivlsomt, og hvis dette skyldes rapporteringsomkostninger, kan der argumenteres for, at jo stærkere virksomheden er økonomisk, jo højere ESG scorer vil den sandsynligvis også have. Præcis dette er omdiskuteret i en publisering af (Doyle, 2018), benævnt som 'company size bias'. Ydermere kan inkompletheden også skyldes manglen på standardiserede rapporteringsmuligheder, hvor nogle virksomheder potentielt ikke ser vigtigheden i at offentligøre deres 'E', 'S' og 'G' mål.

Hernæst er det værd at bemærke at mængden af investorer der benytter sig af ESG ratings er steget markant de seneste år. På den ene side viser denne undersøgelse at manglen på transparens og data, medfører at de konstruerede scorer er mindre valide end tidligere undersøgelser peger på. På den anden side er der en vis usikkerhed i resultaterne, hvorfor en mere omfangsrig undersøgelse kræves for at bakke den påstand op. Det er notesværdigt at investorer flokkes til for at investere bæredygtigt, på baggrund af ESG scorer, uden konkret at vide hvad der ligger bag de ”pæne” ESG mål. Ydermere kan der pointeres til potentiel bias i forhenværende undersøgelse, lavet på baggrund af ESG scorerne.

## 16 ESG-scorer leverandør

Da der er valgt udelukkende at benytte Thomson Reuters ESG scorer og data, skal denne undersøgelse sættes i perspektiv til andre ESG-scoremål. Som benævnt er der i de senere år, kommet flere leverandører af ESG-scorer, hvor Thomson Reuter blot er en i blandt dem. Det kunne tænkes at andre resultater var opnået ved at benytte sig at af en anden leverandør. Dette skal dog ikke tages for givet. Teoretisk burde ESG-score dataen i niveau 4 ikke være forskellig blandt leverandørerne, da det er faktuelt rådata fra virksomhederne. Det kan dog vise sig at Thomson Reuters data-kvalitet er lavere end de andre leverandørers. Dette er ikke valideret. En anden hovedpointe er selve konstruktionen af ESG scorerne, som kan vareiere mellem leverandørerne. Den gennemgåede fremgangsmåde i afsnit 3.4, er unik for netop konstruktion af Thomson Reuters ESG-scorer.

### Konstruktion og percentil industri rangering

Da de konstruerede ESG-scorer er baseret på et percentilt rang-system, vil de virksomheder der offentligøre mest data automatisk blive scoret forholdsmaessigt højere end andre virksomheder. Dette skyldes blandt andet at det percentile rang-system er ”matching” indenfor de enkelte industrier, se afsnit 3.4.

Da der bliver rangeret på industribasis, er de endelige ESG-scoremål også svære at sammenligne på tværs af industrier. På trods af det bruges ESG-scorene sammenligneligt som mål for de enkelt virksomheds performance inden for ’miljø’, ’social’ og ’governance’.

I et studie af (Tamimi and Sebastianelli, 2017) omkring transparens i ESG scoring for S&P500 virksomheder, findes signifikant forskel mellem transparensen blandt forskellige sektorer i indekset. Da denne undersøgelse heller ikke beskæftiger sig med en industriel adskillelse af virksomhederne, kan det potentielt have forværret resultaterne, da sammenligninggrundlagt ikke er unbiased. Da manglen på observerbare mønstre er gennemsyrende for denne undersøgelses del- og hovedresultater, er det vigtigt at pointere at selvom der ikke er fundet nogle indflydelsesrike mønstre i denne analyse, er det ikke i overenstemmelse med at disse ikke eksisterer. Den udeladte robusthedstest af industri-indekseringen, kunne inkludere modellering med sektor/industri ’dummies’ i panelregressionen, for at isolere effekten af ESG-variablene til enkelte industrier/sekター. Dette var dog i mindre grad relevant, da en inddeling på sektorniveau blot ville indeholde få eller ingen virksomheder for de forskellige sektorer.

Udover selve det percentile rang-system, er konstruktionen af ESG-scorene meget følsom. Det kan visualiseres som et isbjerg, hvor alt rådataen ligger under vandoverfladen og ikke er brugt i almen jargon eller investeringsprocesser. Det vil også sige at når største delen af de relevante data er

ufleksibelt udeladt af de konstruerede scorer, er selve scoren meget overfølsom over for både bias og kausalitetsproblemer.

## 17 Afsluttende om bias og kausalitet

Kausalitetsproblemer opstår på baggrund af ovenstående diskussioner. Primært er der to hoved-kausalitetsproblemer med ESG-scorene:

- 'Company size bias' medfører potentielt at godt performende virksomheder får høje ESG score, og ikke omvendt.
- Den omfattende konstruktionsprocess kan medføre uønsket bias i det kausale link: høj ESG score medfører højt afkast.

Kausalitetsproblemer ses som et gennemgående problem i litteraturen allerede gennemgået i afsnit 3. Her ses det at selve metodevalget i de enkelte undersøgelser har haft stor indflydelse på de enkelte resultater. Denne undersøgelse finder derfor sin plads i litteraturen med en yderst forsiktig og vel-dokumenteret tilgang til data og metode. I ovenstående diskussionsafsnit, om ESG-scorer leverandør og manglende data, er der diskuteret følgende bias problemer:

- Industrirangering kan vise sig at give biased resultater. Dette kommer i form af at enkelte industrier har mange virksomheder med gode ESG rådata mål, dog lav ESG-scoring.
- Bias i form af at ESG-data er forskelligt blandt leverandører.
- Bias på baggrund af manglende data punkter.

Dette ses også i gennemgangen af litteraturen, hvor flere undersøgelser er yderst prægede af bias. Der menes at ovenstående bias og kausalitetsproblemer har haft mulighed for præge denne undersøgelse, der kan dog ikke kvantificeres den præcise betydning.

## 18 Metodevalg

Denne undersøgelse er lavet med udgangspunkt i to hovedanalytiske principper: univariat tidsserieanalyse og paneldata regressionsanalyse. Disse metodevalg har præget opgavens resulter, delkonklusioner og konklusioner. Det kan derfor diskuteres, præcis hvilken indflydelse disse metodevalg har haft for denne undersøgelse.

### Tidsserie analyse

I første del af opgaven er der, for at besvare hoved- og delhypoteserne, lavet en univariat tidsserieanalyse. En univariat tidsserieanalyse egner sig godt til at modellere data, hvor der tiltænkes at fortidige observationer har en indflydelse på de fremtidige. Inkluderer man også eksterne kovariate er det pludseligt muligt at basere den fremtidige værdi på det fortidige, samt værdierne af det relevant økonomisk data. Denne type analyse egner sig derfor godt til at teste indflydelsen af det valgte økonomiske

data på den afhængige tidsserie. Da ARMA-X/GARCH modellerne ikke egner sig til at modellere ikke-stationært data, er de eksterne kovariate ofte transformert. I denne undersøgelse har transformationer medført at de eksterne kovariate er givet i procentuelle eller absolute termer. Det vil altså sige at et direkte sammenhæng mellem den ikke transformerede eksterne kovariat og den afhængige variable, kan være svært at udlede.

Univariate modeller er dog specielt brugbare når der ønskes at finde dynamikker indenfor enkelte virksomheder. Det betyder også at dynamikker på tværs af virksomheder ikke er mulige at modeller ved brug af disse modeller, hvorfor paneldata regressionanalyse også er benyttet.

### Paneldata analyse

For videre at undersøge resultaterne drevet af den univariate tidsserianalyse, fulgte en paneldata regressionsanalyse. (Hsiao, 2007) belyser både fordele og ulemper netop ved brug af paneldata regressionsanalyse. For det første er paneldata kollektion generelt mere krævende end for tidsserieanalyse. Med henblik på datavalg til denne form af analyse, er det essentielt at kvaliteten af data er ens på tværs af panelet (balanceret), for ikke at opnå biased estimator. Dette har haft en signifikant indflydelse på de endelige datasæt i denne undersøgelse, da mange af de samplingvalg der er truffet i afsnit 6, er foretaget for at forøge den gennemsnitlige datakvalitet til paneldata regressionanalyse.

I paneldata regressionanalyse ses tre hovedfordele i forhold til tidsserieanalyse:

- Mere komplekse modeller, hvor det er muligt at opfange dynamikker på tværs af 'individer'.
- Analyse af ikke stationære tidsserier.
- Kontrol af effekten af udeladte variable.

Paneldata regressionersresultater, kan være sværere at fortolke, og baseres på et flertal af antagelser, som skal overholdes, for at estimaterne er konsistente. Effekten af udeladte variable argumenteres i litteraturen, som den reelle drivkraft bag de fundne resultater (eller mangle på disse), ved korrelation mellem de udeladte variable og de inkluderede forklarende variable (Hsiao, 2007). Opsummerende giver paneldata mere informativt data, mere variation, mindre kolinearitet mellem variable, flere frihedsgrader og mere efficiens (Baltagi, 2015).

På baggrund af de endelig resultater i denne undersøgelse, kunne man argumentere for at en anden tilgang til data kunne have været mere fordelagtig, dette betvivles dog. Motivationen for brug af OLS baseres på Gauss-Markov antagelserne. Det viste sig at indførelsen af robuste standardfejl og ved fjernelse af en enkelt variabel overholdte OLS kriterierne for at være BLUE.

Slutteligt kan der argumenteres for at samhørigheden af de to analytiske tilgænge til data har været med til at validere de endelige resultater.

### Afsluttende bemærkninger om metodevalg

Udover de anvendte metoder til at besvare hoved og del-hypoteser, findes der et utal andre metoder. Blandt andet en mere porteføljekninsk tilgang, som Farma-French's tre-faktor model. Farma-French beskriver, teoretisk, aktie-afkast ved brug af tre primær faktorer: markedsrisiko, differensen mellem små og store virksomheder og differensen mellem høje og lave 'book to market' virksomheder. Der

findes et flertal af udvide Farma-French modeller, hvor ekstra faktorer er tilføjet til modellen, så som ESG-scoren. (Melas et al., 2016) viser blandt andet ved brug af dette at en høj ESG-score karakteriserer en virksomhed med høj markedsværdi, høje dividender og er af god ”kvalitet”. Det kunne tænkes at en Farma-French metodisk tilgæng, hvor de eksterne variable samles i bestemte ”grupper” ville have belyst problemstillingen på en anden vis.

## Part V

### 19 Konklusion

Denne undersøgelse ønskede at besvare to primære problemstillinger. 1: Hvilken grad af indflydelse har de forskellige niveauer af ESG scorer på virksomheders finansielle performance, i form af afkast på det underliggende aktiv? 2: Hvilke underliggende ESG faktorer ligger til grund for indflydelsen? Vi teoretiserede at effekten af ESG scorernes underliggende niveauer er større end i det overordnede niveau, at høje ESG scorer har en positiv indflydelse, og at miljø variablen har en signifikant indflydelse, på virksomheders finansielle performance.

For at kunne besvare vores problemstillinger benyttede vi os af en struktureret og metodisk dataudvælgelses proces. Her valgte vi at analysere virksomheder i S&P500 med henblik på indeksets historiske overlegne datakvalitet. Udover ESG-scorer, valgte vi også at hente miljø- og rådata fra ESG niveau fire. Udvælgelsen af data har forsøgt at minimere bias i det endelige sample. Ydermere muliggjorde den robuste dataindsamlingsprocess nøje studeren af de latente bagvedliggende faktorer i ESG-scorene. Vi minimerede hernæst mængden af manglende datapunkter i en robust andensorteringsprocess. Her blev der metodisk fravalgt én enkelt variabel eller virksomhed en af gangen for at tjekke om det ville forbedre det eksisterende datagrundlag. Det endelige sample er valgt på baggrund af et cut-off niveau på 50% manglende datapunkter for en given variable, per virksomhed. Dette er ikke set før i litteraturen, hvor systematisk validering af kvaliteten af ESG-score data førhen har været udeladt. Denne samplingsproces resulterede i et endeligt datasæt bestående af 45 virksomheder, med deres tilhørende ESG-score, søjle-score, under-score, 21 miljø og rådata variable samt 3 kontrolvariable.

Vi valgte at benytte os af data imputering, for at forbedre det nøje udvalgte datagrundlag yderligere. Her blev flere forskellige imputeringsmetoder overvejet, men slutteligt valgte vi at benytte en variant af multipel imputation. Vi imputerede manglende datapunkter i tre omgange, og kombinerede hernæst disse til ét endeligt datasæt, imputering(1). Herved opnåede vi robuste estimerater af de manglende datapunkter. Hernæst imputerede vi yderlige tre datasæt og kombinerede dem igen for at konstruerede et robust datasæt til kontrol af imputering(1).

Vi viste at alle 45 virksomheders log-afkast ikke fulgte den samme fordeling. På trods af dette valgte vi at modellere univariate tidsserie modeller med normalfordelingsantagelsen intakt, for at danne det bedst mulige sammenligningsgrundlag. Vi viste også at de månedlige log-afkast indeholder autokorrelation og at en signifikant mængde af disse også indeholder heteroskedacitet. Med henblik på at modellere både autokorrelation og heteroskedacitet valgte vi at benytte os af ARMA-X/GARCH modeller, hvor vi specifikt valgte ARMA-X(p,q)/GARCH(1,1) modeller på baggrund af eksisterende litteratur. Hernæst modellerede vi over ti-tusinde ARMA-X(p,q)/GARCH(1,1) for at udvælge de modeller der fittede data bedst. På baggrund af et endeligt grundlag, bestående af 345 modeller tjekkede vi in-sample fits for modeller med udvalgte kovariate.

Da univariate tidsserier ikke kan opfange dynamikker på tværs af virksomhederne, og videre er restrikterede til at modellere ikke stationære tidsserier, valgte vi at undersøge virksomhederne på tværs af både tids- og enhedsdimensionen. Dette gjorde vi ved brugen af panel data regressionsanalyse, der ydermere tillader at håndtere effekterne fra udeladte variable. Vi valgte at undersøge de tre generelle

paneldata regressionsmodeller, pooled, fixed og random effekt for hver af de fire ESG niveauer. På baggrund af overvejeler om udeladte virksomhedsspecifikke effekter og ligeledes udeladte tidseffekter, inkluderede vi to-vejs effekter til de passende modeller. På baggrund af passende F-test, Breusch-Pagan Lagrange multiplier test og Hausman test, viste det sig at en fixed effekts model med to-vejs effekter var foretrukket. Dette var i overensstemmelse med de indledende overvejser om passende modelspecifikationer.

Efterfølgende testede vi Gauss-Markov antagelserne for OLS, hvor vi fandt at indførelsen af en robust kovariansmatrice viste sig nyttig til at tage hånd om autokorrelation og heteroskedaciteten. På baggrund af de indledende modelspecifikationer foretog vi en modeloptimering for at finde de enkeltstående signifikante variable i de forskellige niveauer. Her viste det sig at de fittede modeller udviste mangel på forklaringsevner, hvorfor det ikke videre var muligt at konkludere på effekterne af variablene.

Til sidst testede vi robustheden af de endelige modeller. Her viste vi at de metodiske valg truffet igennem undersøgelsen var robuste mod ændringer i informationskriterie, datafrekvens, imputering, indflydelsesobservationer og valg af tidsinterval.

Vi kan slutteligt afkræfte hovedhypotesen, samt begge underhypoteser. Både vores univariate tidsserie analyse og paneldata regressionsanalyse viste at der ikke eksisterer nogen sammenhæng mellem ESG-variablene og virksomhedernes afkast. Det viste sig derfor ikke muligt at differentierer indflydelesesgraden af de forskellige ESG-variable.

## 20 Fremtidig forskning

I nyere tider er bæredygtig et emne der er på alles læber. Denne undersøgelse forholder sig til bæredygtighedsmål, og data, vedrørende finansielle virksomheder.

Det er vist, hvordan disse bæredygtighedsmål ikke har nogen direkte indflydelse på afkastet af S&P500 virksomheder, grundende er dog uvisse. På baggrund af denne undersøgelse er der et stort grundlag for fremtidig forskning med henblik på, hvorfor de populære ESG-mål ikke kan vises at have en indflydelse på virksomheders finansielle performance. Samtidig danner denne undersøgelse også fundament for metodetekniske valg af fremtidig forskning. Der teoretiseres i denne undersøgelse om, hvorvidt den basale mangel på datakvalitet kan være et validt grundlag til at invalidere ESG-score målene. Det tænkes at fremtidig forskning inden for netop dette område, kan baseret til dels på de konklusioner der drages i denne opgave.

Primært kan denne undersøgelse konkludere at fremtidig forskningen inden for andre bæredygtige datamål kræves for at validere de fundne resultater. De problemer denne undersøgelse har prøvet at overkomme med henblik på mangelfuld data, kan der med fordel undgås ved at benytte et opdateret og mere komplet datagrundlag. Her peges der i mod en større proces, hvor aktiv indsamling af data med stor sandsynlighed er krævet. Sluteligt, men ikke mindst, kunne det være spændende at beskæftige sig med multivariete modellering i et tidsserie-setting for at udvide resultaterne af denne undersøgelse.

## References

- Alexander, G. J. and Buchholz, R. A. (1978). Corporate social responsibility and stock market performance. *Academy of Management journal*, 21(3):479–486.
- Auer, B. R. and Schuhmacher, F. (2016). Do socially (ir) responsible investments pay? new evidence from international esg data. *The Quarterly Review of Economics and Finance*, 59:51–62.
- Avetisyan, E. and Hockerts, K. (2017). The consolidation of the esg rating industry as an enactment of institutional retrogression. *Business Strategy and the Environment*, 26(3):316–330.
- Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of accounting research*, pages 159–178.
- Baltagi, B. (2015). *Econometric Analysis of Panel Data*. Wiley.
- Barnard, J. and Meng, X.-L. (1999). Applications of multiple imputation in medical studies: from aids to nhanes. *Statistical methods in medical research*, 8(1):17–36.
- Baumol, W. J. and Blackman, S. A. B. (1991). Perfect markets and easy virtue business ethics and the invisible hand.
- Bebchuk, L. A., Cohen, A., and Wang, C. C. (2013). Learning and the disappearing association between governance and returns. *Journal of financial economics*, 108(2):323–348.
- Becker, B. E., Huselid, M. A., Pickus, P. S., and Spratt, M. F. (1997). Hr as a source of shareholder value: Research and recommendations. *Human Resource Management: Published in Cooperation with the School of Business Administration, The University of Michigan and in alliance with the Society of Human Resources Management*, 36(1):39–47.
- Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. Wiley, New York.
- Bernard, V. L. and Thomas, J. K. (1989). Post-earnings-announcement drift: delayed price response or risk premium? *Journal of Accounting research*, 27:1–36.
- Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3):637–654.
- Blackrock (2020). What is sustainable investing? <https://www.blackrock.com/ch/individual/en/themes/sustainable-investing>. (Besøgt: 15.05.2020).
- Borgers, A., Derwall, J., Koedijk, K., and Ter Horst, J. (2013). Stakeholder relations and stock returns: On errors in investors' expectations and learning. *Journal of Empirical Finance*, 22:159–175.
- Box, G. E. and Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical association*, 70(349):70–79.

- Brammer, S., Brooks, C., and Pavelin, S. (2006). Corporate social performance and stock returns: Uk evidence from disaggregate measures. *Financial management*, 35(3):97–116.
- Breusch, T. S. and Pagan, A. R. (1980). The lagrange multiplier test and its applications to model specification in econometrics. *The review of economic studies*, 47(1):239–253.
- Cavanaugh, J. E. (1997). Unifying the derivations for the akaike and corrected akaike information criteria. *Statistics & Probability Letters*, 33(2):201–208.
- Chua, C. T., Goh, J., and Zhang, Z. (2010). Expected volatility, unexpected volatility, and the cross-section of stock returns. *Journal of Financial Research*, 33(2):103–123.
- Cohen, M. A., Fenn, S., and Naimon, J. S. (1995). *Environmental and financial performance: are they related?* Citeseer.
- Core, J. E., Holthausen, R. W., and Larcker, D. F. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of financial economics*, 51(3):371–406.
- Daszynska-Zygadlo, K., Slonski, T., Zawadzki, B., et al. (2016). The market value of csr performance across sectors. *Engineering Economics*, 27(2):230–238.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22.
- Derwall, J. (2007). *The economic virtues of SRI and CSR*. Number 101.
- Dimson, E., Karakaş, O., and Li, X. (2015). Active ownership. *The Review of Financial Studies*, 28(12):3225–3268.
- Dixon-Fowler, H. R., Slater, D. J., Johnson, J. L., Ellstrand, A. E., and Romi, A. M. (2013). Beyond “does it pay to be green?” a meta-analysis of moderators of the cep–cfp relationship. *Journal of business ethics*, 112(2):353–366.
- Doyle, T. (2018). Ratings that don’t rate: The subjective world of esg ratings agencies. *Report, American Council for Capital Formation*.
- Edmans, A. (2011). Does the stock market fully value intangibles? employee satisfaction and equity prices. *Journal of Financial economics*, 101(3):621–640.
- Efron, B. (1992). Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer.
- Egerod, B. (2016). Hvorfor ols er den bedste estimator.
- EU (2015). Paris Agreement. [https://ec.europa.eu/clima/policies/international/negotiations/paris\\_en](https://ec.europa.eu/clima/policies/international/negotiations/paris_en). (Besøgt: 10.10.2019).
- Faleye, O. and Trahan, E. A. (2011). Labor-friendly corporate practices: Is what is good for employees good for shareholders? *Journal of Business Ethics*, 101(1):1–27.

- Fama and French (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.
- Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636.
- Finlay, Barbara og Agresti, A. (2013). *Statistical methods for the social sciences*. Dellen.
- Fox, J. (2015). *Applied regression analysis and generalized linear models*. Sage Publications.
- Freeman, R. N. (1987). The association between accounting earnings and security returns for large and small firms. *Journal of accounting and economics*, 9(2):195–228.
- Friede, G., Busch, T., and Bassan, A. (2015). Esg and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4):210–233.
- Friedman, M. (1970). A friedman doctrine: The social responsibility of business is to increase its profits. *The New York Times Magazine*, 13(1970):32–33.
- Gallup (2019). Millennials Worry About the Environment - Should Your Company? <https://www.gallup.com/workplace/257786/millennials-worry-environment-company.aspx>. (Besøgt: 10.10.2019).
- Gelman, A. and Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press.
- Gompers, P., Ishii, J., and Metrick, A. (2003). Corporate governance and equity prices. *The quarterly journal of economics*, 118(1):107–156.
- Greene, W. H. (2018). *Econometric Analysis*. Pearson, New York, NY.
- Gu, L. and Hackbarth, D. (2013). Governance and equity prices: Does transparency matter? *Review of finance*, 17(6):1989–2033.
- Guenster, N., Bauer, R., Derwall, J., and Koedijk, K. (2011). The economic value of corporate eco-efficiency. *European Financial Management*, 17(4):679–704.
- Halbritter, G. and Dorfleitner, G. (2015). The wages of social responsibility—where are they? a critical review of esg investing. *Review of Financial Economics*, 26:25–35.
- Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a garch (1, 1)? *Journal of applied econometrics*, 20(7):873–889.
- Hazilla, M. and Kopp, R. J. (1990). Social cost of environmental quality regulations: A general equilibrium analysis. *Journal of Political Economy*, 98(4):853–873.
- HBR (2019). The Investor Revolution. <https://hbr.org/2019/05/the-investor-revolution>. (Besøgt: 10.10.2019).

- Honaker, J. and King, G. (2010). What to do about missing values in time-series cross-section data. *American journal of political science*, 54(2):561–581.
- Hong, H. and Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1):15–36.
- Hsiao, C. (2007). Panel data analysis—advantages and challenges. *Test*, 16(1):1–22.
- Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
- Hyndman, R. J., Khandakar, Y., et al. (2007). *Automatic time series for forecasting: the forecast package for R*. Number 6/07.
- Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.
- Johnston, R., Jones, K., and Manley, D. (2018). Confounding and collinearity in regression analysis. *Quality & quantity*, 52(4):1957–1976.
- Kempf, A. and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, 13(5):908–922.
- Kim, Y., Park, M. S., and Wier, B. (2012). Is earnings quality associated with corporate social responsibility? *The accounting review*, 87(3):761–796.
- Klassen, R. D. and McLaughlin, C. P. (1996). The impact of environmental management on firm performance. *Management science*, 42(8):1199–1214.
- Kropko, J. and Kubinec, R. (2020). Interpretation and identification of within-unit and cross-sectional variation in panel data models. *PloS one*, 15(4):e0231349.
- Maddala, G. S. and Lahiri, K. (1992). *Introduction to econometrics*, volume 2. Macmillan New York.
- Manuel, A. (1987). Computing robust standard errors for within group estimators. *Oxford bulletin of Economics and Statistics*, 49(4):431–434.
- Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*, volume 16. John Wiley New York.
- McGuire, J. B., Sundgren, A., and Schneeweis, T. (1988). Corporate social responsibility and firm financial performance. *Academy of management Journal*, 31(4):854–872.
- McQuarrie, A. D. and Tsai, C.-L. (1998). *Regression and time series model selection*. World Scientific.
- McWilliams, A. and Siegel, D. (2000). Corporate social responsibility and financial performance: correlation or misspecification? *Strategic management journal*, 21(5):603–609.
- Melas, D., Nagy, Z., and Kulkarni, P. (2016). Factor investing and esg integration. *MSCI Research Insight*.

- Merton, R. C. (1987). A simple model of capital market equilibrium with incomplete information. *The journal of finance*, 42(3):483–510.
- Millo, G. (2017). Robust standard error estimators for panel models: a unifying approach. *Journal of Statistical Software*, 82(3).
- Mooij, S. (2017). The esg rating and ranking industry; vice or virtue in the adoption of responsible investment? *Vice or Virtue in the Adoption of Responsible Investment*.
- Morningstar (2020). Sustainable Fund Flows in 2019 Smash Previous Records. <https://www.morningstar.com/articles/961765/sustainable-fund-flows-in-2019-smash-previous-records>. (Besøgt: 15.01.2020).
- Nielsen (2018). Was 2018 the year of the influential sustainable costumer. <https://www.nielsen.com/us/en/insights/article/2018/was-2018-the-year-of-the-influential-sustainable-consumer/>. (Besøgt: 10.11.2019).
- Orlitzky, M., Schmidt, F. L., and Rynes, S. L. (2003). Corporate social and financial performance: A meta-analysis. *Organization studies*, 24(3):403–441.
- Palmer, K., Oates, W. E., and Portney, P. R. (1995). Tightening environmental standards: the benefit-cost or the no-cost paradigm? *Journal of economic perspectives*, 9(4):119–132.
- Porter, M. E. and Van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives*, 9(4):97–118.
- Refinitiv (2019). Environmental, social and governance (esg) scores from refinitiv. June.
- Ribando, J. M. and Bonne, G. (2010). A new quality factor: Finding alpha with asset4 esg data. *Starmine Research Note, Thomson Reuters*, 31.
- Roberts, M. R. and Whited, T. M. (2013). Endogeneity in empirical corporate finance1. In *Handbook of the Economics of Finance*, volume 2, pages 493–572. Elsevier.
- Ruppert, D. (2010). *Statistics and Data Analysis for Financial Engineering*. Springer Texts in Statistics. Springer New York.
- Schafer, J. L. and Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate behavioral research*, 33(4):545–571.
- Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.
- Statman, M. and Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 65(4):33–46.
- Stock, J. H. and Watson, M. W. (2010). *Introduction to econometrics*.
- Tamimi, N. and Sebastianelli, R. (2017). Transparency among s&p 500 companies: An analysis of esg disclosure scores. *Management Decision*.

- Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, 44(12-part-1):1650–1661.
- Torres-Reyna, O. (2010). Getting started in fixed/random effects models using r. *Data & Statistical Services. Princeton University*.
- Tsay, R. S. (2010). *Analysis of financial time series*. John wiley & sons.
- Ullmann, A. A. (1985). Data in search of a theory: A critical examination of the relationships among social performance, social disclosure, and economic performance of us firms. *Academy of management review*, 10(3):540–557.
- Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.
- Velte, P. (2016). Women on management board and esg performance. *Journal of Global Responsibility*.
- Waddock, S. A. and Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic management journal*, 18(4):303–319.
- Weisberg, S. (2005). *Applied linear regression*, volume 528. John Wiley & Sons.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: journal of the Econometric Society*, pages 817–838.
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the Econometric Society*, pages 1–25.
- Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30(1):79–82.
- Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.
- Yadav, P. L., Han, S. H., and Rho, J. J. (2016). Impact of environmental performance on firm value for sustainable investment: Evidence from large us firms. *Business strategy and the environment*, 25(6):402–420.
- Yao, Q. and Brockwell, P. J. (2006). Gaussian maximum likelihood estimation for arma models. i. time series. *Journal of time series analysis*, 27(6):857–875.

## Part VI

# Appendix

## A Dataarbejde

### Imputerings tabel - Robusthed

	Middelværdi	Median	std	Min.	Max.
CO2 Udledning - Total	3148943.6464	1024056.5	6856306.5032	-22828240.3891	68284041
Energi - Totalt Forbrug	34582017.8936	10471402.3092	91167312.4601	-259599245.9641	868311006
CO2 Udledning - Estimeret	1e-04	1e-04	2e-04	-2e-04	0.0011
Vandforbrug - Total / USD	0.0033	4e-04	0.0132	-0.0135	0.1172
Affald - Genbrug / Total	0.002	0	0.0129	-0.0173	0.1341
Affald - Total Genbrug	109626.411	44168.18	249814.1182	-1098028.5528	3152000
Affald - Total	14030846.8456	79467.3725	90334732.0579	-129585650.7454	885588319.9549
Vand - / Genbrug	56.2896	60.684	25.5669	-38.9248	102.1311
Vand - Total Tilbagetrækning	43251583.5079	8538500	157798773.4746	-124519475.0785	1.738e+09
Affald - Total / USD	0.5643	0.6061	0.2558	-0.3601	1.0024
Bestyrrelsen - Gns. Siddetid	8.5379	8.7273	2.2939	1.9286	14.8636
Bestyrrelsen - Kønsfordeling	20.5791	20	8.2323	0	45.4545
Bestyrrelsen - Kompensation	2984284.7868	2853389	1342636.4332	-617013.8403	14685739
Senior Executive - Kompensation	38715121.0523	33770530	21938161.056	1564816	189004246
Executive - Kønsfordeling	15.7039	15.3846	10.9109	0	50

Table 34: Imputtering (2) - data summary

### Korrelationsmatricer

Tabellerne ses på næste side

Table 35: Korrelations matrix

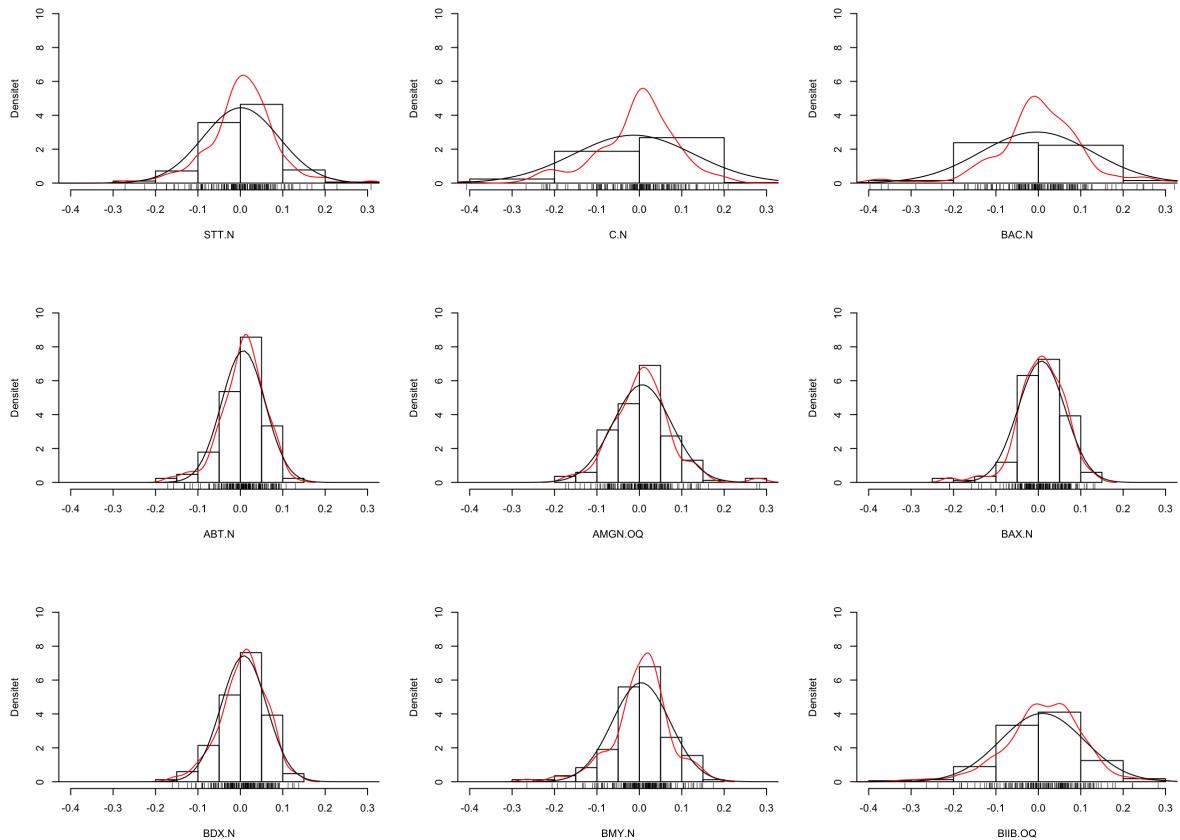
Variabel	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$
CO2 Udledning - Total ( $x_1$ )	1	<b>0.8364</b>	0.6480	0.2606	0.1671	0.5588	0.5194	-0.2599	0.6876	-0.2600	0.0692	0.1259	0.0735	0.1611	-0.066
Energi - Totalt Forbrug ( $x_2$ )	<b>0.8364</b>	1	0.5451	0.2416	0.1561	0.4163	0.4424	-0.2928	0.6105	-0.2938	0.0866	0.0940	0.1117	0.1471	-0.0382
CO2 Udledning - / USD ( $x_3$ )	0.6480	0.5451	1	0.5418	0.4681	0.2188	0.4775	-0.4641	0.5202	-0.4626	0.0997	-0.0843	0.0172	-0.0871	-0.1526
Vandforbrug - / USD ( $x_4$ )	0.2606	0.2416	0.5418	1	0.7262	0.0941	0.4808	-0.3417	0.5343	-0.3386	-0.0369	-0.0189	-0.0312	-0.1100	-0.0560
Affald - Genbrug / Total ( $x_5$ )	0.1671	0.1561	0.4681	0.7262	1	0.0163	0.5680	-0.3403	0.3619	-0.3343	-0.0808	-0.0047	-0.0676	-0.1469	-0.0972
Affald - Total Genbrug ( $x_6$ )	0.5588	0.4163	0.2188	0.0941	0.0163	1	0.5973	0.2066	0.5748	0.2080	0.1201	0.0523	-0.1259	-0.1012	0.0176
Affald - Total ( $x_7$ )	0.5194	0.4424	0.4775	0.4808	0.5680	0.5973	1	-0.2645	0.6871	-0.2609	0.0582	-0.0487	-0.1474	-0.1470	-0.1375
Vandforbrug - / Genbrug ( $x_8$ )	-0.2599	-0.2928	-0.4641	-0.3417	-0.3403	0.2066	-0.2645	1	-0.2222	<b>0.9954</b>	0.0242	0.0976	-0.1361	-0.0878	0.2216
Vand - Total Tilbagetrækning ( $x_9$ )	0.6876	0.6105	0.5202	0.5343	0.3619	0.5748	0.6871	-0.2222	1	-0.2192	0.0558	0.1171	-0.0176	0.0379	-0.0846
Affald - Total / USD ( $x_{10}$ )	-0.2600	-0.2938	-0.4626	-0.3386	-0.3343	0.2080	-0.2609	<b>0.9954</b>	-0.2192	1	0.0200	0.0951	-0.1407	-0.0908	0.2169
Bestyrreisen - Gns. Siddetid ( $x_{11}$ )	0.0692	0.0866	0.0997	-0.0369	-0.0808	0.1201	0.0582	0.0242	0.0558	0.0200	1	-0.0264	-0.0917	-0.1099	0.0359
Bestyrreisen - Kønsfordeling ( $x_{12}$ )	0.1259	0.0940	-0.0843	-0.0189	-0.0047	0.0523	-0.0487	0.0976	0.1171	0.0951	-0.0264	1	0.1336	0.1874	0.2820
Bestyrreisen - Kompenstation ( $x_{13}$ )	0.0735	0.1117	0.0172	-0.0312	-0.0676	-0.1259	-0.1474	-0.1361	-0.0176	-0.1407	-0.0917	0.1336	1	0.5338	0.1413
Senior Executive - Komp. ( $x_{14}$ )	0.1611	0.1471	-0.0871	-0.1100	-0.1469	-0.1012	-0.1470	-0.0878	0.0379	-0.0908	-0.1099	0.1874	0.5338	1	0.1976
Executive - Kønsfordeling ( $x_{15}$ )	-0.066	-0.0382	-0.1526	-0.056	-0.0972	0.0176	-0.1375	0.2216	-0.0846	0.2169	0.0359	0.2820	0.1413	0.1976	1

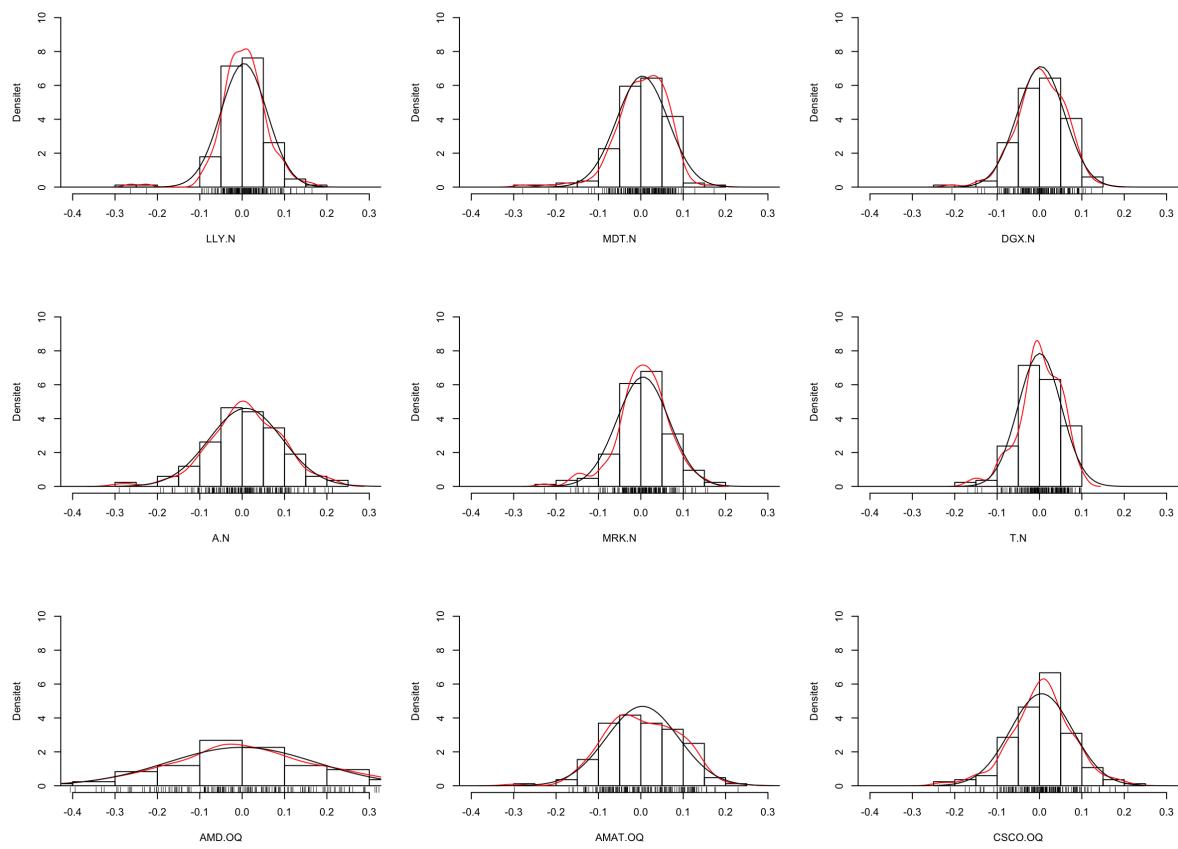
Table 36: Korrelations matrix for relevante transformationer

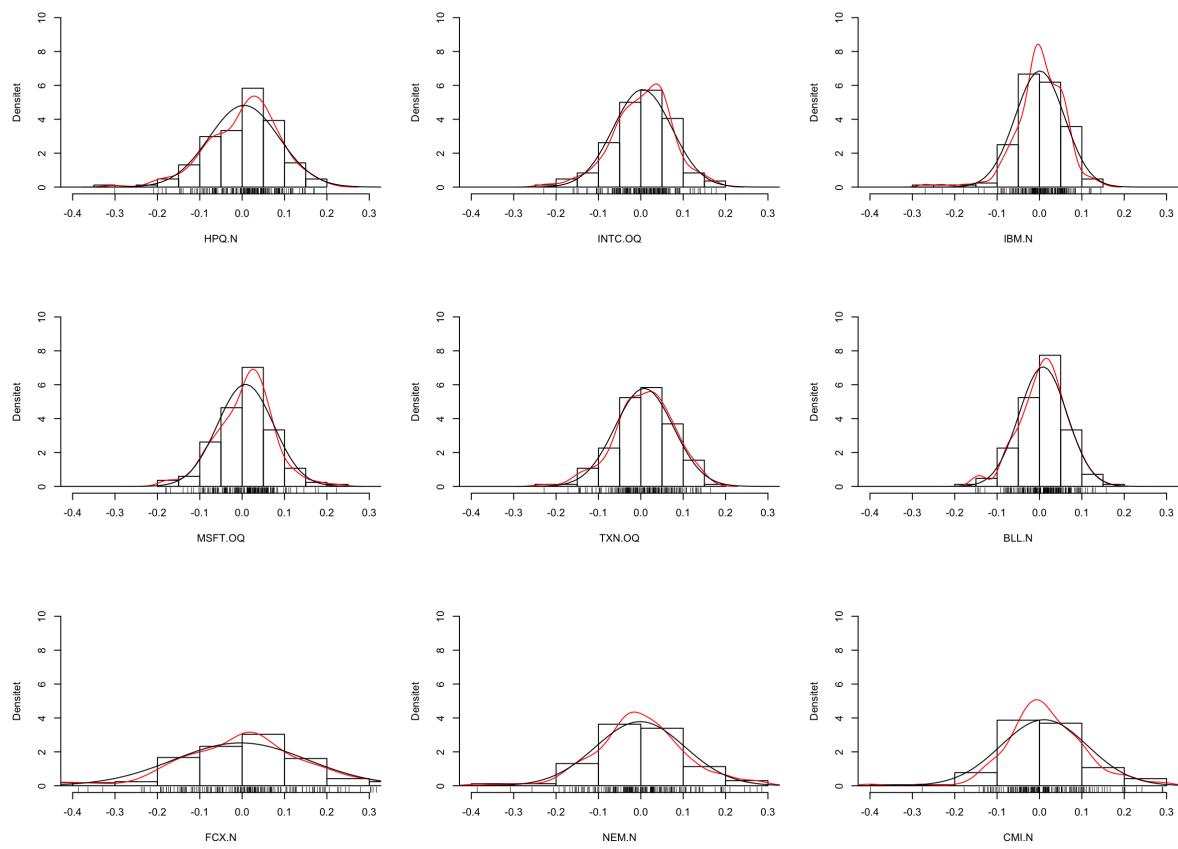
Variabel	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$
CO2 Udledning - Total ( $x_1$ )	1	<b>0.9441</b>	0.4924	0.1241	0.0386	0.1127	0.0482	-0.2823	0.2554	-0.2849	-0.1017	0.0148	0.1580	0.1625	-0.0843
Energi - Totalt Forbrug ( $x_2$ )	<b>0.9441</b>	1	0.4022	0.1037	0.0334	0.0879	0.0460	-0.2716	0.2277	-0.2737	-0.1161	0.0201	0.1727	0.1618	-0.0684
CO2 Udledning - / USD ( $x_3$ )	0.4924	0.4022	1	0.5418	0.4681	0.1155	0.4665	-0.4641	0.4571	-0.4626	0.0997	-0.0843	0.0429	0.0751	-0.1526
Vandforbrug - / USD ( $x_4$ )	0.1241	0.1037	0.5418	1	0.7262	0.0318	0.7070	-0.3417	<b>0.8743</b>	-0.3386	-0.0369	-0.0189	0.0043	-0.0711	-0.0560
Affald - Genbrug / Total ( $x_5$ )	0.0386	0.0334	0.4681	0.7262	1	0.0138	<b>0.9640</b>	-0.3403	0.3630	-0.3343	-0.0808	-0.0047	-0.0582	-0.0916	-0.0972
Affald - Total Genbrug ( $x_6$ )	0.1127	0.0879	0.1155	0.0318	0.0138	1	-0.0027	0.2236	0.0462	0.2262	0.0515	-0.0263	-0.1150	-0.1308	0.0069
Affald - Total ( $x_7$ )	0.0482	0.0460	0.4665	0.7070	<b>0.9640</b>	-0.0027	1	-0.3586	0.3845	-0.3537	-0.0884	0.0148	-0.0495	-0.0796	-0.1023
Vandforbrug - / Genbrug ( $x_8$ )	-0.2823	-0.2716	-0.4641	-0.3417	-0.3403	0.2236	-0.3586	1	-0.2888	<b>0.9954</b>	0.0242	0.0976	-0.1413	-0.1031	0.2216
Vand - Total Tilbagetrækning ( $x_9$ )	0.2554	0.2277	0.4571	<b>0.8743</b>	0.3630	0.0462	0.3845	-0.2888	1	-0.2872	-0.0148	9e-04	0.0821	0.0274	0.0014
Affald - Total / USD ( $x_{10}$ )	-0.2849	-0.2737	-0.4626	-0.3386	-0.3343	0.2262	-0.3537	<b>0.9954</b>	-0.2872	1	0.0200	0.0951	-0.1462	-0.1075	0.2169
Bestyrreisen - Gns. Siddetid ( $x_{11}$ )	-0.1017	-0.1161	0.0997	-0.0369	-0.0808	0.0515	-0.0884	0.0242	-0.0148	0.0200	1	-0.0264	-0.0930	-0.1434	0.0359
Bestyrreisen - Kønsfordeling ( $x_{12}$ )	0.0148	0.0201	-0.0843	-0.0189	-0.0047	-0.0263	0.0148	0.0976	9e-04	0.0951	-0.0264	1	0.0735	0.1366	0.2820
Bestyrreisen - Kompenstation ( $x_{13}$ )	0.1580	0.1727	0.0429	0.0043	-0.0582	-0.1150	-0.0495	-0.1413	0.0821	-0.1462	-0.0930	0.0735	1	0.5205	0.1403
Senior Executive - Komp. ( $x_{14}$ )	0.1625	0.1618	-0.0751	-0.0711	-0.0916	-0.1308	-0.0796	-0.1031	0.0274	-0.1075	-0.1434	0.1366	0.5205	1	0.1854
Executive - Kønsfordeling ( $x_{15}$ )	-0.0843	-0.0684	-0.1526	-0.056	-0.0972	0.0069	-0.1023	0.2216	0.0014	0.2169	0.0359	0.2820	0.1403	0.1854	1

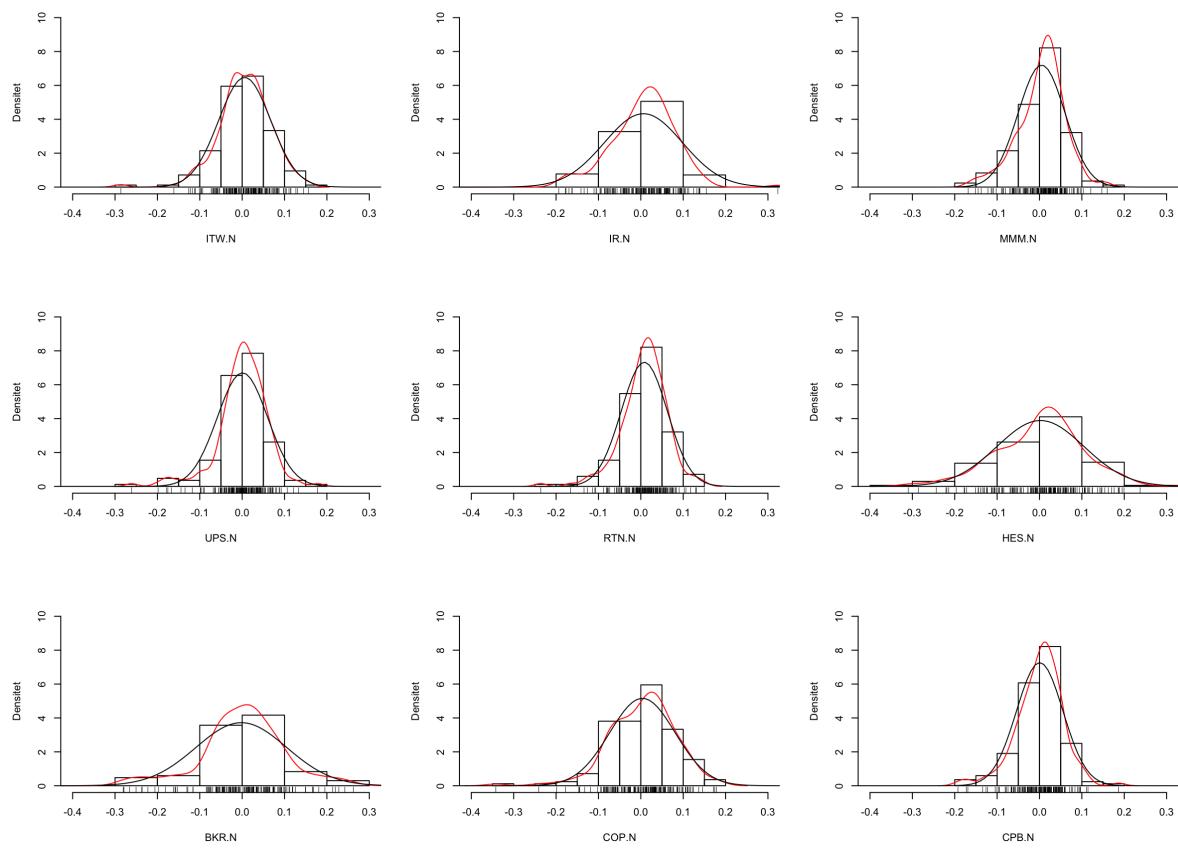
## B Tidsserieanalyse

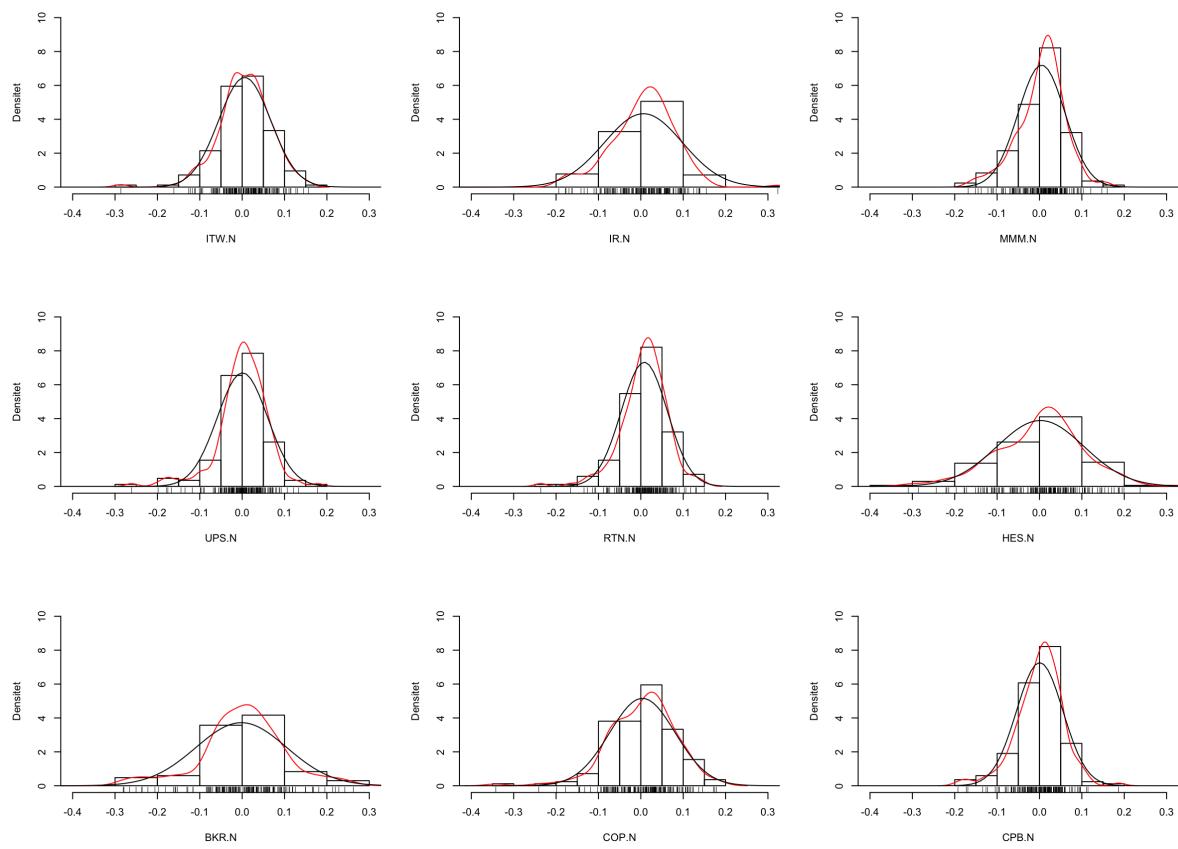
### Densitets plot



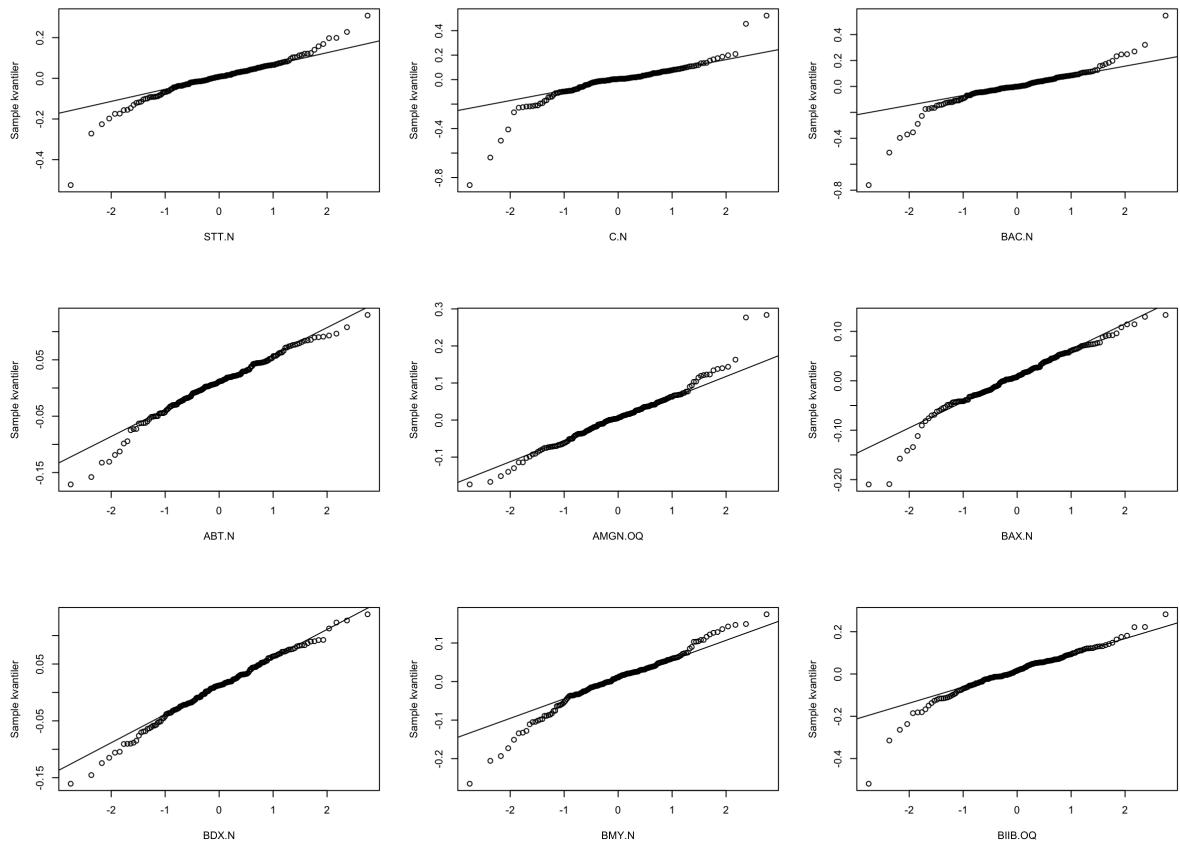


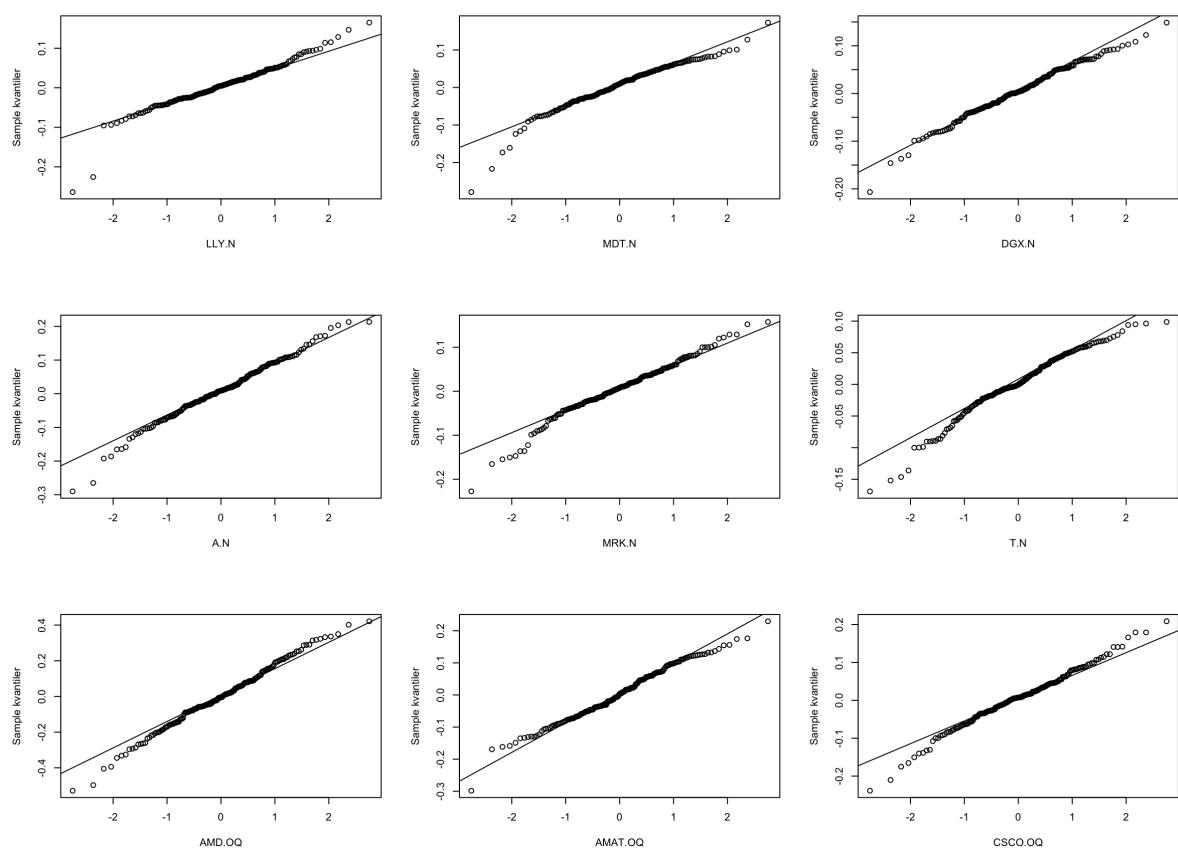


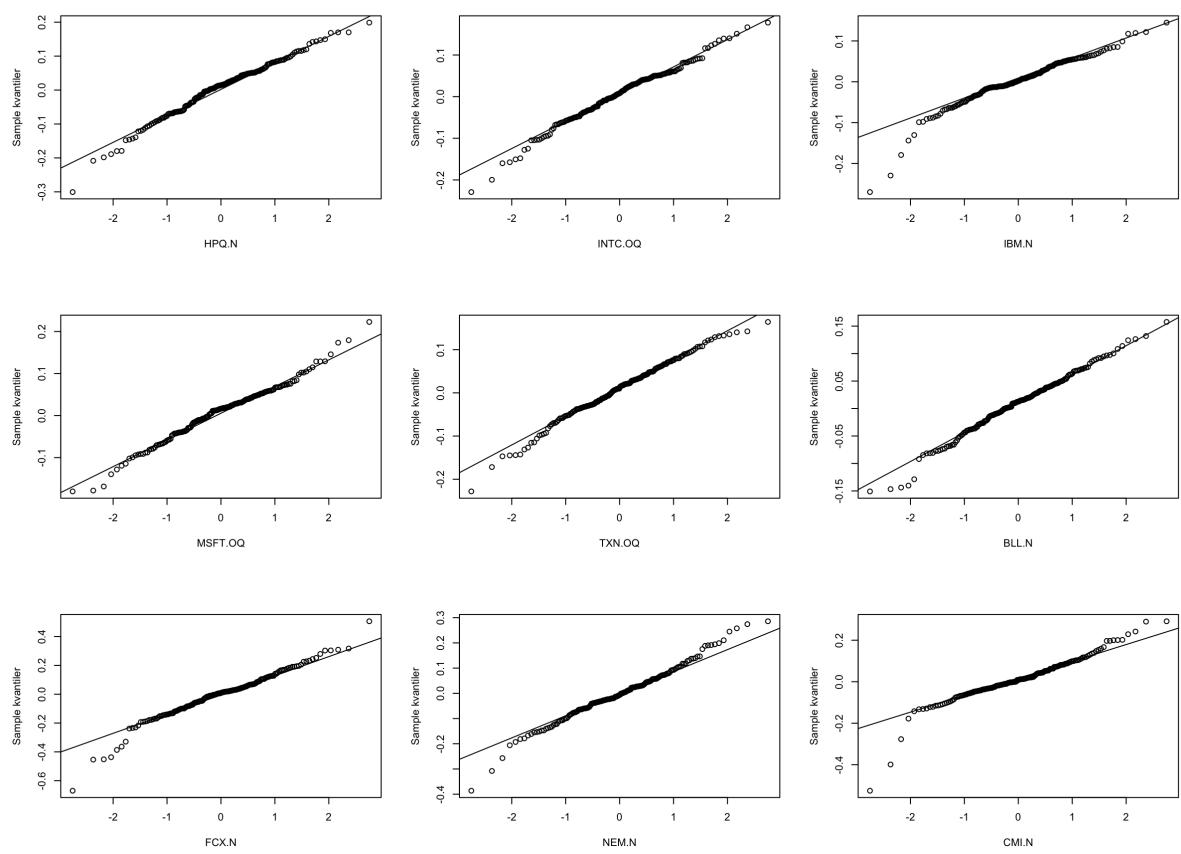


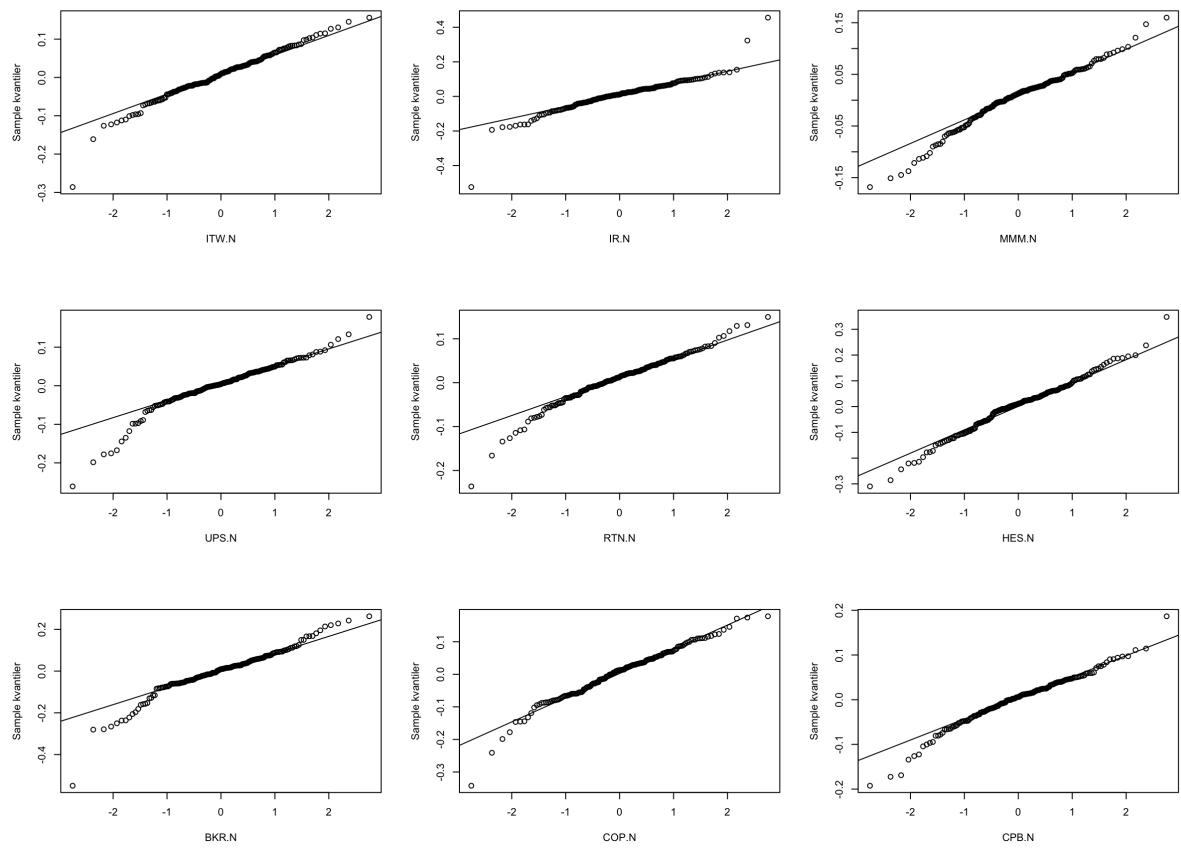


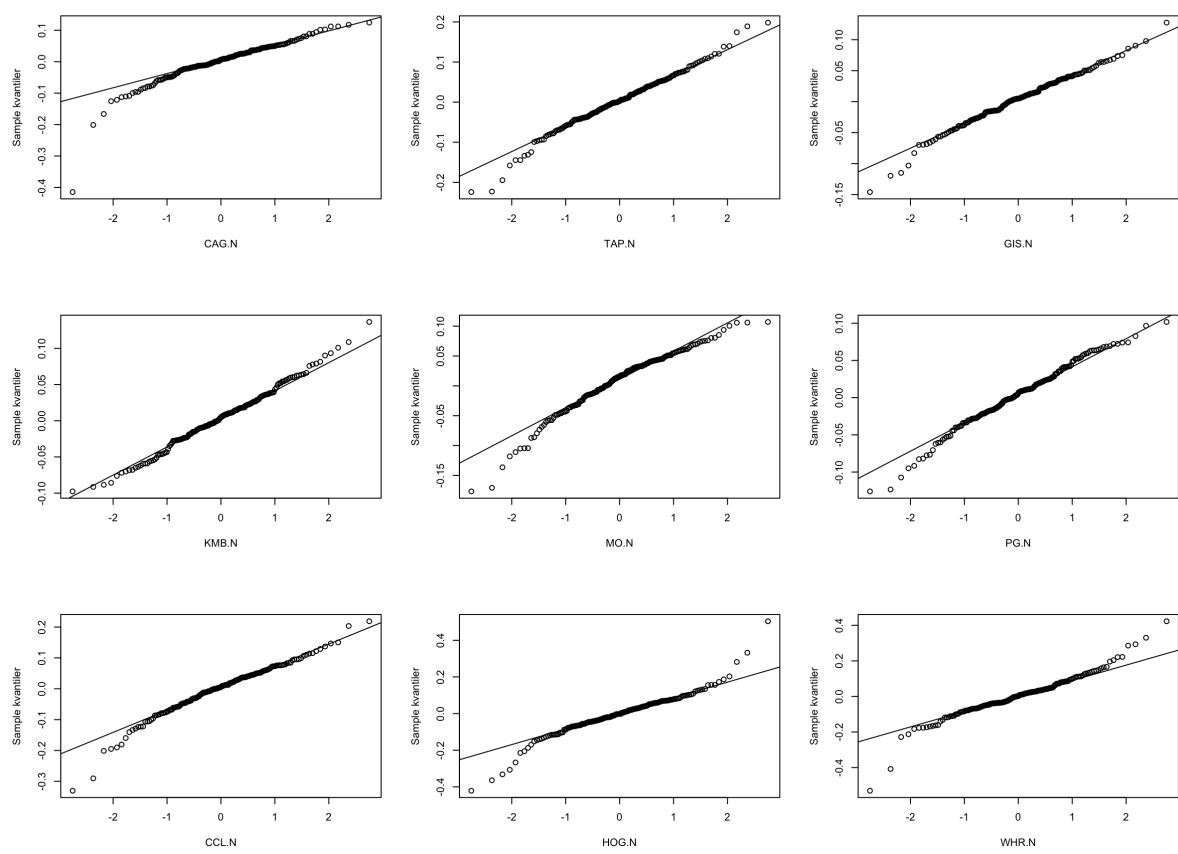
## Sample kvantiler af log-afkast











## Jarque-Bera test

	STT.N	C.N	BAC.N	ABT.N	AMGN.OQ	BAX.N	BDX.N
Test statistik	407.95827	908.99496	563.26314	18.18026	39.6476	50.22648	4.86225
P-værdi	0***	0***	0***	0.00011***	0***	0***	0.08794
	BMY.N	BIIB.OQ	LLY.N	MDT.N	DGX.N	A.N	MRK.N
Test statistik	31.18041	197.73352	149.34756	91.8482	6.93475	5.73977	18.1751
P-værdi	0***	0***	0***	0***	0.0312*	0.05671	0.00011***
	T.N	AMD.OQ	AMAT.OQ	CSCO.OQ	HPQ.N	INTC.OQ	IBM.N
Test statistik	14.20898	0.74651	0.39934	5.39465	9.842	6.5412	120.47843
P-værdi	0.00082***	0.68849	0.819	0.06739	0.00729**	0.03798*	0***
	MSFT.OQ	TXN.OQ	BLL.N	FCX.N	NEM.N	CMI.N	ITW.N
Test statistik	4.05312	4.75403	3.03199	43.50191	7.78664	219.66651	50.56163
P-værdi	0.13179	0.09283	0.21959	0***	0.02038*	0***	0***
	IR.N	MMM.N	UPS.N	RTN.N	HES.N	BKR.N	COP.N
Test statistik	574.99295	10.52256	100.39971	60.2066	4.23378	134.19975	42.70065
P-værdi	0***	0.00519**	0***	0***	0.12041	0***	0***
	CPB.N	CAG.N	TAP.N	GIS.N	KMB.N	MO.N	PG.N
Test statistik	25.2237	844.51632	9.09596	10.09768	0.43995	26.78634	4.4628
P-værdi	0***	0***	0.01059*	0.00642**	0.80254	0***	0.10738
	CCL.N	HOG.N	WHR.N				
Test statistik	41.13082	117.31543	109.46074				
P-værdi	0***	0***	0***				

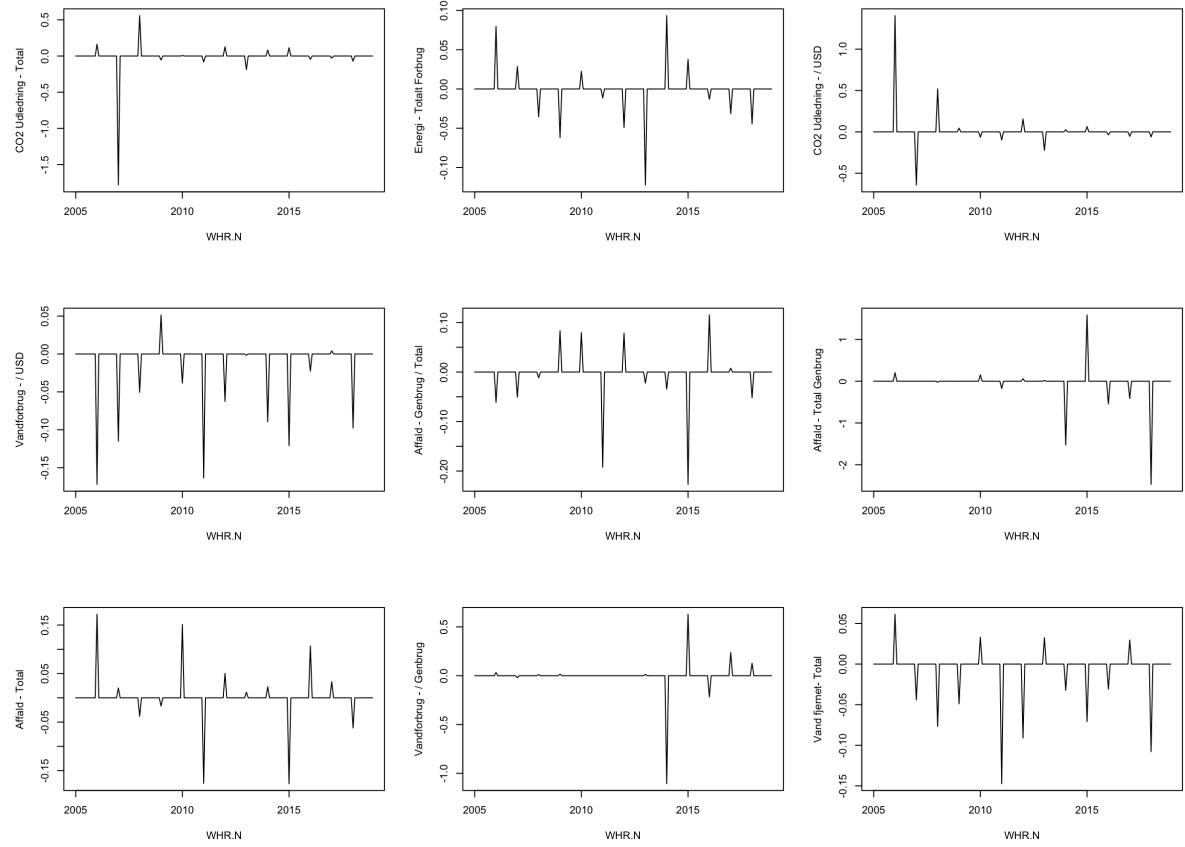
Table 37: Jarque Bera Test

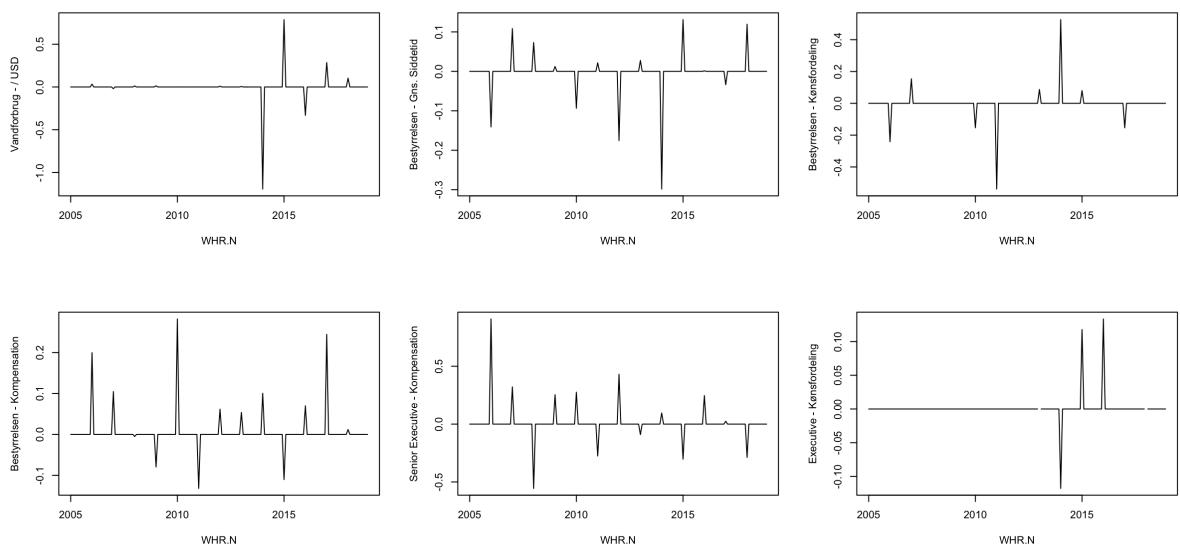
## ADF test

	STT.N	C.N	BAC.N	ABT.N	AMGN.OQ	BAX.N	BDX.N
Test statistik	-5.64322	-5.68215	-6.67227	-5.75303	-6.38138	-5.68072	-5.37103
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	BMY.N	BIIB.OQ	LLY.N	MDT.N	DGX.N	A.N	MRK.N
Test statistik	-5.49277	-6.10625	-4.95759	-6.44815	-5.9492	-5.99297	-4.88313
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	T.N	AMD.OQ	AMAT.OQ	CSCO.OQ	HPQ.N	INTC.OQ	IBM.N
Test statistik	-4.88502	-5.34345	-4.38215	-5.90415	-4.49458	-6.04304	-6.49697
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	MSFT.OQ	TXN.OQ	BLL.N	FCX.N	NEM.N	CMI.N	ITW.N
Test statistik	-5.77112	-5.77383	-6.56982	-5.47973	-5.27875	-6.89463	-5.50999
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	IR.N	MMM.N	UPS.N	RTN.N	HES.N	BKR.N	COP.N
Test statistik	-5.43017	-5.41112	-6.01074	-4.32553	-5.10134	-5.33302	-4.52331
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	CPB.N	CAG.N	TAP.N	GIS.N	KMB.N	MO.N	PG.N
Test statistik	-5.08179	-3.40404	-5.26424	-5.70012	-6.0558	-5.5405	-5.90608
P-værdi	0.01**	0.05617**	0.01**	0.01**	0.01**	0.01**	0.01**
	CCL.N	HOG.N	WHR.N				
Test statistik	-5.87414	-5.83024	-6.06144				
P-værdi	0.01**	0.01**	0.01**				

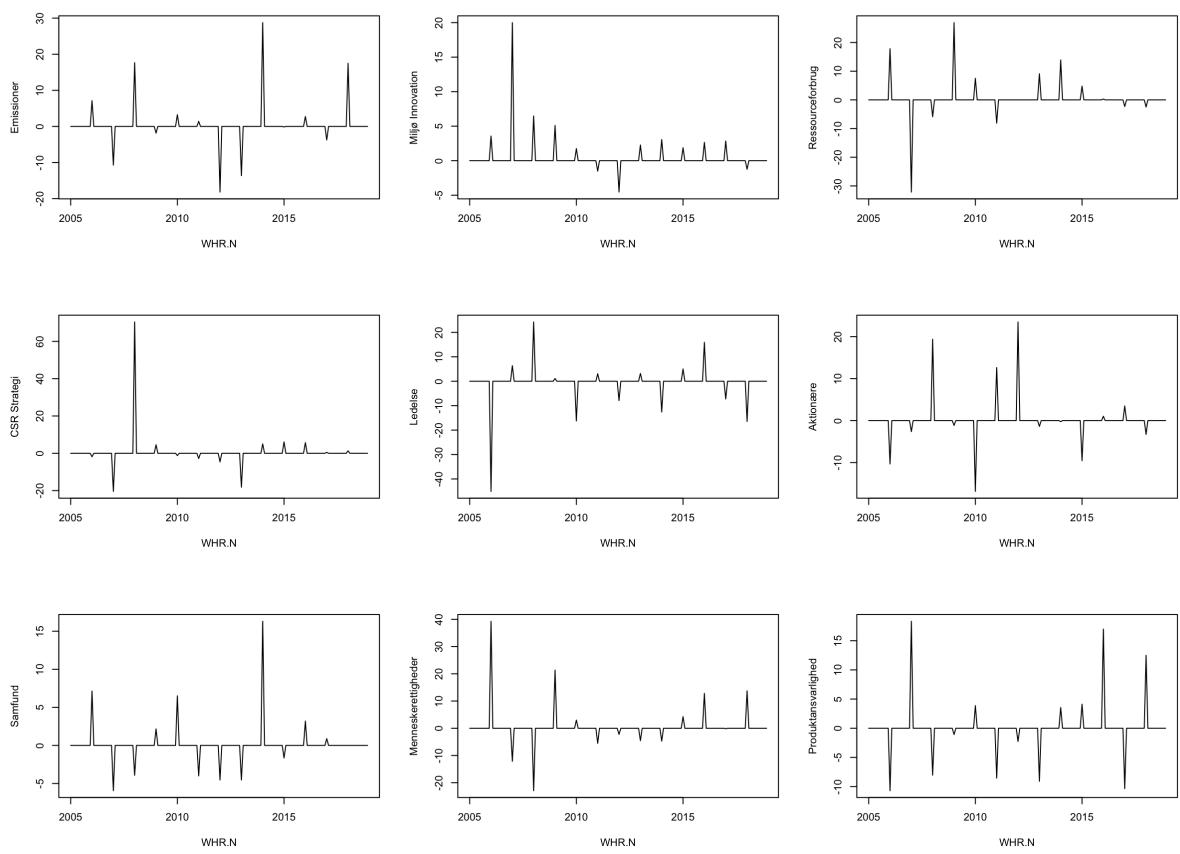
Table 38: Augmented Dickey-Fuller Test

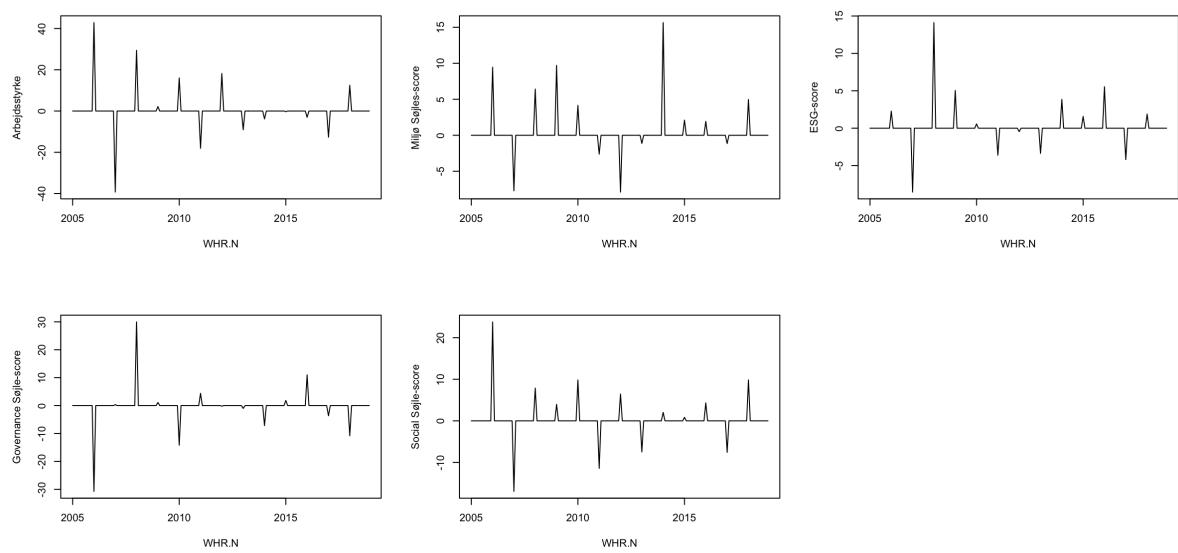
## Transformerede score variable





## Transformerede miljø- og rådata variable





**ADF Test, WHR - score data**

Table 39: WHR - score data

	Emissioner	Miljø Innovation	Ressourceforbrug	CSR Strategi	Ledelse	Aktionære	Samfund
Test statistik	-5.13073	-5.82921	-5.10404	-5.12908	-5.14972	-5.08814	-5.10623
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	Menneskerettigheder	Produktansvarlighed	Arbejdstil	Miljø Søjlescore	ESG-score	Governance Søjle-score	Social Søjle-score
Test statistik	-5.13061	-5.14185	-5.15704	-5.2763	-5.13555	-5.1473	-5.1473
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**

Table 40: WHR - Miljø- og rådata

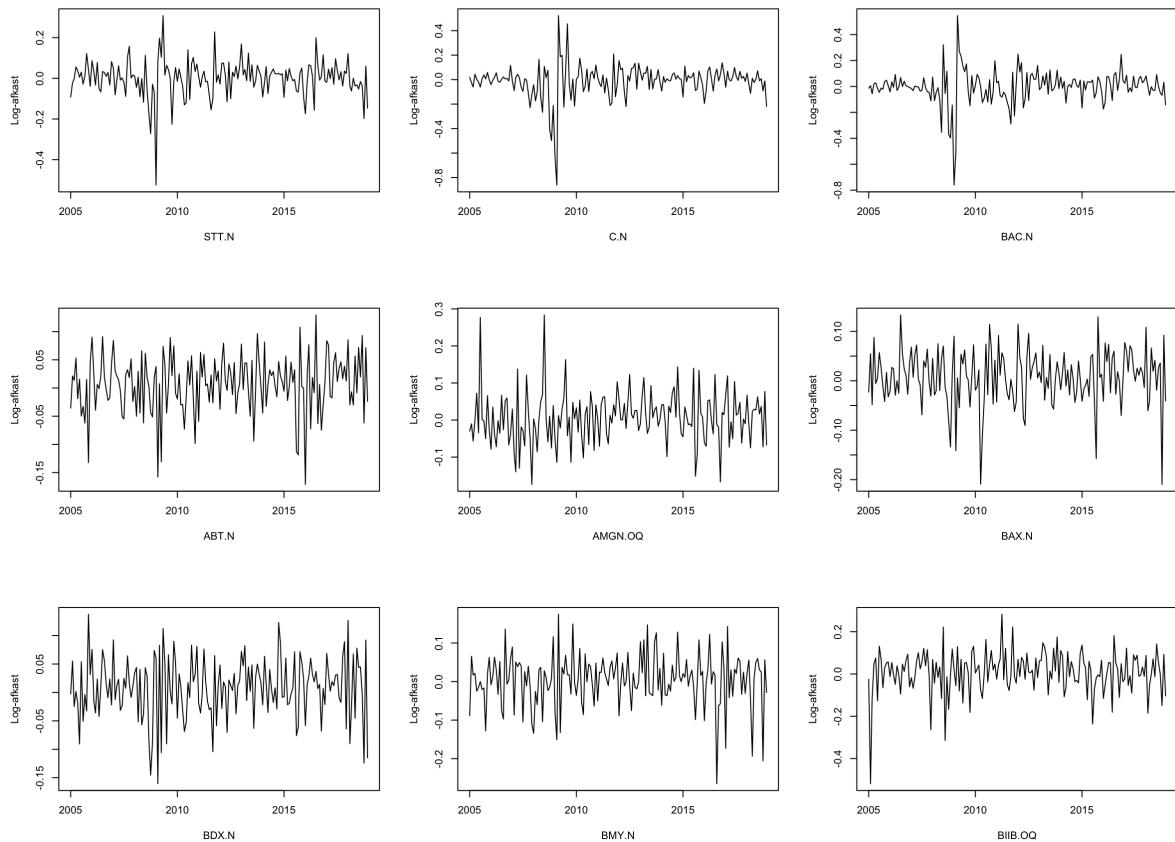
	CO2 Uddeling - Total	Energi - Total Forbrug	CO2 Uddeling - / USD	Vandforbrug - / USD	Affald - Genbrug / Total	Affald - Total Genbrug	Affald - Total
Test statistik	-5.17955	-5.15787	-5.29732	-6.10591	-5.13564	-5.39145	-5.17025
P-værdi	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	Vandforbrug - / Genbrug	Vand fyrenet - Total	Vandforbrug - / USD	Bestyrelsen - Gns. Sidstefid	Bestyrelsen - Kønsfordeling	Bestyrelsen - Kompenstation	Executive - Kompenstation
Test statistik	-5.07264	-5.55024	-5.07147	-5.11183	-5.12437	-5.38497	-5.2742
P-værdi	0.01**	0.01**	0.01**	0.01*	0.01**	0.01**	0.01**

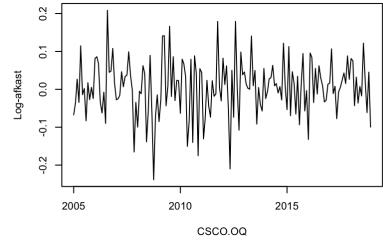
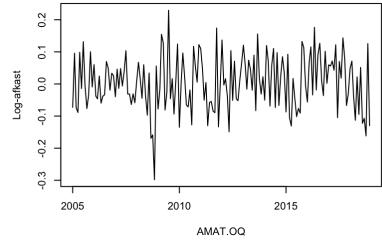
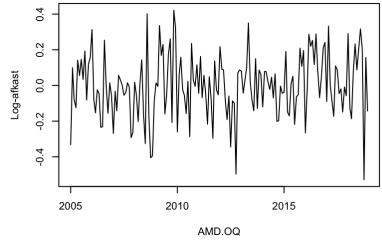
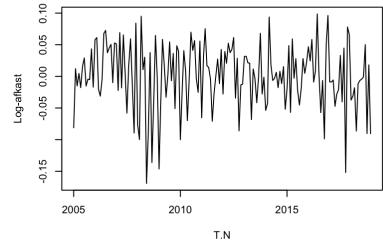
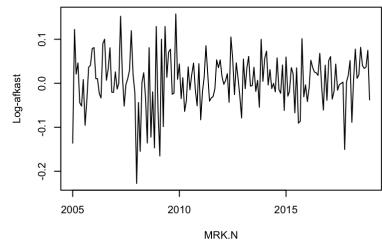
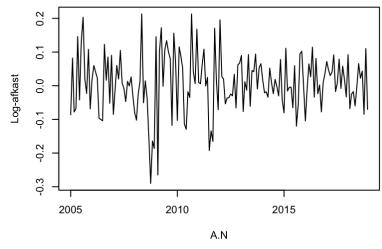
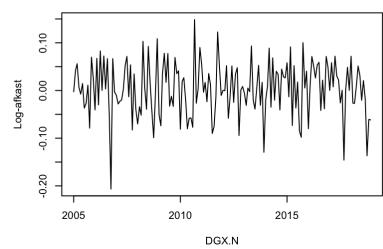
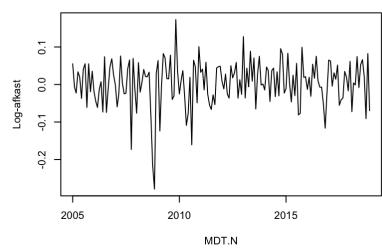
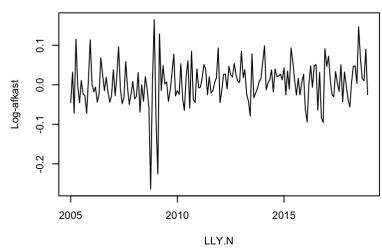
## Ljung-Box test

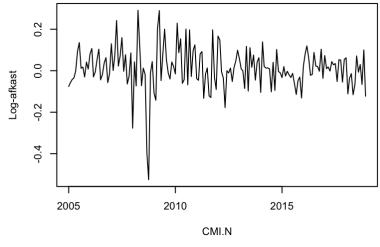
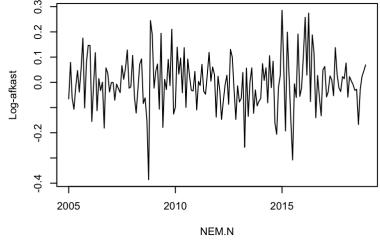
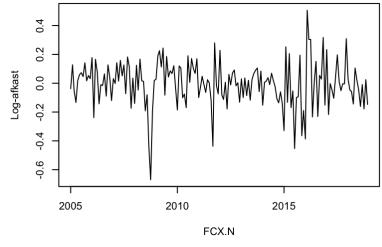
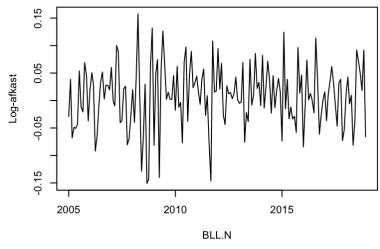
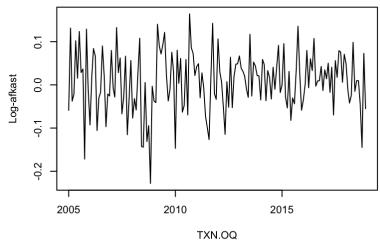
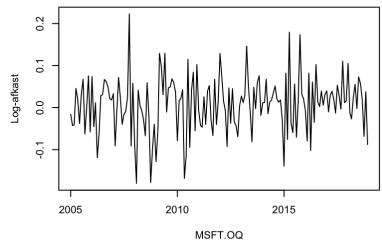
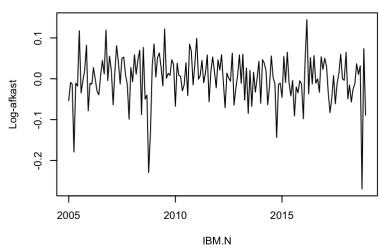
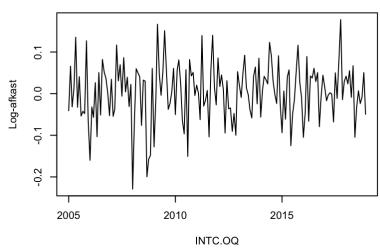
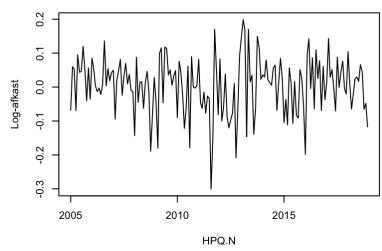
	STT.N	C.N	BAC.N	ABT.N	AMGN.OQ	BAX.N	BDX.N
Test statistik	8.1439	20.66635	30.50988	11.49741	3.55422	2.69623	13.8729
P-værdi	0.41954	0.00809**	0.00017***	0.17508	0.89494	0.95195	0.08514
	BMY.N	BIIB.OQ	LLY.N	MDT.N	DGX.N	A.N	MRK.N
Test statistik	5.69857	9.13943	14.6975	9.17681	8.66526	10.41236	7.96276
P-værdi	0.68095	0.33067	0.0653	0.3276	0.37132	0.23727	0.43712
	T.N	AMD.OQ	AMAT.OQ	CSCO.OQ	HPQ.N	INTC.OQ	IBM.N
Test statistik	11.68464	10.76018	7.84839	7.02917	6.96896	5.14903	6.51173
P-værdi	0.16584	0.21566	0.44842	0.53349	0.53999	0.74154	0.59011
	MSFT.OQ	TXN.OQ	BLL.N	FCX.N	NEM.N	CMI.N	ITW.N
Test statistik	6.70084	13.08074	18.31652	11.78549	7.53089	24.67635	11.5066
P-værdi	0.56923	0.1091	0.01897*	0.16103	0.48058	0.00176**	0.17461
	IR.N	MMM.N	UPS.N	RTN.N	HES.N	BKR.N	COP.N
Test statistik	21.83125	6.28158	18.62729	4.57773	7.57869	9.02292	6.55697
P-værdi	0.00524**	0.61572	0.01699*	0.80161	0.47567	0.34037	0.5851
	CPB.N	CAG.N	TAP.N	GIS.N	KMB.N	MO.N	PG.N
Test statistik	7.5257	7.23104	6.12962	6.96879	9.00079	8.71225	7.42046
P-værdi	0.48111	0.51192	0.63272	0.54	0.34223	0.36715	0.49202
	CCL.N	HOG.N	WHR.N				
Test statistik	11.9074	21.73455	24.15234				
P-værdi	0.15538	0.00543**	0.00216**				

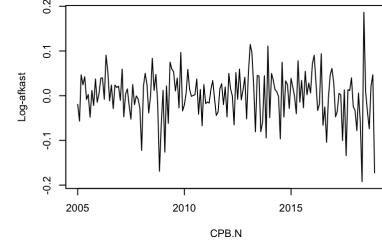
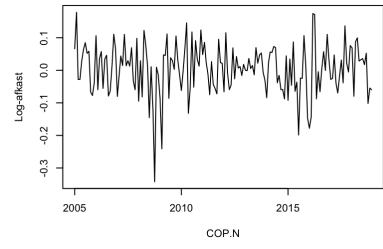
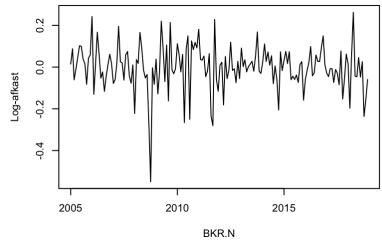
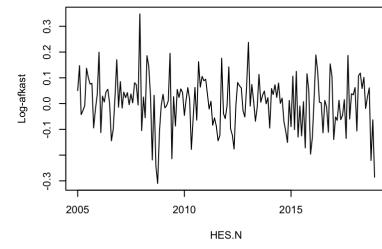
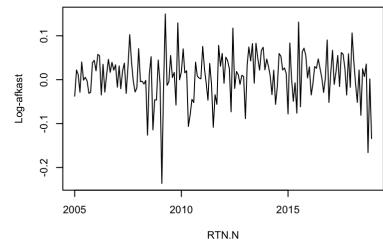
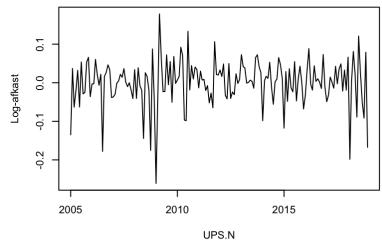
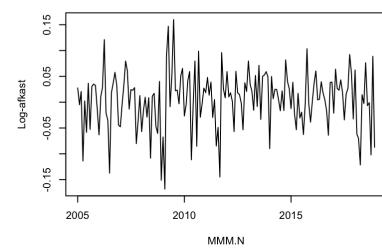
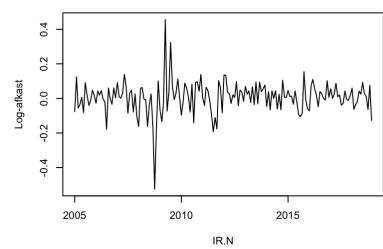
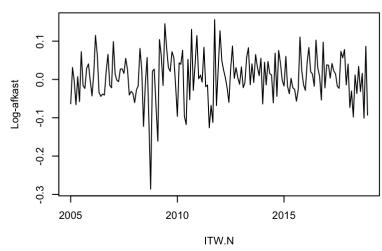
Table 41: Ljung-Box Test af log-afkast

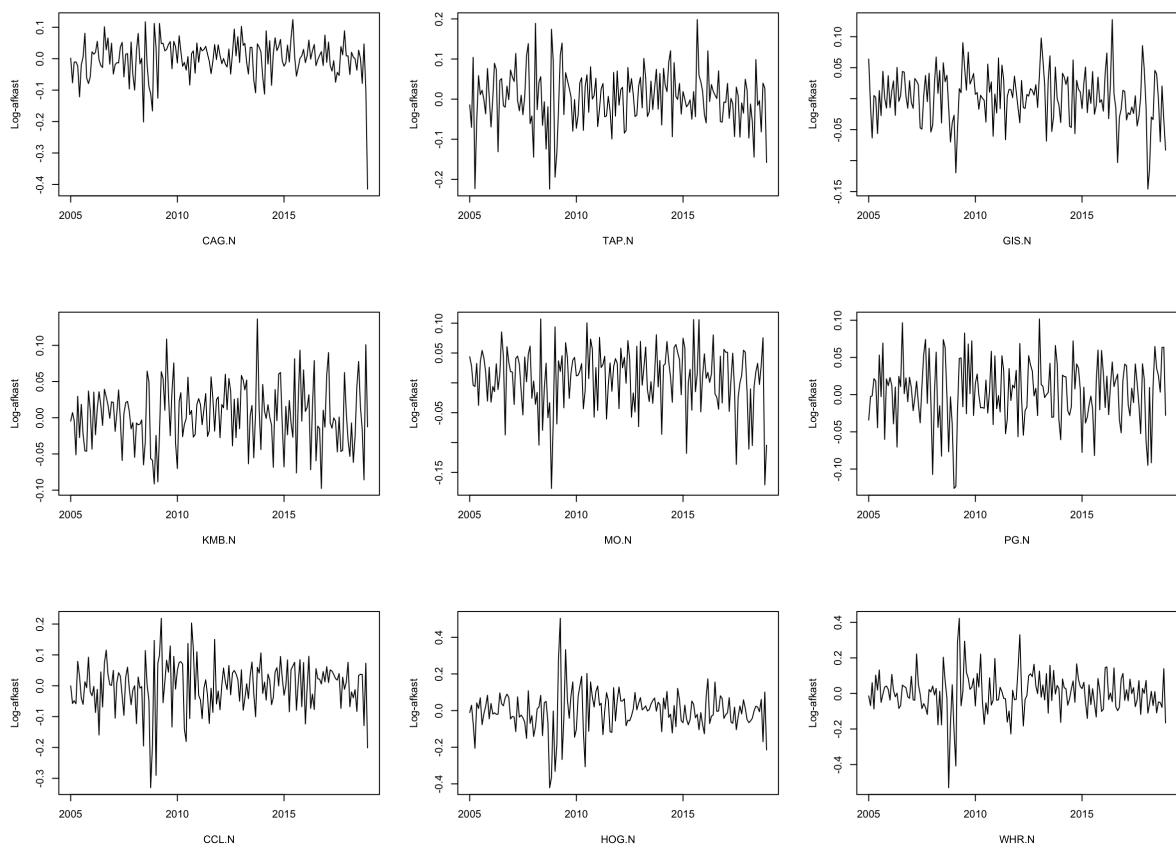
## Log-afkast plots











### AICc for WHR

Lags	q = 0	q = 1	q = 2	q = 3
p = 0	-1.0931156	-0.2018113	1.155405	3.009714
p = 1	-0.9030935	0.1010880	1.579726	3.557690
p = 2	-0.6837709	0.3761763	2.003282	4.104698
p = 3	-0.4482640	0.7735597	2.481576	4.736010

Table 42: AICc for ARMA(p,q)/GARCH(1,1) modeller

Samtlige modeller med en signifikant eksterne kovariate fra ESG-niveau to

Table 43: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med søjle-score

Model	$\mu$	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	Søjle-score	$\omega$	$\alpha_1$	$\beta_1$
Governance - AMD.OQ	-0.10434	-0.31031	-1.01791	0.4484	1.11782	0.11846	0.00133***	0.00951	0.14787	0.48312	0.01408
Governance - BDX.N	-0.03416	0.38774	0.40134	-0.64412	-0.4239	0.06855	0.00046***	0	0.05	0.9	
Governance - BMY.N	0.00857	1.19483	-0.50136	-1.58321	0.90796	-0.27795	-6e-05*	1e-05	0	0.999	0.11391
Governance - CAG.N	-0.02343	-0.65122	0.0619	0.83262	-0.05489	-0.29348	0.00038***	0.00058	0.43312	0.56328	0.18687
Governance - GIS.N	-0.01697	0.36792	0.9495	-0.24872	-1.06713	-0.42745	0.00028***	1e-05	0.02	0.97834	0.27888
Governance - HES.N	0.06267	0.6804	0.45285	-0.69555	-0.35627	-0.20459	-0.00117***	3e-05	0.00048	0.99725	
Governance - HOG.N	0.04432	1.02445	-1.18349	-0.98103	0.05422	-0.00095***	0.00048	0.00048	0.34619	0.65381	
Governance - HPQ.N	0.0589	1.16795	-0.98704	-0.93909	0.76676	0.11729	-0.00094***	0.00127	0.10255	0.68963	0.27964
Governance - IBM.N	0.0483	0.38511	0.50954	-0.49309	-0.56499	-0.0956	-0.00063***	0	0.0506	0.90001	
Governance - INTC.OQ	0.02427	-0.50166	-0.97279	0.65024	1.0792	0.16014	-0.00035***	0	0.01258	0.98353	
Governance - ITW.N	0.00823	0.79238	-1.15608	-0.62965	1.08802	0.30634	-3e-05***	0.00085	0.34941	0.42075	
Governance - MDT.N	0.03137	0.90557	-1				-0.00038***	0.00122	0.40289	0.29611	
Governance - MMM.N	-0.0762	-0.65367	-0.81092	0.72274	1.02399	0.0095***	0.00042	0.106	0.74186		
Governance - MO.N	0.09134	0.07486	0.9338	-0.06438	-0.98444	-0.00062***	1e-05	0.01086	0.98787	0.19299	
Governance - MRK.N	0.0127	-0.18204	-0.60454	-0.08492	0.86111	-0.46806	-0.00012***	0.00013	0.07593	0.87995	0.50867
Governance - RTN.N	0.12375	1.11571	-1.10771	-1.01446	1.07316	0.20697	-0.00141***	0.00039	0.06345	0.80015	
Governance - T.N	0.02659	0.71618	0.46085	-0.08225	-0.34275	-0.71791	-0.00043***	0.00038	0.1289	0.69127	0.87997
Miljø - AMAT.OQ	-0.03498	1.55935	-1.04728	-1.68117	1.24675	0.09188	0.00051***	0.00053	0.11006	0.80014	-0.09863
Miljø - AMGN.OQ	0.04692	-0.45404	-0.60239	0.31058	0.52597	0.41676	-0.00043***	0	0	0.99811	-0.65779
Miljø - BDX.N	0.03501	-0.15652	-0.94108	-0.05784	1.06767	-0.09637	-0.00045***	0.00082	0.13944	0.49527	-0.23558
Miljø - CPB.N	0.02333	0.23569	0.73838	-0.38746	-0.68674	-0.00025	-3e-04***	0.00023	0.11118	0.82487	
Miljø - DGX.N	-0.00386	0.86117	-1			0.00015***	0	0	0	0.999	
Miljø - FCX.N	-0.01455	0.2765	0.57136	-0.27618	-0.48694	0.00036***	2e-05	0.05	0.9		-0.31595
Miljø - GIS.N	-0.06129	1.07139	-0.322	-0.91871	0.10672	0.02637	0.00086*	0	0	0.999	
Miljø - HES.N	0.13369	0.64633	0.28863	-0.70558	-0.08097	-0.00159***	1e-05	0.05617	0.93507		-0.27093
Miljø - IBM.N	0.08478	1.03434	0.48781	-1.23656	-0.28919	-0.76461	-0.00095*	0.00022	0	0.9223	0.74151
Miljø - LLY.N	-0.04166	0.67872	0.26226	-0.84645	-0.44405	-9e-05***	1e-05	0.00126	0.32324	0.26251	
Miljø - MO.N	0.04451	-0.16195	-0.66704	-0.05438	0.88412	-0.44405	-0.00051***	0.00011	0.07078	0.89407	0.49686
Miljø - MRK.N	0.04421	-0.14289	-0.66704	-0.05438	-0.08684	-1.00577	-0.0018***	0.0011	0.30871	0.60142	
Miljø - STT.N	0.13214	0.10588	0.87915	-0.08684	-0.10577	0.00094***	0.00041	0.34343	0.58292	0.16635	
Miljø - UPS.N	-0.07862	0.12407	0.55969	-0.49616	-0.71085	-0.85812	-0.00147***	6e-05	0.11869	0.86102	1.03318
Social - A.N	0.13719	-0.31223	-0.18394	0.38944	0.23228	0.56376	-0.00118***	0.00171	0.04171	0.88952	-0.37653
Social - AMD.OQ	0.10377	0.07606	-0.80715	0.14552	0.95731	-0.74796	-0.00017***	0	0	0.999	-1.04257
Social - BAX.N	0.02167	-0.73696	0.56236	0.64824	-0.74796	0.87498	-0.00023***	0.00041	0.09231	0.74272	
Social - BDX.N	0.23019	-0.49585	-0.82978	0.49064	1.03476	-0.31119	-3e-04***	1e-05	0.99875	0.16385	
Social - BMY.N	0.02226	1.17772	-0.46102	-1.53502	0.83331	-0.21642	-0.00044***	3e-04	0.3308	0.66791	-0.26494
Social - CAG.N	0.04034	-1.611175	-0.4624	1.77501	0.58557	-0.00082***	0.00048	0.27645	0.66523		
Social - COP.N	0.07421	0.95479	-1			0.00079***	0.00012	0.05103	0.92339		
Social - CSCO.OQ	-0.05245	-1.37573	-0.47286	1.26605	0.24115	4e-04**	0	0.05	0.9		-0.08273
Social - DGX.N	-0.01622	0.50207	0.19943	-0.7124	-0.25787	0*	0.00248	0.00073	0.75025		
Social - HES.N	0.02502	0.40161	0.50399	-0.38252	-0.53774	0*	0.00088***	0	0.999	0.94724	
Social - IBM.N	0.08478	-1.72843	-1.45598	1.82	1.71696	-0.74122	-0.00058***	0.00128	0.33194	0.29171	0.68829
Social - MDT.N	0.11054	-0.73824	-0.4321	0.72638	0.29801	-0.57592	-0.00174***	1e-04	0.09301	0.87448	-1.03733
Social - MMM.N	-0.06756	0.23749	-0.23249	-0.29798	0.2771	0.68717	0.00085***	0.00014	0.03991	0.95867	0.13466
Social - MO.N	0.08255	0.04308	0.9398	0.00532	-0.98653	-0.43099	-0.00063***	0.00014	0.07786	0.8696	0.46162
Social - MRK.N	0.06493	-0.09068	-0.65571	-0.13152	0.88102	-0.43099	-0.00036***	0	0.07582	0.91551	
Social - PG.N	0.03212	0.85069	-0.86357	-0.15585	-0.2052	-0.79832	-0.00044***	0.00174	0.52663	0.14211	0.18574
Social - TAP.N	0.03213	1.11846	-0.34661	-1.6094	0.79832	-0.2052	-0.00044***	0.00174			

## Samtlige modeller med en signifikant eksterne kovariate fra ESG-niveau fire

	STT.N	C.N	BAC.N	ABT.N	AMGN.OQ	BAX.N	BDX.N
Test statistik	26.69503	14.94094	17.30819	12.64838	15.70676	14.78674	14.84812
P-værdi	0.14406	0.77978	0.63288	0.89196	0.73464	0.78848	0.78503
	BMY.N	BIIB.OQ	LLY.N	MDT.N	DGX.N	A.N	MRK.N
Test statistik	7.78629	12.95516	13.40993	18.83964	18.24454	29.35948	24.46048
P-værdi	0.99319	0.8793	0.85911	0.53227	0.5713	0.08092	0.22286
	T.N	AMD.OQ	AMAT.OQ	CSCO.OQ	HPQ.N	INTC.OQ	IBM.N
Test statistik	28.54864	15.16962	17.12414	24.41851	16.8377	19.9152	14.5004
P-værdi	0.09703	0.76662	0.6449	0.22459	0.66349	0.46325	0.80425
	MSFT.OQ	TXN.OQ	BLL.N	FCX.N	NEM.N	CMI.N	ITW.N
Test statistik	15.93252	23.84784	14.74569	6.82886	15.00967	18.00698	23.83671
P-værdi	0.7208	0.2491	0.79077	0.99721	0.77585	0.58695	0.2496
	IR.N	MMM.N	UPS.N	RTN.N	HES.N	BKR.N	COP.N
Test statistik	21.14323	24.14312	15.90852	11.25535	12.97757	20.33148	11.65246
P-værdi	0.38875	0.2362	0.72228	0.93933	0.87834	0.43737	0.92752
	CPB.N	CAG.N	TAP.N	GIS.N	KMB.N	MO.N	PG.N
Test statistik	21.497	19.09044	15.21888	13.9119	13.90444	22.52337	18.2682
P-værdi	0.36842	0.51595	0.76375	0.83493	0.83531	0.3128	0.56975
	CCL.N	HOG.N	WHR.N				
Test statistik	20.89708	6.37211	13.27207				
P-værdi	0.40321	0.99829	0.86541				

Table 51: Breusch-Godfrey Test - ESG Model

Table 44: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med under-score

Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
Aktionær - AMAT.OQ	0.03981	1.31579	-0.6872	-0.12917	-1.46456	0.91907	0.127	-0.00062***	0.00092	0.16386
Aktionær - BDX.N	-0.00979	0.29717	0.38163	0.18129	-0.62491	-0.42541	0.00222***	0	0.05348	0.90125
Aktionær - BIIB.OQ	-0.0297				-0.11752	-0.03222	-0.09595	0.0081*	0	0.99778
Aktionær - CAG.N	-0.02368	-0.55838	-0.87949	-0.16206	0.43419	1.06836	-0.02663	0.0035***	0.00668	0.665523
Aktionær - CMN.N	-0.06065	-1.14336	-0.1262		1.18287	0.12825	0.00105***	0.00031	0.17602	0.80278
Aktionær - CFB.N	-0.00572	0.32812	0.6285	-0.49291	-0.603	0.02402	0.0021***	0.00023	0.1068	0.82561
Aktionær - CSCO.OQ	0.0154	-1.58082	-0.64254		1.50323	0.28612	-0.20205	-0.00017***	7e-05	0.03826
Aktionær - DGX.N	-0.00119	0.25877	-0.52023	0.75579	-0.45974	0.43968	-1.04814	1e-04***	0	0.99899
Aktionær - GIS.N	-0.00184	0.24766	0.94008	-0.2344	-0.00552	-1.10928	0.0011***	1e-05	3e-05	0.99206
Aktionær - HFS.N	-0.03293	0.27447	0.71875	-0.17149	-0.35587	-0.70858	0.012***	1e-05	0.05016	0.90011
Aktionær - IBM.N	-0.01275	0.09229	0.86436	0.05611	-0.25283	-0.81331	0.0026***	1e-05	0.00791	0.99066
Aktionær - MDT.N	-0.01099	-0.41867	-0.15486		0.2213		0.0035*	0.00144	0.46435	0.20462
Aktionær - MMM.N	-0.01834	0.38111	-0.1988	0.8154	-0.31099	0.2748	-1.05204	3e-04***	0.00028	0.11073
Aktionær - MO.N	0.04216	0.04737	0.96601		-0.01876	-1.03717	0.11554	-0.00016***	0	0.10641
Aktionær - MRK.N	0.00743	-0.16623	-0.65043	-0.46672	-0.06845	0.80523	0.51137	-5e-05***	0.00016	0.06632
Aktionær - NEM.N	-0.0291	0.71802	-0.6856	0.95008	-0.81968	0.77503	-1.04088	4e-04***	0.00131	0.16985
Aktionær - PG.N	0.00246	0.93509	-0.07122		-1		1e-05*	0.00062	0.12419	0.50967
Aktionær - STT.N	0.00482	0.12254	0.88787		-0.09517	-1.01282	-0.00016***	0.00108	0.26175	0.63828
Arbejdskraft - A.N	0.07454	-0.34098	-0.23542	-0.88281	0.34763	0.25126	1.04394	-0.00076***	1e-05	0.13031
Arbejdskraft - AMAT.OQ	0.13763	1.36537	-0.74872	-0.10958	-1.4874	0.91597	0.12724	-0.00155***	0.00048	0.13161
Arbejdskraft - AMD.OQ	0.10485	0.00272	-0.84555	0.4996	0.20987	0.97327	-0.33215	-0.00125***	0.01478	0.13136
Arbejdskraft - BDX.N	-0.01288	0.38742	0.44363	0.09521	-0.61403	-0.4559	0.00017***	0	0.05	0.9
Arbejdskraft - BMY.N	0.01219	-0.28392	0.60255		0.40545	-0.75345	-0.06737	-0.00014*	1e-05	0.999
Arbejdskraft - C.N	0.04182						-0.00074*	0.000087	0.34318	0.62084
Arbejdskraft - CAG.N	0.02717	-0.05632	1.03788		0.15999	-1.08655	-0.16108	-0.00045***	5e-04	0.42021
Arbejdskraft - CCL.N	-0.00379	0.32059	0.68755		-0.49035	-0.58578	-0.00024***	3e-04	0.22028	0.76924
Arbejdskraft - CSCO.OQ	-0.021	-1.37641	-0.4792		1.26023	0.23461	0.00035***	1e-04	0.0479	0.92977
Arbejdskraft - DGX.N	0.00134	0.60215	-0.72566	0.74297	-0.75588	0.76136	-1.07723	4e-05***	1e-04	0.99664
Arbejdskraft - FCX.N	0.03294	0.4352	0.59123		-0.54197	-0.34795	-0.17414	-0.00013***	0.00692	0.34013
Arbejdskraft - HES.N	-0.01765	0.073	0.65034		-0.10959	-0.7044	-0.27086	0.00063***	1e-05	0.05
Arbejdskraft - IBM.N	0.03836	-1.38154	-0.72615	0.14506	1.62694	1.08535	0.00164	-0.00037***	1e-05	0.99803
Arbejdskraft - ITW.N	0.004	0.81662	-1.17133	0.34439	-0.62621	1.09634	4e-05***	0.00098	0.33328	0.34686
Arbejdskraft - MMM.N	-0.04072	0.21999	-0.21061	0.65043	-0.31396	0.26503	-1.05078	0.00058***	0.00016	0.08557
Arbejdskraft - MO.N	0.02985	0.59912	0.47699	-0.92307	-0.56658	-0.67692	0.102938	-0.00032**	0.00055	0.1678
Arbejdskraft - MRK.N	0.4746	-0.0509	-0.69427	-0.34042	-0.1146	0.91964	0.41294	-0.00486***	2e-05	0.114
Arbejdskraft - T.N	0.02205	1.18	-0.22316	-0.37606	-1.43822	0.32469	0.47818	-5e-04***	0.00023	0.06078
Arbejdskraft - TAP.N	0.03304	0.75858	0.12433	0.07528	-1		-4e-04***	0.00174	0.24904	0.38156
Arbejdskraft - UPS.N	0.06201	0.76485			-0.9865	-0.04566	-0.00073***	0.00013	0.15192	0.84644
CSR - A.N	-0.13066	0.50924	-0.86217		-0.61414	1.02183	0.00149***	0.00044	0.24768	0.69871
CSR - AMAT.OQ	0.21354	1.26994	-0.99318		-1.47999	1.28845	-0.16621	-0.00218***	5e-04	0.16702
CSR - AMGN.OQ	0.33667	-0.49926	-0.52246	0.44543	0.27534	0.39507	-0.75817	-0.00352***	0.00011	0.05051
CSR - BAC.N	0.37702	1.97031	-0.98694		-1.99297	1.00677	-0.0373***	0.00065	0.41693	0.58206
CSR - BAX.N	0.0145	-0.80046	0.52089	1.04998	0.69422	-0.69451	-1.14757	-8e-05***	3e-05	0.98864
CSR - BDX.N	0.13269	0.05206	-0.94931	0.12869	-0.19805	1.03531	-0.38112	-0.00132***	0.00155	0.19572
CSR - CAG.N	-0.04189	-1.27405	-1.134	-0.75687	1.41471	1.47285	0.98714	0.0057***	0.00048	0.52282
CSR - CCL.N	5e-05	-0.215	0.48845	0.68214	0.13618	-0.44087	-0.78472	-1e-04***	0.00048	0.29991
CSR - CSCO.OQ	0.05096	-1.40484	-0.37446	0.10972	1.57129	0.55461	-0.00053***	0.00014	0.0504	0.92172
CSR - DGX.N	0.00148	0.68294	-0.76533	0.8091	-0.7526	0.76662	-1.09021	4e-05***	0	0.999
CSR - GIS.N	0.1299	-0.03533	0.72949	-0.19085	0.21893	-0.86473	-0.00136***	0	1e-04	0.99873

Table 45: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med under-score

Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
CSR - HES.N	0.09688	0.27447	0.70049	-0.32347	-0.61063	-0.12634	-0.00084***	3e-05	0.01127	0.98698
CSR - HOG.N	0.03572	0.70989	-0.46068	0.74125	-0.61546	-0.25731	-0.68804	-0.00088***	0.00043	0.29909
CSR - HPQ.N	0.0446	0.44577	0.42529	0.03311	-0.45976	-0.60506	-0.00027***	1e-05	0.05	0.69905
CSR - MDT.N	0.1775	0.9509	-0.18169	0.78107	-0.30082	0.27635	-1.03823	-9e-05***	0.00116	0.43617
CSR - MMM.N	0.00935	0.31574	-0.187957	-0.83201	-0.98043	-0.93455	1.07831	0.00117***	0.00035	0.28572
CSR - MO.N	-0.09708	0.82397	0.87957	-0.83201	0.07321	1.03221	-0.27641	-0.00084***	0.00011	0.04182
CSR - MRK.N	0.06023	-0.17446	-0.90926	0.20596	-0.19157	-0.07009	1.06691	0.00071***	0.00083	0.15656
CSR - NEM.N	-0.06154	0.05423	0.07953	-0.83742	-1	-0.0021***	-0.00021***	0.00065	0.15385	0.76031
CSR - PG.N	0.02225	0.84995	-1	-0.0021***	-0.00021***	0.00065	0.13131	0.69215	0.46154	0.46154
CSR - TAP.N	0.00833	-0.41968	0.39372	0.91488	0.29082	-0.52625	-0.8995	-0.00011***	0.00075	0.13131
CSR - UPS.N	-0.01526	0.51545	0.34551	-0.70339	-0.33739	2e-04***	-0.00056	0.22573	0.59347	0.59347
Emission - STT.N	0.05414	0.56634	-0.46877	-0.55401	-0.76681	0.58196	0.4603	-0.00053***	0.00065	0.32902
Emissioner - A.N	-0.00668	-0.91491	-0.99779	0.07963	0.99361	1.08535	-0.00024***	0.00041	0.157	0.78382
Emissioner - AMAT.OQ	-0.00966	1.55518	-1.03355	0.08422	-1.68239	1.27451	-0.12847	0.00015***	0.00064	0.13556
Emissioner - AMD.OQ	0.10497	-0.67396	0.13985	0.74683	-0.24339	0.10499	-0.00124***	0.00472	0.14658	0.68921
Emissioner - AMGN.OQ	-0.05969	0.71838	-	-0.85758	-0.85758	0.00076***	0	0	0	0.99837
Emissioner - BAX.N	-0.00466	-0.78026	0.52488	1.01854	0.67415	-0.68995	-1.11104	0.00018***	1e-05	0.00646
Emissioner - CAG.N	0.00943	-1.61602	-0.49884	0.17992	1.81141	0.66453	-0.21915	-4e-05***	2e-04	0.28935
Emissioner - COP.N	0.13396	-0.27112	0.57897	0.53914	0.33367	-0.6813	-0.74554	-0.0015***	0.00035	0.17058
Emissioner - CPB.N	0.01374	0.30428	0.60785	-0.48587	-0.57191	-	-0.00017***	0.00013	0.08419	0.88286
Emissioner - CSCO.OQ	-0.02267	-1.36166	-0.35215	0.08728	1.56831	0.55195	0.00037***	2e-05	0.0392	0.95368
Emissioner - DGX.N	0.00194	0.35004	-0.51052	0.83181	-0.47595	0.43773	-1.05527	4e-05***	1e-05	0.00121
Emissioner - GIS.N	-0.02146	-0.44525	0.37458	-0.13466	0.77866	-0.28681	0.00027***	0	3e-05	0.9989
Emissioner - HES.N	0.02408	0.5666	0.43165	-0.55778	-0.39538	-0.10534	-4e-05***	1e-05	0.05953	0.93003
Emissioner - HPQ.N	0.04758	0.70958	0.29846	-0.03213	-0.68522	-0.35004	-0.00046***	2e-05	0.03312	0.9642
Emissioner - KMB.N	0.03831	0.16078	0.68035	-1	-0.17103	-0.88191	-0.00046***	1e-05	0.08624	0.91153
Emissioner - MDT.N	0.12635	-0.3392	-0.18312	0.12888	-	-0.00127*	0.00127	0.56079	0.19225	0.19225
Emissioner - MMM.N	0.10547	-0.65264	-0.8166	0.71808	1.02408	-	-0.00139***	0.00041	0.10853	0.74141
Emissioner - RTN.N	0.02277	-0.09257	0.0703	-0.84741	0.05841	-0.01463	1.08809	-0.00013***	0.00018	0.09591
Emissioner - WHR.N	0.07746	0.91149	-0.00532	-	-	-	-0.00118***	0.00098	0.17413	0.75749
Ledelse - A.N	0.00995	-0.33076	-0.2071	-0.87636	0.35212	0.25325	1.04386	9e-05***	0.00041	0.16882
Ledelse - AMAT.OQ	0.00362	1.72792	-1.27055	0.22721	-1.82879	1.52534	-0.29027	-1e-05***	0.00048	0.12511
Ledelse - AMGN.OQ	0.03146	-0.37551	-0.52358	0.5646	0.2111	0.39876	-0.7915	-0.00032***	0	0.99829
Ledelse - BDX.N	-0.01918	0.39465	0.43953	0.07213	-0.61465	-0.44514	0.00026***	0	0.05	0.9
Ledelse - BIB.B.Q	0.01298	0.07206	0.90649	-0.19869	-0.86754	-	-0.00029***	0	0	0.9981
Ledelse - CSCO.OQ	-0.00557	-1.40256	-0.49258	-	1.29725	0.27227	-	0.00022***	1e-04	0.05162
Ledelse - DGX.N	-0.00954	0.74182	0.24514	-	-0.78072	-0.23411	-0.0607	0.00036***	1e-05	0.00028
Ledelse - GIS.N	-0.0051	0.59803	0.34133	-	-0.5137	-0.50768	-0.0219	0.00015***	1e-05	0.03801
Ledelse - HES.N	0.02933	0.44002	0.52193	-0.15575	-0.61156	-0.44795	-	-0.00058***	1e-05	0.09872
Ledelse - HOG.N	0.02169	1.03338	-0.11689	-1	-	-	-0.00056***	0.00046	0.31897	0.68003
Ledelse - IR.N	0.03913	1.06037	-0.96591	-	-0.94194	0.81749	0.24979	-	0.05162	0.9257
Ledelse - MDT.N	0.02379	0.89398	-	-1.06469	0.03744	-	-0.00026***	4e-05	0.10699	0.99751
Ledelse - MMM.N	-0.06003	0.34821	-0.13354	0.7911	-0.30382	0.27167	-1.04693	0.00067***	0.00025	0.10281
Ledelse - MO.N	0.00786	-0.87803	-0.74233	-0.75501	0.91766	0.67598	0.88033	3e-05***	0.00011	0.02043
Ledelse - MRK.N	0.00136	-0.11491	-0.63496	-0.42903	-0.11422	0.87141	0.47037	7e-05***	0.00011	0.06004
Ledelse - RTN.N	0.03446	-0.10879	0.0629	-	-0.85784	0.06498	-0.02149	1.09102	-0.00027***	2e-04
Ledelse - UPS.N	-0.00995	-0.09091	-0.20058	-0.12243	-	-	-0.00029***	0.00071	0.25359	0.57289
Menneskerettigheder - A.N	0.01726	-0.30381	-0.17528	-0.86058	0.348	0.24565	1.04635	-0.00011***	1e-05	0.11442
Menneskerettigheder - ABT.N	0.15968	-0.05938	-0.64703	-0.68402	-0.1827	0.82154	0.45015	-0.00159***	0	0.99899
Menneskerettigheder - AMAT.OQ	-0.02121	1.60182	-1.10024	0.12218	-1.71014	1.30179	-0.13412	0.00036***	0.00085	0.15773

Table 46: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med under-score

Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
Menneskerettigheder - AMD.OQ	0.10107	0.00915	-0.86879	0.47596	0.15375	0.97933	-0.32637	-0.00113***	0.00637	0.12508
Menneskerettigheder - AMGN.OQ	0.05974	-0.43338	-0.55957	0.45595	0.29459	0.49795	-0.68426	-0.00558***	0	0.6562
Menneskerettigheder - BAX.N	0.01484	-0.43412	0.41068	0.80427	0.32646	-0.54741	-0.90857	-9e-05***	1e-05	0.09817
Menneskerettigheder - BDX.N	0.05567	0.34176	0.45388	0.08593	-0.56879	-0.48243	-0.00663***	0	0.00077	0.99772
Menneskerettigheder - BIB.OQ	0.07321	0.70323	0.49998	-0.94594	-0.90876	-0.41583	0.92879	-7e-04***	0	0.9
Menneskerettigheder - CAG.N	0.04176	0.51367	-0.43884	0.75108	-0.57503	0.59211	-1.10246	-0.0064***	0.00066	0.67889
Menneskerettigheder - CPB.N	0.02193	0.50032	0.45718	-0.75369	-0.30782	-0.0028***	0.00018	0.10289	0.84012	0.31943
Menneskerettigheder - CSCO.OQ	0.02187	-1.52815	-0.59615	1.43052	0.28377	-0.12659	-0.00023***	7e-05	0.04959	0.93613
Menneskerettigheder - DGX.N	0.01184	0.33237	0.52073	-0.48628	-0.55883	-0.00022***	1e-05	0.00156	0.99633	0.99774
Menneskerettigheder - HES.N	0.072	1.05712	-0.18305	-1.11858	0.23514	-0.14551	-0.00084***	2e-05	0.00014	0.99788
Menneskerettigheder - HQ.Q.N	0.07708	1.01527	-1.03704	0.1093	-0.98651	1.00798	-0.00099*	4e-04	0.07405	0.86183
Menneskerettigheder - IBM.N	0.0832	-1.26236	-0.58702	0.23442	1.55184	1.00155	-0.06253	-0.00099***	1e-05	0.00063
Menneskerettigheder - INTC.OQ	0.03524	-0.28429	0.11785	0.86299	0.21334	-0.27517	-1.02402	-5e-04***	0	0.00418
Menneskerettigheder - IR.N	0.02881	1.51275	-1.22313	0.17355	-1.47347	1.10776	-0.00019***	0.00046	0.24263	0.99253
Menneskerettigheder - KMB.N	-0.22588	0.01907	-0.18143	0.0241***	0.00241***	1e-04	0.13221	0.8223	0.9	0.72721
Menneskerettigheder - MO.N	-0.05646	-0.26356	0.573	0.26475	-0.72288	0.12694	0.00105***	0	0.05002	0.9
Menneskerettigheder - MRK.N	-0.00421	-0.15462	-0.70417	-0.43266	-0.08652	0.8947	0.47727	0.00016***	0.00012	0.8769
Menneskerettigheder - PG.N	0.01908	0.86008	-0.91719	-0.10565	-0.0021***	0	0.00021***	0	0.07132	0.91894
Menneskerettigheder - STT.N	0.01403	0.11552	0.88025	-0.0974	-0.99212	-0.0023***	0.00109	0.25837	0.62414	0.62414
Menneskerettigheder - UPS.N	-0.01539	0.80788	-0.09037	-1	-0.36209	0.24843	1.04207	4e-05***	0.00042	0.31939
Miljø Innovation - A.N	0.00421	-0.35213	-0.24718	-0.86033	-1.86783	1.55675	-0.30448	0.00017***	5e-04	0.15449
Miljø Innovation - AMAT.OQ	-0.00535	1.7909	-1.39621	0.2928	0.94463	-0.38067	0.00115***	0.00427	0.11793	0.7983
Miljø Innovation - AMD.OQ	-0.09335	0.07044	-0.83234	0.54357	0.11729	0.46732	-0.73191	-0.00046***	0	0.8011
Miljø Innovation - AMGN.OQ	0.04759	-0.37706	-0.52722	0.46897	0.23567	-0.86256	-1.03925	-0.00043***	0	0.99806
Miljø Innovation - BAX.N	0.04536	-0.89744	0.71145	0.9666	0.79093	-0.38212	1.13966	-0.21042	1e-05	0.99871
Miljø Innovation - BMY.N	0.02878	0.11312	-0.94386	-0.04018	-0.38212	-0.21042	-0.00064***	0.00035	0.07462	0.83629
Miljø Innovation - CAG.N	0.01638	0.52713	0.45554	-0.42289	-0.56111	-0.07525	-0.00033***	8e-04	0.47066	0.47956
Miljø Innovation - CPB.N	0.02347	-1.05579	-0.92123	0.91909	0.85232	-0.14988	-4e-04*	0.00033	0.16861	0.73312
Miljø Innovation - DGX.N	0.00943	0.89253	-1	-0.00671	-1.05711	-0.00019*	0	0	0.01121	0.98686
Miljø Innovation - GIS.N	-0.00587	0.26194	0.93551	-0.2509	-0.73278	-0.19352	-0.00021***	1e-05	0.07238	0.91287
Miljø Innovation - HES.N	0.04403	0.7704	0.21665	-0.15838	-0.6293	-0.03905	-0.00194*	0.00069	0.28779	0.67431
Miljø Innovation - HOG.N	0.04443	0.73516	-1.14163	-0.1021	1.82035	1.39097	0.20744	-0.00031***	7e-05	0
Miljø Innovation - IBM.N	0.02586	-1.65055	-1.00949	0.06582	0.82878	1.17667	0.13969	-0.00018***	0	0.97319
Miljø Innovation - INTC.OQ	0.01945	-0.63349	-0.4514	0.78618	0.5759	-0.65309	-1.09103	-4e-05***	0.00013	0.97982
Miljø Innovation - LLY.N	-0.00885	-1.15492	-0.93078	1.24543	1.0248	0.0031*	0.00097	0.36624	0.29048	0.29048
Miljø Innovation - MO.N	0.03244	0.08965	0.91946	-0.54244	-0.02991	0.85351	0.58635	3e-05***	0	6e-05
Miljø Innovation - MRK.N	9e-04	-0.27332	-0.68283	-0.89635	-0.83839	0.76257	-1.01848	0.00017***	5e-04	0.06498
Miljø Innovation - NEM.N	0.00455	0.67322	-0.63849	-0.11728	0.4889	1.55097	0.12422	-0.04***	0.00058	0.27786
Miljø Innovation - PG.N	0.00423	-0.55124	0.78618	0.5759	-0.65309	-1.09103	-0.00081***	0.00011	0.05966	0.85827
Miljø Innovation - RTNN.N	0.0207	-0.07401	0.08236	-0.87983	0.06067	-0.02022	1.09459	-0.00013***	0.00018	0.09505
Produktantsvarlighed - UPS.N	-0.02638	0.81256	-0.00893	-0.00689	-1	0.003905	0.00039***	5e-04	0.34946	0.54945
Produktantsvarlighed - AMAT.OQ	0.24241	1.30687	-0.65398	-0.16479	-1.43401	0.82544	0.18369	0.00012***	0.00012	0.89616
Produktantsvarlighed - CAG.N	-0.00701	-1.29143	0.11728	0.93304	0.77762	-0.87272	-1.06004	-0.00271***	0.00058	0.54536
Produktantsvarlighed - COP.N	0.04511	-0.77333	0.8026	-0.16479	1.29403	0.26848	-0.5134	1e-04***	0.00058	0.31457
Produktantsvarlighed - CSCO.OQ	0.02554	-1.41117	-0.5044	0.4889	0.77762	-0.87272	-1.06004	-0.00024***	9e-05	0.11567
Produktantsvarlighed - GIS.N	-0.00252	0.17008	0.96002	-0.18153	0.01039	-1.07036	9e-05***	0	0.04653	0.93163
Produktantsvarlighed - HES.N	1e-05	0.54424	0.44732	-0.53679	-0.37423	-0.1491	0.00036***	2e-05	0.0028	0.99477
Produktantsvarlighed - IBM.N	0.08475	-0.04389	1.00036	-0.01949	-1.07182	-0.00092***	2e-04	0	0.93522	0.93522
Produktantsvarlighed - KMB.N	0.03861	0.25965	0.55899	0.15381	-0.29371	-0.77497	-6e-04***	0	0.05004	0.9
Produktantsvarlighed - MMM.N	-0.05792	0.25399	-0.22789	0.72846	-0.30156	0.27375	-1.0446	7e-04***	1e-04	0.09504
										0.87653

Table 47: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med under-score

Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
Produkttransvarlighted - MO.N	-0.00651	0.09393	0.9175	-0.02195	-0.102993	0.11888	0.00054***	1e-05	0.00095	0.99763
Produkttransvarlighted - MRK.N	0.02394	-0.13861	-0.65262	-0.47964	-0.08422	0.86328	0.51182	-0.00021***	0.00013	0.86862
Produkttransvarlighted - NEM.N	0.00204	0.68997	-0.67649	0.91035	-0.82724	0.76821	-1.05387	-0.00012***	0.00342	0.31543
Produkttransvarlighted - PG.N	0.00194	0.85553			-0.88711	-0.14071	1e-05***	0	0.07144	0.91923
Produkttransvarlighted - UPS.N	-0.06735	0.42459	-0.3492	-0.22812	-0.8544	0.34695	0.24297	1.02743	-0.00066***	0.00048
Ressource - A.N	0.06522	-0.07853	1.01104	-0.00538	-0.53054	-0.5392	-0.16048	0.00047***	0.00064	0.49053
Ressource - CAG.N	0.02756	0.35047	0.6759	-0.87992	-0.46336	0.44833	-1.10048	5e-05***	0.00049	0.12172
Ressource - CPB.N	0.00176	0.28267	-0.51166	-0.11487	0.81228	-0.25704	0.00041***	0	0	0.999
Ressource - DGX.N	-0.03878	-0.51025	0.34302	-0.10399	-0.54469	-0.55061	-0.00245***	2e-05	0.00021	0.99879
Ressource - GIS.N	0.22895	0.51118	0.64386	0.19098	0.89276	0.107988	-0.00245***	0	1e-05	0.99671
Ressource - HES.N	0.19473	-0.71258	-0.88559	-0.17438	-0.59743	-0.46235	-0.06377	-0.00251***	9e-05	0.99868
Ressource - INTC.OQ	0.23784	0.90598	-0.81219	-0.29673	-0.12854	-0.89886	0.02393	1.02393	0.00109***	0.00058
Ressource - MMM.N	-0.09573	-0.64785	0.76899	-0.17191	0.15748	0.00735	0.17091	-0.86595	0.07941	0.16579
Ressource - MO.N	0.03057	-0.17438	-0.2869	-2.14774	-0.8234	2.41883	0.86019	0.51375	-0.00025***	1e-05
Ressource - MRK.N	0.04817	-0.33088	-0.20797	-0.99395	-0.86998	0.33799	0.2499	1.04491	-0.00043***	0.00012
Ressource - MSFT.OQ	-0.13821	0.28801	-0.74762	-0.99347	-0.99395	-1.49557	0.130421	-0.17565	0.00013***	0
Ressource - RTN.N	0.00548	-0.29673	-0.15748	-0.20432	0.33923	1.68141	0.39	-0.36787	-0.00039	0.09871
Ressource - T.N	0.03954	-0.47772	-0.80946	0.54962	0.14866	0.95888	-0.35427	-0.00102***	0.00021	0.82746
Ressource - TAP.N	0.00956	-0.42736	-0.87124	0.13468	1.56351	1.08168	0.00055***	0.000133	0.20303	0.99697
Sanfund - A.N	0.38275	-0.33098	-0.20797	-0.86998	0.33799	0.2499	1.04491	-0.000399***	0.00039	0.11193
Sanfund - AMAT.OQ	-0.00542	1.27347	-0.99395	-0.71945	-0.15745	-1.65689	0.130421	-0.17565	0.00013***	0.00021
Sanfund - AMD.OQ	0.06961	0.0751	-0.62449	-0.53851	-0.56151	0.89469	0.26367	-0.00046***	0.00024	0.06127
Samfund - BAC.N	-0.03362	0.16187	0.33429	0.49675	0.10758	-0.57053	-0.49375	-0.00164***	0	0.16247
Samfund - BDX.N	0.04013	1.35454	-0.20432	0.33923	1.68141	0.39	-0.36787	-0.000399***	0.00039	0.791
Samfund - BMY.N	0.03222	-1.47013	0.25007	-0.62449	-0.53851	-0.56151	0.89469	0.26367	-0.00046***	0.00052
Samfund - CAG.N	0.03283	-0.13303	2.05925	-1.90588	0.80236	2.13477	0.20945	-0.97917	0.00158***	0.00037
Samfund - CCL.N	-0.01999	-1.38239	-0.47479	1.27306	0.24762	-1	-0.00034***	1e-04	0.04628	0.55598
Samfund - CSCCO.OQ	0.04017	0.87319	-0.86074	-0.92611	0.80305	0.94044	-0.30385	-0.191	-0.00069***	0
Samfund - DGX.N	0.36353	-0.61447	-0.40643	-0.90025	1.73203	0.82716	-0.00457	-0.00039***	0.01013	0.99569
Samfund - FCX.N	0.06258	-1.91426	-1.58089	-1.06325	1.90151	0.89229	1.35495	0.17726	-0.00065***	1e-05
Samfund - GFS.N	0.05592	-0.03714	-0.65241	-0.83607	0.22096	0.86927	1.06832	0.00121***	5e-05	0.99717
Samfund - HES.N	0.13476	-0.28409	-0.18659	-0.90025	-0.00622	1.1108	1.00457	0.00044*	0.00026	4e-05
Samfund - IBM.N	-0.09911	-0.26376	-0.64346	-0.80553	0.06943	0.72275	1.02621	-0.00138***	0.00111	0.98192
Samfund - IR.N	0.03181	0.0739	-0.08709	-1.32188	0.25597	-1.52919	1.1829	-0.04882	-0.00027***	0.00048
Samfund - ITW.N	-0.02459	-1.02769	-0.90025	-0.92611	0.00378	0.84197	0.57026	-0.00115***	0.00012	0.1276
Samfund - MDT.N	0.11644	-0.25241	-0.64116	-0.51705	-0.92705	-0.96112	-0.21495	-0.00055***	1e-05	0.80685
Samfund - MMM.N	0.00772	0.85628	-0.91045	-0.90687	0.03058	-0.81819	0.87079	-0.00122***	0.0013	0.98192
Samfund - MO.N	0.10893	0.96271	-0.03712	-1	-	-	-	-	-	0.99875

Table 48: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med miljø- og rådata

Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
Bestyrelsen gns. siddetid - A.N	0.03775	-0.05541	0.05474	-0.93526	0.11499	0.01882	1.0852	-0.00087***	0.00021	0.15021
Bestyrelsen gns. siddetid - AMD.OQ	-0.02025	0.08712	-0.0586	0.15137				0***	3e-05	0.05
Bestyrelsen gns. siddetid - AMGN.OQ	-0.00893	-0.27425	-0.46933	0.53894	0.13329	0.38965	-0.82088	0.00082*	0	0.99842
Bestyrelsen gns. siddetid - BDX.N	-0.00149	0.23618	0.4912	0.10199	-0.55567	-0.50916		0.00032***	0	0.9
Bestyrelsen gns. siddetid - BLB.QQ	-0.05233	0.07958	0.89541		-0.20749	-0.87006	NA	0.00249***	5e-05	0.99275
Bestyrelsen gns. siddetid - CAG.N	-0.01021	0.04244	-0.06481		-0.42826	-0.64261		0***	0	0.05
Bestyrelsen gns. siddetid - CCL.N	-0.01551	0.24846	0.68106		1.42743	0.20411	-0.20697	-0.0012***	2e-05	0.04024
Bestyrelsen gns. siddetid - CSCO.OQ	0.03898	-1.53611	-0.59453		-1.46252	0.42979	0.01468	0.00019***	1e-05	0.95303
Bestyrelsen gns. siddetid - DGX.N	0.00014	1.37041	-0.45359		0.14418	1.07709	0.19455	-0.00016***	0	0
Bestyrelsen gns. siddetid - GHS.N	0.00819	0.08868	-1.036		-0.23927	-0.60663	-0.22317	-0.01215***	2e-05	0.00101
Bestyrelsen gns. siddetid - HES.N	0.18159	0.22449	0.66019		0.00276	-1.06543		-9e-04***	0.00038	0.00073
Bestyrelsen gns. siddetid - IBM.N	0.01708	-0.07821	0.97693		-0.9187	0.80764	0.28094	-0.00056***	0.00068	0.00068
Bestyrelsen gns. siddetid - IR.N	0.02527	1.04759	-0.97584		0.32355	-0.62464	1.09219	NA	0.00055***	0.38739
Bestyrelsen gns. siddetid - ITW.N	-0.0016	0.80622	-1.16323		0.7639	-0.36691	0.26973	-1.0488	-0.0012***	0.00017
Bestyrelsen gns. siddetid - MMM.N	0.03028	0.28334	-0.22998		0.14464	-0.98179	-0.00817	-0.00368***	0	0.999
Bestyrelsen gns. siddetid - MO.N	0.08492	0.06211	0.92405		0.21914	-1.02144	1.09503	-0.00107***	0.00028	0.00028
Bestyrelsen gns. siddetid - RTN.N	0.0274	1.10693	-1.12339		-0.00271	-1.07737		-0.00248***	0.00083	0.26229
Bestyrelsen gns. siddetid - STT.N	0.04045	0.04239	0.94107		-1.73054	0.87293	0.08236	-0.00166***	0.00038	0.86355
Bestyrelsen gns. siddetid - T.N	0.04713	1.64196	-0.86419		0.30051	-1.2846	1.26315	-0.1591	-0.00277***	1e-04
Bestyrelsen gns. siddetid - TXN.QQ	0.07118	1.38709	-1.32744		-1			0***	0.00039	0
CO2Total - DGX.N	0.00453	0.90169			-0.23289	0.84979		0.00032***	0.00033	0.13833
Senior Executive Kompenstation - ABT.N	0.00995	-0.01813	0.799		-1.5229	1.35011	-0.19689	-5e-03***	0.00046	0.19465
Senior Executive Kompenstation - AMAT.QQ	0.00906	1.25856	-0.98099					***	0.75901	
Senior Executive Kompenstation - AMD.QQ										
Senior Executive Kompenstation - BAX.N	0.00608	-0.78026	0.63669	0.87857	0.74695	-0.80899	-1.06831	0.00015***	0	0.99891
Senior Executive Kompenstation - BDX.N	0.01751	0.39285	0.41928	0.06689	-0.6377	-0.42535		-0.00164***	0	0.05001
Senior Executive Kompenstation - BMY.N	0.00968	0.08349	0.92425	0.01534	-0.06822	-1.13206	0.08328	-0.00049***	1e-05	3e-05
Senior Executive Kompenstation - CCL.N	-0.0122	0.60876	0.36965		-0.73915	-0.27555		-2e-05***	6e-04	0.21244
Senior Executive Kompenstation - CPB.N	0.01789	0.2733	0.68635	0.03401	-0.48114	-0.60385		-0.00049***	0.00027	0.70214
Senior Executive Kompenstation - DGX.N	0.00413	1.14537	-0.22932		-1.2793	0.25189		-5e-05***	0	0.10632
Senior Executive Kompenstation - GIS.N	0.00518	0.67948	0.33382	-0.03955	-0.53215	-0.47412		-0.00022***	0	0.80398
Senior Executive Kompenstation - HES.N	0.02592	-0.611447	0.62073	0.81058	0.56039	-0.73561	-1.01024	-0.00297***	0.0089	0.01031
Senior Executive Kompenstation - HPQ.N	-0.00134	1.67886	-0.55256	-0.15462	-1.83115	0.82208		0.00051***	0.00062	0.117
Senior Executive Kompenstation - IBM.N	0.00385	-0.09215	0.9984		0.02865	-1.0788		-0.00166***	5e-05	0.98242
Senior Executive Kompenstation - INTC.OQ	0.01349	-0.45038	-0.95482	0.15799	0.61318	1.08613		-0.00115***	0	0.00797
Senior Executive Kompenstation - KMB.N	-0.0041	-0.4999	0.45036	0.86724	0.58352	-0.61553	-1.09251	0.00033***	0.00013	0.98867
Senior Executive Kompenstation - LLY.N	0.02684	-1.15384	-0.9275		1.24648	1.03088		-0.00136*	0.00098	0.12789
Senior Executive Kompenstation - MRK.N	0.01456	-0.07867	-0.65394	-0.40575	-0.16146	0.89252	0.43396	-0.00059***	0.00015	0.34059
Senior Executive Kompenstation - TAP.N	0.0124	-1.17777	-0.24535	NA	0.97581	-0.00102*	0.00282	0.47808	0	0.30182

Table 49: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med miljø- og rådata

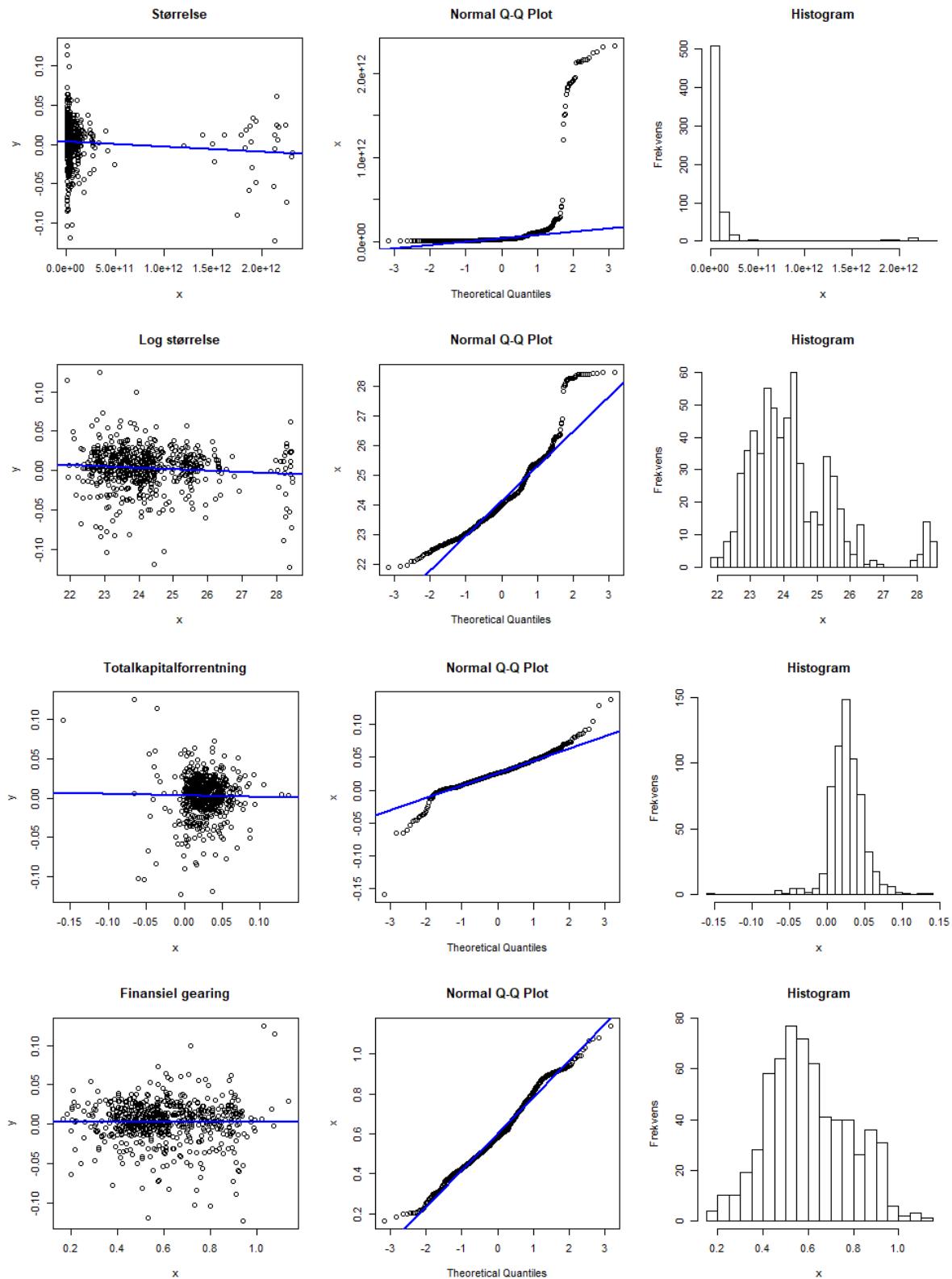
Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
CO2 udledning/USD - BIIB.OQ	-0.00169	0.63072		1.222283	0.12814 NA	86.87967***	1e-05		1e-04	0.7664
CO2 udledning/USD - CAG.N	0.01052	-0.17579	0.90587	0.18063	-1.09952	-0.20701	-61.86242***	0.00031	0.25076	0.67587
CO2 udledning/USD - PG.CN	-0.00385	0.38325	0.56239	-0.26627	-0.62598	-0.21378	84.63013***	2e-05	0.03457	0.95342
CO2 udledning/USD - T.N	-0.0029	1.72014	-1.27436	0.23864	-1.82003	1.47014	-0.24536	-0.0043***	0.00021	0.09561
Affald total/USD - AMAT.OQ	0.00586							0***	0.00073	0.12366
Affald total/USD - AMD.OQ	-0.07254							0***	3e-05	0.9
Affald total/USD - CAG.N	-0.02251	-1.07255	-0.29358	0.36343	1.13881	0.22189	-0.58227	0.00182***	0.00023	0.31327
Affald total/USD - COP.N	0.06633	-0.15501	0.72982	0.36037	0.22775	-0.83436	-0.55739	-0.01231***	4e-04	0.22099
Affald total/USD - CSCO.OQ	0.0354	0.01706	0.78979		-0.02424	-1.05053		-0.00358***	1e-05	0.03168
Affald total/USD - DGX.N	-0.00819	0.24045	-0.52519	0.73532	-0.47388	0.4309	-1.05343	0.00168***	0	0.0015
Affald total/USD - GIS.N	-0.02058	-0.11608	0.70593	-0.15346	0.29502	-0.79812		0.00225***	0	2e-05
Affald total/USD - HES.N	-0.00525	0.2886	0.74388		-0.26991	-0.70869	-0.12793	0.00297***	1e-05	0.05814
Affald total/USD - IBM.N	0.00726	-0.10123	1.00563		0.02965	-1.09814		-0.0089***	0.00026	0
Affald total/USD - IR.N	-0.06814	0.21069	0.62137		-0.27164	-0.69359	-0.096411	0.00729***	0.00083	0.43953
Affald total/USD - MDT.N	0.03442	-0.36775	-0.1674	0.15949				-0.0292*	0.0013	0.56847
Affald total/USD - MMM.N	-0.01586	-0.65891	-0.81489		0.72244	1.0185		0.00252***	0.00047	0.11795
Affald total/USD - MO.N	0.02593	0.05984	0.95532		0.0243	-1.0942		-0.00132***	0.00019	0.05568
Affald total/USD - MRK.N	-0.06102	-0.10233	-0.67208	-0.44903	-0.12916	0.8763	0.47409	0.00851***	9e-05	0.05757
Affald total/USD - NEM.N	-0.14771	-0.11378	-0.21984	-0.99497	0.22732	0.20569	1.13003	0.01501***	0.00105	0.28989
Affald total/USD - PGN	0.00953	0.93951	-0.0787		-1			-0.0081**	0.00063	0.12153
Affald total/USD - RTN.N	-0.02505	1.13476	-1.12281	0.22383	-1.01999	1.09859		0.00404***	2e-04	0.06347
Affald total/USD - TAP.N	0.03547	-0.00129	0.89063 NA	-0.17593	-0.95847	0.25822	-0.00557*	0.00182	0.42959	0.22028
Affald total/USD - TXN.OQ	0.17898	-0.3072	-0.95735	0.28437	0.93852	-0.0777	-0.01729**	0.00017	0.0555	0.63817
Affald total/USD - UPS.N	0.00951	-0.0274	NA	-0.07953	-0.15784	-0.12375	-0.0031***	0.0064	0.20671	0.63838
Vandforbrug total/USD - AMAT.OQ	0.03639	1.23614	-0.5579	-0.20899	-1.38547	0.7366	0.27526	-30.98856***	0.00058	0.08702
Vandforbrug total/USD - BLL.N	-0.00513	0.09418	-0.12559	0.92826	-0.11908	0.00263	-1.03103	33.75637***	0.00046	0.17836
Vandforbrug total/USD - BMY.N	0.009	1.14222	-0.3935	-0.35719	-1.51417	0.78064	0.20255	-41.80475***	1e-05	0
Vandforbrug total/USD - CAG.N	0.01752	0.17732	0.68461	-0.11712	-0.80737	-0.17037	-0.17037	-6.39061***	0.00054	0.38549
Vandforbrug total/USD - CSCO.OQ	0.0072	0.33801	0.49373	-0.52399	-0.66754	0.12184	-0.15317***	6e-05	0.05422	0.93095
Vandforbrug total/USD - DGX.N	0.0053	0.3256	-0.50336	0.81635	-0.48674	0.42747	-1.06878	-0.16133***	0	0
Vandforbrug total/USD - FCX.N	-0.0047				0.07914	0.18367		0***	2e-05	0.05
Vandforbrug total/USD - HES.N	0.01622	0.20914	0.79525	-0.07615	-0.29574	-0.75931	-0.04828	4.58967***	2e-05	0.00296
Vandforbrug total/USD - HPQ.N	0.00054	0.93654	-0.79291	-0.09701	-0.74545	0.53465	0.52718	19.72513***	0.00065	0.08583
Vandforbrug total/USD - IBM.N	-0.00108	-0.07713	1.01974		-0.00068	-1.1166		-0.45105***	1e-05	3e-05
Vandforbrug total/USD - INTC.OQ	-0.00019	-0.22163	0.19975	0.95216	0.18789	-0.31187	-1.05929	-0.67086***	0	4e-05
Vandforbrug total/USD - IR.N	0.03701	1.52565	-1.25738	0.21716	-1.48996	1.12049		-59.72262***	4e-04	0.27677
Vandforbrug total/USD - KMB.N	-0.00125	-0.46825	0.50104	0.87948	0.56084	-0.65789	-1.11801	-0.10084***	0.00016	0.10917
Vandforbrug total/USD - MMM.N	0.02231	-0.64333	-0.7804		0.7185	1.03468		-9.40028***	0.00044	0.13869
Vandforbrug total/USD - MO.N	0.01369	0.72999	0.86339	-0.76824	-1.0197	-0.94465	1.12392	-2.25161***	0.00039	0.12341
Vandforbrug total/USD - MRK.N	0.00267	-0.16852	-0.92677	0.20922	0.05398	1.04313	-0.28184	40.08709***	0.00017	0.05027
Vandforbrug total/USD - RTN.N	0.00926	-0.29535	0.16149	-0.55706	0.28992	-0.09929	0.77483	0.90695***	0.00022	0.09172
Vandforbrug total/USD - T.N	-0.00068	0.25316	0.38709	-0.86773	-0.24571	-0.53403	0.91663	0.81744***	0.00019	0.0926

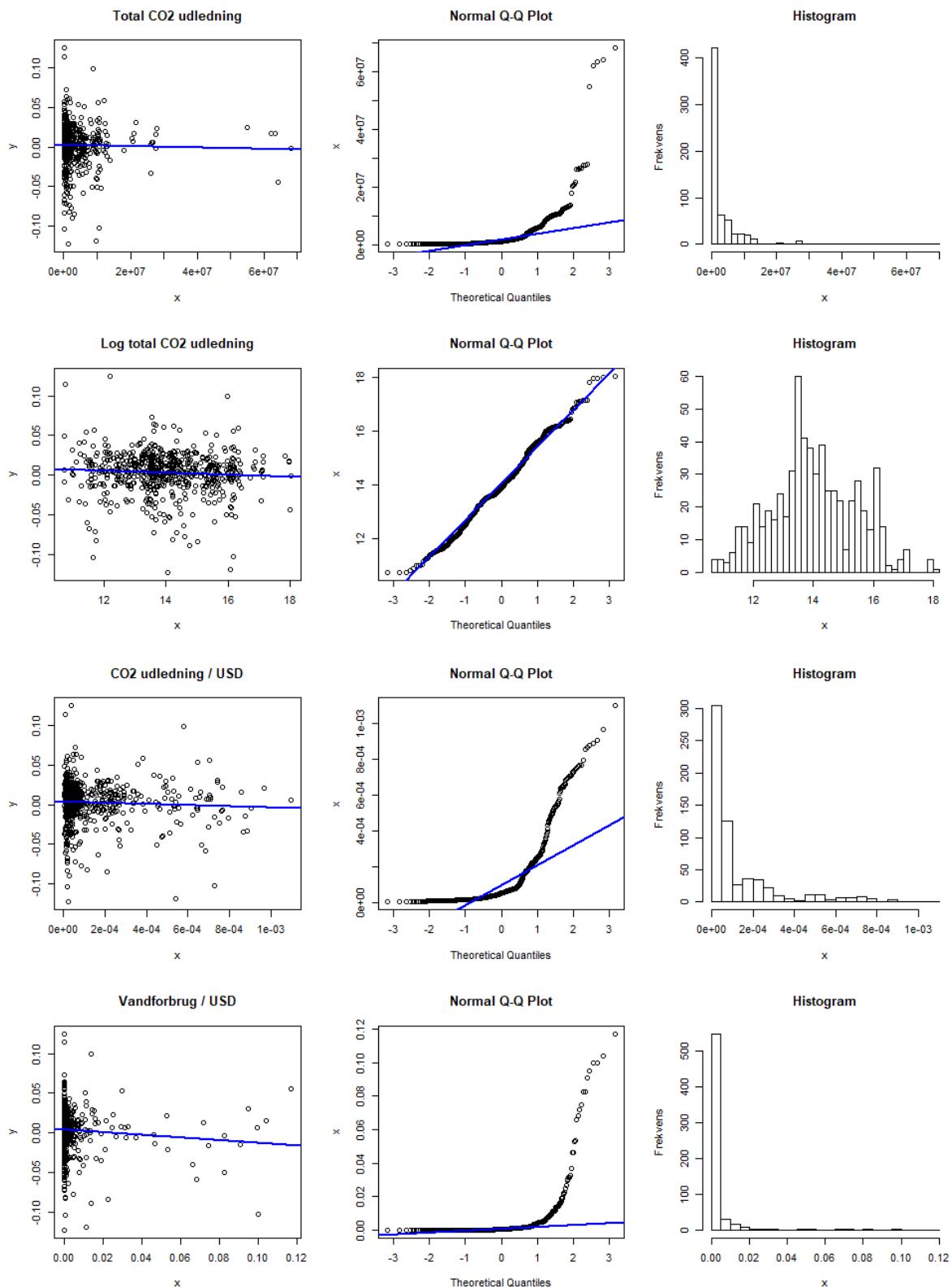
Table 50: Signifikante ARMA-X(p,q)/GARCH(1,1) modeller med miljø- og rådata

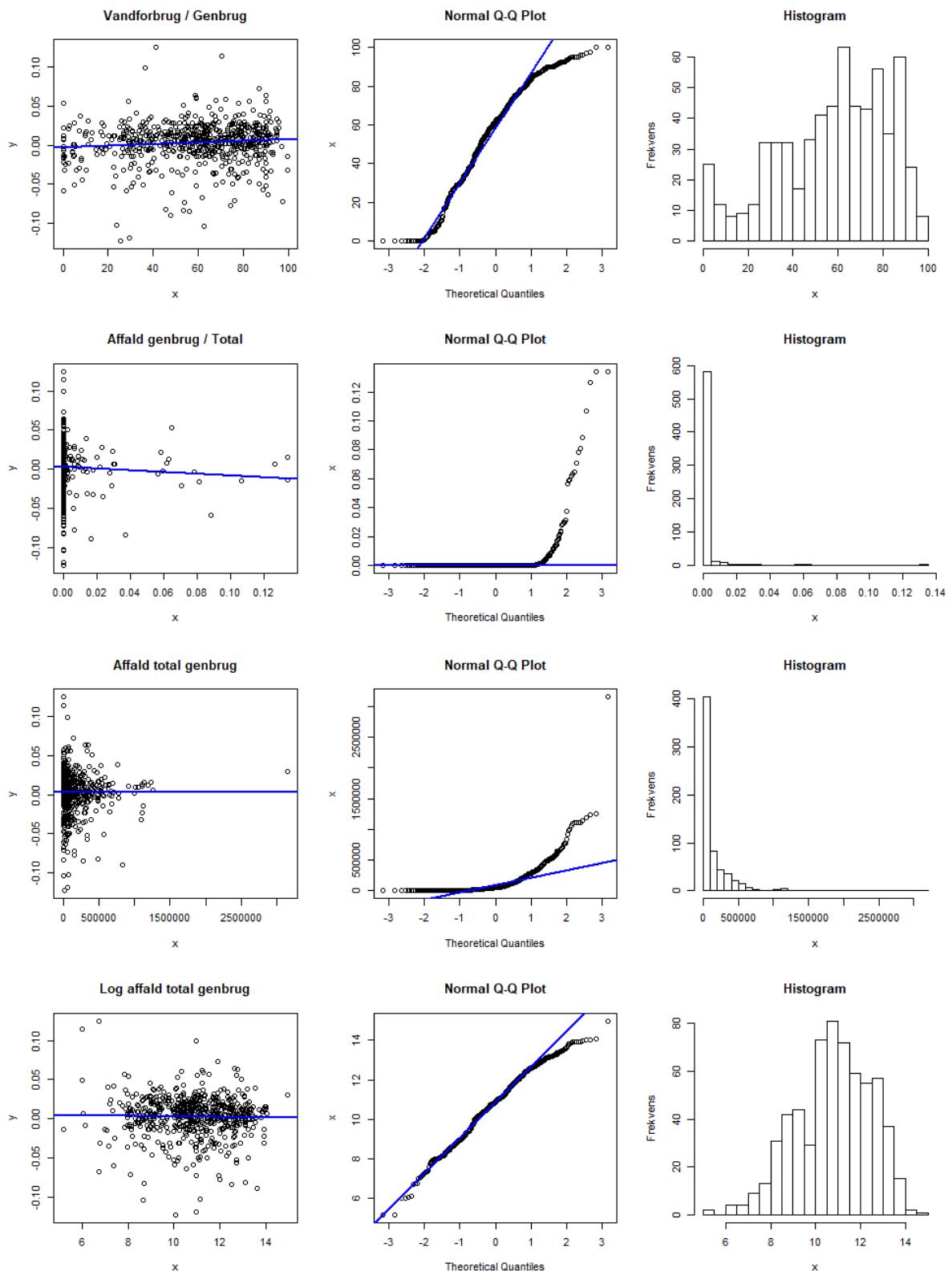
Model	intercept	xreg	$\phi_1$	$\phi_2$	$\phi_3$	$\theta_1$	$\theta_2$	$\theta_3$	$\theta_4$	$\theta_5$
Affald genbrugt - AMD.OQ	0.00482	0.09005	-0.05603	0.15375					3e-05	0.05
Affald genbrugt - AMGN.OQ	0.04212	-0.30643	-0.39112	0.68349	0.16476	0.28023	-0.84398	0*	0	0.99889
Affald genbrugt/total - AMD.OQ	0.0356	-0.42397	-1.05006	0.58342	1.13322	0.1333	-1e-05***	0.00307	0.08433	0.79687
Affald genbrugt/total - AMGN.OQ	0.04088	-0.41161	-0.46152	0.52227	0.20994	0.35708	-0.76842	0***	0	0.99837
Affald genbrugt/total - DGX.N	0.0044	0.90488		-1					0.00038	0
Affald genbrugt/total - RTN.N	0.01325	1.63238	-1.56532	0.80368	-1.70218	1.72044	-0.8849	0***	0.00023	0.09119
Affald genbrugt/total - SFT.N	0.04358	-0.30972	-0.31249	-0.16922	-0.848	0.34628	0.24143	1.03471	0.00021***	0.84339
Affald totalt - A.N	-0.00743	-0.31249	-0.16922						5e-05	0.05
Affald totalt - CCL.N	-0.05209				-0.12289				0.00189***	0.85271
Affald totalt - CML.N	0.057	0.00468	0.79927		-0.20135	-0.8741			-6e-04***	0.7069
Affald totalt - COP.N	0.03065	-0.74387	0.80931	0.88381	0.77746	-0.86987	-1.04736		-0.00108***	0.38731
Affald totalt - CSCO.OQ	0.0355	-0.00707	0.70627	0.04175	-0.08936	-0.99235			-0.00084***	0.58695
Affald totalt - DGX.N	-0.00164	0.34941	-0.46194	0.82393	-0.46361	0.42627	-1.04769		1e-04***	0.9
Affald totalt - HES.N	-0.00518	0.06969	0.62011		-0.10061	-0.68111	-0.29153		0.00054***	0.00538
Affald totalt - HPQ.N	0.04738	1.23726	-1.07205	0.16358	-1.0292	0.86122	0.19336		-0.00062***	0.99216
Affald totalt - IBM.N	0.04244	-1.22845	-0.49975	0.26532	1.53504	0.94241	-0.09316		-0.0064***	0.9
Affald totalt - MMM.N	-0.02117	0.29864	-0.23319	0.74011	-0.32075	0.2684	-1.05151		0.0046***	0.87592
Affald totalt - MO.N	0.10185	-0.70423	0.19765		0.77762	-0.24194	0.06771		-0.0129***	0
Affald totalt - MRK.N	-0.0022	-0.18712	-0.611981	-0.47723	-0.06821	0.85903	0.52048		2e-04***	0.04763
Affald totalt - NEM.N	0.13283	0.70592	-0.69537	0.9266	-0.82039	0.7757	-1.03554		-0.0145***	0.95124
Vand total tilbagetrækning - A.N	-0.0554	-1.57824	-0.98976		1.55915	0.93751	-0.08264		0.08176**	0.77045
Vand total tilbagetrækning - ABT.N	0.01821	0.34656	-0.86908	-0.06344	-0.7116	1.24807	-0.30757		-0.01611***	0.75899
Vand total tilbagetrækning - AMAT.OQ	0.00554	1.69249	-1.21192	0.19619	-1.79583	1.45633	-0.237		-0.00343***	0.21172
Vand total tilbagetrækning - BMY.N	0.0086	1.13274	-0.43356	-0.32405	-1.5488	0.85699	0.16429		-0.00437***	0.12402
Vand total tilbagetrækning - CAG.N	-0.01184	0.11322	0.70182		-0.14589	-0.88986	-0.06887		0.00016***	0.80416
Vand total tilbagetrækning - CCL.N	-0.05862	0.81214	0.16795	-0.97442					0	0.999
Vand total tilbagetrækning - COP.N	0.03073	-0.74781	0.81087	0.89304	0.77756	-0.87436	-1.05251		2e-04***	0.69706
Vand total tilbagetrækning - CSCO.OQ	0.03394	0.06115	0.78379		-0.16083	-1.06265	0.13196		-0.0032***	0.30159
Vand total tilbagetrækning - DGX.N	-0.00088	0.30427	-0.50298	0.81924	-0.48025	0.43183	-1.07038		0.00021***	0.93654
Vand total tilbagetrækning - FCX.N	-0.04444	-0.58883	0.16885		0.50247				0.00516***	0.99302
Vand total tilbagetrækning - GIS.N	0.01391	-0.58728	0.32199	-0.07772	0.8208	-0.25058			-0.01454***	0
Vand total tilbagetrækning - HES.N	-0.01221	0.50238	0.64228	-0.20289	-0.48559	-0.60807			4e-05	0.64055
Vand total tilbagetrækning - IBM.N	-0.02842	-0.55934	0.56697	0.95778	0.50686	-0.62884	-1.06503		0.09118***	0.99648
Vand total tilbagetrækning - INTC.OQ	-0.01953	-0.41855	0.2216	0.84738	0.27064	-0.38273	-1.04676		2e-05	0.99542
Vand total tilbagetrækning - IR.N	0.005	1.53774	-1.20238	0.21551	-1.49769	1.14584			0.04572***	0.00101
Vand total tilbagetrækning - MMM.N	-0.02141	0.27046	-0.24209	0.81654	-0.33553	0.25435	-1.07003		0.03545***	0.97616
Vand total tilbagetrækning - MO.N	0.06852	0.07269	0.93681		0.17336	-0.99323	-0.04429		-0.03537***	0
Vand total tilbagetrækning - MRK.N	-0.00669	-0.16493	-0.64968	-0.48047	-0.06653	0.8558	0.52052		0.02445***	0.999
Vand total tilbagetrækning - NEM.N	-0.01045	0.75003	-0.02323	0.15542	-0.91472	0.03532	-0.19737		9e-05	0.8947
Vand total tilbagetrækning - TAP.N	0.03181	-1.82687	-0.83302 NA	1.7407	0.54447	-0.19135	-0.03366***	0.00231	0.00142	0.68633
									0.28788	0.2348

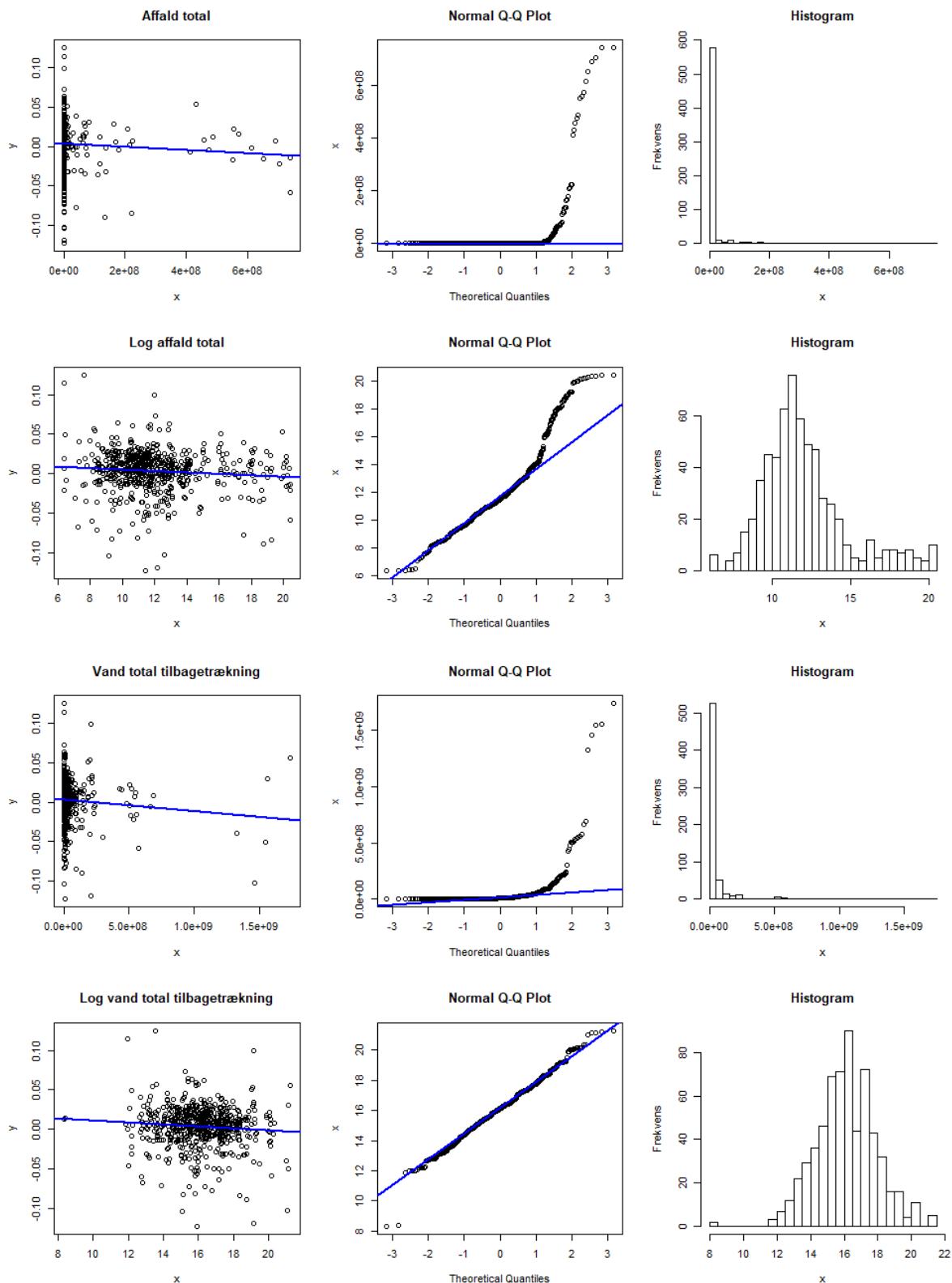
## C Regressionsanalyse

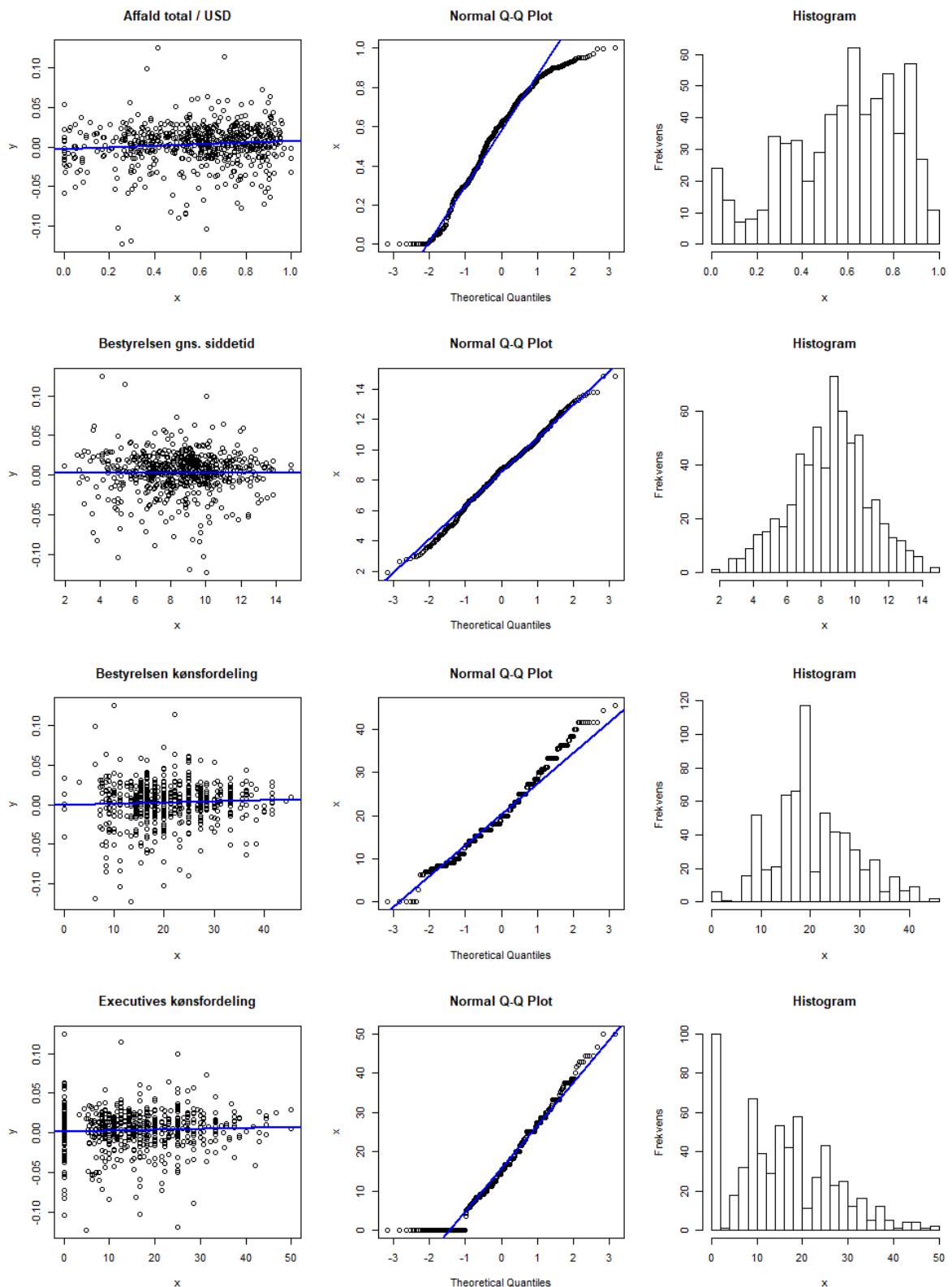
### Transformering af variable

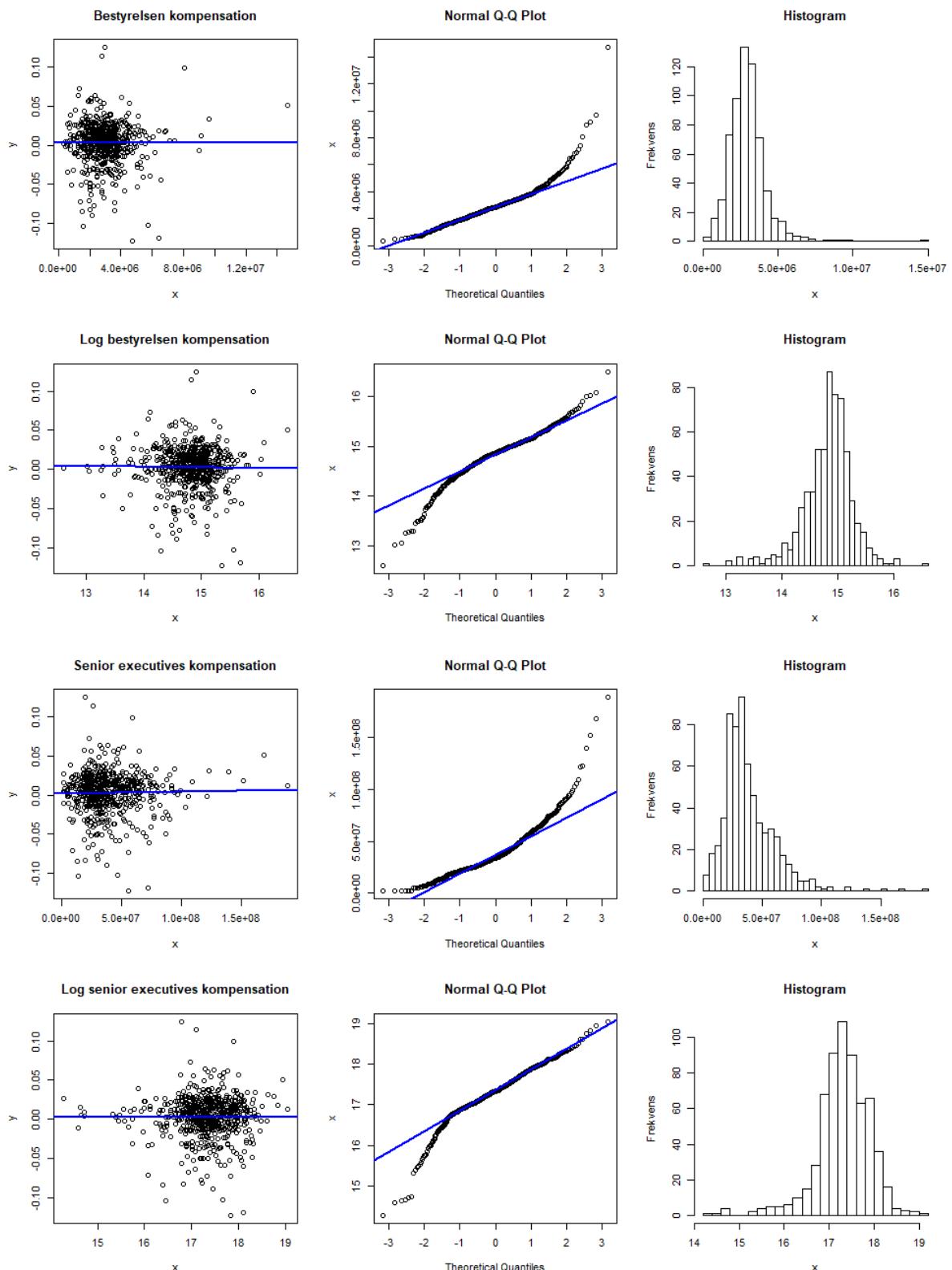












## Regressions resultater

	<i>Respons variabel:</i>		
	Logaritmisk afkast		
	Pooled	Random	Fixed
ESG Score	0.0001* (0.0001)	0.0001 (0.0001)	0.00005 (0.0001)
Totalkapitalforrentning	-0.019 (0.031)	-0.031 (0.028)	-0.093*** (0.031)
Fi nsiel gearing	0.004 (0.006)	0.002 (0.005)	0.032*** (0.011)
Størrelse	-0.003*** (0.001)	-0.002*** (0.001)	-0.012*** (0.003)
Intercept	0.054*** (0.018)	0.054*** (0.016)	
Observationer	7,560	7,560	7,560
Virksomheder	45	45	45
R <sup>2</sup>	0.001	0.002	0.005
Adjusted R <sup>2</sup>	0.001	0.001	-0.025
F Statistic	2.685** (df = 4; 7555)	12.640**	8.403*** (df = 4; 7344)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 52: Regressions Resultater

	<i>Respons variabel:</i>		
	Logaritmisk afkast		
	Pooled	Random	Fixed
Environmental	-0.00002 (0.0001)	-0.00005 (0.0001)	-0.00004 (0.0001)
Social	0.0001 (0.0001)	0.0001 (0.0001)	0.0001 (0.0001)
Governance	0.0001 (0.0001)	0.00004 (0.0001)	-0.00004 (0.0001)
Totalkapitalforrentning	-0.022 (0.031)	-0.034 (0.028)	-0.095*** (0.031)
Finansiell gearing	0.004 (0.006)	0.002 (0.005)	0.034*** (0.011)
Størrelse	-0.003*** (0.001)	-0.002*** (0.001)	-0.012*** (0.003)
Intercept	0.052*** (0.018)	0.053*** (0.016)	
Observationer	7,560	7,560	7,560
Virksomheder	45	45	45
R <sup>2</sup>	0.002	0.002	0.005
Adjusted R <sup>2</sup>	0.001	0.001	-0.025
F Statistic	2.002* (df = 6; 7553)	14.679**	5.959*** (df = 6; 7342)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 53: Regression Results

	Respons variable:		
	Logaritmisk afkast		
	Pooled	Random	Fixed
CO2 Udledning - Total	-0.001 (0.002)	-0.0004 (0.002)	-0.001 (0.002)
CO2 Udledning / USD	8.774 (10.337)	-1.379 (9.039)	-12.836 (13.359)
Energi - Totalt Forbrug	0.001 (0.001)	0.001 (0.001)	0.0003 (0.001)
Affald - Genbrug / Total	0.074 (0.133)	0.154 (0.115)	-0.090 (0.204)
Affald - Total Genbrug	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Affald - Total	-0.001 (0.001)	-0.001 (0.001)	-0.0003 (0.001)
Affald - Total / USD	0.034 (0.041)	0.002 (0.036)	-0.001 (0.038)
Vandforbrug - Total / USD	-0.143 (0.128)	-0.164 (0.111)	-0.176 (0.145)
Vandforbrug / Genbrug	-0.0003 (0.0004)	0.00002 (0.0004)	0.0001 (0.0004)
Vand - Total Tilbagetrækning	-0.001 (0.001)	-0.0001 (0.001)	-0.001 (0.001)
Bestyrrelsen - Gns. Siddetid	-0.0003 (0.0004)	-0.0003 (0.0004)	-0.001 (0.001)
Bestyrrelsen - Kønsfordeling	0.0002* (0.0001)	0.0002* (0.0001)	0.0003 (0.0002)
Bestyrrelsen - Kompensation	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)
Senior Executive - Kompensation	0.004 (0.002)	0.003* (0.002)	0.002 (0.002)
Executive - Kønsfordeling	0.00005 (0.0001)	0.00003 (0.0001)	-0.00004 (0.0001)
Grønne Bygninger	0.0004 (0.002)	-0.001 (0.002)	0.00000 (0.003)
Miljø Distributionskæde Politik	-0.004 (0.003)	-0.004 (0.003)	-0.002 (0.004)
Miljø Kontroverser	-0.009 (0.006)	-0.009* (0.005)	-0.012** (0.006)
Affaldsreducerende Initiativer	-0.003 (0.004)	-0.002 (0.003)	-0.007* (0.004)
Vedvarende og Grøn Energiforbrug	0.002 (0.003)	0.003 (0.002)	0.007** (0.003)
Reduktion af Giftige Kemikalier	0.003 (0.002)	0.003 (0.002)	0.004 (0.003)
Totalkapitalforrentning	-0.038 (0.033)	-0.056* (0.029)	-0.112*** (0.032)
Fi nsiel gearing	0.004 (0.006)	0.001 (0.005)	0.027** (0.011)
Størrelse	-0.003*** (0.001)	-0.003*** (0.001)	-0.012*** (0.004)
Intercept	0.023 (0.042)	0.037 (0.039)	
Observationer	7,560	7,560	7,560
Virksomheder	45	45	45
R <sup>2</sup>	0.004	0.005	0.008
Adjusted R <sup>2</sup>	0.001	0.002	-0.024
F Statistic	1.416* (df = 24; 7535)	41.313**	2.537*** (df = 24; 7324)

Note:

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 54: Regression Results  
I42

Variable	VIF: Y4
<b>ESG Niveau 4</b>	
CO2 Udledning - Total	<b>12.5265</b>
CO2 Udledning - / USD	3.225401
Energi - Totalt Forbrug	<b>10.5134</b>
Affald - Genbrug / Total	<b>33.6917</b>
Affald - Total Genbrug	1.221911
Affald - Total	<b>19.3216</b>
Affald - Total / USD	<b>113.728</b>
Vandforbrug - / USD	<b>35.0570</b>
Vandforbrug - / Genbrug	<b>114.361</b>
Vand - Total Tilbagetrækning	<b>19.9511</b>
Bestyrrelsen - Gns. Siddetid	1.155025
Bestyrrelsen - Kønsfordeling	1.215500
Bestyrrelsen - Kompenstation	1.481047
Senior Executive - Kompenstation	1.976378
Executive - Kønsfordeling	1.270043
Grønne Bygninger	1.630694
Miljø Distributionskæde Politik	1.640844
Miljø Kontroverser	1.351833
Affaldsreducerende Initiativer	1.403452
Vedvarende og Grøn Energiforbrug	1.218735
Reduktion af Giftige Kemikalier	1.299405
<b>Kontrol Variable</b>	
Totalkapitalforrentning	1.135778
Finansiell gearing	1.292478
Størrelse	1.982509

Table 55: VIF: Y4 før relevante transformationer

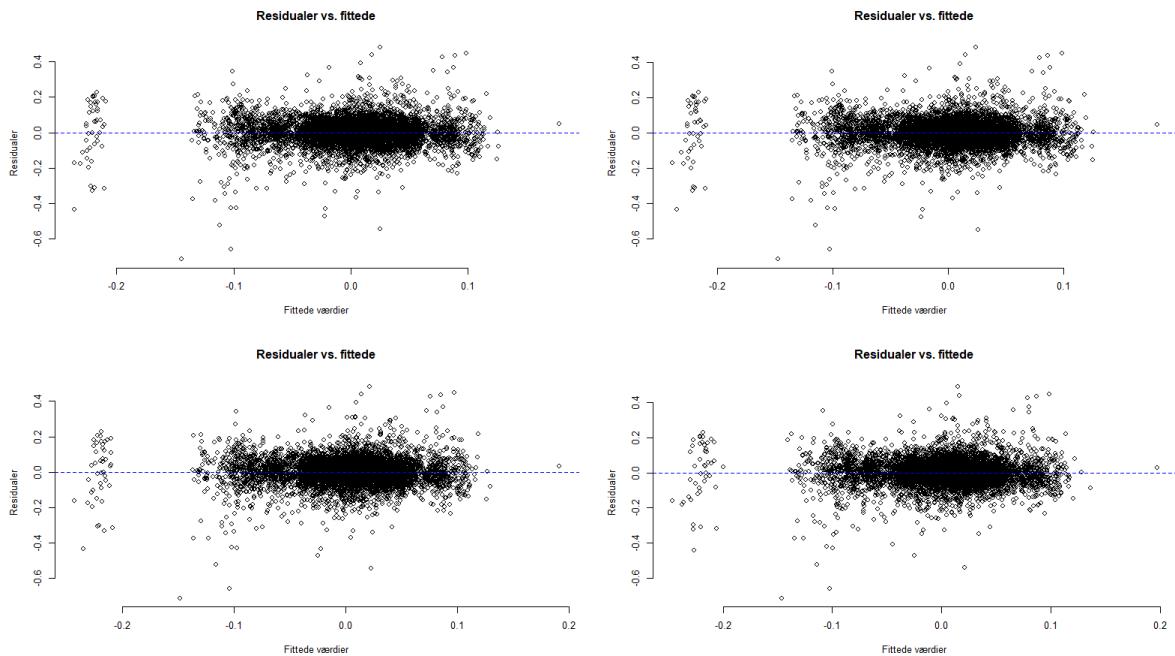


Figure 13: Residual vs. fittede

## D Robustheit

### Residualer ARMA

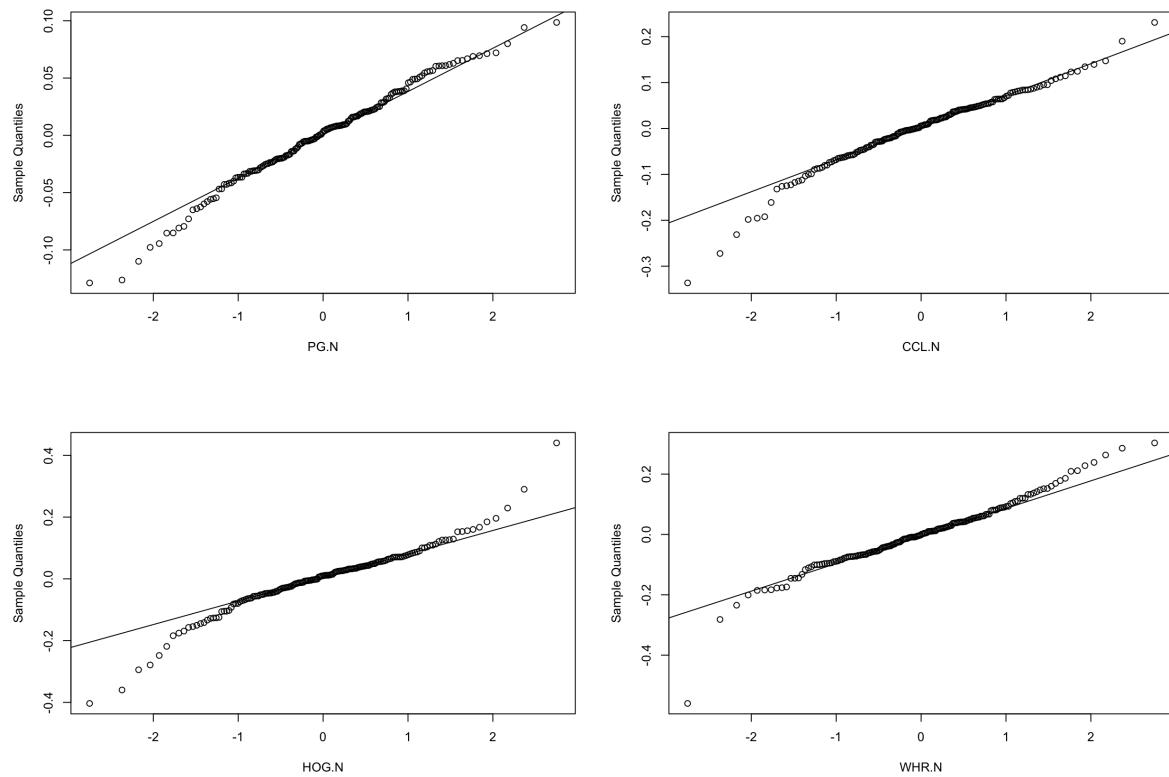


Figure 14: QQ Plot - ARMA residualer

## Fitted mod log-aftkast

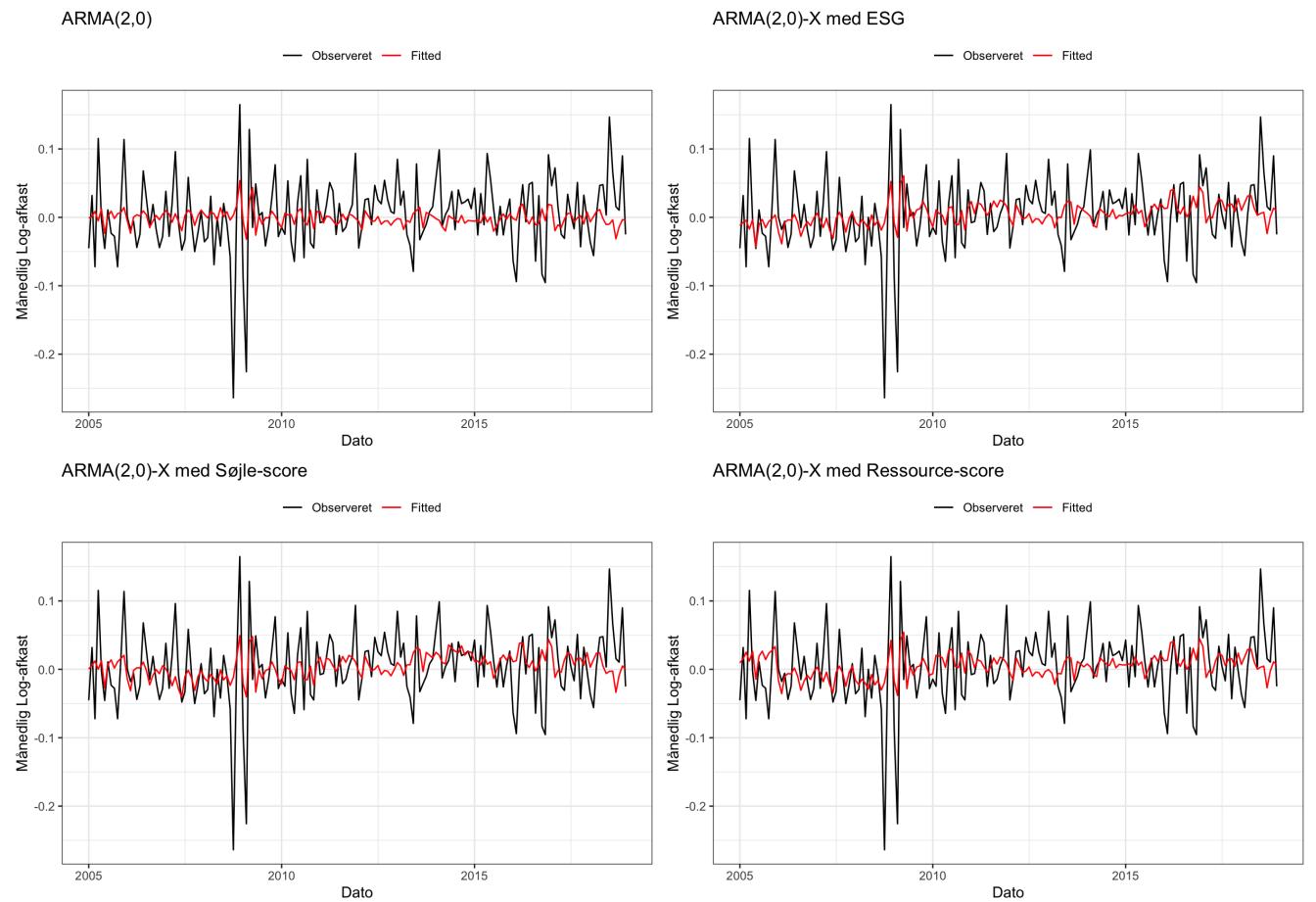


Figure 15: In-Sample fits af LLY

## BG test - ESG Robusthed

	STT.N	C.N	BAC.N	ABT.N	AMGN.OQ	BAX.N	BDX.N
Test statistik	26.73198	15.00696	17.25404	12.22287	15.6759	14.74112	14.80178
P-værdi	0.14297	0.77601	0.63642	0.90819	0.73652	0.79102	0.78764
	BMY.N	BIIB.OQ	LLY.N	MDT.N	DGX.N	A.N	MRK.N
Test statistik	7.81373	12.74214	13.81251	18.97349	18.3809	29.89164	24.4079
P-værdi	0.99303	0.88817	0.83987	0.52355	0.56233	0.07163	0.22503
	T.N	AMD.OQ	AMAT.OQ	CSCO.OQ	HPQ.N	INTC.OQ	IBM.N
Test statistik	28.80609	14.70047	16.83804	25.09789	16.84272	21.81733	14.70842
P-værdi	0.09165	0.79328	0.66346	0.19771	0.66316	0.3505	0.79284
	MSFT.OQ	TXN.OQ	BLL.N	FCX.N	NEM.N	CMI.N	ITW.N
Test statistik	13.41866	23.88445	14.77149	6.318	15.03731	18.46963	23.90204
P-værdi	0.85871	0.24748	0.78933	0.99839	0.77427	0.5565	0.2467
	IR.N	MMM.N	UPS.N	RTN.N	HES.N	BKR.N	COP.N
Test statistik	13.13815	22.39677	16.60065	11.36212	7.83443	20.11031	11.94212
P-værdi	0.87138	0.31937	0.67873	0.93629	0.99291	0.45105	0.91805
	CPB.N	CAG.N	TAP.N	GIS.N	KMB.N	MO.N	PG.N
Test statistik	21.55944	19.15632	15.36025	13.92296	13.93873	22.61581	18.60191
P-værdi	0.36489	0.51169	0.75544	0.83438	0.83359	0.30805	0.54782
	CCL.N	HOG.N	WHR.N				
Test statistik	20.74748	6.30961	13.29367				
P-værdi	0.41213	0.99841	0.86443				

Table 56: Breusch-Godfrey Test - ESG

## Opsummeringstabeller

Model	ESG-score	Pillar-score	Under-score	Miljø- og Rådata
ESG-score	1			
Pillar-score		0		
CSRStratScore			2	
EmissScore			5	
HumanRightScore			3	
MgmtScore			4	
RessourceScore			5	
ShareholderScore			2	
ProductResponsScore			8	
WorkforceScore			1	
Bestyrelsen gns. siddetid				3
BoardGendDiv				2
BoardMemberComp				6
CO2Total				10
EnergyUseTotal				11
ExecGenderDivPercent				1
Senior Executive Kompensation				8
CO2 udledning/USD				5
Affald total/USD				3
Affald genbrugt				7
Affald genbrugt/total				3
Affald totalt				7
WaterRecycledRatio				3
Vand total tilbagetrækning				6

Table 57: Tidsserier med signifikante regressors

Model Spec	CSR	Emmision	HR	Ledelse	Produkt	Ressource	Aktionærers	Arbejd
ARMA(0,0)	1	2		1	3	1		
ARMA(0,1)						2		
ARMA(0,2)		1				1		
ARMA(0,3)				1				
ARMA(0,4)			1	1				1
ARMA(1,0)			1					
ARMA(1,1)	1				2			
ARMA(1,2)		1						
ARMA(2,0)				1		1		1
ARMA(2,2)		1						
ARMA(3,0)			1					
ARMA(3,2)						1	1	
Gns. Effekt	-0.00038	-0.00076	0.00122	-0.0007	-0.00083	0.00121	0.00135	0.00069

Table 58: Tidsserier med signifikante regressors, Under-score

Model Spec	Bestyrelse Komp.	CO2 Total	Energiforbrug	SE Komp.	CO2 Emiss./USD	Genbrugt	Affalld	Affald total	WWT
ARMA(0,0)	1	4	7	3	1	4	2	3	
ARMA(0,1)		1				1			
ARMA(0,2)						1		1	
ARMA(0,3)				1					
ARMA(0,4)									
ARMA(1,0)				1					
ARMA(1,1)		1	1				2		
ARMA(1,2)	1			1					
ARMA(1,3)			1						
ARMA(2,0)	1	1		1	1				1
ARMA(2,2)			1		1				
ARMA(2,3)					1		1		
ARMA(3,0)		1			1				
ARMA(3,2)	3	2	1	1			2	2	
Gns. Effekt	-0.0555	0.04444	-0.03124	-0.03782	-249.21701	-0.01145	-0.04217	-0.06328	

Table 59: Tidsserier med signifikante regressors, Miljø- og Rådata

### Regressionsanalyse

<i>Respons variable:</i>	
Logaritmisk afkast	
Affaldsreducerende Initiativer	-0.008** (0.003)
CO2 udlening total / USD	-24.277** (11.921)
Vandforbrug total / USD	-0.217*** (0.054)
Vand genbrug / Total	0.0001** (0.00004)
Vand total tilbagetrækning	-0.001* (0.001)
Bestyrelsen kønsfordeling	0.0003** (0.0001)
Totalkapitalforrentning	-0.114** (0.046)
Fi nsiel gearing	0.025* (0.013)
Størrelse	-0.012*** (0.004)
Observationer	7,560
Virksomheder	45
R <sup>2</sup>	0.007
Adjusted R <sup>2</sup>	-0.024
F Statistic	5.569*** (df = 9; 7339)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Table 60: Regressions resultater imputering (2)