



Handle ansvarlig eller fremstå ansvarlig?

En kvantitativ studie av ansvarlige investeringer

Eskild Tuft Thorgeirsson og Mathias Woldseth Thorstensen

Veileder: Øystein Gjerde

Masteroppgave, økonomi og administrasjon, finansiell økonomi

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer innestår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Sammendrag

Tema for denne oppgaven er om ansvarlige investeringer kan forsvares fra et økonomisk perspektiv, eller kun med bakgrunn i etiske hensyn. Dette undersøkes ved å analysere effekten ansvarlige investeringer har på avkastning. For å identifisere ansvarlige selskaper benytter vi ESG-score og kontroversscore. ESG-score er en score utregnet på bakgrunn av informasjon om selskapets håndtering av miljø, sosiale forhold og selskapsstyring. En kontroversscore er utregnet ut ifra antall negative hendelser knyttet til ESG som blir tatt opp i media.

Vi benytter to ulike analytiske metoder for å studere forholdet mellom ansvarlige investeringer og avkastning. Først benytter vi klassiske faktormodeller til å kontrollere ulike portefølgers avkastning mot kjente risikofaktorer. Deretter utfører vi en paneldatanalyse, hvor vi ser på sammenhengen mellom ansvarlighet og avkastning på selskapsnivå.

Hovedsakelig viser resultatene ingen sammenheng mellom ansvarlighet og avkastning. Vi ser noen tegn til at kontroversscore tidligere hadde en negativ effekt på avkastning. Ingen resultater indikerer at denne sammenhengen fortsatt eksisterer. Vi ser også noen tegn til at selskapene med lavest ESG-score, har lavere risikojustert avkastning. Ingen av de signifikante funnene er entydige på tvers av de analytiske metodene. Vi kan derfor ikke konkludere hverken med at ansvarlige forhold har eller ikke har en sammenheng med avkastning.

Forord

Denne masterutredningen er et selvstendig arbeid utført i forbindelse med gjennomføringen av masterstudiet i økonomi og administrasjon ved Norges Handelshøyskole (NHH). Oppgaven ble gjennomført høsten 2018.

Oppgaven er en kvantitativ studie av ansvarlige investeringer. Motivasjonen bak arbeidet var å gjøre en kvantitativ studie i praksis og lære mer om dette. Vi anså ansvarlige investeringer som et spennende og aktuelt tema som kunne anvendes til dette. I denne sammenheng ble vi nysgjerrige på om mål på ansvarlighet som er mer synlige for folk flest, har større innvirkning på avkastning, enn klassiske mål på faktisk underliggende ansvarlighet.

Vi vil gjerne takke veileder Øystein Gjerde for gode innspill, og for å ha tilpasset seg endringer i oppgaven underveis. Vi vil også takke Fredrik Andersen Kavli for hjelp i forbindelse med uthenting av datamaterialet.

Innholdsfortegnelse

SAMMENDRAG	2
FORORD	3
INNHOLDSFORTEGNELSE	4
1. INTRODUKSJON	6
1.1 BAKGRUNN	6
1.2 MOTIVASJON	7
1.3 OPPGAVENS OPPBYGNING	8
2. ANSVARLIGE INVESTERINGER	9
2.1 IMPLEMENTERING AV ANSVARLIGE INVESTERINGER	9
2.2 INTUISJON FOR EFFEKT PÅ AVKASTNING.....	10
3. TIDLIGERE STUDIER PÅ ANSVARLIGE INVESTERINGER	14
3.1 ESG	14
3.2 KONTROVERSER	16
4. TEORI	18
4.1 GRUNNLEGGENDE ANTAKELSE – MARKEDSEFFISIENS	18
4.2 PORTEFØLJETEORI.....	19
4.3 RISIKOFAKTORER	21
5. DATASETT	24
5.1 INNDATA	24
5.2 FINANSIELLE MÅLTALL	24
5.3 MÅLTALL PÅ ANSVARLIGHET	26
5.4 DATAVASKING	29
6. METODE	30
6.1 MOTIVASJON FOR VALG AV METODE.....	30

6.2	PORTEFØLJER	30
6.3	SELSKAPER	32
6.4	MODELLENES FORUTSETNINGER	35
7.	EMPIRISK ANALYSE.....	37
7.1	PORTEFØLJENIVÅ	37
7.2	INDIVIDUELT NIVÅ	52
7.3	DISKUSJON AV RESULTATER	59
7.4	SVAKHETER	62
8.	KONKLUSJON.....	63
9.	LITTERATURLISTE.....	65
10.	APPENDIKS.....	71
10.1	RESIDUALPLOT – PORTEFØLJER	71
10.2	RESIDUALPLOT – SELSKAPER	77
10.3	LINEARITETSPLOT – PORTEFØLJER	79
10.4	LINEARITETSPLOT – SELSKAPER	83
10.5	UTVALGTE REGRESJONER	84

1. Introduksjon

I dette kapitlet ser vi først på dagens situasjon knyttet til ansvarlige investeringer, og hvor utbredt dette er. Vi introduserer deretter begrepene ESG-score og kontroversscore, som er begreper som gjerne brukes i forbindelse med hvor ansvarlig et selskap er. Når begrepene er introdusert utledes motivasjonen bak og formålet med denne masteroppgaven. Avslutningsvis presenteres oppgavens struktur og oppbygning.

1.1 Bakgrunn

Ansvarlige investeringer har røtter tilbake til det som kalles "etiske investeringer", som handler om å unngå investeringer som strider mot moralske eller etiske verdier. Fagfeltet ansvarlige investeringer er i dag bredere og inneholder også miljømessige og sosiale forhold som kan påvirke den finansielle verdien av investeringene på sikt. I denne oppgaven bruker vi to ulike mål på ansvarlighet. Det første er ESG-score, som er et mål på hvor ansvarlig et selskap er. ESG står for "Environmental, Social and Corporate Governance". En ESG-score er altså en score utregnet med bakgrunn i et selskaps håndtering av miljø, sosiale forhold og selskapsstyring (Norsif, 2016). Det andre målet er kontroversscore, som er en score som utregnes basert på antall kontroverser knyttet til ESG et selskap har vært involvert i. Med slike kontroverser menes oppmerksomhet i media, og påfølgende oppmerksomhet blant interessenter, som følge av manglende eller sviktende ESG-arbeid (Aouadi og Marsat, 2016). Eksempler kan være avsløringer om bruk av barnearbeid eller korrupsjon (Thomson Reuters, 2018). En kontroversscore kan muligens gi et bedre bilde av hvordan selskaper fremstår for interessenter enn ESG-score, som i større grad ser på summen av de underliggende forholdene.

ESG-faktorer danner i dag grunnlaget for FNs prinsipper for ansvarlige investeringer (Principles for Responsible Investment, 2018). Så langt har over 2000 organisasjoner undertegnet på at de skal følge disse prinsippene. Disse organisasjonene forvalter tilsammen verdier på omtrent 89 billioner USD. Når så mange ønsker å investere ansvarlig er det interessant å se nærmere på hvorfor de ønsker å gjøre dette. Tidligere studier har funnet flere årsaker til at investorer ønsker å investere ansvarlig. Blant disse kan det i hovedsak trekkes frem tre (UBS, 2015). Den første årsaken er et ønske om å tilpasse sine investeringer til sine personlige verdier, slik at man kan bevare sin gode samvittighet. Den andre årsaken er et ønske

om å gjøre en forskjell ved å investere i selskaper en selv mener vil bidra til å gjøre verden til et bedre sted. Den tredje, og kanskje mest kyniske årsaken, er et ønske om å oppnå høyere avkastning.

At høyere avkastning er en vanlig årsak til at investorer ønsker å investere ansvarlig, gjør det interessant å undersøke om det faktisk er en slik sammenheng. Dette skal vi se nærmere på i denne masteroppgaven. Påvirker ansvarlighet avkastning, eller er dette en feiloppfatning som gjør at ansvarlige investeringer kun kan begrunnes med etiske hensyn?

1.2 Motivasjon

Å undersøke hvordan ansvarlige investeringer påvirker avkastning og risiko er på ingen måter noe nytt. I litteraturgjennomgangen (avsnitt 3.1) vises det til flere metastudier som omfatter flere tusen studier av hvordan ansvarlige investeringer påvirker avkastning og risiko. Selv med så mange studier, er det likevel ingen klar konsensus om den faktiske effekten. Denne oppgaven henter noen elementer fra tidligere studier, i tillegg til at vi legger til nye elementer for å forsøke å få ny innsikt.

Tidligere studier av ansvarlige investeringer har hatt mye fokus på effekten ESG har på avkastning, og lite oppmerksomhet har blitt viet kontroverser (ref. avsnitt 3.2). Aouadi og Marsat (2016) argumenterer for at ESG og kontroverser relatert til ESG, ikke er direkte motsetninger. Vi mener derfor at å inkludere kontroverser i en analyse av ansvarlige investeringer kan bidra til ny informasjon. Kanskje kan kontroversscore være en indikator på forhold som er mer egnet til å belyse effekten på avkastning. Vi ønsker derfor å undersøke nærmere hvilken effekt ESG og kontroverser har på avkastning, og om dagens økende fokus på ansvarlige investeringer kan forsvares ut ifra et økonomisk perspektiv.

Vårt bidrag blir å benytte et oppdatert datasett som strekker seg helt frem til 2018, ta i bruk moderne prisingsmodeller og ta hensyn til flere forhold relatert til det å investere ansvarlig. Denne oppgaven begrenser seg til europeiske selskaper i perioden 2004 til 2018. Vi inkluderer ikke observasjoner før dette, fordi før 2004 reduseres antall selskaper som har informasjon om ansvarlige forhold drastisk. Vi begrenser oss til Europa fordi tidligere studier har vist at det er geografiske forskjeller når det kommer til effekten ansvarlige investeringer har på avkastning

(Friede, Busch og Bassen, 2015). Oppgaven kommer til å fokusere på nivået av ESG- og kontroversscorer. Det kunne vært interessant å også undersøke hvordan en endring i score påvirker avkastningen, men det blir utenfor denne oppgavens rammer.

Problemstillingen som oppgaven tar for seg er:

Hvilken effekt har hensyn til ESG og kontroverser knyttet til ESG på aksjeavkastning i Europa?

1.3 Oppgavens oppbygning

Oppgaven består av i alt 10 kapitler hvorav dette er første og innledningen. Kapittel 2 ser på hvordan ansvarlige investeringer kan implementeres som investeringsstrategi, og rasjonale bak hvordan ESG og kontroverser kan tenkes å påvirke avkastning. Kapittel 3 tar for seg tidligere publiserte studier innen feltet ansvarlige investeringer, hvordan disse er gjennomført og hva de kommer frem til. I kapitlet går vi inn på funn relatert til både ESG-score og kontroversscore. I kapittel 4 presenteres relevant teori, både generell porteføljeteori og faktormodeller. I kapittel 5 presenteres datamaterialet, og det forklares hvordan dette er innhentet og satt sammen. Kapittel 6 er metodedelene. Her beskrives det hvordan analysen gjennomføres rent metodisk, og motivasjonen bak metodevalgene. I kapittel 7 presenteres og drøftes resultater fra analysen. Deretter diskuteres implikasjoner av funnene. Kapittel 8 inneholder konklusjonen, og det pekes også her på hva som kan være interessant å undersøke videre. Kapittel 9 og 10 inneholder henholdsvis kilder og appendiks.

2. Ansvarlige investeringer

I dette kapitlet forklarer vi hvordan ansvarlige investeringer kan implementeres som investeringsstrategi. Deretter diskuterer vi hva som kan være intuitjonen for at ESG-score og kontroversscore kan henge sammen med avkastning.

2.1 Implementering av ansvarlige investeringer

Det finnes flere måter å implementere ansvarlige investeringer. Hovedsakelig deles det inn i tre ulike metoder: effektinvestering, eksklusjon og integrering. Av disse metodene er eksklusjon mest utbredt, deretter kommer integrering, og til slutt kommer effektinvestering (UBS, 2015).

Effektinvestering betyr at man gjør spesifikke investeringer med mål om å løse spesifikke samfunns- og miljøproblemer (US SIF, 2016). Eksklusjon baserer seg på at man unngår å investere i selskaper som deltar i en gitt aktivitet. Oljefondet benytter blant annet denne metoden gjennom sin utelukkelse av selskaper som bryter visse krav (Norges Bank Investment Management, 2018). Integrering går ut på å kombinere informasjon om ESG med tradisjonell finansiell informasjon. Dette er den mest avanserte formen for ansvarlige investeringer, og krever spesifikk informasjon. Heldigvis har ESG-informasjon blitt mer tilgjengelig de siste årene. Selskaper rapporterer stadig mer informasjon, og informasjonen har blitt mer standardisert som følge av ulike standardiseringsinitiativer, som for eksempel Sustainability Accounting Standards Board (SASB) og Global Reporting Initiative (GRI). Det finnes også flere kommersielle leverandører av ESG-informasjon. Den enkleste formen for ESG-integrering er screening på bakgrunn av ESG-score. Denne metoden er i utgangspunktet veldig lik vanlig eksklusjon, men man benytter seg av en ESG-score. Dette gjøres ved å sette et minimumskrav til ESG-score for at selskaper skal være en del av investeringsuniverset. Enten setter man en fast grense for alle selskaper, eller så kan man sette individuelle grenser for ulike sektorer. Å sette én grense er enklest, men det kan føre til at en portefølje blir vridd mot enkelte sektorer og bort fra andre. Ved å sette ulike grenser vil man fortsatt inkludere selskaper som er relativt gode på ESG i en generelt dårlig sektor, mens man utelukker dem som er relativt dårlige i en generelt god sektor. Den mest avanserte måten å integrere ESG på, er å implementere ESG-score som et vurderingspunkt i en analyse av ulike investeringsmuligheter,

på lik linje med andre finansielle måltall. Da er det også mulig å benytte ESG-score som er bransjevektet. På samme måte kan man også integrere en kontroversscore i analysen. Enkelte måltall som benyttes i en ESG-score, er trolig allerede indirekte tatt hensyn til i en tradisjonell investeringsanalyse. Ved spesifikt å integrere ESG-score i analysen, vil man imidlertid sikre at den fulle effekten av ESG tas hensyn til (US SIF, 2016).

2.2 Intuisjon for effekt på avkastning

Bruk av ESG i forbindelse med ansvarlige investeringer handler altså om å vurdere selskaper ut fra en rekke kriterier knyttet til miljø, sosiale forhold og selskapsstyring. Det er en betydelig mengde litteratur som påstår at å ta hensyn til ESG kan gi meravkastning. Intuisjonen bak denne meravkastningen varierer mye, og det er også forskjellige oppfatninger om hvilken vei intuisjonen peker. De fleste argumenterer for at ved å benytte ESG-score som en positiv indikator, vil man kunne oppnå meravkastning (Friede, Busch og Bassen, 2015). Noen er derimot av motsatt oppfatning, og mener at man vil kunne oppnå meravkastning ved å unngå selskaper med høy ESG-score (Zembrowski, 2018). Vi skal nå se nærmere på hvordan det kan argumenteres for at ESG og kontroverser kan påvirke avkastning.

2.2.1 ESG

Giese et al. (2017) sammenfatter funnene i artiklene til Eccles, Ioannou og Serafeim (2014), El Ghoul et al. (2011) og Gregory, Tharyan og Whittaker (2014), om hvordan ESG-faktorer påvirker den systematiske risikoen til selskaper. Totalt sett hevder de at høy ESG-score leder til høyere verdsettelse. Et selskap som er høyt priset vil ha lavere avkastning enn et som er lavt priset, alt annet like. Det økonomiske rasjonale det vises til, er at høy ESG-score gjør selskaper mindre utsatt for systematisk risiko. Dersom selskaper med høy ESG-score har lavere systematisk risiko, vil det føre til et lavere avkastningskrav, som gir lavere kapitalkostnad for disse selskapene. Lavere kapitalkostnad for et selskap med høy ESG-score, gjør at selskapet har høyere verdsettelse enn et tilsvarende selskap med dårligere ESG-score, alt annet likt. El Ghoul et al. (2011) identifiserer flere årsaker til at selskaper med høy ESG-score kan ha lavere systematisk risiko. De argumenterer blant annet for at disse selskapene er mindre eksponert for ulike typer markedssjokk. For eksempel vil selskaper med effektivt energiforbruk, være mindre eksponert mot svingninger i energiprisene.

Hvis ESG er en risikofaktor, og den allerede er priset inn, så vil den ikke kunne gi risikojustert meravkastning. Hvis markedet derimot ikke er effisient, så kan det ta tid før markedet er priset riktig. I denne perioden før markedet er riktig priset, kan man oppnå risikojustert meravkastning ved å tilte en portefølje mot høy ESG-score (Zhang, 2017). Et annet forhold som kan påvirke den risikojusterede avkastningen, er en potensiell samvittighetspremie. Hvis enkelte investorer verdsetter at selskaper er ansvarlige, kan de bidra til å presse prisen overdrevent opp. Dette kan føre til at selskaper med høy ESG-score blir kunstig høyt priset. I perioden hvor ansvarlige investorer presser prisen opp, kan andre investorer som allerede holder aktivumet oppnå risikojustert meravkastning. Når selskapet først er kunstig høyt priset, vil investorer ha lavere risikojustert avkastning. Ettersom det blir vanligere at selskaper er ansvarlige, vil en samvittighetspremie trolig gradvis forsvinne. Da vil prisen til de selskapene som tidligere skilte seg ut ved å være ansvarlige, igjen gå mot fundamentalverdien. Hvis en slik effekt eksisterer, blir spørsmålet hvor i prosessen markedet og ulike selskaper er i dag. Denne intuisjonen er imidlertid kompleks, da den forutsetter at ikke-ansvarlige investorer ikke holder prisene nede ved å selge seg ut, når ansvarlige investorer presser prisene opp (Revelli og Viviani, 2015).

Det er imidlertid ikke kun den systematiske risikoen som påvirker en porteføljes totale risiko. Hvis man velger å utelukke selskaper som har lav ESG-score, vil dette føre til redusert investeringsunivers og dermed også reduserte diversifiseringsmuligheter (Zhang, 2017). Dette kan føre til at den totale risikoen man er eksponert for blir større, noe som vil gi lavere risikojustert avkastning. Giese et al. (2017) så også på hvordan ESG-score henger sammen med usystematisk risiko. De tar utgangspunkt i resonnetet til Godfrey et al. (2009), Jo og Na (2012) og Oikonomou, Brooks og Pavelin (2012). Dette resonnetet, for hvordan selskaper med høy ESG-score kan ha redusert usystematisk risiko, baserer seg på at slike selskaper kjennetegnes av å ha over gjennomsnittlig god risikostyring. Dette gjelder både for selskapet spesifikt og for leverandørkjeden. God risikostyring medfører videre at selskaper med høy ESG-score, sjeldnere har problemer med alvorlige hendelser som korrupsjon og bestikkelser. Færre slike alvorlige hendelser resulterer videre i at slike selskaper har lavere risiko for å møte uforutsette tap og problemer. Giese et al. (2017) fant ytterligere empiriske bebreftelser for disse antakelsene ved å se på volatilitet og antall konkurser for selskaper med høy ESG-score relativt til selskaper med lav ESG-score. Selv om man i teorien ikke kompenseres for usystematisk risiko, fordi dette kan diversifisere bort, er det nok i praksis

ikke like lett å gjøre dette for alle investorer. Hvis en investor ikke benytter markedsporteføljen, eller lignende tilnærmet totaldiversifisert benchmark, kan det være en fordel å tilte en portefølje mot selskaper med høy ESG, om dette reduserer risikoen for negative ESG-relaterte hendelser. Da kan de muligens redusere porteføljens totale risiko, og oppnå høyere risikojustert avkastning.

Gregory et al. (2014) så på hvordan ulik ESG-score kan påvirke kontantstrømmen til selskaper, både på lang og kort sikt. De finner at selskaper som klassifiseres som "grønne" (selskaper med styrker innenfor ESG), har høyere langsiktig vekstpotensiale enn selskaper med lav ESG-score. Dette er konsistent med deres hypotese om at grønne selskaper har forventet lengre vedvarende konkurransefortrinn, og presterer bedre enn selskaper med lav ESG-score. På sikt vil dette ifølge forfatterne, føre til høyere utbytte i selskaper med høy ESG-score. På den annen side argumenteres det for at å øke og opprettholde ESG-score medføre direkte kostnader (Waddock og Graves, 1997). Høyere kostnader fører til redusert profitt. Dette var også oppfatningen til Friedman (2007), som argumenterer for at investeringstiltak for å bedre ESG er verdiødeleggende for eierne. Her må vi imidlertid være oppmerksomme på at om et selskap tjener godt eller dårlig ikke er det som i hovedsak avgjør om det er en "god" investering, og om man kan forvente høy avkastning. Det kommer an på prisingen av selskapet relativt til risikoen og den underliggende fundamentalverdien.

Til syvende og sist er det også mulig at ESG ikke har noen sammenheng med avkastning. Et vanlig argument for dette, er at det er så mange andre variabler som kommer mellom ESG og finansielle resultater, at det er urealistisk å forvente noen sammenheng (Waddock og Graves, 1997). Dessuten kan det hende at fordelene og ulempene utjevner hverandre. Å innhente og behandle informasjon om ESG er kostbart. Det er for eksempel mulig at den potensielt lavere risikoen man kan oppnå ved å integrere ansvarlige investeringer, blir utjevnet av de økte kostnadene investor påtar seg ved å gjøre dette, gitt at disse kostnadene ikke er priset.

2.2.2 Kontroverser

Da vi diskuterte den usystematiske risikoen til ESG var vi inne på uforutsette hendelser som kan føre til økonomiske tap og fall i verdi. Et eksempel på en slik hendelse kan være kontroverser knyttet til ESG. Ikke bare kan slike kontroverser være dyre i seg selv, som følge

av for eksempel erstatninger, men de kan også føre til redusert tillit til selskapet og dermed forventninger om ytterligere problemer eller flere kontroverser (Parnassus Investments, 2017).

Hvis selskaper faller i verdi som følge av kontroverser, åpner dette for to potensielle måter en kontroversscore kan benyttes til å oppnå meravkastning. En mulighet er at man kan være i stand til å gjøre et godt kjøp. En kontrovers kan tenkes å presse prisen overdrevent ned, slik at den blir lavere enn selskapets fundamentalverdi. Hvis man kjøper selskaper som har vært involvert i kontroverser, kan man tjene godt på dette hvis forventningene til selskapet i ettertid normaliseres og prisen igjen beveger seg mot fundamentalverdien. Å være involvert i en kontrovers, kan også presse en bedrift til å gjøre forbedringer innen ESG, og samtidig gjøre investorer mer oppmerksomme på disse endringene. Hvis man anser ESG som en systematisk risikofaktor, og forutsetter at markedet er effisient, kan man da utnytte en potensiell sammenheng mellom kontroverser og kommende forbedringer innen ESG. Den reduserte risikoen vil presse prisen opp, og gi avkastning når markedet tilpasser seg. Siden en ESG-score gjerne publiseres med en forsinkelse, kan kontroverser muligens fungere som en ledende indikator på kommende forbedringer innen ESG. Det er altså flere argumenter for at man kan bruke informasjon om kontroverser til å utnytte feilprising og prisjusteringer i markedet.

En annen mulighet for å utnytte informasjon om at et selskap har vært involvert i en kontrovers, baserer seg på at forventningene om ytterligere kontroverser stemmer. Det kan tenkes at økt oppmerksomhet fra media og myndigheter i kjølvannet av en kontrovers, øker sannsynligheten for flere kontroverser, som vil redusere selskapsverdien ytterligere. En annen mulighet er at forhold i enkelte selskaper gjør disse systematisk mer utsatt for kontroverser, og at en kontroversscore kan brukes som en indikator på slik systematisk risiko. Det kan også tenkes at en kontroversscore bedre fanger opp de ESG-forholdene som faktisk har systematisk effekt på risiko.

3. Tidligere studier på ansvarlige investeringer

Vi har allerede vært innom hvordan ESG og kontroverser kan ha effekt på avkastning. Dette kapittelet gjennomgår den eksisterende litteraturen på den faktiske effekten av ESG og relaterte måltall. Først tar vi for oss studier som ser på ESG, deretter ser vi nærmere på studier som ser på kontroverser knyttet til ESG.

3.1 ESG

Det finnes flere tusen studier som tar for seg sammenhengen mellom et selskaps finansielle prestasjoner og ansvarlighet. Vi mener derfor at det blir lite representativt å referere til enkeltstudier. Derfor har vi valgt å fokusere på litteraturstudier som gir en oppsummering av allerede eksisterende forskning.

Revelli og Viviani publiserte i 2015 en metaanalyse hvor de undersøkte 85 tidligere studier og 190 internasjonale eksperimenter knyttet til ansvarlige investeringer de siste 20 årene. De konkluderte med at det globalt ikke er holdepunkt for finansielle fordeler eller ulemper forbundet med ansvarlige investeringer. Videre hevdet de at ulike studiers resultater avhenger av studienes metodevalg og data.

Zhang (2017) publiserte en litteraturstudie med en anbefaling til det tyske pensjonsfondet. Hun hevdet at ansvarlige investeringer ikke hemmer tradisjonelle investeringer, og at ansvarlige investeringer til og med kan generere meravkastning gjennom å utnytte feilprising av ESG-faktorer. Hun hevdet at porteføljer med høy ESG-score har prestert bedre enn deres referanseporteføljer. Dette gjelder spesielt for faktorer innen "corporate governance", "eco-efficiency" og "employee relation". På bedriftsnivå pekte hun på studier som viser at selskaper med høy ESG-score, har sterkere finansielle prestasjoner og lavere finansieringskostnad. Likevel hevdet hun at denne overprestasjonen allerede kan være priset inn i markedet som følge av økende fokus på ESG.

Margolis, Elfenbein og Walsh (2009) undersøkte 214 studier fra 1972 til 2007. De så på "Corporate Social Performance" (CSP), som inneholder mange av de samme faktorene som ESG. De kom frem til at påvirkningen av CSP på regnskapsmessige måltall som kapitalavkastning (ROA) og egenkapitalavkastning (ROE), er større enn den markedsmessige

effekten. Videre viste studiene at høy CSP ikke skader finansiell prestasjon, men at det heller ikke gir en sterk positiv effekt. Deres metaanalyse var imidlertid påvirket av at et lite antall av studiene de undersøkte, med et lite utvalg av selskaper, stod for mye av effekten. De påpeker også at metaanalyser generelt kan være utsatt for en viss bias, fordi studier med signifikante resultater har større sannsynlighet for å bli publisert. Dette kan føre til at metaanalyser potensielt baserer seg på et skjevt utvalg, og derfor kan få kunstig høye signifikansnivåer. De forsøkte å løse dette ved å inkludere så mange upubliserte rapporter som mulig, samtidig som de gjennomførte sensitivitetsanalyser som så på hvor mange upubliserte artikler som må eksistere for å påvirke resultatet.

Friede, Busch og Bassen (2015) undersøkte over 2200 unike studier. De fant ut at omtrent 10 % av de totale studiene fant negativ sammenheng mellom ESG og finansielle prestasjoner, mens det store flertallet av studiene fant en positiv eller ikke-signifikant sammenheng. Når de vurderte finansielle prestasjoner gjorde de ikke forskjell på om de underliggende studiene benyttet regnskaps- eller markedsbaserte måltall. De positive resultatene var særlig tydelige i Nord-Amerika (42,7 %) og fremvoksende markeder (65,4 %). Denne sammenhengen synes å ha vært stabil siden 90-tallet. I utviklede markeder i Europa viste kun 26,1 % av studiene positiv sammenheng, mens 8 % viste negativ sammenheng. De hevdet imidlertid at data-grunnet i Europa kunne ha blitt påvirket av et stort antall porteføljestudier. De påpekte nemlig at porteføljestudier gjerne får nøytrale eller blandede resultater som følge av inkluderte forvaltningskostnader, sammensetningsrestriksjoner og at porteføljene er utsatt for overlappende effekter av systematiske og usystematiske risikofaktorer. Revelli og Viviani (2015) påpekte også at ESG-screening medfører økte kostnader. Disse forholdene forvrenger effekten av ESG i porteføljestudier. Dersom eventuell gevinst av ansvarlige investeringer er liten, vil investor ha problemer med å beholde denne etter implementeringskostnader.

Enkelte studier har også vist at såkalte momentumstrategier, hvor man investerer i selskaper med positiv endring i ESG-score, presterer bedre enn strategier basert på å investere i selskaper med høyt ESG-nivå (NN investment partners, 2016). Slike momentumstudier kan gjøre at litteraturstudiene finner høyere grad av positive sammenhenger mellom ESG og avkastning, enn om de kun hadde sett på nivåstudier.

3.2 Kontroverser

En ESG-score består av mange ulike måltall, og kontroverser knyttet til ESG blir ofte inkludert (Thomson Reuters, 2017). Å trekke kontroverser ut og å se på de gjennom en egen score er en nyere metode. Thompson Reuters begynte for eksempel først med dette etter at de i 2017 erstattet sitt tidligere scoringssystem kalt ASSET4. Vår litteraturgjennomgang tilsier at det følgelig er skrevet lite om bruken av kontroverser som egen score eller eget vurderingskriterium på lik linje med vanlig ESG-score. Hovedtrekkene i det som er kommet frem i publiserte studier oppsummeres her.

Aouadi og Marsat (2016) argumenterer for at ESG og kontroverser knytte til ESG ikke er direkte motsetninger, og studerte derfor effekten av kontroverser i en studie av over 4000 selskaper fra 58 land. Studien er basert på regresjon med Tobins Q som avhengig variabel, og blant annet kontrovers- og ESG-score som uavhengige variabler. De finner at å være involvert i flere kontroverser er forbundet med en signifikant meravkastning når man ser på "high-attention" selskaper. Disse selskapene kjennetegnes av at de er større, gjør det bedre, får mer oppmerksomhet fra investorer og er lokalisert i land med høyere grad av pressefrihet. Dette indikerer en positiv sammenheng mellom kontroverser publisert om et selskap i media og Tobins Q. Denne positive effekten er kanskje lite intuitiv. Det er imidlertid slik at den direkte effekten av kontroversene ikke er signifikant når de ser på ESG-score og kontroverser sammen. Tross den negative omtalen i media, hevder forfatterne at synligheten til selskapet vil øke, og dermed også synligheten av deres ESG-score. Dette hevder de vil forsterke effekten ESG-scoren har på selskapets verdi. For "low-attention" selskaper er ikke effekten av kontroverser signifikant, hverken direkte eller i samspill med ESG-score.

En annen interessant studie ble gjennomført av NN Investment Partners og European Centre for Corporate Engagement ved Universitetet i Maastricht i 2016. De viste at dersom man ekskluderer selskaper som har vært involvert i kontroverser, kan man oppnå en høyere Sharpe-rate, altså høyere risikojustert avkastning (NN investment partners, 2016).

En studie gjennomført av Blackrock (2018) kom frem til at de selskapene som rapporterte å ha flest retningslinjer knyttet til ESG, også var innblandet i flest kontroverser. De som rapporterte relativt få ESG-policyer var utsatt for få kontroverser. Fremtidige kontroverser var også positivt korrelert med størrelse på selskapet og hvor mange eksisterende kontroverser de

allerede var innblandet i. Kontroverser fører også gjerne til implementering av ytterligere retningslinjer knyttet til ESG, men som regel ingen merkbar endring av praksis. I en tidligere studie på vegne av Blackrock, argumenterer Garvey et al. (2016) for at kontroverser har signifikant negativ effekt på avkastning. Denne effekten blir både større og mer signifikant når man kontrollerer for størrelse, bransje og land. De påstår derfor at kontroverser kan brukes som et instrument for omdømme- og reguleringsrisiko.

4. Teori

Dette kapittelet er tredelt. Første del beskriver grunnleggende antakelser om hvordan markedet oppfører seg. Andre del omhandler teorien som ligger bak ulike analytiske metoder, og forutsetninger for disse. Tredje del presenterer ulike risikofaktorer som brukes i oppgavens analyse.

4.1 Grunnleggende antakelse – Markedseffisiens

Konseptet markedseffisiens kan spores tilbake til starten av 1900-tallet. Bachellier (1900) hevdet i sin doktorgradsavhandling at både tidligere, nåværende og fremtidige hendelser reflekteres i markedsprisen til en aksje. Likevel, fordi det er så ekstremt mange faktorer som påvirker børsen, er ikke sammenhengen alltid så lett å se. Dette arbeidet ble i mange år oversett. På midten av 1900-tallet undersøkte Kendall og Hill (1953) britiske aksjer og kom frem til at det var så mye tilfeldigheter i aksjekursenes bevegelser at enhver systematisk sammenheng ble neglisjerbar. Disse resultatene ble til teorien om at aksjemarkedet beveget seg med en såkalt tilfeldig gange. Samuelson (1965) utdypet teorien og fremmet ideen om at et velfungerende marked nødvendigvis vil måtte bevege seg med tilfeldig gange. Han mente at hvis en fremtidig prisendring kunne forutsees, så ville den allerede ha skjedd som følge av at folk utnyttet denne kunnskapen.

Et ideelt marked er et marked hvor prisene gir gode signaler for allokering av ressurser. Altså et marked hvor prisene fullt ut reflekterer all tilgjengelig informasjon (Malkiel og Fama, 1970). Et slikt marked kalles effisient. Malkiel og Fama (1970) mente at en teori basert på at all informasjon gjenspeiles i prisene blir for ekstrem. I forbindelse med testing av teorien delte de effisiens inn i tre former etter hvilken informasjon som gjenspeiles. Dette skulle gjøre det lettere å se hvilket informasjonsnivå som gjør at teorien svikter. Disse tre formene er svak-form, semisterk-form og sterk-form.

Sterk-form effisiens: Markedspriser gjenspeiler all informasjon, både offentlig og privat

Semisterk-form effisiens: Markedspriser gjenspeiler all offentlig informasjon

Svak-form effisiens: Markedspriser gjenspeiler all historisk informasjon

Malkiel og Fama (1970) viste at teorien holder for svak-form og semisterk-form effisiens. De fant bare begrenset bevis for sterk-form effisiens, og konkluderte med at monopolistisk tilgang til informasjon var uvanlig. De få tilfellene hvor monopolistisk tilgang til informasjon kunne bevises, var det snakk om spesialister og personer med koblinger til selskapene som satt på denne informasjonen. Malkiel og Fama (1970) mente derfor at teorien om markedseffisiens var et godt utgangspunkt for en approksimasjon av virkeligheten.

I dag er det få som tror markedet er helt effisient. Blant annet har Grossman og Stiglitz (1980) fremmet en teori om at dette ikke kan være tilfellet. De baserte dette på at det er kostbart å innhente informasjon. Om all informasjon allerede er tatt hensyn til i prisen, vil det ikke være ønskelig for noen å innhente mer informasjon. Men hvis ingen innhenter informasjon, er det ikke mulig at denne informasjonen kan være tatt hensyn til i prisen. Dette paradokset viser at markedet ikke kan være helt effisient. Det vil alltid være lommer av ineffisiens. Aktive investorer vil søke etter disse lommene for å oppnå meravkastning. Denne prosessen vil i seg selv bidra til å gjøre markedet mer effisient.

4.2 Porteføljeteori

Porteføljeteori handler om hvordan man kan konstruere en optimal portefølje som maksimerer forventet avkastning for en gitt risiko. Teorien, som ble introdusert av Markowitz (1952), illustrerer hvordan det er mulig å konstruere en effektiv porteføljefront som gir maksimal avkastning for ulike risikonivåer for en gitt beholdning av aksjer. En forutsetning teorien bygger på er at investorer er risikoaverse. Å være risikoavers vil si at hvis man kan velge mellom to alternativer med lik forventet avkastning, så vil man velge det alternativet med lavest risiko (Bodie, Kane og Marcus, 2011). Selv om alle investorer er risikoaverse, vil de likevel ha forskjellige tilpasning til risiko med bakgrunn i individuelle preferanser og grad av risikoaversjon. For å påta seg mer risiko forutsetter teorien at man da krever kompensasjon i form av økt forventet avkastning.

4.2.1 Risiko

For å kunne gjøre optimale porteføljebeslutninger må man skille mellom to former for risiko: systematisk- og usystematisk risiko. Usystematisk risiko består av selskapsspesifikke faktorer

og kan elimineres ved å spre investeringene. Siden usystematisk risiko kan diversifiseres bort, kompenseres man ikke for denne. Systematisk risiko påvirker hele markedet, ikke bare et spesielt selskap eller industri. Denne formen for risiko kan ikke diversifiseres bort. Derfor er det denne man kompenseres for i form av høyere forventet avkastning. Det er derfor optimalt å spre investeringer i mange forskjellige selskaper og sektorer, slik at denne risikoen forsvinner i størst mulig grad. Jo mer av den usystematiske risikoen man klarer å diversifisere bort, jo større andel av den totale risikoen man utsettes for, vil man bli kompensert for (Ang, 2014).

4.2.2 CAPM

Kapitalverdimodellen (CAPM) ble utviklet med utgangspunkt i arbeidet til Markowitz (1952), gjennom artikler av Sharpe (1964), Lintner (1965) og Mossin (1966). Modellen beskriver forholdet mellom systematisk risiko og forventet avkastning for eiendeler. Modellen bygger på prinsippet om at investorer må kompenseres for to ting når de kjøper aksjer: tidsverdien av penger og risikoen de påtar seg. Risikofri rente representerer en tidskompensasjon. I praksis brukes gjerne avkastning på statsobligasjoner, da dette er det nærmeste man kan komme en risikofri plassering. Kompensasjon for systematisk risiko er gitt ved faktoreksponering (beta) multiplisert med markedspremien. Markedspremien er den meravkastningen markedet i snitt oppnår utover en risikofri plassering. Hvis man investerer i en aksje som er utsatt for høy systematisk risiko (høy beta), vil man kreve kompensasjon i form av høyere forventet avkastning. Hvis aksjen ikke kompenserer for den høye risikoen med høyere forventet avkastning, vil investorer heller investere i et annet aktivum med samme forventede avkastning, men lavere risiko. Kapitalverdimodellen, som kan sees i ligning 1, kan benyttes til å vise sammenhengen mellom forventet avkastning og ulik eksponering mot markedet (beta). I ligningen er $E(r_i)$ forventet avkastning for aktivumet, r_f er risikofri rente, β_i er eksponering mot den forventede markedspremien, $E(r_m)$ er markedets forventede avkastning.

$$E(r_i) = r_f + \beta_i(E(r_m) - r_f) \quad (1)$$

4.2.3 Jensens alfa

Som en videreutvikling av CAPM, introduserte Jensen (1968) et risikjustert mål på hvor mye forvalters evne til å investere i markedet, bidrar til avkastning. Jensen så på muligheten for å estimere systematisk risiko ved å benytte beta fra CAPM. Han så imidlertid et problem med dette hvis en portefølje ble forvaltet av en "flink" forvalter. CAPM inneholder nemlig i teorien et feilledd som har forventning lik null. Jensen mente at en "flink" forvalter systematisk kan oppnå et positivt feilledd. Beregningen av beta vil da bli unøyaktig. Jensen løste dette ved å inkludere en konstant som ikke måtte være lik null. Denne konstanten kalles Jensens Alfa (α), eller bare alfa. Alfa vil fange opp den systematiske avkastningen som ikke forklares av markedsavkastningen. Feilleddet vil da få forventning lik null. Avkastningen kan da uttrykkes som i ligning 2. Her er alfa for aktivum i gitt ved α_i . Feilleddet med forventning null for aktivum i på tidspunkt t er gitt ved $u_{i,t}$.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + u_{i,t} \quad (2)$$

En strategi basert på å kjøpe og holde markedet vil resultere i alfa lik null. En forvalter som klarer å velge aksjer som i snitt slår markedet, vil få en positiv alfa. Hvis forvalter gjør det dårligere enn markedet, vil alfa bli negativ (Jensen, 1968).

4.3 Risikofaktorer

Ang (2014) hevder at det er faktorrisiko som driver risikopremier, og at det er faktoreksponeringen og ikke aktivumet i seg selv som er viktig for investoren. Kapitalverdimodellen viser forventet avkastning ved ulik eksponering mot markedsporteføljen. Markedsporteføljen er et eksempel på en faktor som over tid har vært forbundet med meravkastning i form av kompensasjon for å bære risiko. Eksponering mot markedsporteføljen er imidlertid kun én av flere faktorer som er identifisert og forbundet med meravkastning over tid.

4.3.1 Fama-French trefaktormodell

En faktormodell kan sees på som en utvidelse av kapitalverdimodellen, og skiller seg fra denne ved å inkludere flere faktorer. Ved å inkludere flere risikofaktorer, kan man kontrollere

for flere aspekter av avkastningen som ellers ville blitt fanget opp av alfa. Fama og French (1993) identifiserte fem risikofaktorer, hvorav tre var fra aksjemarkedet og historisk har hatt innvirkning på et aktivums avkastning og risiko. De tre faktorene identifisert i aksjemarkedet bestod av én relatert til selskapsstørrelse, én relatert til verdi/vekst og den allerede kjente markedsfaktoren. Størrelsesfaktoren (SMB) avspeiler at store selskaper gjerne har lavere avkastning enn små selskaper. Faktoren regnes ut ved å ta gjennomsnittlig avkastning for selskaper med lav markedsverdi, og trekke fra gjennomsnittlig avkastning for selskaper med høy markedsverdi. Faktoren relatert til verdi/vekst kalles verdifaktor (HML), og regnes ut ved å se på forskjellen i avkastning mellom selskaper med høy og lav bok/pris. Faktoren brukes fordi man historisk har sett at selskaper med høy bok/pris (verdiaksjer), har utkonkurrert selskaper med lav bok/pris (vekstaksjer). I praksis beregnes faktoren ved at man tar gjennomsnittlig avkastning for de 30 % av selskapene som har høyest bok/pris, og trekker fra den gjennomsnittlige avkastningen til de 30 % av selskapene som har lavest bok/pris. Trefaktormodellen kan sees i ligning 3.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + s_iSMB_t + h_iHML_t + u_{i,t} \quad (3)$$

$r_{i,t}$ er avkastning for aktivum i på tidspunkt t . $r_{f,t}$ er risikofri avkastning på tidspunkt t . β_i er beta, altså eksponeringen for aktivum i mot markedet. $r_{m,t}$ er den forventede markedsavkastningen på tidspunkt t . s_i er eksponeringen til aktivum i mot en etterlignende portefølje som skal fange størrelseseffekten (SMB). h_i er eksponeringen til aktivum i mot en etterlignende portefølje som skal fange verdieffekten (HML).

Størrelsesfaktoren var på sitt sterkeste på starten av 1980-tallet. Siden midten av 1980-tallet har ikke faktoren vært signifikant. I ettertid har det blitt diskutert om de signifikante resultatene man først fant var tilfeldige (Black, 1993), eller om rasjonelle investorer som følge av funnene i ettertid har bydd opp prisen på små selskaper til effekten forsvant (Schwert, 2003). Verdifaktoren derimot, sies fortsatt å være signifikant. Faktorens rasjonale bygger på at verdiselskaper er grunnleggende mer risikable, og gjør det spesielt dårlig i dårlige tider for økonomien. Faktoren har vært signifikant de siste 50 årene, men i perioder som finanskrisen har man tapt på en verdistrategi (Ang, 2014).

4.3.2 Carhart firefaktormodell

Carhart (1997) inkluderte en faktor for momentum i den allerede etablerte trefaktormodellen, og viste at man da kunne forklare enda mer av avkastningsvariasjonen. Faktoren ble inkludert på bakgrunn av tidligere undersøkelser av avkastning for "tapere" og "vinnere" i aksjemarkedet, som viste at man kan tjene på ulike strategier knyttet til momentum (Jegadeesh og Titman, 1993). Faktoren blir utformet ved å lage porteføljer hvor man kjøper aksjer som allerede har prestert godt, og selger aksjer som har gjort det dårlig i samme periode. Denne faktoren kalles gjerne WML, som står for "winners minus losers". Ligning 4 eksemplifiserer en faktorportefølje som inkluderer en momentumfaktor. w_i er eksponeringen til aktivum i mot en etterlignende portefølje for momentum (WML_t).

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + s_iSMB_t + h_iHML_t + w_iWML_t + u_{i,t} \quad (4)$$

4.3.3 Fama-French femfaktormodell

Som følge av nye studier som viste at trefaktormodellen var mangelfull, introduserte Fama og French (2015) to nye faktorer. Faktorene skal fange opp effekter knyttet til lønnsomhet og investeringer. Lønnsomhetsfaktoren RMW, som står for "robust minus weak", konstrueres ved å beregne forskjellen i avkastning mellom diversifiserte porteføljer med sterk og svak lønnsomhet. Lønnsomhet måles ved årlige inntekter minus varekostnader, rentekostnader og salgs- og administrasjonskostnader delt på bokverdi av egenkapitalen. Investeringsfaktoren CMA, som står for "conservative minus aggressive", konstrueres ved å ta forskjellen i avkastning mellom to diversifiserte porteføljer med høye og lave investeringer. Investeringer måles ved vekst i totale eiendeler delt på totale eiendeler ved slutten av fjoråret. I likhet med de tidligere faktorene, forventes også disse to å ha positivt fortegn. Fama og French (2015) estimerer at denne nye modellen forklarer mellom 71 % og 94 % av tversnittvariasjonen i avkastningen til porteføljene de undersøker. De nye faktorene er imidlertid omstridte, da det er satt spørsmål ved deres robusthet og intuisjon (Blitz et al., 2016). Femfaktormodellen er illustrert i ligning 5:

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + s_iSMB_t + h_iHML_t + c_iCMA_t + rw_iRMW_t + u_{i,t} \quad (5)$$

5. Datasett

Dette kapitlet er tredelt. I første og andre del forklares inndata overordnet og de finansielle variablene. Tredje del ser nærmere på de ulike scorene for ansvarlighet; ESG-score, kontroversscore og en score som kombinerer disse to.

5.1 Inndata

Datasettet vårt består av selskaper fra hovedbørsene til de fleste land i Europa, og er hentet fra Thomson Reuters Datastream. Disse selskapene er fra ulike sektorer og bransjer, noe som bør gjøre utvalget vårt representativt. Analysene våre baserer seg på avkastningsdata for 14 år, fra og med juli 2004 til og med juni 2018. For at datasettet skal samsvare mest mulig med de ferdigkalkulerte Fama-French-faktorene benyttes kun ordinære aksjer.

Til sammen finner vi 1338 europeiske selskaper som har ESG-score i perioden vi undersøker. Dette er relativt god dekning i forhold til det som er mulig. Thomson Reuters (2018) oppgir nemlig at de totalt har ESG-score for i overkant 1400 selskaper i Europa. Til sammenligning oppgir de at de har ESG-score for over 2900 selskaper i Nord-Amerika. Det er derfor mulig at amerikanske studier kan ha et bedre utvalg.

5.2 Finansielle måltall

Den månedlige avkastningen er basert på endring i Datastream sin avkastningsindeks (ΔRI_t), som også inkluderer reinvestering av utbytte. Avkastningsindeksen er gitt som følger:

$$RI_t = RI_{t-1} * \frac{P_t}{P_{t-1}} \quad (6)$$

Unntatt på reinvesteringsdatoen for dividenden, hvor RI_t er gitt ved:

$$RI_t = RI_{t-1} * \frac{P_t + D_t}{P_{t-1}} \quad (7)$$

Her er P_t prisen på dag t , og D_t er dividenden assosiert med reinvesteringsdatoen t . Deretter regnes avkastningsindeksen om til avkastning oppgitt i prosentpoeng, (r_t).

$$r_t = \Delta RI_t = \frac{RI_t - RI_{t-1}}{RI_{t-1}} * 100 \quad (8)$$

Vi benytter månedlig avkastning rapportert i slutten av hver måned.

Som markedsportefølje benyttes verdivektet avkastning for alle selskapene som er med i utvalget. Risikofaktorene SMB, HML, RMW, CMA og WML er hentet fra "Fama/French European 5 Factors" og "European Momentum Factor" i Kenneth Frenchs database, som igjen benytter data fra Bloomberg. Den risikofrie renten er hentet fra samme kilde. Disse dataene rapporteres som månedlig tall, og er oppdaterte versjoner av tallene som benyttes i Fama og French (2012).

For å regne ut vektet avkastning for verdivektede porteføljer med ulike ESG-score bruker vi markedsverdien til selskapenes egenkapital (ME), som også hentes fra Datastream. ME oppgis i millioner, og defineres som aksjeprisen multiplisert med antall ordinære aksjer. Til å vekte selskaper under sammensetningen av porteføljer for perioden juli i år t til juni i år $t + 1$ benyttes ME fra slutten av juni år t . For å vekte porteføljens månedsavkastning benyttes ME i starten av måneden. Siden ME som benyttes til å vekte porteføljer og avkastning må være sammenlignbar mellom ulike selskaper, regnes denne om til en felles valuta. Vi benytter Euro som felles valuta siden dette er den vanligste i Europa. Dette gjøres ved hjelp av en innebygd funksjon i Datastream. Bokverdi av egenkapitalen (BE) benyttes til å regne ut bok/pris (BE/ME). I likhet med Fama og French (1992) benytter våre avkastningsobservasjoner som begynner i juli år t og strekker seg til juni år $t + 1$, observasjoner av BE fra regnskapsåret som avsluttes i år $t - 1$. ME som benyttes til å regne ut BE/ME for denne perioden, er basert på observasjoner fra slutten av år $t - 1$.

For analysen på selskapsnivå inkluderer vi også en momentumfaktor. Denne er laget ved å regne ut gjennomsnittlig avkastning for de foregående 12 månedene. Vi krever derfor at alle selskaper må ha hatt 12 måneder med avkastning før de tas med i analysen. For at vi skal ha observasjoner i analysen fra og med juli 2004, har vi derfor hentet avkastningstall fra og med juli 2003.

5.3 Måltall på ansvarlighet

Det finnes flere leverandører av ESG-rangeringer, og metodikken som benyttes varierer i stor grad. Noen kjente er for eksempel Bloomberg, MSCI, Sustainalytics og Thomson Reuters. Vi har valgt å benytte Thomson Reuters fordi det er den leverandøren som vurderer flest indikatorer av alle de store anerkjente leverandørene (Davis Polk, 2017), samtidig som det er data vi har tilgang til gjennom NHH. Thomson Reuters virker også som en av de vanligste leverandørene å benytte i akademisk sammenheng.

5.3.1 ESG-score

Thomson Reuters (2018) bygger opp sin ESG-score basert på data som selskaper rapporterer offentlig. ESG-score rapporteres som årlige verdier i takt med selskapers regnskapsår. Totalt samler de over 400 ulike måltall. Av disse har de valgt ut de 178 mest sammenlignbare og relevante måltallene. Disse blir igjen delt inn i 10 underkategorier. Underkategoriene og fordelingen av måltall kan sees i Tabell 1.

Kategori	Underkategori	Måltall
Miljø	Ressursforbruk	19
	Utslipp	22
	Innovasjon	20
Sosiale forhold	Arbeidsstyrke	29
	Menneskerettigheter	8
	Samfunn	14
Selskapsstyring	Produktansvar	12
	Ledelse	34
	Eiere	12
	CSR-strategi	8

Tabell 1. Oversikt over oppbygning av kategorier i ESG-score.

Måltallene kan være både kvalitative og kvantitative. Kvalitative måltall registreres som "Ja", "Nei" eller "NA". Hvis det kvalitative måltallet ikke besvares så registreres det enten som "Nei" eller "NA". De kvalitative måltallene gis også en numerisk verdi. "Ja" blir tildelt verdien 1, "Nei" blir tildelt verdien 0,5 og "NA" blir tildelt verdien 0. Hvis et selskap mangler data for et kvantitativt måltall, fjernes dette måltallet fra beregningen. Det eneste unntaket er utslipp av drivhusgasser, hvor det i stedet blir beregnet estimerte utslipp. Hvis et måltall ikke er relevant for alle bedrifter, så ekskluderes det for de det ikke er relevant for. For både kvalitative

og kvantitative måltall, vil de numeriske verdiene behandles som negative tall hvis et positivt eller høyt svar kan ansees som et negativt attributt. Eksempler på ulike måltall som benyttes av Thomson Reuters kan sees i Tabell 2.

Miljø (Environmental)	Sosiale forhold (Social)	Selskapsstyring (Governance)
Utslipp av CO2-ekvivalenter	Benytter selskapet barnarbeidere?	Antall styremøter i året
Søppel produsert	Er selskapet med i Initiativ for Etisk Handel (IEH)?	Styrets kjønnsfordeling
Vannforbruk	Har selskapet en policy for fri handel?	Gjennomsnittlig antall år styrets medlemmer har sittet i styret
Mengde vann som blir resirkulert	Producerer selskapet tobakk?	Er daglig leder styremedlem?
Energikonsum	Kjønnsfordeling blant de ansatte	Er daglig leders kompensasjon prestasjonsbasert?

Tabell 2. Ulike måltall i Thomson Reuters beregninger av ESG.

For å regne ut den endelige scoren regnes det først ut en persentilscore for hvert enkelt måltall, ved hjelp av formelen i ligning 9. Denne scoren er basert på rangering. Hvert enkelt selskaps rangering for hvert måltall er da avhengig av hvor mange selskaper som har dårligere verdi, hvor mange som har lik verdi, og hvor mange som i det hele tatt rapporterer en verdi. Hva som benyttes som benchmark, altså hvilke selskaper man skal sammenligne med, varierer. For miljø og sosiale forhold benyttes bransje som benchmark. Dette skyldes at disse forholdene varierer i stor grad med bransje. Selskapsstyring varierer ikke like mye med bransje, så her benyttes i stedet landet selskapets hovedkvarter ligger i som benchmark. Siden scorene benytter rangering har ikke ekstremverdier noen særskilt effekt.

$$score\ selskap\ x = \frac{\# selskaper\ med\ dårligere\ verdi + \frac{\# selskaper\ med\ lik\ verdi\ som\ selskap\ x}{2}}{\# selskaper\ med\ verdi} \quad (9)$$

For hver underkategori regner man så ut en gjennomsnittlig score basert på de ulike måltallene. Deretter benytter man samme formel til å rangere selskapene basert på denne gjennomsnittlige scoren for å finne en persentilscore for underkategorien. Hver kategori får så en score basert på et snitt av underkategoriene vektet etter antall måltall i hver underkategori. Den samlede ESG-scoren er så basert på et snitt av kategoriene vektet etter antall underliggende måltall for hver kategori. Scoren multipliseres avslutningsvis med 100.

5.3.2 Kontroversscore

Kontroversscoren er basert på 23 underliggende måltall knyttet til nylige kontroverser. Alle kontroversene er kvantitative, og knyttet til antall kontroverser publisert i media. Eneste unntaket er et kvalitativt måltall for om et viktig medlem av ledelsen har fått sparken eller

sluttet frivillig av annen årsak enn pensjonering. Bransjer benyttes som benchmark. Kontroversscoren regnes ut ved at alle kontroversene et selskap har vært utsatt for, summeres opp. Deretter benyttes formelen i ligning 9 til å regne ut den endelige kontroversscoren. Om et selskap har vært involvert i mange kontroverser, vil det få en lav kontroversscore. Har det vært involvert i få kontroverser, vil det få en høy score. Hvis ingen selskaper har vært involvert i kontroverser, vil alle selskapene få en score på 50 etter at den er multiplisert med 100.

En kontroversscore gjelder for et gitt regnskapsår. Om det inntreffer nye kontroverser etter at et regnskapsår er avsluttet, og før det neste er ferdig, vil kontroversen inkluderes i forrige avsluttede regnskapsår. Kontroversscoren vil da oppdateres fortløpende. Når et nytt regnskapsår avsluttes, vil de kontroversene som var inkludert i det forrige regnskapsåret, men hendte i ettertid, flyttes frem til det nye avsluttede regnskapsåret. Dermed kan dataene for siste år endre seg fortløpende ettersom kontroverser finner sted, og avhengig av når kontroversscoren for det neste året blir publisert. Dette skaper litt usikkerhet rundt de siste observasjonene (Thomson Reuters, 2018).

5.3.3 Kombinertscore

Kombinertscore er basert på tre ulike scenarioer. Det første scenarioet er at kontroversscoren er over 50. Da har selskapet gjort det relativt bra, med tanke på kontroverser, i forhold til de andre selskapene i benchmark. I et slikt tilfelle vil den kombinerte scoren bli lik den vanlige ESG-scoren. Det andre scenarioet er at kontroversscoren er under 50, men fortsatt større enn ESG-scoren. Da har selskapet gjort det bedre med hensyn til kontroverser enn med faktisk ESG-score. I et slikt tilfelle benyttes også ESG-score som kombinertscore. Det siste tilfellet er hvis kontroversscoren er under 50, samtidig som den er lavere enn ESG-scoren. I et slikt tilfelle vil den kombinerte scoren være gjennomsnittet av ESG-scoren og kontroversscoren (Thomson Reuters, 2018).

Scenario	Kontroversscore	ESG-score	Kombinertscore
Kontroversscore ≥ 50	57	38	38
Kontroversscore < 50 > ESG-score	49	42	42
Kontroversscore < 50 < ESG-score	48	49	48,5

Tabell 3. Kombinertscore i ulike scenarioer.

5.3.4 Egenkonstruerte faktorer for ansvarlighet

For å konstruere faktorer rangeres selskaper etter score. Faktorene baserer seg på en lagget score. Det vil si at for porteføljene som brukes til å lage faktorene fra juli år t til juni år $t + 1$, benyttes score fra regnskapsåret som avsluttes i år $t - 1$. Når selskapene er rangert, deles de inn i tre porteføljer ut fra antall selskaper, med splitt på den 30. og 70. persentilen. Disse porteføljene rebalanseres i starten av juli hvert år. Faktoren utregnes deretter ved å trekke den månedlige vektete avkastningen til porteføljen med liten score fra den med høy score. Dette gjøres for både ESG- og kontroversscore, slik at vi får to ulike faktorer.

5.4 Datavasking

Datasettet hentet fra Datastream består av til sammen 16857 selskaper. Alle selskaper har ikke alle data vi ønsker for alle månedene vi undersøker. For de periodene som mangler nødvendig data, slettes observasjonene for det aktuelle selskapet for den perioden. Dette gjør at vi sitter igjen med 1302 selskaper totalt. Etter at manglende observasjoner er fjernet, winsoriseres avkastning, ME og BE/ME på 0,5% i henhold til Fama og French (1992). Det vil si at de 0,5 % høyeste (laveste) observasjonene reduseres (økes) til de er like den høyeste (laveste) observasjonen som ikke er endret. Dette reduserer effekten av uteliggere, samtidig som vi unngår å miste mange observasjoner.

6. Metode

Første del av metodedelen tar for seg motivasjonen bak metodevalgene. Andre del ser nærmere på hvilke metoder som er benyttet til å analysere porteføljene, og hvordan porteføljene er satt sammen. Tredje del ser på metoden som er anvendt for å analysere avkastning og ansvarlige investeringer ved bruk av paneldata. Fjerde del ser på forutsetninger bak metodene.

6.1 Motivasjon for valg av metode

Å konstruere porteføljer basert på ESG-score, er en av de vanligste tilnærmingene for å undersøke forholdet mellom ansvarlighet og finansielle resultater (Halbritter og Dorfleitner, 2015). Når porteføljene er satt opp som tidsserier er det mulig å benytte faktormodeller.

Ved å benytte panelbaserte selskapsdata kan man undersøke den direkte effekten av ESG på selskapets avkastning. Når man benytter porteføljer, vil man gå glipp av detaljer i den selskapsesifikke informasjonen. Ang, Liu og Schwarz (2017) viser at selv om porteføljer gir mindre usikker faktoreksponering, fører ikke dette til lavere standardfeil for risikopremieestimer. Dette skyldes at man ved å gruppere aksjer inn i porteføljer diversifiserer bort mye av informasjonen i de selskapsindividuelle faktoreksponeringene. Resultatet er at porteføljeanalyser får store effektivitetstap i forhold til analyser av enkeltelskaper. I tillegg til fordelene som følger med hver av de ulike metodene, vil også analysens robusthet styrkes ved å ta i bruk flere metoder.

6.2 Porteføljer

6.2.1 Porteføljekonstruksjon

I slutten av juni hvert år fra 2004 til 2017 rangeres selskapene etter ESG-score og kontrovers-score. For år t benyttes ESG-score fra regnskapsåret som ble avsluttet i år $t - 1$, for å ta hensyn til en forsinket effekt. Dette betyr også at ESG-scoren man benytter er tilgjengelig på det tidspunktet man setter sammen porteføljen, og dermed kan brukes i praksis. Deretter deles selskapene inn i fem porteføljer med like stor markedsverdi. Første porteføljesammensetning finner sted i slutten av juni 2004. Porteføljen holdes så i ett år, før den rebalanseres i slutten

av juni 2005. Slik fortsetter det til siste holdeperiode er ferdig i juni 2018. Månedlig verdivektet avkastning måles i slutten av hver måned. Ved å vurdere avkastningen i disse porteføljene opp mot hverandre, vil man kunne sammenligne avkastning fra porteføljer med ulik score. Ved å konstruere egne porteføljer på denne måten ser vi bort fra forvaltningskostnader og enkelte sammensetningsrestriksjoner som ekte porteføljer kan være utsatt for.

6.2.2 Prisingsmodeller

For å kontrollere de forskjellige porteføljene mot allerede kjente risikopremier introdusert i kapittel 4.3, utfører vi tidsserieregresjoner hvor vi gradvis inkluderer flere faktorer. Ved å introdusere faktorene gradvis, kan vi få en bedre forståelse av de ulike faktorenes effekt. De vil fange opp avkastningen som skyldes porteføljens eksponering mot de aktuelle faktorene. Alfa vil da representere den avkastningen som ikke kan forklares av eksponering mot de systematiske risikofaktorene. Når vi kontrollerer for de systematiske risikofaktorene på denne måten, vil vi ved å sammenligne alfa til porteføljer med ulik ESG- eller kontroversscore, i større grad se den faktiske effekten av ESG og kontroverser på avkastning.

Den første prisingsmodellen som benyttes er den klassiske CAPM, som ble vist i teoridelen, se ligning 2. Alfa består da av avkastningen som ikke kan forklares av eksponering mot markedet. Som proxy for markedet benytter vi det samlede utvalget. Deretter tas det hensyn til porteføljens størrelses- og verdieffekter gjennom å inkludere SMB og HML, presentert i teoridelen (ligning 3). Hvis de konstruerte porteføljene har avkastningsforskjeller som skyldes eksponering mot disse faktorene, vil vi luke vekk denne effekten fra alfa ved å inkludere faktorene i regresjonen. Neste prisingsmodell tar hensyn til momentum (WML) i markedet. Altså forskjellen i avkastning i påfølgende periode for selskaper som har gjort det bra og dårlig i forrige periode. Da faktorene er lett tilgjengelige, har vi også inkludert CMA og RMW, som også ble presentert i teoridelen (ligning 5). Når det tas hensyn til flere faktorer, blir det lettere å vurdere om noen av avkastningsforskjellene skyldes ESG eller kontroverser.

I tillegg til de anerkjente risikofaktorene, har vi også valgt å inkludere egne faktorer basert på ESG og kontrovers. Disse faktorene inkluderes for å kontrollere for samvariasjon mellom de ulike scorene, slik at vi kan identifisere det som er unikt. Hvis for eksempel en portefølje med høy kontroversscore har signifikant alfa også etter at ESG-faktoren er inkludert, tyder det på at kontroversscore kan skape meravkastning som ikke kan forklares av forskjeller i ESG-score.

De ulike regresjonene som utføres er varianter av ligning 10. Dersom en modell inneholder faktorer som porteføljene ikke har signifikant eksponering for, er det mulig den blir utelatt i påfølgende modeller for ikke å miste frihetsgrader.

$$r_{i,t} - r_{f,t} = \alpha_i + \beta_i(r_{m,t} - r_{f,t}) + s_iSMB_t + h_iHML_t + w_iWML_t + c_iCMA_t \quad (10) \\ + rw_iRMW_t + e_iESG_t + k_iKontr_t + u_{i,t}$$

Faktoreksponeringene er gitt ved β_i , s_i , h_i , w_i , c_i , rw_i , e_i og k_i . Avkastning for hver portefølje i , i måned t , etter å ha trukket fra risikofri rente ($r_{f,t}$), er angitt ved $(r_{i,t} - r_{f,t})$.

6.3 Selskaper

Sammenhengen mellom ansvarlighet og avkastning undersøkes også på selskapsnivå. Vi har derfor satt opp dataene på paneldataform. Disse dataene inneholder mer selskapsspesifikk informasjon enn det som benyttes i porteføljeanalysen.

Når man utfører regresjoner, er en antakelse at hver observasjon er en tilfeldig trekning av uavhengige variabler. Men i analyser med gruppestruktur er det ikke nødvendigvis slik at en observasjon er uavhengig av forrige tidsperiode, eller de andre observasjonene i samme tidsperiode. Dette gjør at tradisjonelle regresjoner gir feilaktige beregninger og sammenhenger (Angrist og Pischke, 2009).

I analyser innen finans, hvor dataene gjerne er satt opp som paneldata, er det ikke uvanlig at avhengigheter i dataene kan gi bias i de estimerte standardfeilene. Denne biasen er avhengig av både korrelasjonen i de uavhengige variablene og i feilleddene. Finansdata er ofte utsatt for autokorrelasjon i de uavhengige variablene. Hvis man for eksempel bruker størrelse som en variabel, vil denne trolig være relativt stabil. Hvis man også har en uobservert fast effekt i feilleddet, vil dette gi bias i standardfeilene ved vanlig OLS, siden standard OLS forutsetter at produktet av autokorrelasjonen i feilleddene og de uavhengige variablene er tilnærmet 0. Man vil altså ha et problem med standardfeilene om $Cov(X_{i,t}\varepsilon_{i,t}, X_{i,t-k}\varepsilon_{i,t-k}) \neq 0$. Tilsvarende situasjon har vi for en eventuell tidseffekt. I våre data kan det være rimelig å anta at vi har tidseffekt i residualene som følge av markedssvingninger. En fast tidseffekt kan fjernes ved å inkludere dummyer, men dette fjerner ikke en tidseffekt som ikke er lik for alle selskaper. Hvis

vi i tillegg har en tidseffekt i de uavhengige variablene [$Cov(X_{i,t}\varepsilon_{i,t}, X_{k,t}\varepsilon_{k,t}) \neq 0$], vil dette gi bias i de estimerte standardfeilene. Med mange tidsperioder og et ubalansert paneldatasett blir biasen i standardfeilene enda større (Petersen, 2009).

Petersen (2009) så på ulike måter å håndtere slike avhengigheter. Han undersøkte også tidligere publiserte artikler, og fant at hele 42 % av nylige publiserte artikler innen finans, som inkluderte regresjoner på paneldata, ikke justerte standardfeilene for potensiell bias.

En metode for å håndtere avhengigheter i ulike dimensjoner i paneldata, er å clustre for dimensjonen. Denne metoden fjerner forutsetningen om at observasjonene er uavhengige innad i clustere, og tar hensyn til biasen i standardfeilene. Metoden forutsetter fortsatt at observasjoner er uavhengige på tvers av clustere (Wooldridge, 2010). Hvis vi for eksempel ser på avkastningsdata for selskaper kan en dimensjon være forskjellige selskaper. Å clustre for dimensjonen selskaper tillater da at det kan være avhengigheter i observasjonene for samme selskap, men ikke mellom observasjoner på tvers av selskaper.

Den eksakte formelen for clustrede standardfeil kan sees i ligning 11 (Angrist og Pischke, 2009). Ligningen returnerer varians-kovarians matrisen, og de diagonale elementene angir koeffisientenes cluster-robuste standardfeil. Kovariansen mellom residualene innad i clusteren, estimeres ved å kvadrere summen av $X\hat{\varepsilon}$ for hver cluster. Den kvadrerte summen av dette leddet summeres for alle de ulike clusterene i dimensjonen, M . N er antall observasjoner og K er antall uavhengige variabler i regresjonene, inkludert konstantleddet. X er matrisen med alle observasjoner for de ulike forklaringsvariablene. Hvis vi for eksempel clustrer på tid, så er M antall tidsperioder, og det er korrelasjonen mellom residualene innad i de forskjellige tidsperiodene som summeres for å justere standardfeilen (Angrist og Pischke, 2009).

$$VCV_{cl}(\hat{\beta}_{OLS}) = \left(\frac{M}{M-1} \frac{N}{N-K}\right) (X'X)^{-1} \sum_{i=1}^M (X_M \hat{\varepsilon}_M \hat{\varepsilon}'_M X_M) (X'X)^{-1} \quad (11)$$

I tilfeller der det kun er avhengighet mellom feilleddene i form av en uobservert selskaps-spesifikk effekt, kan vi få korrekte standardfeil ved å clustre på selskap. Dersom det kun er en avhengighet mellom feilleddene i form av en tidseffekt, kan vi få korrekte standardfeil ved å benytte Fama-MacBeth (1973). I dette tilfelle kan vi også få korrekt standardfeil ved å clustre på tid, gitt at vi har tilstrekkelig antall tidsperioder. Dersom det er

avhengigheter i feilleddene både på tvers av år og selskaper, så kan dette løses ved å inkludere dummyer for den ene dimensjonen, og deretter beregne standardavvik ved å clustre på den andre dimensjonen. Dette er gitt at den dimensjonen vi inkluderer dummyer for, har en effekt som er lik for alle observasjoner innad i dimensjonen, altså kun en fast effekt. Eventuelt kan man velge å clustre over flere dimensjoner, som i vårt tilfelle er både over tid og selskaper. Dette forutsetter at vi har tilstrekkelig med clustere i hver dimensjon (Petersen, 2009). Thomsen (2011) argumenterer for at man må ha minst 25 observasjoner per dimensjon for at man kan clustre på den. Der vi clustrer i to dimensjoner benytter vi Thompson (2011) sin metode for å beregne standardfeil. Metoden baserer seg på å beregne varians-kovarians matrisen til koeffisientene som angitt i ligning 12.

$$\hat{V}_{selskap\&tid} = \hat{V}_{selskap} + \hat{V}_{tid} - \hat{V}_{white} \quad (12)$$

$\hat{V}_{selskap}$ er varians-kovarians matrisen når det clustres på selskaper, \hat{V}_{tid} er varians-kovarians matrisen når det clustres på tidsdimensjonen og \hat{V}_{white} er varians-kovarians matrisen som fås når man utfører OLS regresjon med White (1980) standardfeil.

Amir et al. (2016) påpekte at selv om standardfeil korrigeres i henhold til Petersen (2009), kan det fortsatt være bias i koeffisientene. Dette fordi de faste effektene som påvirker standardfeilene, også kan tenkes å påvirke koeffisientene dersom de er korrelerte. Vi antar i utgangspunktet at vi kun har en tidseffekt som vi må ta hensyn til. Denne tidseffekten antar vi kan være både fast og varierende. Altså at en del av tidseffekten påvirker alle selskaper likt, mens en annen del påvirker selskapene ulikt. Disse antakelsene blir diskutert nærmere i avsnitt 7.2.2. Ved å inkludere tidsdummyer tar vi hensyn til den faste tidseffekten på standardfeilene, men vi må likevel clustre på tid for å ta hensyn til den varierende effekten. Vi antar i utgangspunktet at tidseffekten i feilleddet ikke er korrelert med de uavhengige variablene. Om dette likevel skulle være tilfelle, blir biasen i koeffisientene redusert ved å inkludere tidsdummyer.

For hvert selskap inkluderer vi den naturlige logaritmen av markedsverdien til egenkapitalen ($\ln ME$) og bok/pris forholdet ($\ln BE/ME$). I tillegg inkluderer vi gjennomsnittlig avkastning for selskapet de 12 foregående månedene (MOM). Disse kontrollvariablene vil dekke noe av det tilsvarende som SMB, HML og WML gjør i porteføljeanalysene. Selve OLS-regresjonen som utføres på paneldataene er variasjoner av ligning 13. Faktorene vi inkluderer i tillegg til

de nevnt ovenfor er D_t , ESG , $Kontr$ og $Komb$. D_t er en fast tidseffekt på tidspunkt t . I praksis fanges denne tidseffekten ved å inkludere månedsdummyer. De tre siste variablene er henholdsvis ESG-score, kontroversscore og kombinertscore. Alle med unntak av MOM er årlige observasjoner for hvert selskap som benyttes til å forklare månedlig avkastning. De årlige variablene er forsinket slik som forklart i avsnitt 5.2.

$$r_{i,t} = D_t + s \ln ME_{i,t} + h \ln (BE/ME)_{i,t} + m MOM_{i,t} + e ESG_{i,t} + k Kontr_{i,t} + c Komb_{i,t} + \varepsilon_{i,t} \quad (13)$$

Faktoreksponeringene for de ulike faktorene er gitt ved s, h, m, e, k og c . Det vi ønsker å undersøke er om e og k har signifikant påvirkning på avkastning.

6.4 Modellenes forutsetninger

De ulike modellene som er brukt i analysedelen bygger på en del forutsetninger. Forutsetningene for paneldataregresjonene og tidsserieregresjonene er håndtert på forskjellige måter, men noe er også felles. Linearitet ble for begge sett testet ved å plote residualene mot de uavhengige variablene. Når denne forutsetningen er innfridd, vil plottet av residualene være jevnt fordelt langs en horisontal linje, uten noen systematiske variasjoner. Våre data tilfredsstillende dette (se linearitetsplot i appendiks).

For tidsserieregresjonene ble heteroskedastisitet undersøkt med Breusch-Pagan (1979)/Cook-Weisberg (1983) test som forutsetter at residualene er normalfordelte. For de regresjonene som ikke hadde normalfordelte residualer, ble plot av residualer mot estimerte verdier undersøkt visuelt for spor etter heteroskedastisitet. For de regresjonene som hadde problemer med heteroskedastisitet ble det beregnet White standardfeil som er robust mot dette (White, 1980). Autokorrelasjon ble undersøkt med Breusch-Godfrey (1978; 1978) test for høyere-ordens autokorrelasjon. I motsetning til Durbin-Watson (1951) test, forutsetter ikke denne testen at alle variablene er strengt eksogene, og den er heller ikke begrenset til å kun teste for første-ordens autokorrelasjon. Dersom en tidsserieregresjon hadde problemer med autokorrelasjon, har det blitt tatt hensyn til ved å bruke Newey og West (1987) robuste standardfeil.

For paneldatagresjonene håndteres avhengigheter på tvers av observasjoner i en tidsperiode, og avhengigheter på tvers av tid for et selskap gjennom clustering, som drøftet i avsnitt 6.3. Clustering er også robust mot heteroskedastisitet, som vi ser tegn til i residualplottene (se appendiks Figur 10 og Figur 16).

For å unngå spuriøse sammenhenger må dataene være stasjonære. Siden vi ser på avkastning, er det rimelig å forvente at vi har stasjonære tidsserier som diskutert i avsnitt 4.1. Vi har likevel formelt bekreftet dette ved å benytte en Dickey-Fuller-test, eventuelt "augmented" Dickey-Fuller-test, der vi identifiserer problemer med autokorrelasjon (Dickey og Fuller, 1979).

For å teste for multikollinearitet på tvers av variablene benyttes en "variance inflation factor" (VIF)-test (Wooldridge, 2015). Selv om enkelte porteføljevariabler, slik som ESG-faktoren og SMB, har høy korrelasjon, identifiserte vi ingen problemer med multikollinearitet. For paneldataene var dette et problem når både kombinert- og ESG-score var inkludert. Men dette gjaldt kun en regresjon, og er ikke kritisk for analysens konklusjon.

Selv om det ikke kreves for at OLS-estimatoren skal være den beste lineære estimatoren uten bias (BLUE), så vil normalfordelte residualer gjøre det mulig å beregne eksakte t- og F-verdier (Wooldridge, 2015). Om residualene er normalfordelte ble testet med en Shapiro-Wilk-test (Shapiro og Wilk, 1965). En del av porteføljeregresjonene viser seg å ikke ha normalfordelte residualer. Ved å undersøke residualene visuelt observerer vi at de ser akseptable ut, men at enkelte verdier skiller seg litt ut, noe som trolig har hatt stor innvirkning på den formelle testen. I henhold til den formelle testen var heller ikke paneldatagresjonenes residualer normalfordelte. Dette er ikke så uventet, fordi ved store utvalg vil selv små avvik føre til at en formell test forkaster nullhypotesen om at fordelingen er normalfordelt. Ved en visuell inspeksjon ser vi at fordelingen kunne vært bedre, men at den ikke er helt urimelig (se residualplot i appendiks). Ved et stort utvalg sier imidlertid sentralgrenseteoremet at de estimerte koeffisientenes fordeling går mot en normalfordeling, noe som gjør at man likevel ikke trenger normalfordelte residualer (Greene, 2012). Hva som er stort nok utvalg for at man skal kunne benytte seg av sentralgrenseteoremet er situasjonsavhengig, men ofte brukes 30 som en tommelfingerregel. For paneldataene kan vi altså se bort fra denne forutsetningen. For porteføljetidsseriene er det litt mer usikkert, men med 168 observasjoner og residualer som ser relativt normalfordelte ut, antar vi det at dette ikke er et stort problem.

7. Empirisk analyse

Analysekapittelet er delt opp i fire deler. Første del tar for seg selskaper som er satt sammen i porteføljer ut fra ESG- eller kontroversscore. Disse porteføljene analyseres ved hjelp av ulike prisingsmodeller. Første modell er den klassiske CAPM, og deretter benyttes ulike flerfaktormodeller. Andre del av analysen består av paneldataregresjoner som benytter individuelle selskaper. I tredje del diskuteres resultatene, og funnene knyttes opp mot tidligere litteratur. I fjerde del påpekes utfordringer og begrensninger for analysen.

7.1 Porteføljenivå

7.1.1 Innledende deskriptiv statistikk

Deskriptiv statistikk for porteføljer konstruert etter ESG- og kontroversscore kan sees i Tabell 4. *ESG høy* er en verdivektet portefølje som inneholder de selskapene med høyest ESG-score som utgjør 20 % av den totale markedsverdien. *ESG lav* er en verdivektet portefølje som inneholder de selskapene med lavest ESG-score som utgjør 20 % av den totale markedsverdien. Som tidligere nevnt rebalanseres porteføljene årlig. *Kontr høy* og *Kontr lav* er tilsvarende porteføljer, men sortert etter kontroversscore.

ESG høy har gjennomsnittlig 50 selskaper i porteføljen, mens *ESG lav* har 421 selskaper. *Kontr høy* har gjennomsnittlig 273 selskaper i porteføljen, mens *Kontr lav* har 52 selskaper. Disse sammenhengene mellom størrelse og score gjelder også for selskapene som ikke er med i porteføljene med høyest og lavest score. Dette indikerer at det er betydelige størrelsesforskjeller for det gjennomsnittlige selskapet med lav og høy score. For ESG ser det ut til at store selskaper har høy score, mens for kontrovers ser det ut til at store selskaper har lav score. Dette kan også indikere en sammenheng mellom det å ha høy ESG-score og lav kontroversscore, noe som kan virke lite intuitivt, men stemmer overens med funnene til Blackrock (2018).

I Tabell 4 ser vi at markedet har en gjennomsnittlig avkastning på 0,57 % i måneden. *ESG høy* har en gjennomsnittlig avkastning under dette, mens *ESG lav* har lik gjennomsnittlig avkastning som markedet. Selskaper med høy ESG-score har altså lavere gjennomsnittlig avkastning, men dette kan skyldes andre faktorer enn ESG. For porteføljene sortert etter

kontrovers ligger snittavkastning for både høy og lav kontroversscore under markedets gjennomsnitt. Dette indikerer at det hverken er porteføljene med høy eller lav kontroversscore som har høyest avkastning, men heller de selskapene som ligger mer midt på treet. Alle avkastningene er oppgitt etter at risikofri rente er trukket fra.

En kanskje litt uventet observasjon er at HML faktoren i gjennomsnitt er negativ for vårt tidsintervall. Dette endrer seg om tidsintervallet endres litt. Årsaken til dette er at HML generelt har vært lav siden finanskrisen. For de egenkonstruerte faktorene ser vi at ESG-faktoren er negativ, noe som stemmer godt med det første inntrykket vi fikk av porteføljene med høy og lav ESG-score. For kontroversscore er gjennomsnittet positivt, altså at selskaper med høy score har høyere snittavkastning enn selskaper med lav score, noe som også stemmer med inntrykket porteføljene gav oss.

Porteføljen med høyest standardavvik er *ESG høy*, mens den med lavest standardavvik er *Kontr høy*. Riktignok har alle porteføljene høyere standardavvik enn markedet. Dette kan tyde på at porteføljene innehar usystematisk risiko som er diversifisert bort i markedsporteføljen.

Tabell 4. Deskriptiv statistikk for avkastning ($ri-rf$) og benyttede variabler. Ulike porteføljer øverst, ulike kontrollvariabler i midten, og egenkonstruerte faktorer for ansvarlige forhold nederst.

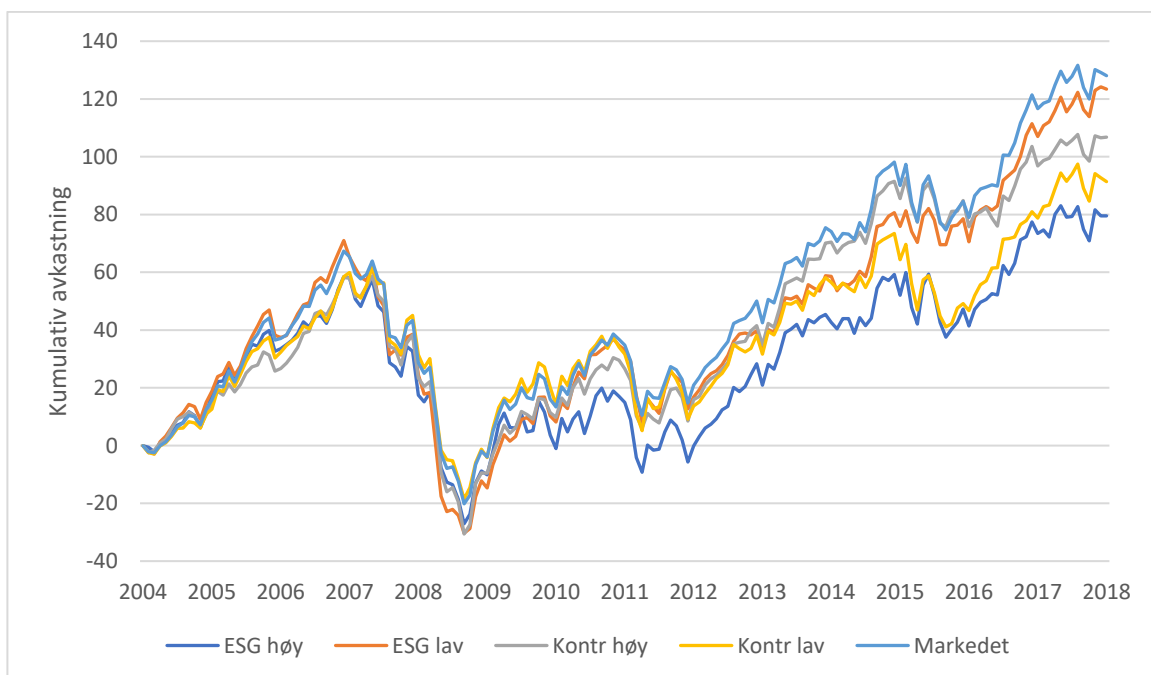
	Gjennomsnitt	Median	Std.avvik	Min	Maks
Markedet	0,570	1,024	3,921	-13,696	12,782
<i>ESG høy</i>	0,454	0,933	4,588	-12,917	14,251
<i>ESG lav</i>	0,570	0,997	4,212	-18,582	15,565
<i>Kontr høy</i>	0,520	0,878	4,137	-17,825	19,657
<i>Kontr lav</i>	0,475	0,937	4,171	-13,437	10,717
RF	0,099	0,010	0,140	0,000	0,440
SMB	0,219	0,215	1,761	-4,660	4,700
HML	-0,003	0,085	2,180	-4,840	7,420
RMW	0,349	0,290	1,438	-3,950	3,890
CMA	0,080	-0,020	1,323	-3,020	5,510
WML	0,827	1,120	3,621	-26,240	10,260
ESG-faktor	-0,041	-0,085	1,866	-6,189	7,291
Kontroversfaktor	0,077	0,111	1,371	-6,134	7,737

Tabell 5 viser korrelasjonen mellom variablene som brukes i porteføljeanalysen. Som beskrevet ovenfor er ESG-faktoren sterkt korrelert med størrelsesfaktoren, SMB. Ellers er også ESG-faktoren og kontroversfaktoren negativt korrelerte, som også ble påpekt ovenfor.

Tabell 5. Korrelasjon mellom ulike variabler som brukes i analysen.

	Markedet	SMB	HML	RMW	CMA	WML	ESG-faktor	Kontroversfaktor
Markedet	1,000							
SMB	-0,105	1,000						
HML	0,495	0,008	1,000					
RMW	-0,329	-0,116	-0,775	1,000				
CMA	-0,278	-0,230	0,285	-0,243	1,000			
WML	-0,420	-0,022	-0,489	0,400	0,158	1,000		
ESG-faktor	-0,095	-0,693	0,101	-0,053	0,485	0,038	1,000	
Kontroversfaktor	0,044	0,209	-0,007	0,008	-0,176	-0,269	-0,392	1,000

Figur 1 viser kumulativ avkastning for de fire porteføljene og markedet over tid. Markedet har høyest vekst, så vidt over *ESG lav*. I de neste avsnittene skal vi undersøke nærmere hva som ligger bak disse forskjellene i avkastning.



Figur 1. Kumulativ avkastning for ulike porteføljer og markedet.

7.1.2 Sortert etter ESG-score

Tabell 6 inneholder regresjoner for porteføljene *ESG høy* og *ESG lav* med ulike forklaringsvariabler. Alfa er signifikant negativ på et 5 prosentnivå for *ESG høy* når vi benytter CAPM. Porteføljen har altså en lavere risikojustert avkastning enn markedet forøvrig. Ettersom det inkluderes flere og flere risikofaktorer reduseres denne risikojusterte avkastningen, og det er kun CAPM som har signifikant alfa. Årsaken til dette er at stadig mer kan forklares av de andre inkluderte risikofaktorene. Dette tyder på at man ikke kan generere meravkastning ved å vekte en portefølje mot selskaper med høy ESG-score, gitt konstant eksponering mot de andre faktorene. I regresjonen der vi inkluderer kontroversfaktorer ser vi at *ESG høy* er negativt eksponert mot denne faktoren. Noe av avkastningen kan altså oppnås tilsvarende ved å investere i selskaper med lav kontroversscore. Dette kan indikere at *ESG høy* i gjennomsnitt består av slike selskaper, altså de som er involvert i mange kontroverser.

Når vi utfører tilsvarende regresjoner for *ESG lav* får alle modellene negativ alfa. Den er imidlertid kun signifikant i modell 2 og 3. Vi ser at *ESG lav* har signifikant positiv eksponering mot SMB og negativ mot HML. Når RMW og CMA inkluderes, er ikke lenger alfa signifikant. Vi ser at CMA blir signifikant og HML går fra å være signifikant til å ikke være signifikant. I modell 5 inkluderes kontroversfaktoren, og selskaper med lav ESG-score har en positiv eksponering mot denne. Dette kan indikere at selskaper med lav ESG-score i gjennomsnitt er utsatt for færre kontroverser, noe vi også så tegn til i den deskriptive statistikken.

Hvis vi ser nærmere på de inkluderte risikofaktorene, kan vi undersøke hva som faktisk skiller porteføljene fra hverandre utover forskjeller i ESG. Størrelsesfaktoren SMB er alltid signifikant på 1 prosentnivå. For *ESG høy* er eksponeringen negativ og for *ESG lav* er den positiv. Igjen antydes sammenhengen mellom størrelse og ESG-score. HML er positiv signifikant for *ESG høy*. Dette taler i utgangspunktet mot hypotesen om at selskaper med høy ESG-score har lavere systematisk risiko. Lavere systematisk risiko tilsier lavere kapitalkostnad og dermed høyere verdivurdering. Selskaper med høy verdivurdering forventes å være negativt eksponert mot HML. *ESG høy* er ikke signifikant eksponert mot RMW, og CMA blir kun signifikant på 10 prosentnivå i modell 5. *ESG lav* er signifikant eksponert mot CMA på et 1 prosentnivå. CMA er negativ for begge porteføljene, noe som kan tyde på at begge porteføljene har overvekt av såkalt aggressive selskaper med høye investeringer.

Tabell 6. Denne tabellen viser koeffisienter og standardfeil fra ulike regresjoner på månedlig avkastning for porteføljer med høy og lav ESG-score. Porteføljene består av henholdsvis de 20 % beste og dårligste selskapene relativt til deres ESG-score. Porteføljene er vektet etter selskapenes markedsstørrelse. Observasjoner angir antall måneder i tidsserien. Gj.snitt. selskaper er antall selskaper i snitt per måned i porteføljen.

<i>ESG høy</i>	(1)	(2)	(3)	(4)	(5)
	CAPM	F&F3	Carhart4	F&F5	F&F5+Kontr
Markedet	1.129*** (0.024)	1.061*** (0.026)	1.055*** (0.027)	1.045*** (0.030)	1.046*** (0.029)
SMB		-0.211*** (0.056)	-0.214*** (0.057)	-0.241*** (0.052)	-0.204*** (0.054)
HML		0.212*** (0.059)	0.196*** (0.057)	0.184** (0.072)	0.194** (0.079)
WML			-0.026 (0.023)		
RMW				-0.114 (0.094)	-0.102 (0.088)
CMA				-0.092 (0.082)	-0.130* (0.074)
Kontroversfaktor					-0.264*** (0.064)
Alfa	-0.189** (0.094)	-0.103 (0.086)	-0.078 (0.089)	-0.040 (0.094)	-0.030 (0.091)
Observasjoner	168	168	168	168	168
Gj.snitt selskaper	50	50	50	50	50
Justert R ²	0.931	0.944	0.944	0.944	0.950
<i>ESG lav</i>	(1)	(2)	(3)	(4)	(5)
	CAPM	F&F3	Carhart4	F&F5	F&F5+Kontr
Markedet	1.031*** (0.037)	1.087*** (0.020)	1.080*** (0.020)	1.044*** (0.023)	1.044*** (0.022)
SMB		0.460*** (0.038)	0.457*** (0.038)	0.414*** (0.040)	0.387*** (0.039)
HML		-0.125*** (0.035)	-0.143*** (0.037)	-0.041 (0.055)	-0.048 (0.053)
WML			-0.030 (0.021)		
RMW				0.017 (0.072)	0.009 (0.069)
CMA				-0.210*** (0.063)	-0.184*** (0.060)
Kontroversfaktor					0.187*** (0.046)
Alfa	-0.017 (0.101)	-0.150** (0.067)	-0.121* (0.070)	-0.105 (0.072)	-0.112 (0.069)
Observasjoner	168	168	168	168	168
Gj.snitt selskaper	421	421	421	421	421
Justert R ²	0.921	0.959	0.959	0.961	0.965

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.1.3 Sortert etter kontroversscore

Tabell 7 viser regresjoner på porteføljer med ulik kontroversscore. For *Kontr høy* har alle de ulike modellene en alfa som ikke er signifikant forskjellig fra null. Det er derfor i utgangspunktet ingen tegn til at man kan oppnå høyere risikojustert meravkastning av å vekte en portefølje mot eller vekk fra selskaper som er utsatt for få kontroverser. Selskaper med høy kontroversscore er i snitt negativt eksponert mot ESG-faktoren. Dette kan tyde på at selskaper som er utsatt for få kontroverser i snitt har lav ESG-score. Dette samsvarer med analysen av høy/lav ESG-score.

Kontr lav har en ikke-signifikant negativ alfa ved bruk av CAPM. Når det inkluderes flere kjente risikofaktorer, går alfa gradvis mot null. Det ser altså ut til at avkastningen også her kan forklares av eksponeringen mot de kjente risikofaktorene. Selskaper med lav kontroversscore er i snitt eksponert positivt mot ESG-faktoren, som også er forventet gitt analysene til nå. Dette kan altså tyde på at selskaper med mange kontroverser i gjennomsnitt har høyere ESG-score.

Ser vi nærmere på risikofaktorene som faktisk kan forklare avkastningen, ser vi at det i stor grad er markedseksponeringen. For *Kontr høy* er det ellers bare WML og den egenkonstruerte ESG-faktoren som er signifikant, og porteføljen er negativt eksponert mot begge. For *Kontr lav* er det flere signifikante faktorer. Blant annet er den signifikant negativt eksponert mot SMB frem til ESG-faktoren inkluderes. En negativ eksponering mot SMB kan indikere at porteføljen generelt består av større selskaper. At større selskaper er mer utsatt for kontroverser kan virke rimelig, da man kan forvente mer oppmerksomhet rundt disse. At SMB ikke lenger er signifikant når vi inkluderer ESG-faktoren, henger trolig sammen med at disse er negativt korrelert, og at ESG-faktoren tar over noe av effekten SMB tidligere hadde. Disse to faktorene er likevel ikke så korrelert at regresjonen ifølge en VIF-test har problemer med multikollinearitet.

Tabell 7. Denne tabellen viser koeffisienter og standardfeil fra ulike regresjoner på månedlig avkastning for porteføljer med høy og lav kontroversscore. Porteføljene består av henholdsvis de 20 % beste og dårligste selskapene relativt til deres kontroversscore. Porteføljene er vektet etter selskapenes markedsstørrelse. Observasjoner angir antall måneder i tidsserien. Gj.snitt. selskaper er antall selskaper i snitt per måned i porteføljen.

<i>Kontr høy</i>	(1)	(2)	(3)	(4)	(5)
	CAPM	F&F3	Carhart4	F&F5	F&F5+WML+ESG
Markedet	1.012*** (0.023)	0.997*** (0.027)	0.963*** (0.035)	0.977*** (0.032)	0.954*** (0.032)
SMB		0.049 (0.052)	0.034 (0.069)	0.024 (0.056)	-0.110 (0.082)
HML		0.062 (0.048)	-0.034 (0.057)	0.087 (0.078)	-0.037 (0.077)
WML			-0.155*** (0.049)		-0.164*** (0.044)
RMW				-0.019 (0.102)	-0.006 (0.096)
CMA				-0.102 (0.088)	0.122 (0.096)
ESG-faktor					-0.222*** (0.069)
Alfa	-0.057 (0.091)	-0.059 (0.092)	0.092 (0.105)	-0.027 (0.102)	0.118 (0.096)
Observasjoner	168	168	168	168	168
Gj.snitt selskaper	273	273	273	273	273
Justert R ²	0.920	0.920	0.933	0.920	0.936
<i>Kontr lav</i>	(1)	(2)	(3)	(4)	(5)
	CAPM	F&F3	Carhart4	F&F5	F&F5+WML+ESG
Markedet	1.030*** (0.025)	0.990*** (0.029)	0.993*** (0.027)	0.986*** (0.027)	1.001*** (0.027)
SMB		-0.146*** (0.051)	-0.145*** (0.050)	-0.169*** (0.048)	-0.019 (0.064)
HML		0.120** (0.049)	0.128** (0.052)	0.042 (0.066)	0.057 (0.067)
WML			0.013 (0.041)		0.026 (0.025)
RMW				-0.173** (0.086)	-0.155* (0.084)
CMA				-0.037 (0.075)	-0.144* (0.080)
ESG-faktor					0.221*** (0.063)
Alfa	-0.112 (0.079)	-0.057 (0.080)	-0.069 (0.093)	0.014 (0.086)	-0.038 (0.086)
Observasjoner	168	168	168	168	168
Gj.snitt selskaper	52	52	52	52	52
Justert R ²	0.937	0.943	0.943	0.944	0.947

Standardfeil i parentes

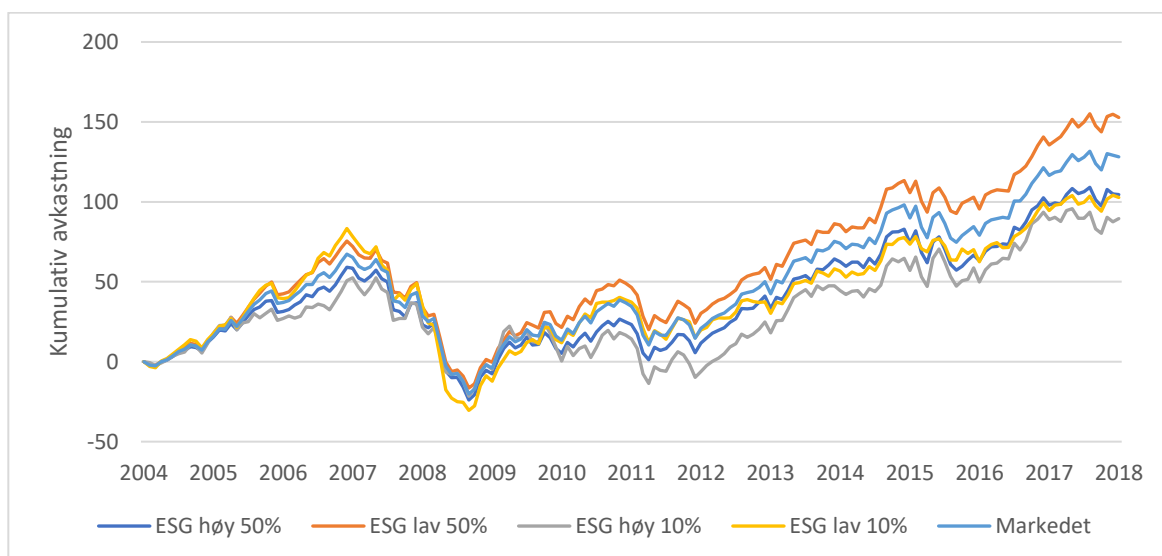
* p < 0.10, ** p < 0.05, *** p < 0.01

7.1.4 Robusthet

For å teste robustheten av resultatene ovenfor har vi i dette kapittelet konstruert porteføljer bestående av 50 % og 10 % av markedet, og gjort en inndeling med like mange selskaper i hver portefølje. I tillegg har vi delt tidsperioden i to for å se om sammenhengen er lik over tid.

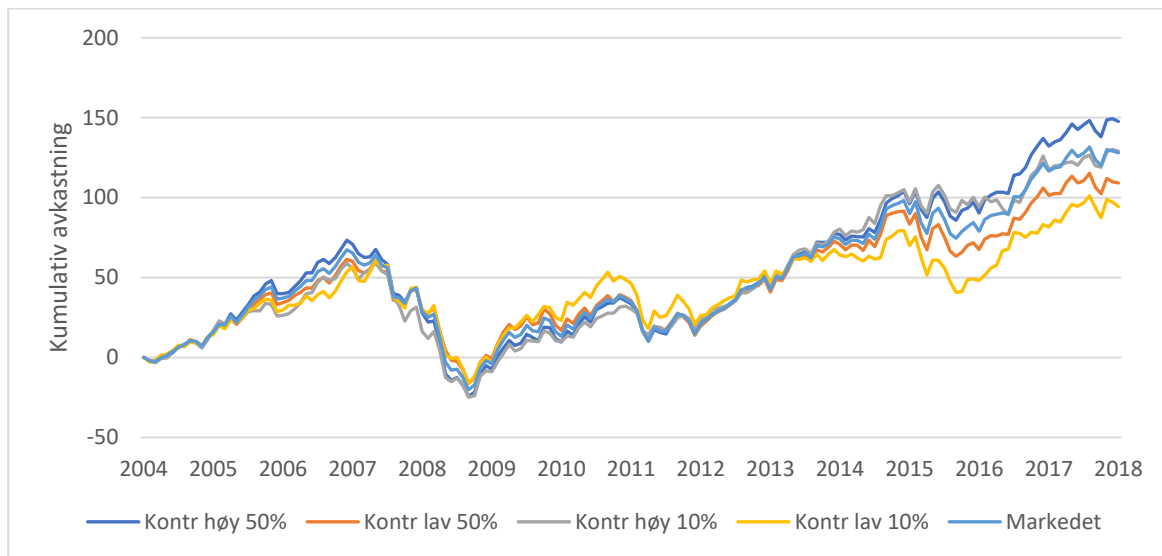
Forskjellige porteføljestørrelser

Figur 2 viser kumulativ avkastning. Man kan se at porteføljen bestående av den halvparten av markedet med lavest ESG-score, *ESG lav 50%*, har høyere kumulativ avkastning enn markedet og selskapene med høy ESG-score. Dette stemmer godt overens med Figur 1, hvor *ESG lav* hadde høyest kumulativ avkastning etter markedet. Det er interessant å observere at *ESG lav 10%* ikke gjør det enda bedre, men derimot omtrent like bra som *ESG høy 50%*. Det ser altså ikke ut til å være en lineær sammenheng mellom ESG-score og avkastning. For porteføljene med høy ESG-score er det derimot indikasjoner på en lineær sammenheng. Her har *ESG høy 10%* en lavere kumulativ avkastning enn den mer inkluderende porteføljen *ESG høy 50%*.



Figur 2. Kumulativ avkastning for ulike ESG-porteføljer og markedet.

For kontrovers ser vi i Figur 3 at halvdelen av markedet med lav score ser veldig lik ut som den med høy ESG-score i Figur 2, og tilsvarende for høy kontroversscore og lav ESG-score. Vi ser tilsvarende sammenheng mellom ESG og kontrovers for de mindre porteføljene, men her gjør *Kontr høy 10%* det litt bedre enn *ESG lav 10%*. I tillegg ser det ut til å være mindre spredning for kontroversporteføljene frem til 2014, enn for ESG-porteføljene.



Figur 3. Kumulativ avkastning for ulike kontroversporteføljer og markedet.

I Tabell 8 kontrollerer vi for de kjente risikofaktorene. Leseren bør være oppmerksom på at *ESG lav 10%* i gjennomsnitt består av hele 259 selskaper, mens *ESG høy 10%* i gjennomsnitt består av kun 21 selskaper. Det er tilsvarende få selskaper i *Kontr lav 10%*. Med så få selskaper i porteføljene, kan det stilles spørsmål om tilfeldigheter kan ha hatt innvirkning, og dermed om hvor robuste resultatene er. De bør derfor tolkes med varsomhet.

For *ESG høy 50%* og *ESG lav 50%* blir alfa tilnærmet lik null, noe som samsvarer med inndelingen på 20 % i Tabell 6. For de mindre porteføljene har *ESG høy 10%* fortsatt ikke-signifikant alfa, mens *ESG lav 10%* får en negativ alfa som er signifikant på 5 prosentnivå. Dette kan tyde på at selskapene med aller lavest ESG-score har noe lavere avkastning, selv etter at det tas hensyn til hvor utsatt de er for de kjente risikofaktorene. Igjen finner vi at selskaper med høy og lav ESG-score er eksponert med ulikt fortegn mot flere av de forskjellige risikofaktorene, og at dette ser ut til å forklare det meste av avkastningsforskjellen.

For kontroversporteføljene er samtlige faktorer signifikante for minst en av modellene. Alfa derimot, er aldri signifikant, selv om *Kontr lav 10%* er markant nærmere enn de andre porteføljene. Det ser altså ikke ut til å være noe meravkastning å hente av å tilte investeringer systematisk mot høy/lav kontroversscore. Begge inndelingene viser igjen at selskaper med mange kontroverser ser ut til å være store, og motsatt for selskaper med få kontroverser.

Tabell 8. Tabellen viser koeffisienter og standardfeil fra regresjoner med Fama & French 5 faktorer og Charharts momentumfaktor på månedlig avkastning for porteføljer med høy og lav ESG-score og porteføljer med høy og lav kontroversscore. Porteføljene er inndelt etter både de 50 % og de 10 % beste og dårligste selskapene relativt til deres ESG-/Kontroversscore. Porteføljene er vektet etter selskapenes markedsstørrelse. Observasjoner angir antall måneder i tidsserien. Gj.snitt. selskaper er antall selskaper i snitt per måned i porteføljen.

ESG	(1)	(2)	(3)	(4)
	ESG høy 50%	ESG lav 50%	ESG høy 10%	ESG lav 10%
Markedet	1.000*** (0.012)	1.000*** (0.012)	1.094*** (0.046)	1.061*** (0.029)
SMB	-0.159*** (0.021)	0.157*** (0.021)	-0.219*** (0.081)	0.556*** (0.050)
HML	0.029 (0.034)	-0.029 (0.033)	0.173 (0.117)	-0.001 (0.072)
WML	-0.041*** (0.013)	0.040*** (0.013)	-0.010 (0.044)	-0.003 (0.028)
RMW	-0.023 (0.038)	0.023 (0.037)	-0.089 (0.146)	0.064 (0.091)
CMA	0.045 (0.042)	-0.044 (0.042)	-0.032 (0.132)	-0.321*** (0.082)
Alfa	0.015 (0.039)	-0.013 (0.039)	-0.031 (0.149)	-0.199** (0.093)
Observasjoner	168	168	168	168
Gj.snitt selskaper	189	691	21	259
Justert R ²	0.988	0.987	0.883	0.945
Kontrovers	(1)	(2)	(3)	(4)
	Kontr høy 50%	Kontr lav 50%	Kontr høy 10%	Kontr lav 10%
Markedet	1.041*** (0.021)	0.961*** (0.021)	1.005*** (0.035)	0.919*** (0.039)
SMB	0.151*** (0.028)	-0.148*** (0.027)	0.124** (0.062)	-0.326*** (0.068)
HML	-0.064* (0.038)	0.060 (0.036)	-0.328*** (0.089)	-0.006 (0.098)
WML	-0.012 (0.017)	0.008 (0.016)	-0.134*** (0.034)	-0.009 (0.037)
RMW	0.014 (0.043)	-0.016 (0.041)	0.140 (0.111)	-0.402*** (0.123)
CMA	-0.012 (0.039)	0.015 (0.037)	0.343*** (0.101)	0.012 (0.112)
Alfa	0.004 (0.026)	-0.001 (0.026)	0.002 (0.114)	0.179 (0.126)
Observasjoner	168	168	168	168
Gj.snitt selskaper	669	212	147	23
Justert R ²	0.984	0.984	0.889	0.884

Standardfeil i parentes

* p < 0.10, ** p < 0.05, *** p < 0.01

Forskjellige tidsinndelinger

Tabell 9 viser resultatene av regresjoner med to ulike tidsperioder. Porteføljene er konstruert på samme måte som i Tabell 6 og Tabell 7, men tidsperioden er delt på midten. Det første tidspunktet er fra starten av juli 2004 til slutten av juni 2011 og det andre tidspunktet er fra starten av juli 2011 til slutten av juni 2018. Kun en av porteføljene har signifikant alfa. Det er *ESG høy* i første periode, som har signifikant negativ alfa på 10 prosentnivå. At denne ikke lenger er signifikant i andre periode, kan tyde på at en eventuell avkastningsforskjell er blitt priset inn. For å teste om de ulike periodene er signifikant forskjellige har vi også utført regresjoner der vi har latt alle koeffisientene variere i de to periodene. Dette har vi gjort ved å multiplisere alle variabler med en tidsdummy for andre periode, samt inkludert en tidsdummy i regresjonen som også lar konstantleddet (alfa) variere. Denne dummyen viser at kun *ESG høy* har signifikant forskjellig alfa i de to periodene, og det på et 10 prosentnivå. Dette samsvarer med Tabell 9. For de resterende porteføljene kan vi altså ikke forkaste en nullhypotese om at periodene har lik alfa.

Tabell 9. Tabellen viser koeffisienter og standardfeil fra regresjoner med Fama & French 5 faktorer og Charharts momentumfaktor på månedlig avkastning for porteføljer med høy og lav ESG-score og porteføljer med høy og lav kontroversscore. Alle porteføljene er delt i to tidsperioder, tid1 og tid2. Porteføljene består av henholdsvis de 20 % beste og dårligste selskapene relativt til deres ESG-/Kontroversscore. Porteføljene er vektet etter selskapenes markedsstørrelse. Observasjoner angir antall måneder i tidsserien. Gj.snitt selskaper er antall selskaper i snitt per måned i porteføljen.

Ulik tidsinndeling	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Tid 1 ESG høy	Tid 2 ESG høy	Tid 1 ESG lav	Tid2 ESG lav	Tid1 Kontr høy	Tid2 Kontr høy	Tid1 Kontr lav	Tid2 Kontr lav
Markedet	0.969*** (0.040)	1.087*** (0.041)	1.089*** (0.033)	0.972*** (0.035)	0.978*** (0.048)	0.954*** (0.039)	1.026*** (0.032)	0.972*** (0.038)
SMB	-0.176** (0.081)	-0.276*** (0.073)	0.431*** (0.055)	0.349*** (0.064)	-0.022 (0.100)	0.117* (0.070)	-0.202*** (0.041)	-0.153** (0.068)
HML	0.488*** (0.106)	-0.024 (0.096)	-0.096 (0.088)	-0.022 (0.084)	-0.072 (0.129)	-0.018 (0.092)	-0.030 (0.090)	0.070 (0.090)
WML	-0.021 (0.023)	0.019 (0.043)	-0.009 (0.028)	-0.005 (0.038)	-0.187*** (0.055)	-0.063 (0.041)	0.024 (0.030)	-0.023 (0.041)
RMW	0.118 (0.186)	-0.338*** (0.120)	-0.000 (0.108)	-0.025 (0.104)	-0.126 (0.159)	0.091 (0.114)	0.027 (0.102)	-0.337*** (0.112)
CMA	0.017 (0.077)	-0.190 (0.116)	-0.116 (0.089)	-0.315*** (0.100)	-0.020 (0.139)	0.108 (0.111)	-0.011 (0.062)	-0.121 (0.108)
Alfa	-0.237* (0.126)	0.041 (0.118)	-0.072 (0.112)	-0.057 (0.102)	0.168 (0.151)	-0.029 (0.113)	-0.013 (0.089)	0.044 (0.110)
Observasjoner	84	84	84	84	84	84	84	84
Gj.snitt selskaper	46	53	364	478	212	334	48	56
Justert R ²	0.950	0.950	0.969	0.942	0.939	0.928	0.949	0.949

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Porteføljer med like mange selskaper

Tabell 10 viser regresjoner på porteføljer sortert etter ESG- og kontroversscore, og inndelt etter antall selskaper istedenfor samlet markedsverdi. Hver portefølje inneholder 20 % av alle selskapene som er med i totalen det året. Det er i gjennomsnitt 176 eller 177 selskaper hvert år i en portefølje. For tre av fire modeller er ikke alfa signifikant. For *ESG lav* er alfa nå signifikant negativ på 1 % nivå. Dette samsvarer med funn fra *ESG lav 10%*, om at de med spesielt lav ESG-score har lavere risikojustert avkastning. Nå som denne porteføljen har enda færre selskaper og dermed er enda mer spisset mot selskaper med lav ESG-score, ser vi at den negative alfaverdien blir enda mer signifikant.

Tabell 10. Tabellen viser koeffisienter og standardfeil fra regresjoner med Fama og French 5 faktorer og Charharts momentumfaktor på månedlig avkastning for porteføljer med høy og lav ESG-score og porteføljer med høy og lav kontroversscore. Porteføljene består av henholdsvis de 20 % beste og dårligste selskapene relativt til deres ESG-/Kontroversscore. Porteføljene er konstruert med like mange selskaper i hver portefølje. Observasjoner angir antall måneder i tidsserien. Gj.snitt. selskaper er antall selskaper i snitt hver måned i porteføljen.

Likt antall selskaper	(1)	(2)	(3)	(4)
	<i>ESG høy</i>	<i>ESG lav</i>	<i>Kontr høy</i>	<i>Kontr lav</i>
Markedet	0.995*** (0.014)	1.064*** (0.029)	0.968*** (0.035)	0.953*** (0.017)
SMB	-0.181*** (0.022)	0.558*** (0.050)	0.113* (0.064)	-0.167*** (0.028)
HML	0.028 (0.034)	0.022 (0.073)	-0.266*** (0.075)	0.037 (0.041)
WML	-0.050*** (0.014)	0.016 (0.028)	-0.136*** (0.042)	0.020 (0.016)
RMW	-0.029 (0.040)	0.135 (0.091)	0.115 (0.107)	-0.043 (0.049)
CMA	0.047 (0.041)	-0.123 (0.082)	0.271*** (0.094)	-0.004 (0.059)
Alfa	0.030 (0.041)	-0.329*** (0.093)	0.037 (0.103)	-0.012 (0.046)
Observasjoner	168	168	168	168
Gj.snitt selskaper	176	177	176	177
Justert R^2	0.987	0.941	0.913	0.980

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.1.5 Sammenfatning porteføljeanalyse

For porteføljer konstruert etter ESG-score startet vi med en innledende analyse som viste at selskaper med lav ESG-score over tid så ut til å utkonkurrere selskaper med høy ESG-score. Etter å ha kontrollert *ESG lav* mot kjente risikofaktorer, så imidlertid avkastningen ut til å bli

forklart av andre forhold enn ESG. Denne trenden så i stor grad ut til å vedvare gjennom hele observasjonsperioden, med unntak av *ESG høy* som hadde signifikant negativ alfa på 10 prosentnivå i første periode. Robusthetsanalysen som inkluderte færre selskaper i porteføljene, viste imidlertid at det kan virke som om selskapene med aller lavest ESG-score har signifikant negativ alfa.

Anta en hypotese om at ESG kan anses som en risikofaktor, hvor selskaper med høy ESG-score har systematisk lavere risiko, og de med lav ESG-score har systematisk høyere risiko. Om dette var priset inn, burde man forvente at en portefølje med lav ESG-score har høyere forventet avkastning som kompensasjon for økt risiko, og motsatt for en med høy ESG-score. Dette kan det være tegn til for *ESG høy* i første periode. Men om dette skyldtes en bevisst prising av risiko i markedet, er det uventet at effekten er blitt borte i andre periode. Vi så heller ikke motsatt effekt for *ESG lav*. En mulig samvittighetspremie kan også forklare mindre avkastning, men det er fortsatt merkelig at effekten er borte i andre periode når stadig flere investorer ønsker å investere ansvarlig.

En annen observasjon som tyder på at markedet ikke priser inn ESG som en risikofaktor, er at vi ser tegn til at de med lavest ESG-score har negativ alfa. Hvis markedet anser selskaper med lav ESG som mer risikable, burde vi sett en positiv alfa. Det er vanskelig å argumentere for at ESG kan anses som en risikofaktor hvor det å ha lav ESG-score, gjør investeringen mindre risikabel. Dette gjør det vanskelig å forklare en negativ alfa ut fra en hypotese om at ESG er priset inn i markedet som en risikofaktor. Hvis ESG ikke er en risikofaktor, og markedet er priset riktig, burde vi ikke fått en signifikant alfa. Hvis ESG derimot er en risikofaktor, og markedet ikke har priset dette inn, så kan det påvirke avkastningen til selskaper med lav ESG-score negativt. Da kan en negativ alfa forklares. Den negative alfaverdien kan derfor implisere at ESG-score fanger opp effekten av underliggende risiko som påvirker de med dårligst ESG negativt, og at dette ikke er priset inn i markedet.

En annen interessant observasjon er forholdet mellom ESG-score og størrelsen på selskaper. ESG-faktoren og SMB har en korrelasjon på -0,693. Det at disse faktorene korrelerer i så stor grad kan skyldes at de består av mange av de samme selskapene. Altså at de som benyttes til å konstruere SMB, også er mange av de samme selskapene som vi benytter til å konstruere ESG-faktoren, og som vi også finner i *ESG høy* og *ESG lav*. Hvis det er vanskelig å skille

disse effektene, kan det bli vanskelig å identifisere en ESG-effekt uavhengig av størrelseeffekten.

Gjennomgående i alle regresjoner har vi negativ eksponering mot vår egenkonstruerte kontroversfaktor for porteføljen med høy ESG-score, og en positiv eksponering for porteføljen med lav ESG-score. Dette er litt merkelig, ettersom det virker rimelig å anta at de som er gode på ESG også unngår kontroverser. En mulig årsak kan som sagt være at de med høy ESG er større, og dermed har mer oppmerksomhet rundt seg. En annen årsak kan være at det stilles høyere forventninger til de som fremmer seg som et selskap som er gode på ESG. Da kan det være mer interessant for media å ta det opp når disse selskapene gjør noe galt. Uavhengig av hva årsaken er, ser vi at selv om faktoren er sterkt signifikant endrer den ikke signifikansnivået til alfa. Faktoren ser heller ut til å bli signifikant fordi den forklarer variasjon som tidligere ble forklart av andre faktorer. Inkludering av de egenkonstruerte faktorene gir oss derfor ikke så mye ny innsikt.

For porteføljer konstruert etter kontroversscore, startet vi med innledende analyser som viser at både høy og lav portefølje har noe lavere avkastning enn markedet. Modeller som benytter kjente risikofaktorer gir ikke-signifikant alfa for alle modellene. Oppdeling i tidsperioder gir heller ikke signifikant alfa for noen av periodene. Når porteføljene konstrueres med like mange selskaper i hver portefølje, eller med en 50 % og 10 % inndeling, får vi igjen tilsvarende resultater med ikke-signifikant alfa. Vi ser altså ingen tegn til at man kan bruke kontroversscore til å utnytte feilprising i kjølvannet av kontroverser, eller at kontroversscore har en effekt på avkastning. Vi ser heller ikke tegn til at markedet hverken anser eller bør anse kontroverser relatert til kontroversscore som en indikator på underliggende systematisk risiko. Markedet fremstår altså ganske effisient med hensyn til kontroverser.

Porteføljeanalysen gir oss økt innsikt i effekten av ESG og kontroverser. Enkelte resultater er imidlertid litt uklare. Samtidig har vi enkelte brudd på normalitetsforutsetningen som skaper noe usikkerhet rundt inferensen. I del to av analysen skal vi benytte oss av en ny innfallsvinkel hvor vi ser på enkelt-selskaper, og ikke på porteføljer. Da får vi forhåpentligvis enda større innsikt og forståelse.

7.2 Individuelt nivå

Deskriptiv statistikk for paneldataene på selskapsnivå er gitt i Tabell 11. Gjennomsnittlig avkastning av å holde et selskap i én måned er 0,8 %, og det samme er medianavkastningen. Standardavviket for avkastningen er stort. Samtlige avkastningstall i denne delen oppgis uten å trekke fra risikofri rente. For scorene knyttet til ansvarlighet, er gjennomsnittet og medianen for kontroversscore mest forskjellig. Dette må skyldes noen observasjoner med betydelig lavere score, som drar gjennomsnittet ned. Dette støttes opp av svært lav minsteobservasjon.

Tabell 11. r_i er månedlig avkastning og er oppgitt i prosentpoeng. ME er markedsverdien av egenkapitalen og er oppgitt i millioner, BE/ME er bok/pris-raten. MOM er en momentumfaktor. ESG-, Kontrovers- og kombinertscore er scorene fra Datastream. Selskaper per måned er antall selskaper per måned i datasettet, observasjoner per selskap er antall månedsobservasjoner per selskap.

	Gjennomsnitt	Median	Std.avvik	Min	Maks
r_i	0,8	0,8	9,597	-35,01	38,86
ME	8789,425	3078,24	16380,55	44,39	112819,1
BE/ME	0,737	0,522	0,848	-0,676	7,22
MOM	0,777	1,043	3,075	-26,552	20,456
ESG-score	55,984	56,61	16,127	6,66	96,39
Kontroversscore	49,045	58,33	20,921	0,09	87,5
Kombinertscore	49,941	48,26	15,513	5,39	93,92
Selskaper per måned	867,607	930,5	165,019	416	1058
Observasjoner per selskap	111,949	124	53,75	2	168

Tabell 12 viser korrelasjonen mellom ulike faktorer som brukes som forklaringsvariabler. Faktorene som korrelerer mest er ESG-score mot $\ln(\text{ME})$ og ESG-score mot kombinertscore. Det at ESG-score korrelerer med $\ln(\text{ME})$, som er en proxy for størrelse, samsvarer med funn fra porteføljeanalysen.

Tabell 12. Korrelasjon mellom ulike forklaringsvariabler.

	ln(ME)	ln(BE/ME)	MOM	ESG-score	Kontroversscore	Kombinertscore
ln(ME)	1,000					
ln(BE/ME)	-0,197	1,000				
MOM	0,070	-0,026	1,000			
ESG-score	0,455	-0,014	0,005	1,000		
Kontroversscore	-0,348	-0,034	-0,004	-0,313	1,000	
Kombinertscore	0,137	-0,033	0,003	0,701	0,416	1,000

7.2.1 Paneldataanalyse

I Tabell 13 vises resultatene fra 9 regresjoner på paneldatanivå. Standardfeilen er konstruert ved å clustre på tidsdimensjonen. Avhengig variabel i samtlige modeller er månedlig avkastning for selskap i . Regresjonene viser hva som kan forklare avkastningen og inkluderer gradvis flere forklaringsvariabler. Den faste tidseffekten er fjernet ved å inkludere tidsdummyer for alle perioder. En innledende observasjon er at variablene for ansvarlighet ikke er signifikante i noen av regresjonene.

I modell 4 er alle kontrollvariabler inkludert. Her ser vi at ME er negativ, men ikke signifikant. Sammenheng mellom størrelse og avkastning er diskutert av flere, og den mest utbredte oppfatning er at størrelse ikke lenger har effekt på avkastning (Ang, 2014). En litt merkelig sammenheng som kan observeres i dataene er at BE/ME er negativ. Koeffisienten er riktignok ikke signifikant, men for enkelte av modellene er det ikke mye som skiller den fra et signifikansnivå på 10 prosent. I prisingsmodeller er denne sammenhengen vanligvis positiv. For å se nærmere på denne negative koeffisienten, utførte vi også regresjonene uten finansselskaper, siden disse ofte har en annen sammenheng mellom BE/ME og avkastning enn andre selskaper. Disse regresjonene gav BE/ME-koeffisienter som var nærmere null og enda mindre signifikante (se appendiks Tabell 17). Den siste kontrollvariabelen er MOM, som er gjennomsnittlig avkastning for selskapet de foregående 12 månedene. Denne er signifikant positiv i samtlige modeller.

I modell 5 inkluderes ESG-score. Koeffisienten er positiv, noe som peker i retning av at høyere ESG-score er forbundet med høyere avkastning. Koeffisienten er imidlertid svært lav, og ikke signifikant. Det kan derfor ikke vektlegges at den er positiv. I modell 6 ser vi at kontroversscore har en negativ koeffisient, men heller ikke denne er signifikant. Vi kan altså ikke se noen

klar sammenheng med avkastning, hverken for ESG-score eller kontroversscore. Modell 7 benytter kombinertscore. Denne vil fungere litt som et mer avansert interaksjonsledd, hvor kontroversscoren kan få innvirkning om den er lav, men ikke om den er høy. Vi ser at denne variabelen alene ikke er signifikant, og dermed ikke er assosiert med hverken økning eller reduksjon i avkastningen. I modell 8 er både ESG- og kontroversscore inkludert. Selv om begge variablene er inkludert, ser vi at koeffisientene og standardfeilene er uforandret relativt til å være eneste inkluderte variabel for ansvarlige forhold. Følgelig er begge fortsatt ikke-signifikante. Den største effekten av å inkludere faktorene ser ut til å være at koeffisienten til ME øker. Likevel er den fortsatt utenfor et signifikansnivå på 10 prosent. Dette kan ha å gjøre med korrelasjonen mellom koeffisientene som sees i Tabell 12, og det faktum at selskaper med høy ESG-score ser ut til å være store. For eksempel vil et stort selskap nå få marginalt lavere avkastning som følge av stor ME, men høy ESG-score vil trekke avkastningen marginalt opp igjen. I regresjon 9 er alle variablene for ansvarlighet med. Det å inkludere kombinertscoren ser ut til å ha høy effekt på de andre variablene for ansvarlighet. Kombinertscore er som vi ser i Tabell 12, i stor grad korrelert med ESG- og kontroversscore, noe vi også ser i utslagene på standardfeilene til de estimerte koeffisientene. En VIF-test på denne modellen viser at vi har problemer med multikollinearitet. Resultatene fra den regresjonen bør derfor tolkes med varsomhet.

Siden vi i porteføljeanalysen fant sammenheng mellom ESG og avkastning for de selskapene med aller lavest ESG-score, utførte vi også regresjonen i modell 8 i Tabell 13 på selskaper med ESG-score under ulike nivåer. Disse regresjonene ga heller ingen systematiske sammenhenger eller signifikante resultater for ansvarlighet. Regresjonene er vedlagt i appendiks Tabell 18.

Tabell 13. Koeffisienter og standardfeil fra regresjoner av månedlig aksjeavkastning på markedsverdi (ME), bok-pris (BE/ME), momentum (MOM), ESG-score (ESG), Kontroversscore (Kontr) og kombinertscore (Komb). Alle uavhengige variabler med unntak av MOM er årlige observasjoner for hvert selskap som benyttes til å forklare månedlig avkastning. BE i BE/ME er fra regnskapsåret som avsluttes i år t-1 og benyttes for avkastningsobservasjoner fra juli år t til juni år t+1. For samme periode er ME som benyttes i BE/ME målt i slutten av år t-1. ME som benyttes alene er målt i slutten av juni år t. MOM er gjennomsnittlig avkastning de foregående 12 månedene før måned t. ESG-, Kontrovers- og Kombinertscore er hentet fra regnskapsåret som ble avsluttet i år t-1. For å unngå ekstremverdier er 0,5 % av de høyeste (laveste) observasjonene for ME, BE/ME og avkastning redusert (økt) til de er like den høyeste (laveste) observasjonen som ikke endres. For modellene er det også inkludert tidsdummyer for samtlige måneder.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	ME	BE/ME	MOM	Samlet	ESG	Kontrovers	Kombinert	ESG + Kontrovers	Interaksjon
ln(ME)	-0.007 (0.052)			-0.043 (0.049)	-0.054 (0.052)	-0.048 (0.053)	-0.043 (0.049)	-0.057 (0.055)	-0.060 (0.055)
ln(BE/ME)		-0.151 (0.111)		-0.140 (0.105)	-0.142 (0.104)	-0.142 (0.104)	-0.140 (0.105)	-0.143 (0.104)	-0.144 (0.104)
MOM			0.150** (0.058)	0.148** (0.058)	0.149** (0.058)	0.149** (0.058)	0.148** (0.058)	0.149** (0.058)	0.149** (0.058)
ESG					0.002 (0.003)			0.002 (0.003)	0.007 (0.008)
Kontr						-0.001 (0.001)		-0.001 (0.001)	0.002 (0.004)
Komb							0.000 (0.002)		-0.005 (0.008)
Observasjoner	145758	142949	135091	132501	132501	132501	132501	132501	132501
Justert R ²	0.222	0.225	0.223	0.226	0.226	0.226	0.226	0.226	0.226

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.2.2 Robusthet

Valg av statistisk metode

Den første regresjonen i Tabell 14 er en OLS-regresjon med White (1980, 1982) standardfeil. De andre modellene i tabellen sammenliknes i utgangspunktet med denne. Videre er det én OLS-regresjon hvor vi har clustret på selskap, og én hvor vi har clustret på tid. Modell 4 er dobbeltclustret, altså både på selskapsdimensjonen og tidsdimensjonen. Hensikten med å se på disse ulike modellene, er å undersøke om det er ulikheter i standardfeilene på tvers av modellene. Det vil i så fall være et resultat av avhengigheter mellom observasjonene i modellene, som det må tas hensyn til. Når vi sammenligner de ulike clusteringmetodene med modell 1, som benytter White standardfeil, vil eventuelle forskjeller skyldes avhengighet innad i clusterene, og ikke heteroskedastisitet (Petersen, 2009). Modell 5 viser en vanlig Fama-Macbeth (1973) regresjon som kontroll for clusteringen på tid. Modell 6 er en Fama-Macbeth regresjon med Newey-West (1987) standardfeil, som er robust for autokorrelasjon og heteroskedastisitet, hvor vi har inkludert 12 lags. Fama-Macbeth regresjon er en metode som består av to steg. Man gjennomfører først tverrsnittregresjoner for hvert tidspunkt, og deretter regner man ut gjennomsnittet av koeffisientene fra tverrsnittregresjonene (Fama og MacBeth, 1973). Denne metoden er i utgangspunktet robust ved tverrsnittavhengighet, men vil gi bias i standardfeilene ved autokorrelasjon (Petersen, 2009).

I første regresjon, hvor vi bruker White standardfeil, tas det ikke hensyn til eventuelle selskapseffekter eller varierende tidseffekter. Tidsdummyene vil ta hensyn til en eventuell fast tidseffekt. I modell 2, der det clustres på selskap, tillates heller ikke varierende tidseffekt, men fast og varierende selskapseffekt tillates gjennom clustering på selskap. Dersom vi sammenlikner standardfeilen i modell 1 og modell 2, ser vi at standardfeilene er omtrent identiske. Dette tyder på at det ikke eksisterer en selskapseffekt (Petersen, 2009). I modell 3 clustrer vi på måned, noe som tar hensyn til både fast og varierende tidseffekt. Denne modellen tillater ikke selskapsavhengigheter, men fra sammenligningen av modell 1 og 2 vet vi at denne effekten er liten, om den i det hele tatt eksisterer. I denne modellen ser vi at standardfeilen øker mer i forhold til modell 1, enn det modell 2 gjorde. Nå er standardfeilen mer enn to ganger større for BE/ME og MOM. Denne forskjellen tyder på at vi har betydelig varierende tidseffekt i modellen, altså en tidseffekt som påvirker selskapene ulikt.

I modell 4 clustrer vi etter begge dimensjoner, både over tid og på selskap. Standardfeilen til denne modellen er tilnærmet lik modell 3. Dette kan skyldes enten at det ikke er noen

selskapseffekt, eller at det er få selskaper per måned. Men vi har et betydelig antall selskaper per måned (fra 416 selskaper og oppover), noe som styrker vår tidligere mistanke om at det ikke eksisterer noen betydelig selskapseffekt. Fama-MacBeth regresjonen i modell 5 tar også hensyn til tidsavhengigheter. Fra modellen ser vi at standardfeilen (og koeffisientene) endres noe for enkelte av variablene relativt til modell 1. Forklaringsvariablene ME, BE/ME og MOM er ikke uavhengige over tid, og Petersen (2009) argumenterte for at dette er noe som kan gjøre at Fama-MacBeth standardfeil blir for lave. Ved å rapportere standardfeil justert for autokorrelasjon og heteroskedastisitet konsistent med Newey-West (1987) i modell 6, ser vi at standardfeilen blir noe større, men ikke drastisk. Denne modellen gir resultater som er relativt like de vi ser i modell 3 og 4, noe som styrker antakelsen om at modellen har en fast og varierende tidseffekt, men ikke en selskapseffekt. At koeffisientene er noe forskjellige for OLS regresjonene og Fama-MacBeth regresjonene er naturlig, da de behandler observasjoner ulikt. Fama-Macbeth vekter alle tidsperioder likt, uavhengig av antall observasjoner i hver enkelt tidsperiode.

Tabell 14. Avhengig variabel i samtlige modeller er månedlig avkastning for selskapene. Modell 1-4 viser OLS koeffisienter og standardfeil for månedlige regresjoner. OLS-regresjonene inneholder månedsdummier. Modell 1 benytter White standardfeil, modell 2 har standardfeil clustret etter selskap, modell 3 har standardfeil clustret etter tid, modell 4 har standardfeil clustret etter både tid og selskap. Modell 5 og 6 inneholder koeffisienter og standardfeil estimert etter Fama-MacBeth metode. I modell 6 er standardfeilene robuste i henhold til Newey-West metode.

	(1)	(2)	(3)	(4)	(5)	(6)
	White	Selskap	Tid	Selskap+Tid	FMB	FMB-NW
ln(ME)	-0.057** (0.023)	-0.057** (0.025)	-0.057 (0.055)	-0.057 (0.056)	-0.071 (0.052)	-0.071 (0.056)
ln(BE/ME)	-0.143*** (0.030)	-0.143*** (0.033)	-0.143 (0.104)	-0.143 (0.104)	-0.115 (0.083)	-0.115 (0.115)
MOM	0.149*** (0.014)	0.149*** (0.016)	0.149** (0.058)	0.149** (0.058)	0.175*** (0.044)	0.175*** (0.045)
ESG	0.002 (0.002)	0.002 (0.002)	0.002 (0.003)	0.002 (0.003)	0.002 (0.002)	0.002 (0.002)
Kontr	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Observasjoner	132501	132501	132501	132501	132501	132501
Justert R^2	0.226	0.226	0.226	0.226		

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Vi utførte også regresjonsanalyser hvor vi clustret etter industri i stedet for selskap, og hvor vi dobbeltclustret etter industri og måned. Denne metoden så ut til å fange opp litt mer avhengighet mellom observasjonene enn det clustring etter selskap gjorde. Dette tyder på at

de uavhengige variablene og residualene kan være noe mer korrelert innad i en industri, enn innad i individuelle selskaper. Selv om denne avhengigheten virker større, er den fortsatt ikke stor nok til å konkludere med at det er en industrieffekt i dataene. I tillegg kan disse regresjonene ha problemer med at antall clustere er i minste laget, noe som reduserer troverdigheten til resultatene. Å clustre etter industri gav uansett ingen endringer i resultatene med tanke på effekt av ansvarlige investeringer.

To tidsperioder

Tabell 15 viser de samme regresjonene som i modell 5, 6 og 8 i Tabell 13, men for to ulike tidsperioder. De to tidsperiodene er de samme som vi brukte i porteføljeanalysen; de første 84 månedene i Tid1, og de siste 84 månedene i Tid2. Kontrollvariabelen ME er ikke signifikant i noen av periodene. MOM er signifikant i andre periode, men ikke i første. Dette kan ha å gjøre med store og uvanlige svingninger under og rett etter finanskrisen. BE/ME ser ut til å ha litt større effekt i andre periode, og standardfeilen er også mindre, selv om koeffisienten fortsatt ikke er signifikant. Dette stemmer overens med at HML faktoren har gjort det dårlig i tiden etter finanskrisen (Ang, 2014).

Ingen av regresjonene har signifikant ESG-score. Kontroversscore er signifikant på 5 prosentnivå i første tidsperiode, både i modellen med og uten ESG inkludert. For andre tidsperiode er hverken ESG- eller kontroversscore signifikant. For modellene utførte vi også en Chow-test (Chow, 1960) for å teste om koeffisientene for ESG- og kontroversscore er forskjellige i de to periodene. For ESG alene kan vi ikke forkaste at den er lik i de to periodene. For kontrovers kan vi forkaste at koeffisientene er like på et 5 prosent signifikansnivå. For ESG og kontrovers samlet, kan vi også forkaste en nullhypotese om at koeffisientene er like i de to periodene med et signifikansnivå på 5 prosent.

I modell 3 kan man se at kontroversscore i første periode har en koeffisient på -0,006. En forskjell i score på 87,41, som er forskjellen på den minste og største av våre observasjoner, vil da ifølge modellen føre til en avkastningsforskjell på omtrent 0,5 prosentpoeng månedlig. En mer normal forskjell i kontroversscore på 50, vil ifølge modellen gi en avkastningsforskjell på 0,3 prosentpoeng månedlig. Dette er en relativt markant forskjell.

Tabell 15. Avhengig variabel i samtlige modeller er månedlig avkastning for selskapene. Regresjonene tar for seg to ulike tidsperioder (Tid1 og Tid2), heholdsvise de første og siste 84 månedene i observasjonsperioden. Første periode blir da juli 2004 til juni 2011, og andre periode blir juli 2011 til juni 2018. Standardfeil er konstruert ved å clustre på tid. For begge tidsperioder er det tre regresjoner, en med ESG-score, en med kontroversscore og en med ESG- og kontroversscore. Samtlige regresjoner inneholder månedsdummyer.

	(1)	(2)	(3)	(4)	(5)	(6)
	Tid1	Tid2	Tid1	Tid2	Tid1	Tid2
	ESG	ESG	Kontr	Kontr	ESG + Kontr	ESG + Kontr
ln(ME)	-0.134 (0.118)	-0.026 (0.056)	-0.135 (0.113)	-0.012 (0.059)	-0.152 (0.122)	-0.017 (0.059)
ln(BE/ME)	-0.108 (0.206)	-0.138 (0.123)	-0.111 (0.207)	-0.133 (0.123)	-0.114 (0.206)	-0.134 (0.122)
MOM	0.047 (0.124)	0.196*** (0.062)	0.046 (0.124)	0.195*** (0.062)	0.047 (0.124)	0.195*** (0.062)
ESG	0.005 (0.006)	0.001 (0.003)			0.004 (0.005)	0.001 (0.003)
Kontr			-0.006** (0.002)	0.002 (0.002)	-0.005** (0.002)	0.002 (0.002)
Observasjoner	44543	87958	44543	87958	44543	87958
Justert R ²	0.310	0.169	0.310	0.169	0.310	0.169

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

7.2.3 Oppsummert enkeltselskaper

Vi ser altså fra Tabell 13 at for hele perioden er både ESG- og kontroversscore ikke-signifikante. For inndelingen i ulike tidsperioder konkluderte vi med at kontroversscore er signifikant negativ i første periode, og ikke i andre periode. ESG-score er ikke signifikant i noen av periodene. Signifikant negativ kontroversscore vil si at høy kontroversscore, altså å være utsatt for få kontroverser, er assosiert med lavere avkastning.

7.3 Diskusjon av resultater

I begge analysene finner vi altså i utgangspunktet at en avkastningseffekt av ESG- og kontroversscore ikke lenger er signifikant når det inkluderes et større antall kontrollvariabler, og når hele tidsperioden legges til grunn. Når analysene spesifiseres i robusthetsdelen, både i mindre porteføljer og over ulike tidsrom, blir noen sammenhenger signifikante. Man kan se tegn til at selskapene med absolutt dårligst ESG-score har lavere avkastning. Det kan også ha vært en negativ sammenheng mellom kontroversscore og avkastning i første periode, som ikke lenger er signifikant i andre periode. Når vi først finner signifikante sammenhenger, er disse likevel ikke entydige på tvers av de analytiske metodene, noe som svekker vår tillit til disse

funnene. En forklaring på at vi ikke finner tilsvarende resultater i porteføljeanalysen og i selskapsanalysen kan være ulikt datagrunnlag. I porteføljeanalysen ser vi på ulike utvalg fra datasettet, og observasjoner utenfor porteføljer blir ikke tatt med. I selskapsanalysen er alle selskaper vi har tilstrekkelig informasjon om med i hver periode. Et annet element som spiller inn, er at i porteføljeanalysen er avkastningen verdivektet, noe den ikke er i selskapsanalysen. ESG og kontroverser hadde klare relasjoner til størrelse. Når dette behandles forskjellig i analysedelen, kan dette kanskje forklare noen av forskjellene i resultatene.

Våre resultater ser ut til å støtte Revelli og Viviani (2015) sine funn om at det ikke er noen reelle fordeler eller ulemper forbundet med ansvarlige investeringer. De hevder at studiers resultater avhenger av metodevalg og data. Deres studier så riktignok på ESG, men vi finner det samme også for kontroversscore. Dette er riktignok før man tar hensyn til eventuelle kostnader assosiert med å gjennomføre en analyse av selskapers ansvarlighet. Resultatene våre ser også ut til å samsvare med det Friede, Busch og Bassen (2015) fant i deres metastudie. De viser at den store majoriteten av tidligere studier, hverken finner positiv eller negativ sammenheng mellom ansvarlige investeringer og avkastning i Europa. At vi likevel finner noen signifikante resultater, kan være med på å forklare hvorfor noen studier konkluderer med at det er signifikant positiv eller negativ sammenheng.

I litteraturgjennomgangen ble det nevnt at Zhang (2017) hevdet det ikke var noen ulemper ved å investere ansvarlig. Hun hevdet derimot at det muligens kunne gi merverdi, men at dette kanskje allerede var priset inn. Om dette var priset inn, burde risikokompensasjon sørge for at porteføljene med høy ESG-score fikk negativ alfa og de med lav ESG-score fikk positiv alfa. Vi ser som sagt ikke sterke indikasjoner på at dette er tilfellet. Det vi derimot ser, er at de som er aller dårligst med hensyn til ESG, har noe lavere avkastning. Dette kan tyde på at ESG representerer en risikofaktor som særlig påvirker de med aller lavest score negativt, og at dette ikke er priset inn. Selskapsanalysen viser imidlertid at ESG ikke er assosiert med avkastning, heller ikke når vi kun ser på de med lavest ESG-score. Hvis ESG var assosiert med risiko, burde vi sett en positiv sammenheng mellom ESG-score og avkastning, gitt at risikoen ikke var priset inn, og en negativ sammenheng om den var priset inn. Andre potensielle sammenhenger enn den mellom risiko og ESG, burde også vært synlige i selskapsanalysen. Dette tyder for eksempel på at det ikke er feilprising som følge av en samvittighetspremie, da det også burde ført til en sammenheng mellom avkastning og ESG.

For kontroverser relatert til ESG, som det er relativt få studier på, viser våre analyser på porteføljer ingen signifikante resultater. Dette gjelder både på aggregert nivå, og på et mer spesifikt nivå. Det eneste signifikante resultatet for kontroverser er i første periode for enkeltelskaper, hvor høyere kontroversscore, altså å være utsatt for få kontroverser, er assosiert med negativ avkastning, signifikant på 1 prosentnivå. Dette funnet er i strid med Garvey et al. (2016), som hevder at kontroverser var assosiert med lavere avkastning. Aouadi og Marsat (2016) konkluderte med at kontroverser er forbundet med høyere avkastning for "high-attention" selskaper, men ikke når andre forhold for ansvarlighet kontrolleres for. Våre signifikant negative resultater, som illustrerer en positiv sammenheng mellom avkastning og kontroverser, er derimot også signifikant når vi kontrollerer for ESG-score. Vi har heller ikke utelukket "low-attention" selskaper, noe som tyder på at sammenhengen kan være sterkere enn det Aouadi og Marsat fant.

I avsnitt 2.2.2 så vi på flere mulige årsaker til at kontroverser kan gi meravkastning. Eksempelvis feilprising i kjølvannet av kontroverser, og påfølgende endringer i ESG som øker selskapsverdien i etterkant av kontroverser. Siden vi ikke har funnet indikasjoner på at ESG er priset inn som en systematisk risikofaktor, er det lite trolig at årsaken skyldes endringer i ESG-score. Dette gjør det rimelig å anta at sammenhengen kan skyldes justering tilbake mot fundamentalverdi etter feilprising i kjølvannet av kontroverser. Gitt en slik tolkning kan det også tenkes at årsaken til at sammenhengen forsvinner i andre periode, skyldes at markedet er blitt bedre til å unngå slik feilprising. En tredje mulig årsak er at kontroverser systematisk hender oftere for enkelte selskaper enn for andre. Blackrock (2018) hevdet for eksempel at antall nye kontroverser var positivt korrelert med antall tidligere kontroverser. Det kan argumenteres for at investorer bør kompenseres for risiko for slike hendelser. En negativ sammenheng mellom kontroversscore og avkastning, kan forsvares av en slik teori om kompensasjon for risiko. Hvis dette er tilfellet er det imidlertid merkelig at effekten er borte i andre periode. Det bør også påpekes at når vi ikke finner noen signifikant effekt for porteføljene, hverken i første eller andre periode, så svekker det troverdigheten til resultatene og skaper tvil om det faktisk har eksistert en merkbar sammenheng. For porteføljene kontrollerte vi riktignok for flere risikofaktorer under periodeinndelingen, som kan tenkes å fange opp effekten om det er korrelasjon mellom disse faktorene og kontroversscore.

7.4 Svakheter

Når vi utelukker selskaper som ikke oppgir informasjon om ansvarlige forhold, går vi fra å ha data om 16857 selskaper, til å ha data om 1338 selskaper. Utelukkes også de der det ikke oppgis nok data om ME, BE og avkastning i Datastream, reduseres antallet til 1302 selskaper. Vi mister altså et stort antall selskaper som følge av manglende data, og da særlig som følge av manglende ESG- og kontroversdata. For å undersøke om dette har en stor effekt på resultatene, utførte vi også regresjoner på paneldata av selskapene i Europa før vi fjernet de selskapene som ikke har ESG- og kontroversscore i Datastream. Dette gav veldig forskjellige resultater (se appendiks Tabell 16) sammenlignet med modell 4 i Tabell 13. Dette kan tyde på at selskapene vi benytter i våre analyser ikke er representative for det europeiske markedet. Hvis vårt utvalg ikke kan gjenskape kjente sammenhenger mellom risikofaktorer og avkastning, så skaper det tvil om vi er i stand til å vise hvordan ESG og kontroverser faktisk påvirker avkastning. Hvis det er tilfelle kunne vi fått helt andre resultater dersom vi hadde hatt et mer representativt utvalg. Utvalget vårt består av de fleste selskapene i Europa som Datastream har ESG-data for. Hvis dette utvalget ikke er representativt, skaper det ikke bare tvil om våre resultater, men også om tidligere studier på ESG og avkastning. I hvert fall for studier som har basert seg på data fra Datastream. Dette problemet kan også være tilstede andre steder enn i Europa. Thomson Reuters har riktignok et større og muligens mer representativt utvalg i Nord-Amerika, men det utelukker likevel ikke at man kan ha lignende problemer der også.

En forklaring på forskjellene kan være at selskaper som rapporterer nok data til at det kan settes sammen til en ESG-score, er systematisk forskjellige fra de som ikke gjør det. En annen forklaring er kapasitetsbegrensninger når det kommer til det å samle selskapsdata for å konstruere slike scorer. I så fall kan forklaringen på forskjellene være at de skyldes tilfeldigheter, eller at Datastream fokuserer på selskaper som ikke representerer et tilfeldig utvalg fra det totale markedet.

En annen svakhet med analysen er at vi kun har årlige verdier for forhold knyttet til ansvarlighet. Spesielt knyttet til kontroverser kan det være misvisende at det kan være lang tid mellom den kontroversscoren som blir brukt for et selskap i en måned, og når denne kontroversen faktisk ble kjent. Hvis det hadde vært hyppigere målinger, kunne man blant annet fått et klarere bilde av utviklingen i avkastning som følge av kontroverser.

8. Konklusjon

I denne oppgaven har vi undersøkt hvilken effekt ansvarlige forhold har hatt på selskapers avkastning i Europa de siste 14 årene. Effekten har blitt undersøkt både ved å se på enkeltelskaper og ved å se på egenkonstruerte porteføljer. Det er benyttet både ulike porteføljeinndelinger og ulike tidsperioder for å se om disse forholdene kan påvirke resultatene. Som mål på et selskaps ansvarlige forhold har vi benyttet ESG- og kontroversscore utarbeidet av Thomson Reuters. Spesifikt har vi undersøkt effekten av ulike score-nivåer, og ikke endring i scoren eller om det har noen effekt å i det hele tatt ha en score. Datasettet vårt er omfattende for å være en studie av ansvarlige investeringer i Europa, med over 1300 selskaper i utvalget. Det utelukkes likevel mange selskaper, da majoriteten av selskapene i Europa ikke har en ESG-score eller kontroversscore.

For å analysere eventuelle sammenhenger benyttes det flere metoder og beregninger for å ta høyde for potensielle problemer i datasettene. Porteføljekonstruksjonen, datavaskingen og porteføljeanalysen følger standarder i litteraturen. Paneldataanalysen tar hensyn til nyere økonometriske diskusjoner knyttet til analyser og publiserte artikler innen finans.

Vi fant hovedsakelig to sammenhenger der ansvarlige forhold så ut til å ha innvirkning på avkastning. Den første var at selskaper med spesielt lav ESG-score var assosiert med lavere avkastning. Dette var riktignok en effekt som kun ble fanget opp i porteføljeanalysen. Denne sammenhengene var vanskelig å forklare gitt risikokompensasjon, men en mulig årsak kan være at selskapene med lavest ESG-score har høyere risiko, og at dette ikke er priset inn. Den andre sammenhengene så vi i analysen av enkeltelskaper, og viste at kontroverser har hatt en positiv effekt på avkastning i perioden juli 2004 til juni 2011. Denne sammenhengene ble ikke observert i porteføljeanalysen. En mulig forklaring på dette funnet, kan være at det er kompensasjon for risiko. En annen kan være at eksponering for kontroverser fører til midlertidig fall i prisen under fundamentalverdi, slik at den forsinkede kontroversscoren fanger opp en justering i prisen tilbake mot fundamentalverdi. Analysene var uansett entydige på at denne sammenhengene ikke har eksistert i perioden etter juni 2011 og frem til i dag.

De signifikante sammenhengene vi finner mellom ansvarlighet og avkastning er altså ikke entydige på tvers av de ulike metodene. Mangel på konsistens gjør det urimelig å konkludere med at det er en effekt av ESG eller kontroverser på avkastning. Det at vi faktisk finner noen

signifikante resultater gjør det imidlertid vanskelig å utelukke helt at det eksisterer, eller har eksistert, en form for sammenheng.

Videre vil vi anbefale andre som ønsker å se nærmere på ansvarlige investeringer, å undersøke om selskapene med lavest ESG-score faktisk er mer risikable og har lavere risikjustert avkastning. Det kan også være interessant å undersøke hva effekten vil bli på en porteføljes diversifisering om disse selskapene utelukkes. Et annet moment det bør sees nærmere på, er forskjellene mellom selskapene som rapporterer om ansvarlige forhold og har en score for dette, og resten av selskapene i markedet. Hvorfor ser det ut som de har forskjellige eksponeringer mot de kjente risikofaktorene? Å ha en ESG-score kan være en faktor som gjør flere investorer oppmerksomme på de ansvarlige forholdene. Vil dette kunne føre til en annen sammenheng mellom ansvarlighet og avkastning enn for resten av markedet?

9. Litteraturliste

Amir, E., Carabias, J. M., Jona, J. & Livne, G. (2016). Fixed-effects in empirical accounting research.

Ang, A. (2014). *Asset Management: A Systematic Approach to Factor Investing*. Oxford University Press.

Ang, A., Liu, J. & Schwarz, K. (2017). Using individual stocks or portfolios in tests of factor models.

Angrist, J. D. & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press Princeton.

Aouadi, A. & Marsat, S. (2016). Do ESG controversies matter for firm value? Evidence from international data. *Journal of Business Ethics*, 1-21.

Bachelier, L. (1900). Theory of speculation. *The random character of stock market prices*.

Black, F. (1993). Estimating expected return. *Financial analysts journal*, 36-38.

Blackrock. (2018). *Sustainable investing: a "why not" moment - Environmental, social and governance investing insights*. Retrieved from Blackrock.com:

<https://www.blackrock.com/institutions/en-us/literature/whitepaper/bii-sustainable-investing-may-2018-us.pdf>

Blitz, D., Hanauer, M. X., Vidojevic, M. & van Vliet, P. (2016). Five Concerns with the Five-Factor Model.

Bodie, Z., Kane, A. & Marcus, A. J. (2011). *Investment and portfolio management*: McGraw-Hill Irwin.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers*, 17(31), 334-355.

Breusch, T. S. & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the econometric society*, 1287-1294.

- Carhart, M. M. (1997). On persistence in mutual fund performance. *The journal of finance*, 52(1), 57-82.
- Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica: Journal of the econometric society*, 591-605.
- Cook, R. D. & Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, 70(1), 1-10.
- Davis Polk. (2017). *ESG Reports and Ratings: What They Are, Why They Matter?* Retrieved from davispolk.com:
- Dickey, D. A. & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a), 427-431.
- Durbin, J. & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, 38(1/2), 159-177.
- Eccles, R. G., Ioannou, I. & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835-2857.
- El Ghoul, S., Guedhami, O., Kwok, C. C. & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, 35(9), 2388-2406.
- Fama, E. F. & French, K. R. (1992). The cross-section of expected stock returns. *The journal of finance*, 47(2), 427-465.
- Fama, E. F. & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.
- Fama, E. F. & French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of financial economics*, 105(3), 457-472.
- Fama, E. F. & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1), 1-22. doi:<https://doi.org/10.1016/j.jfineco.2014.10.010>
- Fama, E. F. & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3), 607-636.

Friede, G., Busch, T. & Bassen, 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. (2007). The social responsibility of business is to increase its profits. In *Corporate ethics and corporate governance* (pp. 173-178): Springer.

Garvey, G. T., Kazdin, J., Nash, J., Lafond, R. & Safa, H. (2016). *A Pitfall in Ethical Investing: ESG Disclosures Reveal Vulnerabilities, Not Virtues*. Retrieved from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2840629

Giese, G., Lee, L.-E., Melas, D., Nagy, Z. & Nishikawa, L. (2017). FOUNDATIONS OF ESG INVESTING.

Godfrey, L. G. (1978). Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables. *Econometrica: Journal of the econometric society*, 1293-1301.

Godfrey, P. C., Merrill, C. B. & Hansen, J. M. (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk management hypothesis. *Strategic management journal*, 30(4), 425-445.

Greene, W. H. (2012). *Econometric analysis*. Boston; London: Pearson.

Gregory, A., Tharyan, R. & Whittaker, J. (2014). Corporate social responsibility and firm value: Disaggregating the effects on cash flow, risk and growth. *Journal of Business Ethics*, 124(4), 633-657.

Grossman, S. J. & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, 70(3), 393-408.

Halbritter, G. & Dorfleitner, G. (2015). The wages of social responsibility—where are they? A critical review of ESG investing. *Review of Financial Economics*, 26, 25-35.

Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The journal of finance*, 48(1), 65-91.

- Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The journal of finance*, 23(2), 389-416.
- Jo, H. & Na, H. (2012). Does CSR reduce firm risk? Evidence from controversial industry sectors. *Journal of Business Ethics*, 110(4), 441-456.
- Kendall, M. G. & Hill, A. B. (1953). The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1), 11-34.
- Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, 20(4), 587-615.
- Malkiel, B. G. & Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of finance*, 25(2), 383-417.
- Margolis, J. D., Elfenbein, H. A. & Walsh, J. P. (2009). Does it pay to be good... and does it matter? A meta-analysis of the relationship between corporate social and financial performance.
- Markowitz, H. (1952). Portfolio selection. *The journal of finance*, 7(1), 77-91.
- Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the econometric society*, 768-783.
- Newey, W. K. & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 777-787.
- NN investment partners. (2016). *The materiality of ESG factors for equity investment decisions: academic evidence*. Retrieved from yourSRI Socially Responsible Investing: https://yoursri.com/media-new/download/ecce_project_the_materiality_of_esg_factors_for_equity_investment_decisions.pdf
- Norges Bank Investment Management. (2018). Observasjon og utelukkelse av selskaper.
- Norsif. (2016). *Guide til ansvarlige investeringer*. Retrieved from <http://norsif.org/ressurser/#ansvarligeinvesteringer>:

Oikonomou, I., Brooks, C. & Pavelin, S. (2012). The impact of corporate social performance on financial risk and utility: A longitudinal analysis. *Financial Management*, 41(2), 483-515.

Parnassus Investments. (2017). *ESG Controversies May Indicate Widespread Company Problems*. Retrieved from Parnassus Investments: <https://www.parnassus.com/our-firm/highlight/221>

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies*, 22(1), 435-480.

Principles for Responsible Investment. (2018). *Annual Report 2018*. Retrieved from <https://www.unpri.org/annual-report-2018>: <https://www.unpri.org/annual-report-2018>

Revelli, C. & Viviani, J. L. (2015). Financial performance of socially responsible investing (SRI): what have we learned? A meta-analysis. *Business Ethics: A European Review*, 24(2), 158-185.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial management review*, 6(2).

Schwert, G. W. (2003). Anomalies and market efficiency. *Handbook of the Economics of Finance*, 1, 939-974.

Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3), 425-442.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of financial economics*, 99(1), 1-10.

Thomson Reuters. (2017). *Thomson Reuters ESG Scores*. Retrieved from http://www.esade.edu/itemsweb/biblioteca/bbdd/inbdd/archivos/Thomson_Reuters_ESG_Scores.pdf

Thomson Reuters. (2018). *Thomson Reuters ESG Scores*. Retrieved from <https://financial.thomsonreuters.com/content/dam/openweb/documents/pdf/financial/esg-scores-methodology.pdf>

UBS. (2015). *To integrate or to exclude - Approaches to sustainable investing*. Retrieved from [ubs.com: https://www.ubs.com/content/dam/WealthManagementAmericas/documents/to-integrate-or-to-exclude-2015-3Q-sustainable-investing.pdf](https://www.ubs.com/content/dam/WealthManagementAmericas/documents/to-integrate-or-to-exclude-2015-3Q-sustainable-investing.pdf)

US SIF. (2016). *2016 Global Sustainable Investment Review*. Retrieved from [ussif.org: https://www.ussif.org/files/Publications/GSIA_Review2016.pdf](https://www.ussif.org/files/Publications/GSIA_Review2016.pdf)

Waddock, S. A. & Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic management journal*, 18(4), 303-319.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the econometric society*, 817-838.

White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica: Journal of the econometric society*, 1-25.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*: MIT press.

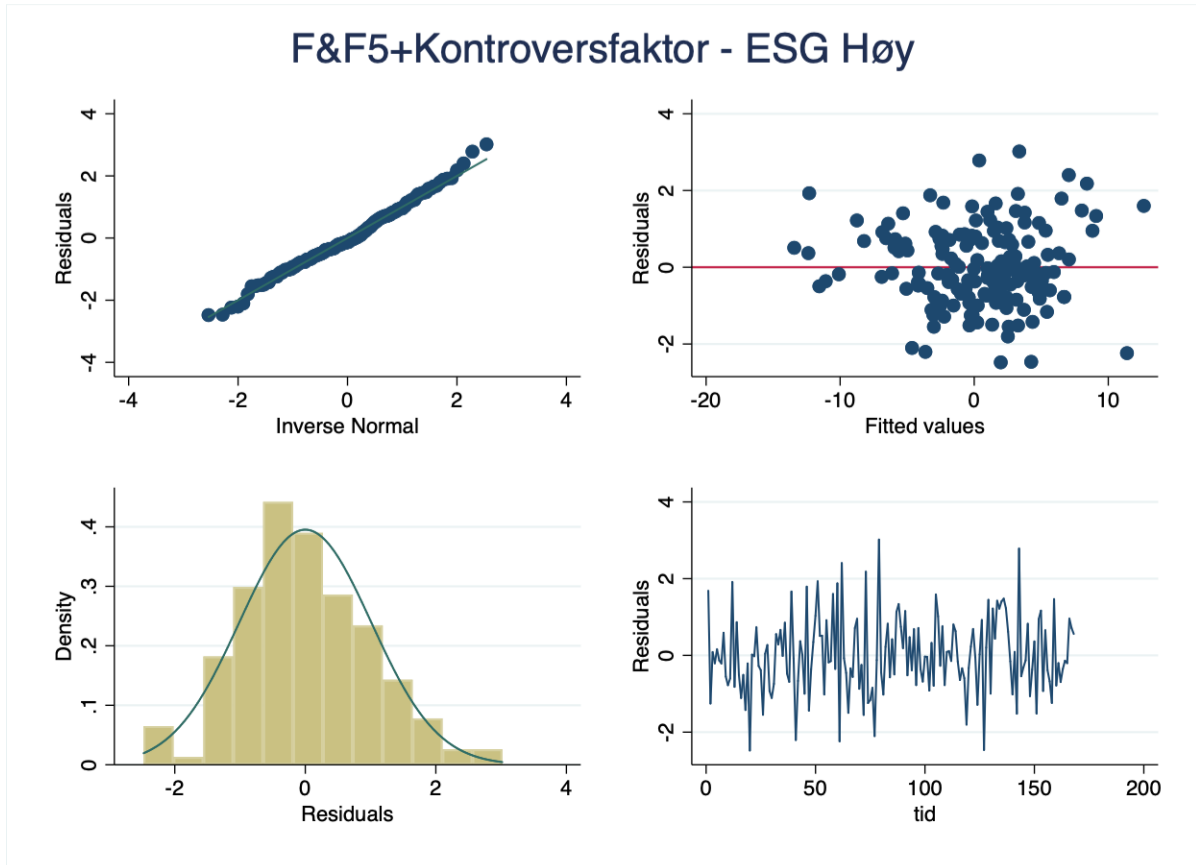
Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*: Nelson Education.

Zembrowski, P. (2018). *Do ESG scores matter for performance?* Retrieved from FSA: <https://fundselectorasia.com/esg-scores-matter-performance/>

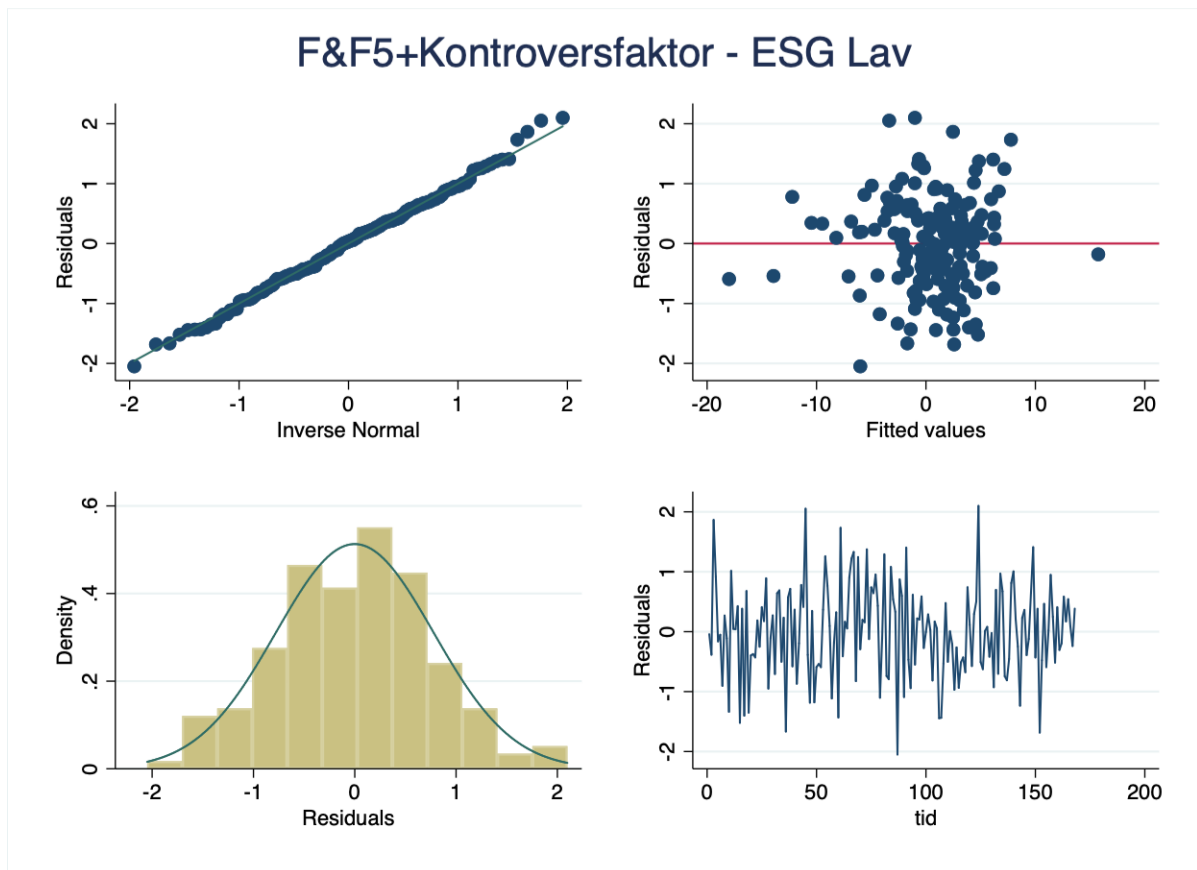
Zhang, L. (2017). *The Financial Return of Responsible Investing*. Retrieved from Sustainable Pension Investments Lab: <https://spilplatform.com/wp-content/uploads/2017/02/SPIIL-The-Financial-Return-of-Responsible-Investing.pdf>

10. Appendiks

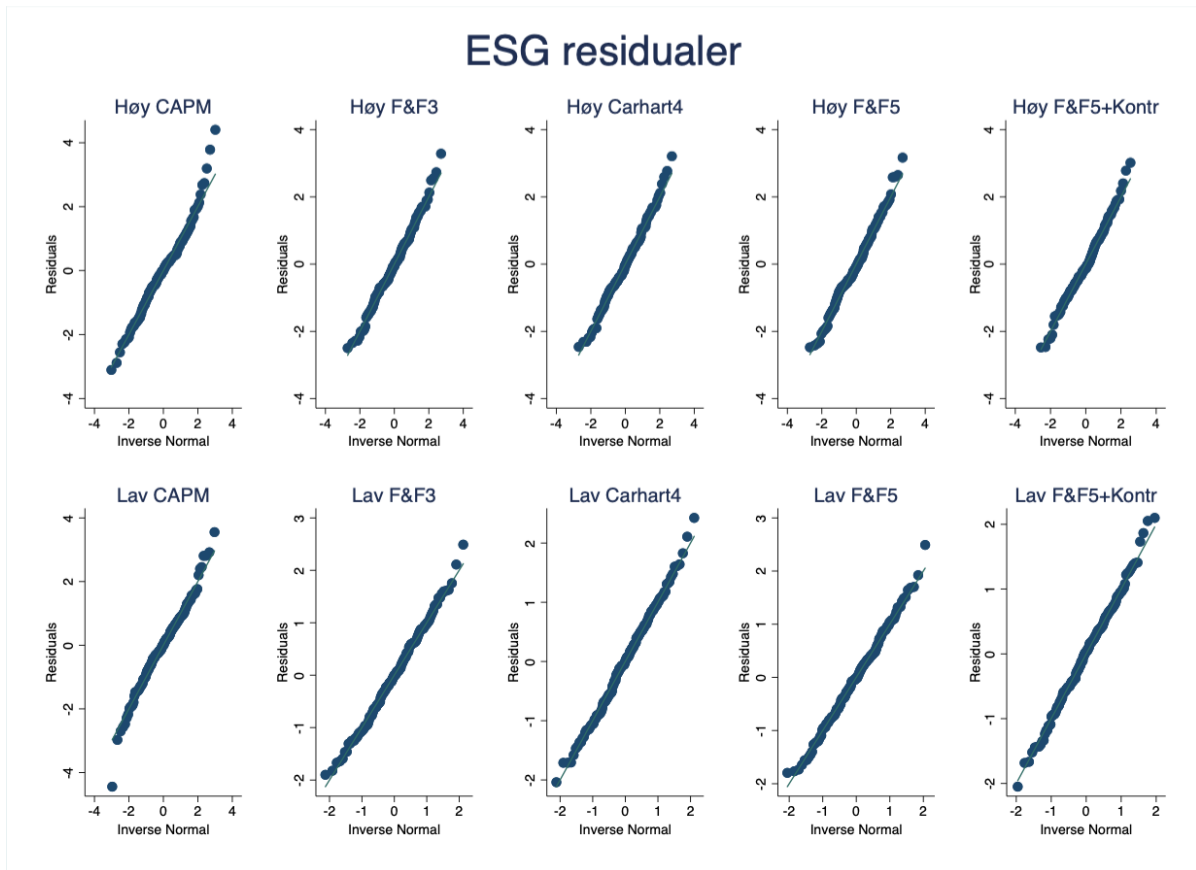
10.1 Residualplot – porteføljer



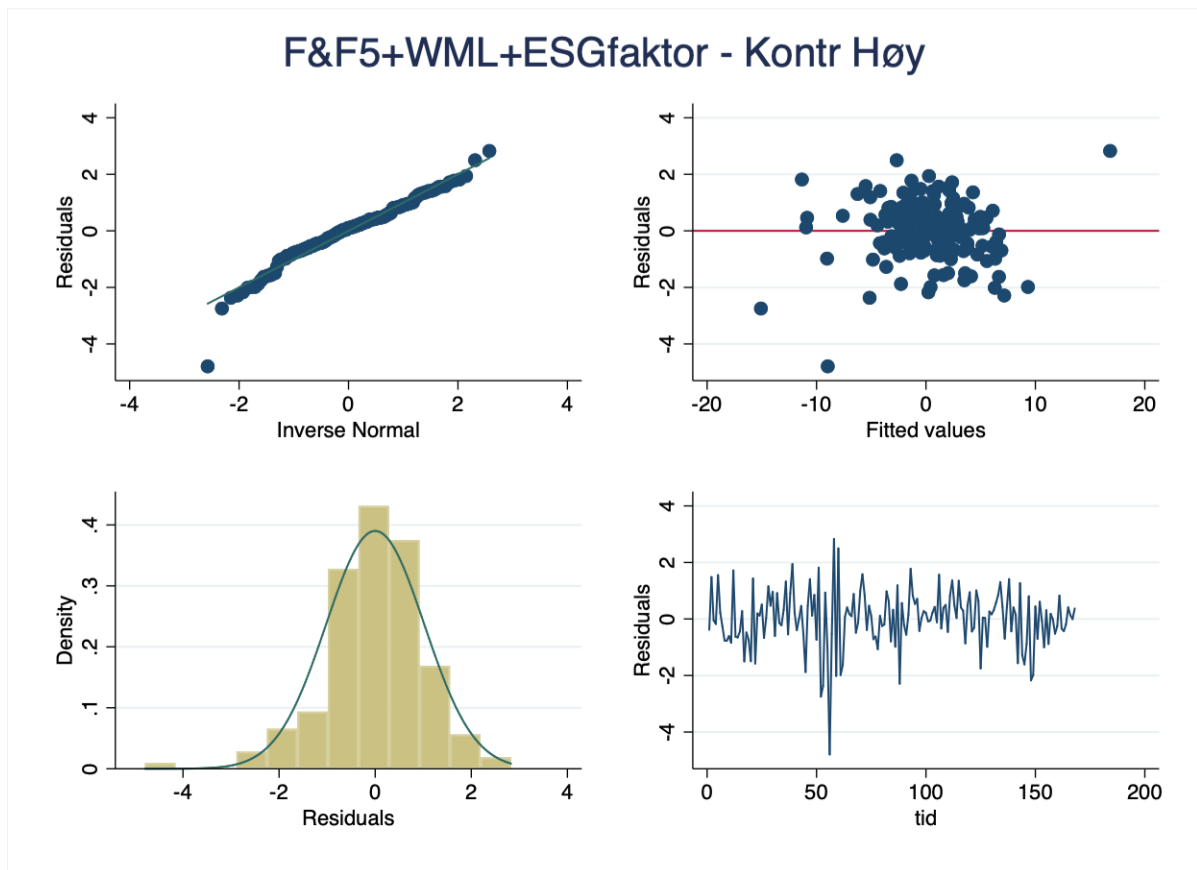
Figur 4. Figuren viser fire forskjellige residualplot for ESG høy portefølje mot F&F5 og den egenkonstruerte kontroversfaktoren. Figurene er normalplot, residualer mot estimerte verdier, residualhistogram, residualer over tid. Regresjonen kan sees i modell 5 i Tabell 6.



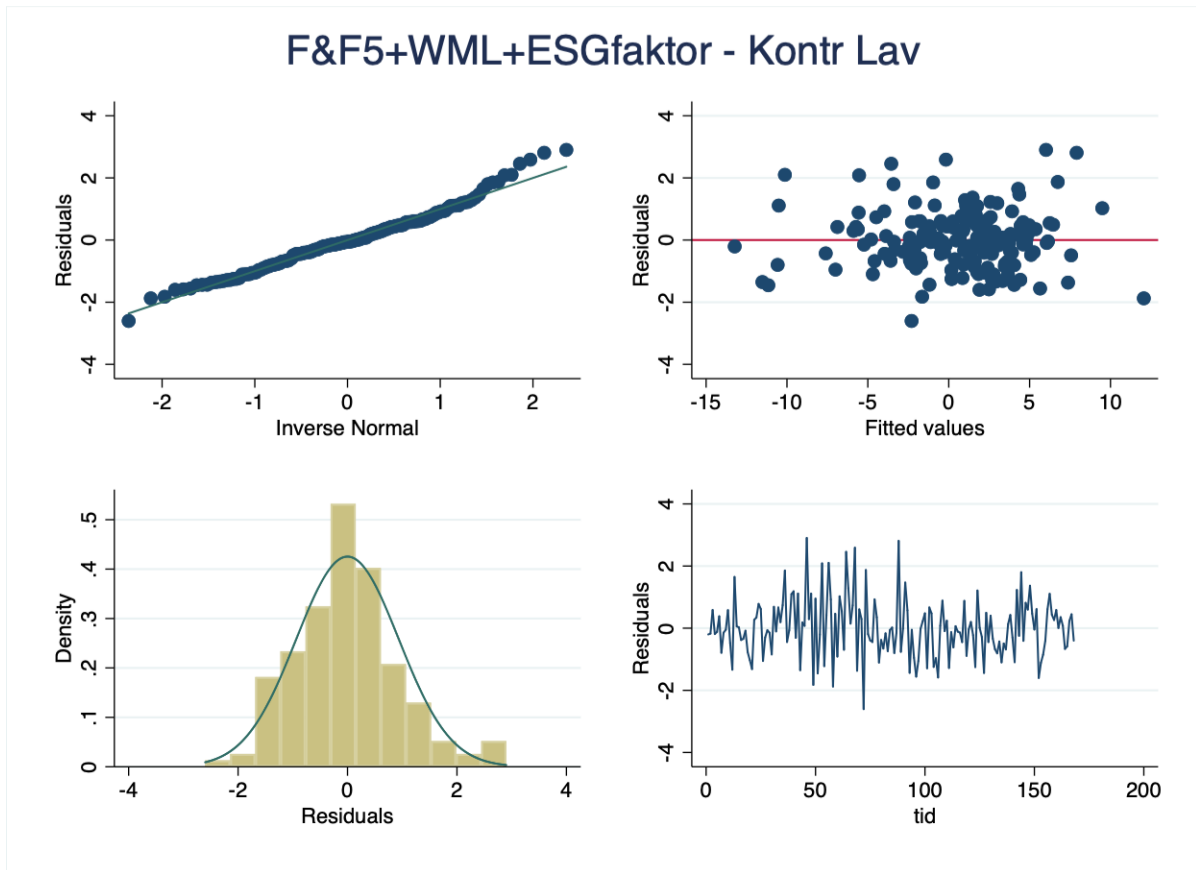
Figur 5. Figuren viser fire forskjellige residualplot for ESG lav portefølje mot F&F5 og den egenkonstruerte kontroversfaktoren. Figurene er normalplot, residualer mot estimerte verdier, residualhistogram, residualer over tid. Regresjonen kan sees i modell 5 i Tabell 6.



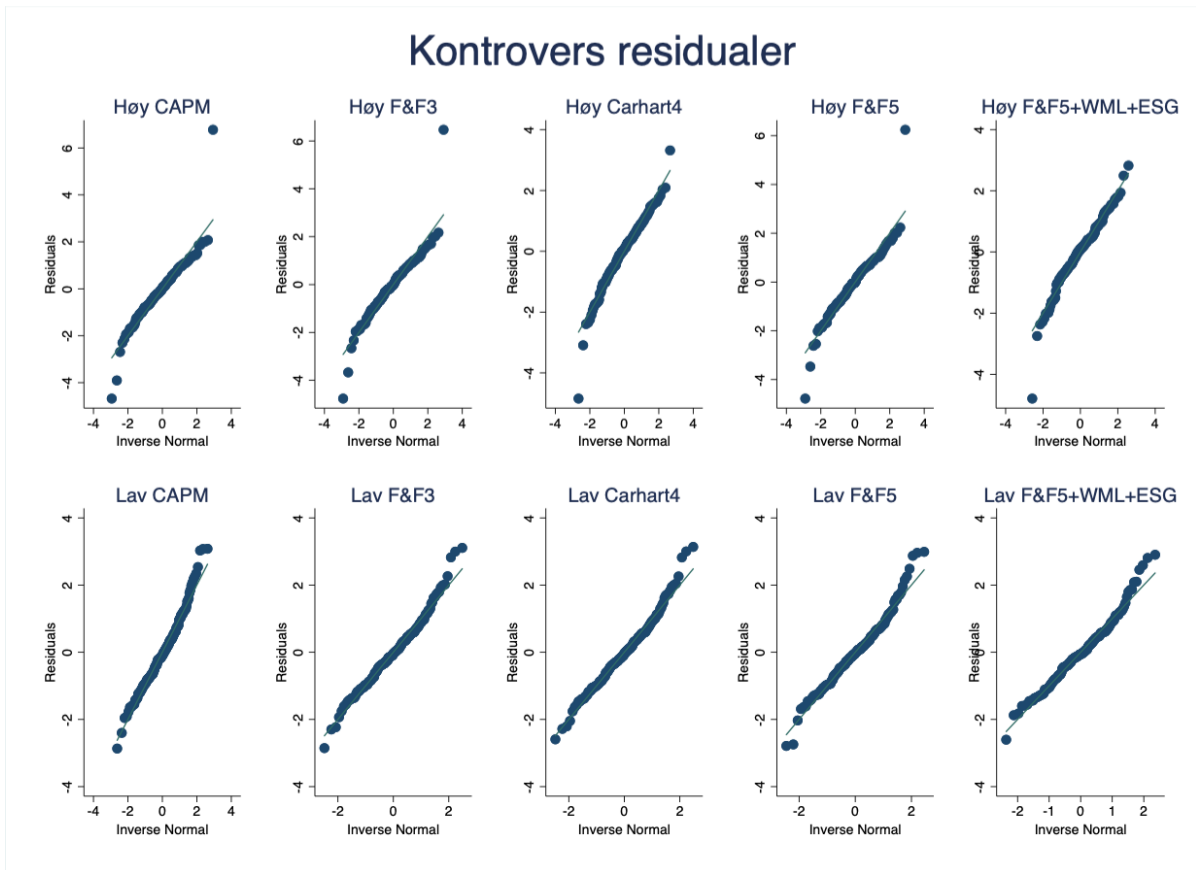
Figur 6. Figuren angir residualplot for samtlige regresjoner i Tabell 6.



Figur 7. Figuren viser fire forskjellige residualplot for Kontr høy portefølje mot F&F5 og den egenkonstruerte ESG-faktoren. Figurene er normalplot, residualer mot estimerte verdier, residualhistogram, residualer over tid. Regresjonen kan sees i modell 5 i Tabell 7.

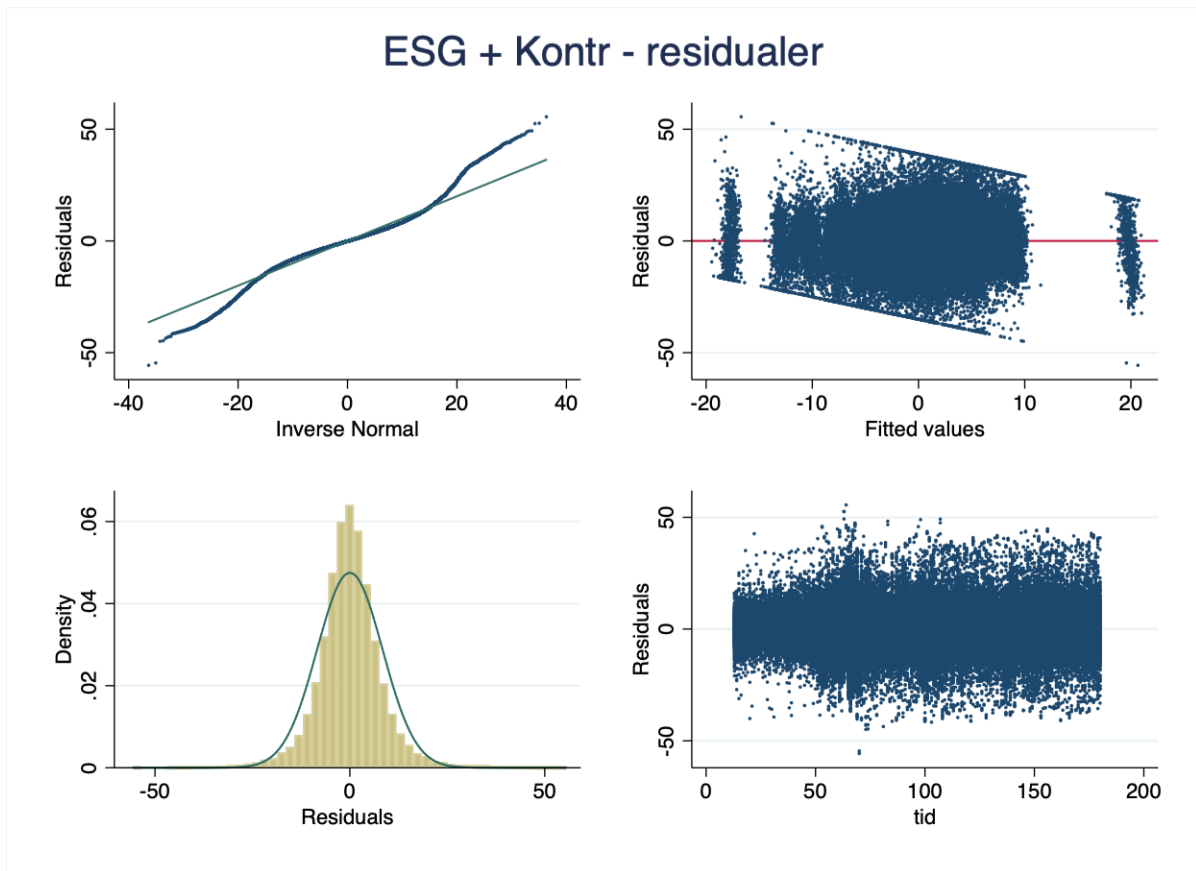


Figur 8. Figuren viser fire forskjellige residualplot for Kontr lav portefølje mot F&F5 og den egenkonstruerte ESG-faktoren. Figurene er normalplot, residualer mot estimerte verdier, residualhistogram, residualer over tid. Regresjonen kan sees i modell 5 i Tabell 7.

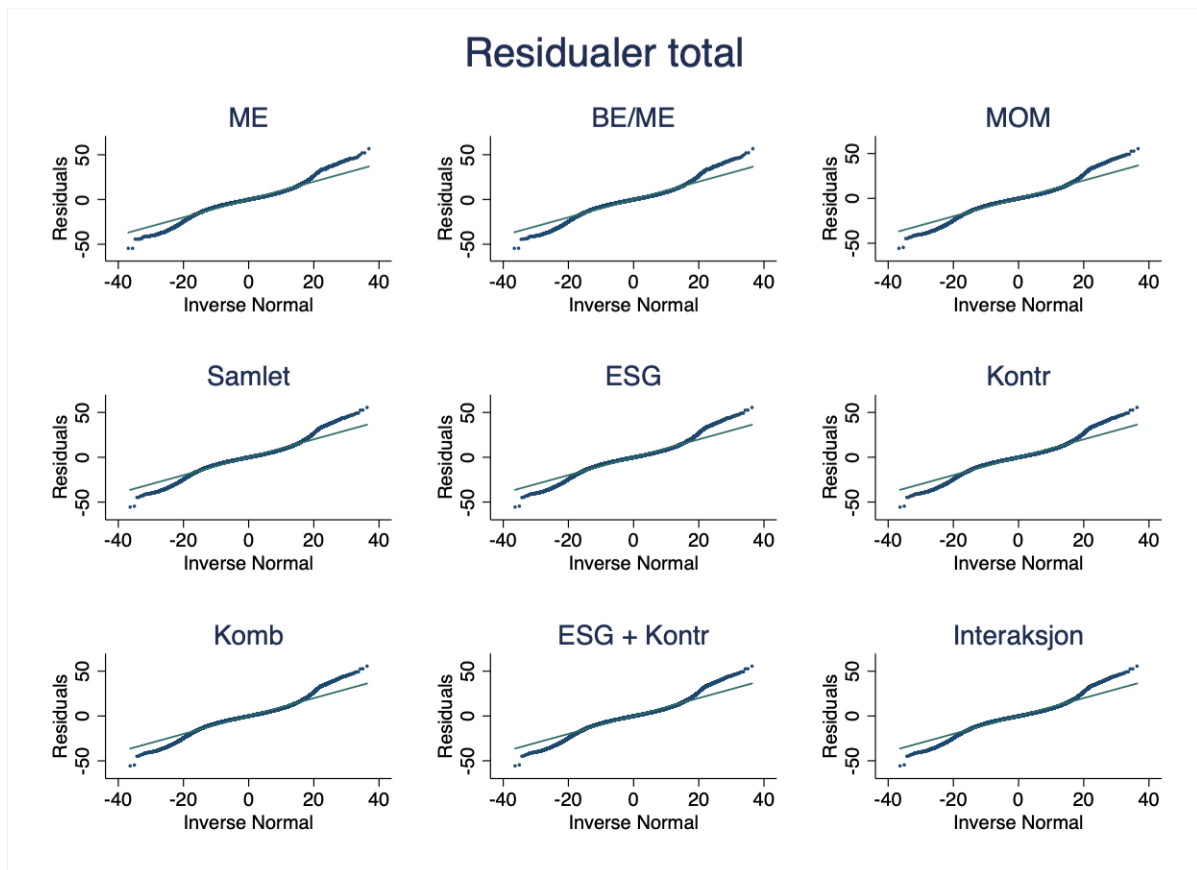


Figur 9. Figuren angir residualplot for samtlige regresjoner i Tabell 7.

10.2 Residualplot – selskaper

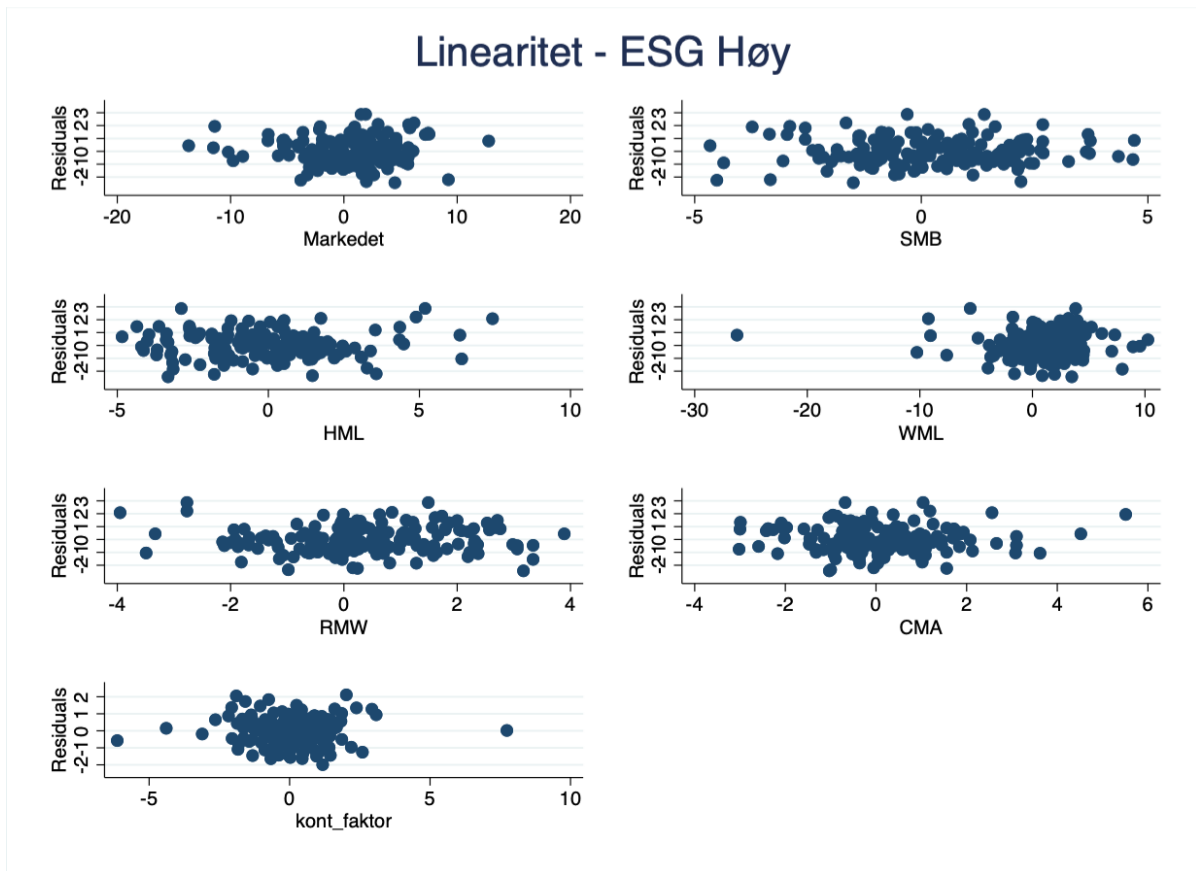


Figur 10. Figuren viser fire forskjellige residualplot for paneldatasettet som inkluderer ESG og Kontr i tillegg til kontrollvariablene. Figurene er normalplot, residualer mot estimerte verdier, residualhistogram, residualer over tid. Hoppene i modellen med estimerte verdier er en følge av dummy under finanskrisen og innhentingene i etterkant.

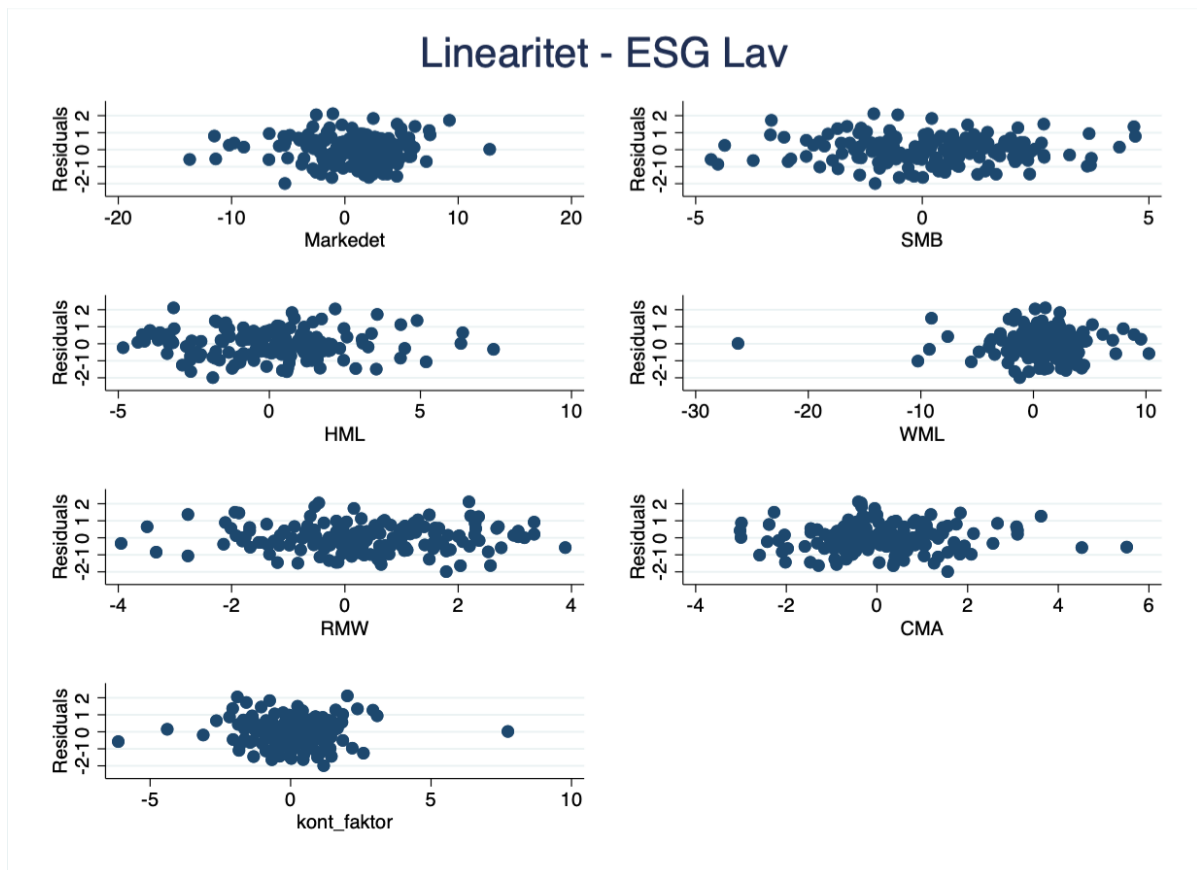


Figur 11. Figuren angir residualplot for samtlige regresjoner i Tabell 13.

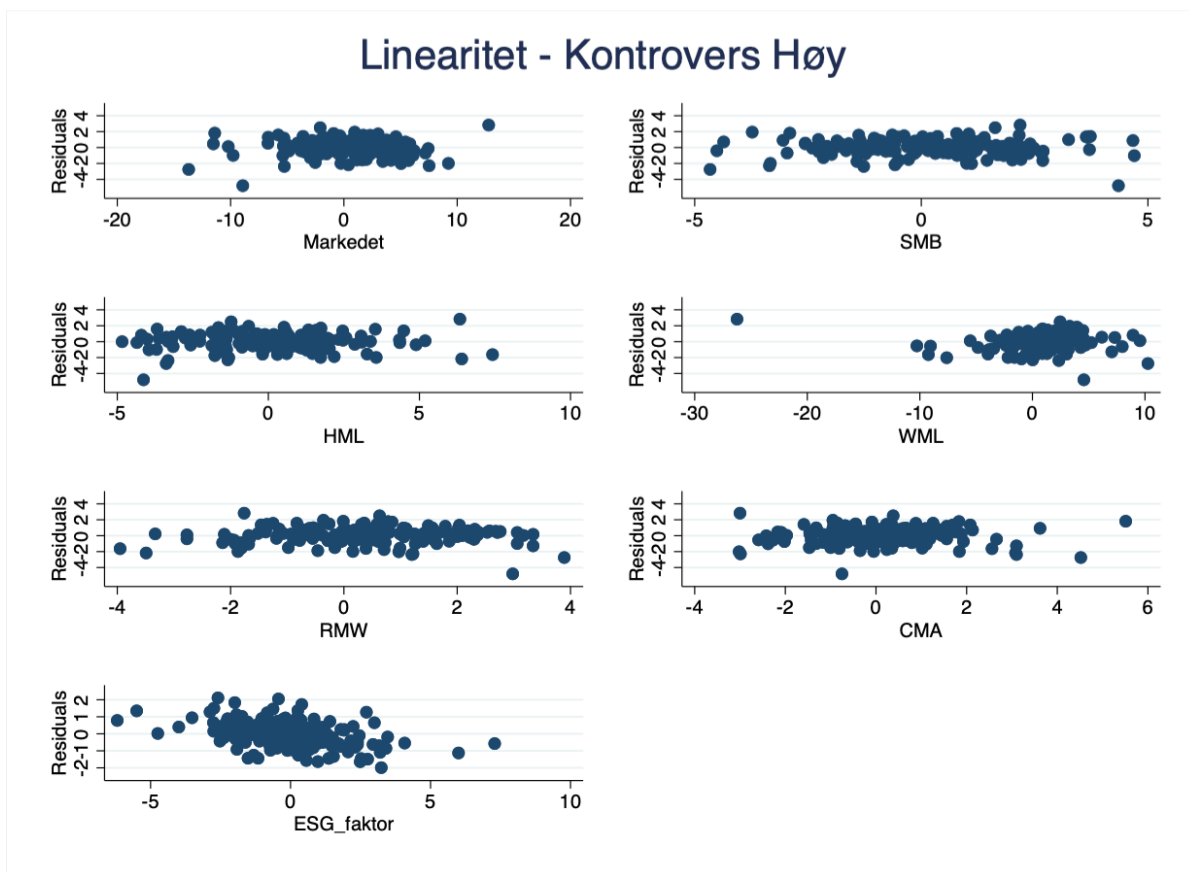
10.3 Linearitetsplot – porteføljer



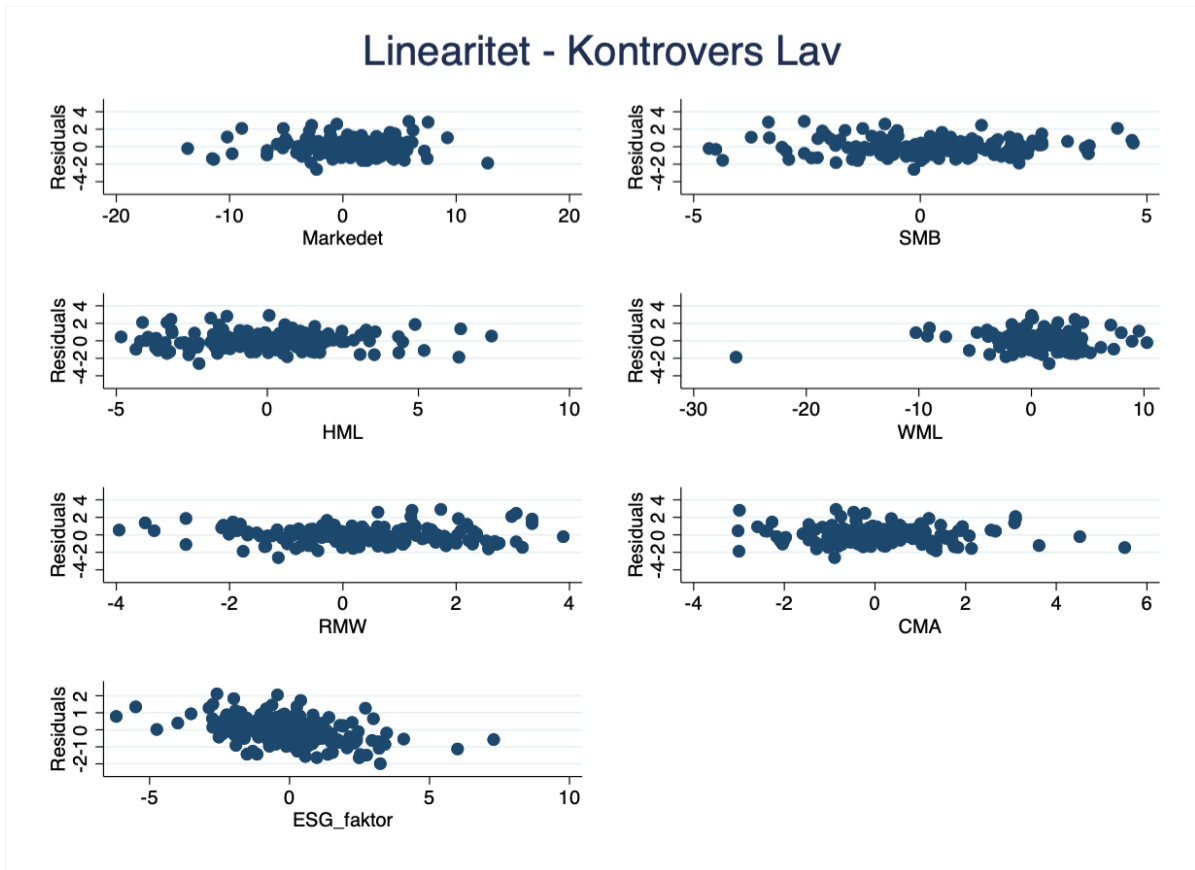
Figur 12. Figuren angir linearitetsplot. Residualene er plottet mot de ulike uavhengige variablene.



Figur 13. Figuren angir linearitetsplot. Residualene er plottet mot de ulike uavhengige variablene.

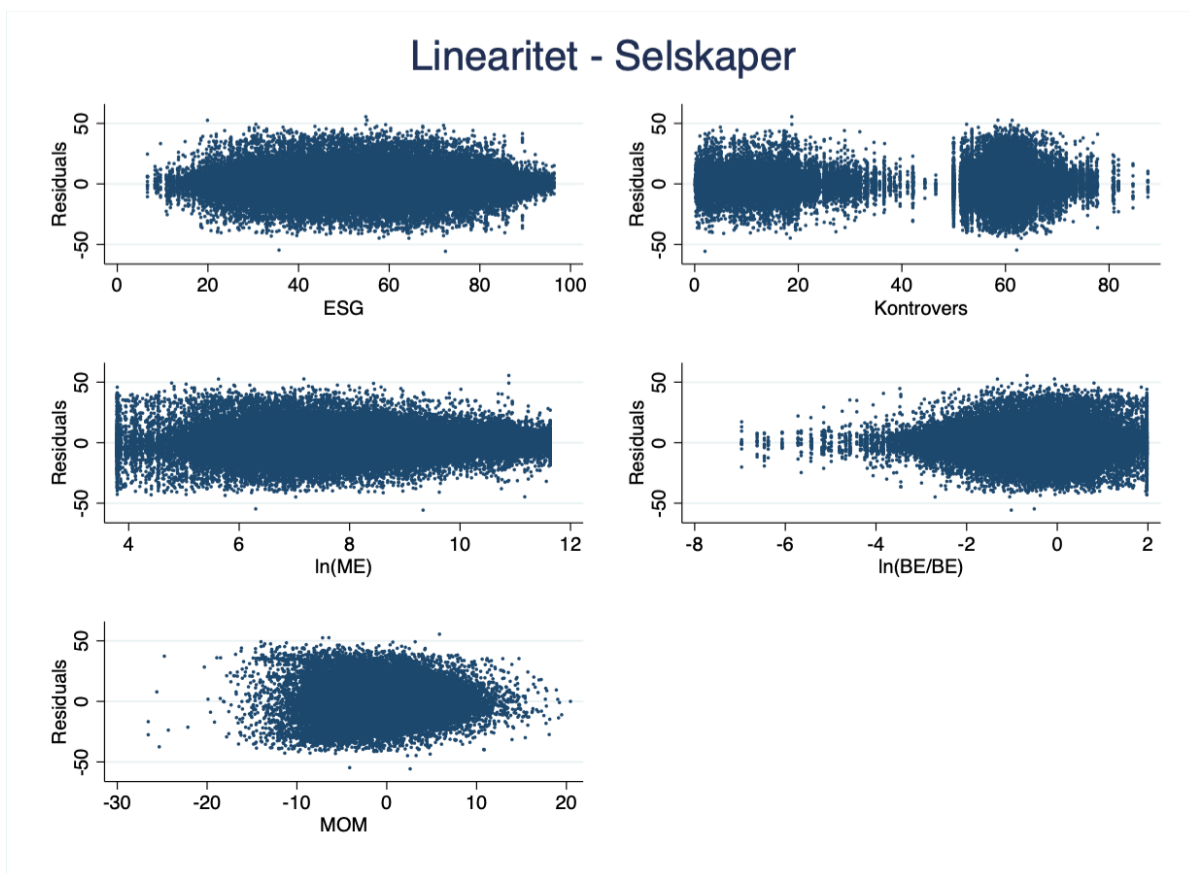


Figur 14. Figuren angir linearitetsplot. Residualene er plottet mot de ulike uavhengige variablene.



Figur 15. Figuren angir linearitetsplot. Residualene er plottet mot de ulike uavhengige variablene.

10.4 Linearitetsplot – selskaper



Figur 16. Figuren angir linearitetsplot. Residualene er plottet mot de ulike uavhengige variablene.

10.5 Utvalgte regresjoner

Regresjoner med alle selskaper i markedet

Tabell 16. Tabellen inneholder tilsvarende regresjoner som nr. 1-4 i Tabell 13. Observasjonene her er alle selskaper på de største børsene i Europa. Årsaken til at antall observasjoner varierer på tvers av modellene, er manglende observasjoner av selskapstall.

	(1)	(2)	(3)	(4)
	ME	BE/ME	MOM	Samlet
ln(ME)	-0.727*** (0.207)			-0.372*** (0.096)
LN(BE/ME)		1.243*** (0.375)		1.031*** (0.356)
MOM			0.002 (0.001)	0.001 (0.004)
Observasjoner	1224958	1070089	1110130	999062
Justert R ²	0.000	0.002	0.000	0.003

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regresjoner uten finansselskaper

Tabell 17. Tabellen inneholder tilsvarende regresjoner som nr. 1-4 i Tabell 13. Observasjonene er samme utvalg som i tabellen, men finansselskaper er utelatt. Årsaken til at antall observasjoner varierer på tvers av modellene, er manglende observasjoner av selskapstall.

	(1)	(2)	(3)	(4)
	ME	BE/ME	MOM	Samlet
ln(ME)	0.016 (0.066)			-0.035 (0.056)
LN(BE/ME)		-0.049 (0.099)		-0.053 (0.090)
MOM			0.147*** (0.053)	0.150*** (0.051)
Observasjoner	113009	110392	104830	102432
Justert R ²	0.209	0.212	0.210	0.213

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Regresjoner på selskapsdata for selskaper med lav ESG-score

Tabell 18. Tabellen inneholder tilsvarende regresjoner som nr. 8 i Tabell 13. Observasjonene er samme utvalg som i tabellen, men de ulike modellene inneholder selskaper med gitte ESG-scorer. Første modell har kun selskaper med score under 30, andre har kun selskaper med score under 20, og siste har kun selskaper med score under 15.

	(1)	(2)	(3)
	ESG<30	ESG<20	ESG<15
ln(ME)	-0.179 (0.122)	0.026 (0.240)	-0.365 (0.371)
LN(BE/ME)	-0.062 (0.144)	0.267 (0.324)	-0.184 (0.633)
MOM	0.131* (0.076)	0.272 (0.170)	0.536** (0.226)
ESG	-0.004 (0.021)	-0.010 (0.091)	0.057 (0.320)
Kontr	0.006 (0.012)	-0.008 (0.030)	-0.013 (0.054)
Observasjoner	7602	1236	360
Justert R^2	0.201	0.198	0.219

Standardfeil i parentes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$