

*En studie av avkastning og risiko ved
handelsstrategien pairs trading basert
på kointegrasjon.*

Andreas Mikkelsen og Petter Størksen

BE305E Finansiering og Investering
Våren 2012

Abstract

Pairs trading is a speculative investment strategy existing since the 1980s, originally exclusive to investment banks and hedge funds because of their economies of scale. The rapid development of ICT tools and internet based brokers has recently made the strategy a viable option for small scale investors.

The idea behind the strategy is to find two securities that move similarly over time. We can thus assume that they are in equilibrium. When the securities diverge from the equilibrium we go long in the relative underpriced security and go short in the overpriced. The idea is that the securities will eventually converge to the equilibrium and this will yield us a profit.

We have tried to follow the methodology of earlier studies of Gatev, Rouwenhorst & Goetzmann (2006) and Do & Faff (2009) to some degree. To start with we chose roughly 1500 stocks from NYSE and NASDAQ. We use a co-integration approach to find pairs and have two groups of securities. One group that was significant at 5 % level or better and one group at 10 % level. The indicator to trade is calculated using the mispricing expressed in standard deviations using a 50 day simple moving average.

The 5 % group made an annual return of 0,42 %, which is in line with the findings of Do & Faff (2009), the 10 % group did however perform better with an annual return of 7,17 %. Both portfolios had much lower risk metrics than the market portfolio, and the beta value indicated that they were market neutral.

Forord

Denne masteroppgaven er skrevet som en avsluttende studie i spesialiseringen Finansiering og Investering ved Handelshøgskolen i Bodø. Tidsrommet for studien strekker seg over hele vårsemesteret 2012 og oppgaven har et omfang på 30 studiepoeng. Temaet til oppgaven kunne velges fritt innenfor spesialiseringens grenser, med den hensikt at vi skal lære å praktisere teoretisk kunnskap.

Vi vil takke vår veileder, professor Frode Sættem ved Norges Handelshøyskole for god hjelp og oppmuntrende tilbakemeldinger underveis i arbeidet med oppgaven. Vi ønsker også å rette en stor takk til førsteamanuensis Berner Larsen ved Universitetet i Nordland for uvurderlig hjelp med programmering i R og statistiske utfordringer.

Avslutningsvis vil vi nevne at eventuelle feil og mangler i oppgaven er undertegnede alene ansvarlige for.

Handelshøgskolen i Bodø, 23.5.2012

Andreas Mikkelsen

Petter Størksen

Sammendrag

Pairs trading er en spekulativ handlestrategi som har eksistert siden 1980-tallet. Opprinnelig var det investeringsbanker og hedgefond på Wall Street som benyttet seg av strategien. Men utviklingen innen IT-løsninger og nettbaserte meglere har gjort at også privatpersoner kan benytte seg av denne. Utgangspunktet for pairs trading er å finne to aksjer som over tid beveger seg likt. Dette gjør at man antar at det finnes en likevekt mellom aksjene. Når aksjene avviker fra denne likevekten kjøper man den aksjen som relativt sett er underpriset, og selger den aksjen som er overpriset. Tanken er da at når aksjene går tilbake til likevekten, løser man inn posisjonene og innkasserer en hyggelig gevinst. For å definere likevekten kan man benytte seg av ulike mål. Eksempler på slike mål er korrelasjon eller kointegrasjon. Korrelasjon er et velkjent begrep, men ulempen er at den bare indikerer en kortsiktig likevekt. Kointegrasjon forutsetter derimot en langsiktig likevekt.

Vi har i vår studie forsøkt å følge metodikken til tidligere studier av Gatev et al. (2006) samt Do & Faff (2009) for å kunne sammenlikne vår studie med resultatene fra deres studier. Begge studiene har benyttet seg av kointegrasjon for å finne aksjer som beveger seg likt, og av den grunn besluttet vi å gjøre det samme.

Vi startet med et utvalg på i overkant av 1500 aksjer fordelt på fem supersektorer fra børsene New York Stock Exchange og NASDAQ. Vi testet bare aksjer innenfor undersektorer mot hverandre for kointegrasjon, fordi at det skulle være størst mulig sjanse for at parene var kointegrerte basert på at de var økonomiske substitutter. Vår framgangsmåte var først å teste alle mulige kombinasjoner av par innen en undersektor med en Johansen kointegrasjonstest. Deretter ble de parene som var statistisk signifikante testet på nytt for kointegrasjon med en Augmented Dickey Fuller (ADF) test og Phillips-Perron (PP) test. De parene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg signifikante på 5 % nivå i ADF og PP testene ble plassert i en gruppe med benevnelsen 5 %. De parene som var signifikante på minimum 10 % nivå i alle testene ble sortert i en gruppe vi kalte 10 %. De øvrige parene ble forkastet.

Videre har vi benyttet oss av en handelsindikator basert på 50 dagers glidende gjennomsnitt. Indikatoren uttrykker feilprisingen i standardavvik av spreaden mellom aksjene. Dersom aksjene divergerte to standardavvik fra likevekten åpnet vi en posisjon. For å gjøre det så realistisk som mulig benyttet vi en form for risikostyring ved å lukke posisjonene dersom

parene fortsetter å divergere. Dette ble gjort ved å programmere indikatoren til å løse inn posisjonene dersom parene divergerte ut over tre standardavvik fra likevekten. Vi lukket også posisjonene dersom parene konvergente tilbake til likevekten, eller når handelsperioden var over. Etter å ha testet par for kointegrasjon satt vi igjen med 161 par i gruppen 5 % i tillegg til 161 par i gruppen 10 %. Vi testet så handelsstrategien over en tidsperiode på seks måneder, fra 3. januar 2011 til 30. juni 2011. Vi lagde porteføljer basert på sektorer, og i tillegg en portefølje med alle parene i en gruppe.

Vi fant at handelsstrategien basert på et 50 dagers glidende gjennomsnitt er veldig handelsintensivt. I gruppen 5 % gjennomførte vi 1772 handler over en seks måneders periode, og tilsvarende 1816 handler for 10 % gruppen. Samtidig viste det seg at parene hadde relativt kort gjennomsnittlig holdingperiode. Holdingperioden for gruppen 5 % var ca. 20 handelsdager. For gruppen 10 % var den noe høyere med i overkant av 23 handelsdager. Dette impliserer at basert på handelsindikatoren brukt i vår studie oppnår man en handelsstrategi med kort tidshorison.

For begge gruppene oppnådde vi en markedsnøytral strategi ettersom beta var nær null uavhengig av hvor mange par vi hadde i porteføljene. Videre viste det seg at vi oppnådde en diversifiseringsgevinst ettersom risiko uttrykt i både standardavvik og Value at Risk ble lavere når flere par ble introdusert i en portefølje. Strategien var gjennomgående mindre volatil sammenliknet med å holde markedsporteføljen.

Strategien oppnådde en årlig avkastning på 0,42 % for gruppen 5 %. Dette var som forventet ettersom Do & Faff (2009) finner en avtagende avkastning for strategien; en avkastning ikke signifikant annerledes enn null i perioden 2003 - 2008. For gruppen 10 % ble den årlige avkastningen 7,17 %, nesten like bra som S&P 500 oppnådde i samme periode med lavere risiko. Vi kan dessverre ikke dra noen konklusjoner på hva grunnen til den gode avkastningen kan være ettersom vi bare har en handelsperiode i vår studie.

Som nevnt er pairs trading en intensiv handelsstrategi, noe som naturlig nok medfører høye transaksjonskostnader. Det virker ikke som at strategien er robust for transaksjonskostnadene som ble estimert til å være i gjennomsnitt 453,6 basispoeng per par i perioden. Vi finner at ved å øke lengden på glidende gjennomsnitt til 150 dager i indikatoren reduserer dette transaksjonskostnadene for strategien med 44 %. Det må kunne sies å være en substansiell reduksjon og dermed ha potensial til å være forskjellen mellom tap og gevinst på lang sikt.

Innholdsfortegnelse

Abstract	II
Forord	III
Sammendrag	IV
Innholdsfortegnelse	VI
Figurliste	IX
Tabelliste	IX
Vedleggsliste	IX
1. Innledning.....	1
1.1 Aktualisering.....	1
1.2 Bakgrunn for oppgaven.....	1
1.3 Avgrensninger	2
1.4 Problemstilling.....	2
1.5 Oppgavens oppbygning.....	3
2. Markedseffisiens	4
2.1 Markedseffisienshypotesen.....	4
2.1.1 Svak form for markedseffisiens.....	6
2.1.2 Semi sterk form for markedseffisiens	7
2.1.3 Sterk form for markedseffisiens.....	8
2.2 Kritikk av hypotesen	8
2.3 Oppsummering.....	10
3. Arbitrasje	11
3.1 Hva er arbitrasje?	11
3.1.1 Ren arbitrasje	11
3.1.2 Relativ arbitrasje	12
3.1.3 Spekulativ arbitrasje.....	12
3.1.4 Aktører	12
3.2 Risiko	13
3.2.1 Effektueringsrisiko.....	13
3.2.2 Konvergeringsrisiko og likviditet	14
3.2.3 Short-squeeze.....	14
3.2.4 Kredittrisiko	15
3.2.5 Long Term Capital Management.....	15

3.3 Implikasjoner av EMH og arbitrasje	16
3.4 Oppsummering.....	16
4. Pairs trading	17
4.1 Hva er pairs trading?	17
4.1.1 Stasjonære og ikke-stasjonære variabler.....	18
4.1.2 Kointegrasjon	20
4.1.3 Noen avklaringer	21
4.1.3.1 Hvorfor metoden hevdes å fungere.....	22
4.2 Markedsnøytral strategi.....	22
4.2.1 Pairs trading med enkel spread som mål.....	22
4.2.2 Pairs trading ved bruk av standardavvik.....	24
4.2.3 Standardavvik eller spread ratio.....	25
4.2.4 Metode ved bruk av beta.....	26
4.3 Pairs trading i praksis	27
4.3.1 Strategiens profitabilitet	27
4.3.2 Fundamentale forhold.....	28
4.3.3 Money management og risiko ved pairs trading	29
4.4 Oppsummering.....	30
5. Metode	31
5.1 Analyseverktøy.....	31
5.2 Utvelgelse av aksjer, sektor og tidsperioder	31
5.3 Utvelgelse av par	32
5.4 Handleregler	34
5.5 Beregninger i analysen	37
Avkastning	37
Risikofri rente	38
Sharpe Ratio (Årlig)	38
Information Ratio	39
Value at Risk (VaR)	39
Expected shortfall.....	40
Max drawdown	40
Beta	40
Standardavvik (Årlig)	40
5.6 Diskusjon, valg av kriterier	41

5.6.1 Indikator, valg av lengde på glidende gjennomsnitt	41
5.6.2 Valg av handleregel	42
5.6.3 Kointegrasjon	43
5.6.4 Risikomål	44
5.6.6 Ordretype	44
5.7 Mulige feilkilder.....	45
5.7.1 Survivorship bias	45
5.7.2 Data snooping bias	46
5.7.3 Look ahead bias.....	46
5.8 Oppsummering.....	47
6. Analyse	48
6.1 Sektorer	49
6.1.1 Basic Materials	49
6.1.2 Conglomerates	51
6.1.3 Consumer goods.....	53
6.1.4 Financials	55
6.1.5 Technology	57
6.1.6 Alle parene	59
6.2 Transaksjonskostnader.....	64
6.2.1 Sensitivitetsanalyse transaksjonskostnader.....	65
6.3 Oppsummering.....	67
7. Avslutning.....	68
7.1 Oppsummering og konklusjon	68
7.2 Kritikk av studiet.....	70
7.3 Forslag til videre forskning	71
8. Litteraturliste.....	72

Figurliste

Figur 1 Oppgavens oppbygning.....	3
Figur 2 Korrelasjon mellom FRO og FUR	18
Figur 3 Tidsserier	19
Figur 4 Autokorrelasjon.....	20
Figur 5 The Sine Wave (Pole, 2007, p. 19).....	23
Figur 6 The Popcorn Process (Pole, 2007, p. 19).....	23
Figur 7 Chart (Grøtte, 2002, p. 513)	26
Figur 8 Delta	35
Figur 9 Sammendrag BVN.HMY.....	36
Figur 10 Oversikt aksjekurser	37
Figur 11 Chart, diverse SMA.....	41
Figur 12 Forskjellige SD	43
Figur 13 Graf, Basic Materials	49
Figur 14 Graf, Conglomerates	51
Figur 15 Graf, Consumer goods.....	53
Figur 16 Graf, Financials	55
Figur 17 Graf, Technology	57
Figur 18 Graf, alle parene.....	59
Figur 19 Fordeling.....	63
Figur 20 Graf, Financials SMA150.....	66

Tabelliste

Tabell 1 Resultater, forskjellige SMA.....	41
Tabell 2 Forskjellige SD.....	43
Tabell 3 Basic Materials.....	49
Tabell 4 Conglomerates.....	51
Tabell 5 Consumer goods	53
Tabell 6 Financials	55
Tabell 7 Technology.....	57
Tabell 8 Alle parene.....	59
Tabell 9 Financials SMA150.....	66

Vedleggsliste

Vedlegg 1 Script for å få data inn i R. Eksempel sektoren Agricultural i Basic Materials.....	A
Vedlegg 2 Johansentest av en hel sektor. Eksempel Property and Casualty Insurance.	B
Vedlegg 3 Eksempel output Johansentest.	B
Vedlegg 4 ADF og PP-test av spredningen mellom to aksjer. Eksempel Aksjeparet HTH.AMSF.....	C
Vedlegg 5 Eksempel output ADF og PP-test.....	C
Vedlegg 6 Script for testing av handelsstrategi. Eksempel fra alle 5 %.....	D
Vedlegg 7 Finne antall par som faktisk handler.	J

1. Innledning

Innledningsvis vil vi gå igjennom bakgrunnen for valg av tema og problemstilling. Vi vil deretter forklare hvorfor dette er et dagsaktuelt tema. Til slutt gjør vi noen avgrensninger for studien, samt går gjennom oppgavens oppbygning.

1.1 Aktualisering

Innen akademia har det vært stor interesse for forskning på aksjemarkedet. Studier av Keim (1983), Lehmann (1990), Lakonishok (1994) og Gatev et al. (2006) omhandler alle ulike fenomener innenfor temaet. I norske aviser¹ har Setterberg (2007) denne våren fått oppmerksomhet med sin studie som påviser at det finnes en momentumeffekt på Stockholmsbørsen.

Blant praktikere har det også blitt brukt store ressurser på utvikling av metoder for å oppnå høy avkastning på deres investeringer. Tradisjonelt har dette vært forbeholdt hedgefond og institusjonelle investorer grunnet deres stordriftsfordeler. Utviklingen innen IT-løsninger og nettbaserte meglere har imidlertid ført til at også privatpersoner har tilgang til markedet til en lav kostnad uten å gå via mellommenn.

Fenomenet pairs trading har blitt benyttet siden 1980-tallet av investeringsbanken Morgan Stanley med stor suksess.

1.2 Bakgrunn for oppgaven

Konsistent positiv avkastning i aksjemarkedet med lavest mulig risiko er svært vanskelig å oppnå. Årsaken til dette skriver seg fra en rekke forhold, men først og fremst kommer det av at det stilles store krav til riktig kunnskap, dedikasjon og erfaring. Et av kjennetegnene på dem som lykkes er deres evne til å tenke i sannsynligheter og hele tiden være trygg på at metoden de til enhver tid bruker har vist seg å være historisk profitabel. Ved å teste strategier på historiske data kan man derfor få kunnskap om hvordan man bør bevege seg i markedet for å redusere risiko og oppnå målet om konsekvent positiv avkastning over tid. Pairs trading er en strategi som over lang tid har vist seg som en fornuftig tilnærming til aksjemarkedet på bakgrunn av eliminert systematisk risiko samt stabil positiv avkastning. Målet i seg selv er derfor ikke å slå markedet til en hver tid, men å unngå den store variansen som man kan oppleve med passiv forvaltning.

¹ <http://www.dn.no/forsiden/borsMarked/article2379758.ece> (Retrieved 13.05.2012)

På bakgrunn av denne tankegangen ble vi nysgjerrige på om pairs trading innehar kvaliteter som lav risiko og positiv avkastning, eller om strategien i årenes løp har mistet disse egenskapene. Et annet poeng er at tilnærmingen til statistisk arbitrasje og algoritmehandel i stor grad er to sider av samme sak, der strategien kan automatiseres og all handel kan utføres mekanisk. Et siste poeng som øker betydningen av studien er at temaet relativt sett er lite utforsket i akademia når man sammenligner med mer kjente og tradisjonelle tilnærminger til aksjemarkedet innen teknisk og fundamental analyse. Det vil derfor være meget interessant å sammenligne nye funn med eldre studier. Summen av dette er bakgrunnen for at vi ønsket å gjennomføre denne studien.

1.3 Avgrensninger

Ettersom vi har hatt begrenset med tid til å gjennomføre studien så vi oss nødt til å avgrense den noe. På grunn av arbeidsmengden hadde vi bare mulighet til å teste par i fem sektorer. Vi mener likevel at vi oppnådde et bredt utvalg ettersom vi hadde prisdata fra i overkant av 1500 aksjer som utgangspunkt for å finne par. Det viste seg at det mest tidkrevende var å teste mulige par for kointegrasjon, da vi ble nødt til å gjøre dette manuelt. Vi har brukt de samme tidsrammer for oppbygging av par og påfølgende handelsperiode som tidligere studier. Dette førte til at vi bare fikk testet strategien over seks måneder.

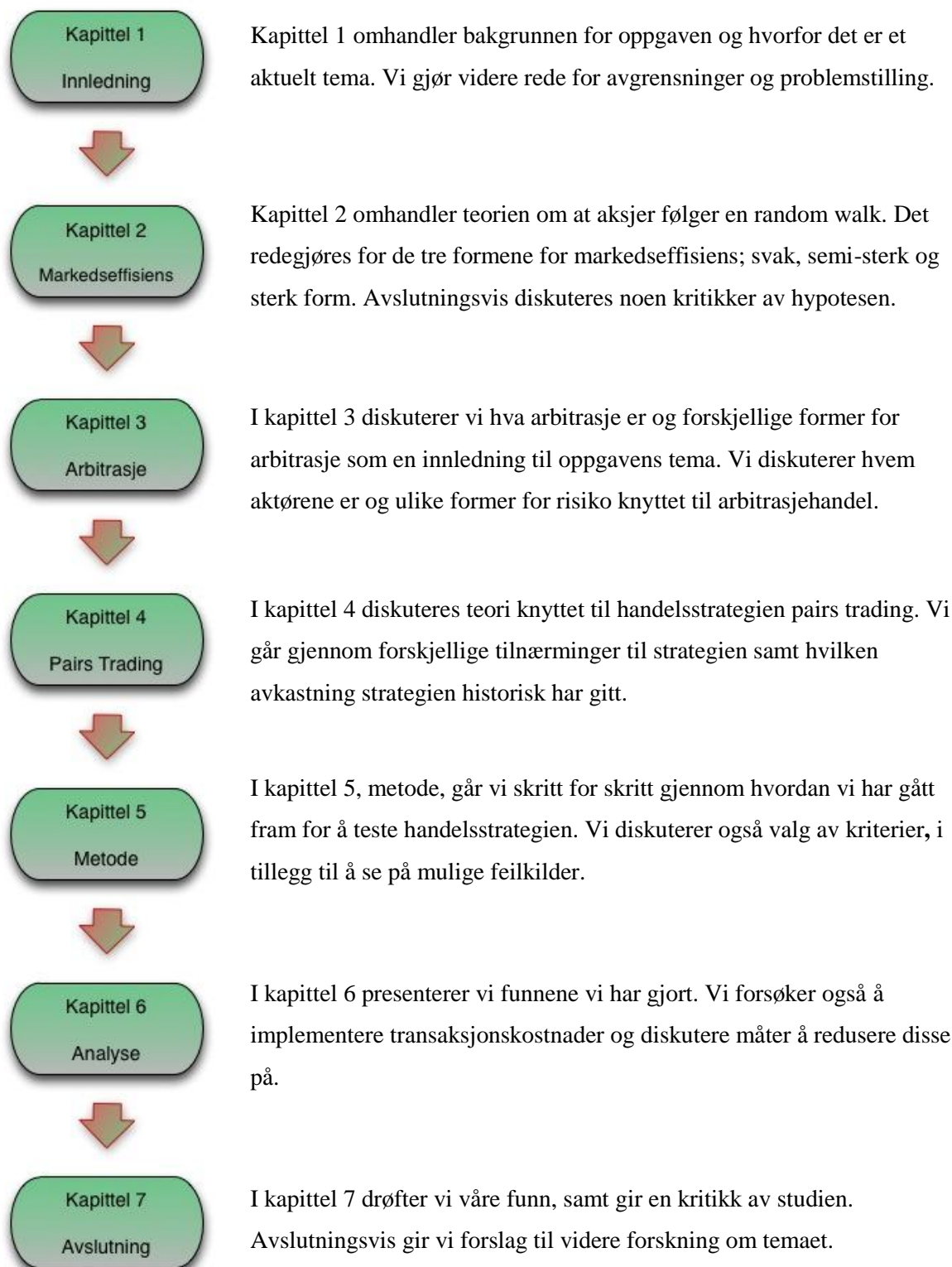
1.4 Problemstilling

Bakgrunnen for oppgaven er at vi ønsket å se på om "vår" pairs trading strategi oppnådde positiv avkastning kombinert med lav risiko justert for transaksjonskostnader. Det var videre naturlig å sammenligne våre resultater med tidligere studier. Testene gjorde vi på historiske data ved hjelp av statistikkprogrammet R.

Problemstillingen er derfor som følger:

"Gir handelsstrategien pairs trading meravkastning kombinert med lav risiko eller er avkastningen avtagende jamfør tidligere forskning?"

1.5 Oppgavens oppbygning



Figur 1 Oppgavens oppbygning

2. Markedseffisiens

2.1 Markedseffisienshypotesen

Etter at IT-teknologi fikk større utbredelse på 1950-tallet ble store mengder data tilgjengelig for forskere til bruk i empiriske studier. Dette inkluderte også studier av prisutviklingen på aksjer. Etter at dataene ble gjennomgått antydte flere studier at prisutviklingen på aksjer fulgte en random-walk modell. Dette stemte godt overens med matematikeren Bacheliers teori om hvordan priser utvikler seg fra 1900-tallet.

"His "fundamental principle" for the behavior of prices was that speculation should be a "fair game"; in particular, the expected profits to the speculator should be zero. With the benefit of the modern theory of stochastic processes, we know now that the process implied by this fundamental principle is a martingale" (Fama, 1970, p. 389).

I følge Björk (1998, p. 34) er egenskapene til en martingale slik at morgendagens forventede verdi er lik dagens observerte verdi, det er ingen drift – trinnvis utvikling i prisen. Fama (1970) beskriver fair game modellen formelt på denne måten, og vi vil kort kommentere den for å få fram noen poenger:

$$E(\tilde{p}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{jt}, \quad (1)$$

E er forventningsverdien. Prisen på aktivum j på tidspunkt t er angitt som p_{jt} . $r_{j,t+1}$ er periodeavkastningen til aktivum j. Symbolet Φ_t angir at all tilgjengelig informasjon på tidspunkt t er bakt inn i prisen. Tødlene over p og r angir at disse er tilfeldige variabler på tidspunkt t.

$$x_{j,t+1} = p_{j,t+1} - E(p_{j,t+1}|\Phi_t) \quad (2)$$

$$E(\tilde{x}_{j,t+1}|\Phi_t) = 0 \quad (3)$$

Likning (2) og (3) angir derfor at prosessen $x_{j,t+1}$ er et fair game med hensyn til informasjonen (Φ) tilgjengelig på tidspunkt t. Ettersom den forventete avkastningen er betinget informasjonen tilgjengelig på tidspunktet t, gjør dette at informasjonen allerede er bakt inn i prisen på tidspunkt t. Dette gjør at man ikke kan oppnå meravkastning utover forventet avkastning, med handelsstrategier basert på Φ .

Fama (1970) mente at antakelsen om at den observerte prisen på et aktivum fullt ut reflekterer all tilgjengelig informasjon, impliserte at suksessive endringer i prisen er

uavhengige. I tillegg var det en vanlig forutsetning at endringene hadde identisk fordeling. Dette førte til en utvidelse av fair game modellen og resulterte i random walk modellen i likning (4).

$$f(r_{j,t+1}|\Phi_t) = f(r_{j,t+1}) \quad (4)$$

Hvis man videre forutsetter at forventningsverdien til et aktivum er konstant over tid får vi likning (5).

$$E(r_{j,t+1}|\Phi_t) = E(r_{j,t+1}) \quad (5)$$

Man kan forstå av uttrykket at gjennomsnittet av distribusjonen til $r_{j,t+1}$ er uavhengig av informasjonen tilgjengelig på tidspunkt t . I tillegg er hele distribusjonen uavhengig av informasjonen (Φ) tilgjengelig på tidspunkt t .

Fama (1970) mener at man bør se på random walk modellen som en utvidelse av fair game modellen. Forskjellen mellom de to er at den førstnevnte tar hensyn til den økonomiske virkeligheten, slik som investorers preferanser og prosesser som gir ny informasjon bidrar til likevekter hvor avkastningsfordelinger repeterer seg over tid. Vi anser det som tilstrekkelig å bruke random walk og fair game som synonymer. Dette vi vil gjøre i resten av oppgaven.

Basert på egen og tidligere forskning publiserte økonomen Fama (1970) teorien om effisiente markeder eller markedseffisienshypotesen (EMH).

I følge markedseffisienshypotesen er alle aktiva korrekt priset i et velfungerende marked. Det sentrale er at prisene fullt ut reflekterer tilgjengelig informasjon. Den eneste grunnen til at man får høyere avkastning på enkelte aktivaklasser kontra andre forklares med høyere risiko og dermed en risikopremie for å holde aktiva med høyere risiko (Fama, 1970).

Etttersom markedseffisienshypotesen er nettopp en hypotese er det derfor ikke overraskende at den som andre økonomiske teorier forsøker å gi et tilstrekkelig, men dog forenklet bilde av virkeligheten. Følgende forutsetninger for hypotesen er gjeldende:

- (i) *there are no transactions costs in trading securities,*
- (ii) *all available information is costlessly available to all market participants, and*
- (iii) *all agree on the implications of current information for the current price and distributions of future prices of each security. In such a market, the current price*

of a security obviously "fully reflects" all available information (Fama, 1970, p. 387).

Selv om dette er ganske urealistiske forutsetninger for hvordan et marked fungerer mener Fama (1970) at dette er tilfredsstillende forutsetninger. Han argumenterer for at selv om det er høye transaksjonskostnader, noe som fører til færre handler og mindre likviditet, vil dette ikke være likestilt med å si at prisene ikke avspeiler all tilgjengelig informasjon. Med hensyn til forutsetningen om at alle markedsaktørene har tilgang til all markedsinformasjon, mener han at det er tilstrekkelig dersom majoriteten av aktørene opptrer rasjonelt.

Vi vil videre gå inn på de tre forskjellige formene for markedseffisiens som beskrevet av Fama (1970) før vi avslutter med noen kritikker av hypotesen.

2.1.1 Svak form for markedseffisiens

Den svake formen for markedseffisiens fokuserer på historiske data som pris og volum. Det vil si at man ikke vil kunne finne under- eller overprisede aksjer basert på historiske data fordi prisen på en aksje allerede reflekterer all tilgjengelig informasjon.

Å få tilgang til denne type data påfører ikke en investor store kostnader i verken ressursbruk eller penger. I følge Bodie et al. (2009) forutsetter man derfor at alle investorer har tilgang til dette, og dersom det dermed fantes muligheter for å utnytte mønster i aksjeprisingen ville denne umiddelbart fått en ny likevekt som følge av tilbud og etterspørsel. I et effisient marked selges alle aktiva til en rettferdig pris.

"Investors get exactly what they pay for when they buy securities, and firms receive exactly what their stocks and bonds are worth when they sell them" (Ross, 2007, p. 357).

I følge Bodie et al. (2009) finnes det derfor ingen måte å forutsi hvilken retning aksjene som er listet på verdens børser vil bevege seg basert på historisk utvikling. De kaller dette for en "random walk". Med dette mener de at pris på forskjellige aktiva kun beveger seg som følge av ny informasjon i markedet. Denne informasjonen er i seg selv uforutsigbar, slik at aksjeprisene umiddelbart stiger ved gode nyheter, og tilsvarende synker ved dårlige nyheter, uavhengig av historisk utvikling.

"The current price of a security "fully reflects" available information was assumed to imply that successive price changes (or more usually, successive one-period returns) are independent" (Fama, 1970, p. 386).

Siden historiske data ikke har noen betydning for framtidig pris vil dette si at man ikke vil kunne oppnå høyere avkastning på å benytte seg av teknisk analyse ifølge hypotesen.

Allikevel bruker fondsforvaltere store summer hvert år på aktiv forvaltning, hvor de leter etter f.eks. underprisede aksjer. De er villige til å bruke betydelige beløp på dette ettersom de bare trenger å få en marginal økning i avkastningen for å øke pengeavkastningen betydelig. I følge EMH skulle det ikke være noen grunn for forvaltere å bruke penger på dette siden aksjene allerede er riktig priset. Det argumenteres for at denne praksisen fører til høyere konkurranse, som i seg selv er med på å gjøre markedene mer effisiente. Fama mener selv at dette ikke motbeviser EMH, men at markedene derimot er tilstrekkelig effisiente.

Spørsmålet er derfor om spekulanter kan finne trender i historiske data som kan gi dem unormal høy avkastning. En måte å identifisere trender i aksjepriser er ved å se på seriekorrelasjonen, som angir tendenser for om avkastningen i aksjepriser kan relateres til tidligere avkastninger. Studier på dette finner en liten sammenheng på ukesbasis, men ikke klart nok til å hevde at det er grunnlag for en god strategi (Bodie et al. 2009).

Andre studier viser at aksjer som enten stiger mye eller faller mye har en tendens til å reversere i etterkant. Et eksempel på dette er en studie som tok for seg de 35 beste og de 35 dårligste aksjene over en fem års periode. Resultatet viste at over de neste tre årene presterte de 35 "dårlige" aksjene gjennomsnittlig 25 % bedre enn de 35 beste. Dette funnet konkluderer med at aksjer overreagerer på relevante nyheter for så å reversere når overreaksjonen er oppdaget (Bodie et al. 2009).

2.1.2 Semi sterk form for markedseffisiens

I den semi sterke formen for markedseffisiens forutsetter en også at all tilgjengelig informasjon om et firma er bakt inn i tilhørende aksjepris. I tillegg til informasjonen man benytter seg av i svak form tas det også hensyn til fundamentale firmaspesifikke data. Dette kan være regnskapspraksis, sammensetning av ledergruppe og immaterielle eiendeler m.m. Siden dette er informasjon som er offentlig tilgjengelig vil dette reflekteres i aksjeprisen, i følge hypotesen.

Fama (1970) konkluderer med at mange typer ny informasjon i markedet underbygger hypotesen. For og støtte hypotesen ser han på prisutviklingen etter at ny informasjon som f.eks. aksjesplitter, kvartalsrapporter og annet har blitt offentliggjort, for så å se hvordan markedet reagerer ved hjelp av event-studier.

“The available semi-strong form evidence on the effect of various sorts of public announcements on common stock returns is all consistent with the efficient markets model ” (Fama, 1970, p. 409).

Overraskende nok viser det seg at enkle fundamentale mål som PE ratio eller børsverdi synes å forutsette unormal risikojustert avkastning. En annen strategi som å kjøpe aksjer etter at de har lagt fram et positivt resultat virker å gi unormal risikojustert avkastning. Dette er i stor kontrast til semi sterk form for EHM som hevder at dette ikke er mulig om markedet er effisient (Bodie et al. 2009).

2.1.3 Sterk form for markedseffisiens

Her forutsettes det at investorer har tilgang til all informasjon som i de to andre formene. I tillegg har de også tilgang til innsideinformasjon om selskapene, og dette skal dermed reflekteres i markedsprisene. Dette er den mest ekstreme formen for markedseffisiens. Det er derfor i følge Bodie et al. (2009) laget reguleringer i markedene som skal forhindre innsidere i å oppnå ekstraordinær avkastning ved handel med innsideinformasjon. De mener derfor at man ikke anser markeder for å ha en sterk form for markedseffisiens.

Selv om man skulle tro at innsidere hadde en fordel med handel i eget selskap, viser en studie av Seyhun i Bodie et al. (2009) at man ikke oppnår unormal avkastning fratrukket transaksjonskostnader ved å følge offentlig kjente innsidekjøp.

I følge Fama (1970) kan sterk form for EHM best sees på som en referanse for å fastslå avvik fra markedseffisiens. Eksempel på slike avvik er i følge Fama (1970) spesialisten på en børs, samt innsidere. Spesialisten har monopol ved å kjenne hele ordreboken, mens innsidere ofte har monopol på informasjon om et selskap. Disse tilfellene er i følge Fama (1970) avvik som kan generere unormalt høy avkastning. Fama (1970) konkluderer videre med at alle andre (unntatt spesialisten og innsidere) må ta til takke med et sterkt effisient marked så lenge det motsatte ikke er bevist.

2.2 Kritikk av hypotesen

Markedseffisienshypotesen har mottatt kritikk, og da særlig fra forskere innen fagfeltet atferdsfinans. Bakgrunnen for kritikken er at man har observert hendelser med unormal avkastning som ikke kan forklares med bakgrunn i modeller, slik som prissettingsmodeller som kapitalverdimodellen og dividendemodellen. Studier har også funnet tilfeller hvor det har vært mulig å forutsi hvilken retning prisene utvikler seg fra en dag til neste, ved hjelp av seriekorrelasjon. Vi vil gjøre rede for noen av de viktigste kritikkene av hypotesen.

Lehmann (1990) viser i sin studie at det er grader av momentum i aksjeprisene på kort sikt. Han mener at markedet overreagerer på nyheter, slik at aksjeprisene stiger for mye ved gode nyheter og tilsvarende faller for mye ved dårlige nyheter. Forklaringen på dette mener han er flokkmentalitet (fads), og at det derfor finnes arbitrasjemuligheter som følge av mean reversal i påfølgende periode. Dersom markedseffisienshypotesen holdt skulle det ikke være muligheter for dette siden alle aksjene er riktig priset. Også andre forskere har i sine studier påvist at det går an å finne pristrender på kort sikt:

"In an efficient market, return predictability from past information should be short-lived and minimal. Given the evidence that such predictability does exist in the short run, understanding its time variation and its relation to other financial market attributes, such as liquidity, are of fundamental importance" (Chordia, Roll, & Subrahmanyam, 2008, p. 266).

Det har også blitt sett på unormal avkastning i enkelte måneder og dager. I følge Malkiel (2003) ble det observert at måneden januar ga unormal høy avkastning for en stund, i tillegg til at mandager også skulle gi meravkastning. Problemet med så banale handlingsregler er at når de blir kjent vil alle investorer prøve å oppnå denne avkastningen slik at handlingsregelen dreper seg selv. Man kan derfor si at for å opprettholde et effisient marked, må markedsaktører jakte etter godbitene slik at de forsvinner like raskt som de oppstår. Dette betegnes som effisiensparadokset.

Lakonishok, et al. (1994) finner i sin forskning at såkalte verdiaksjer gjør det bedre enn vekstaksjer. De viser at ved å benytte seg av motsatt strategi enn den som majoriteten av investorene bruker, er det mulig å oppnå ekstra avkastning, ettersom man kjøper underprisede verdiaksjer og selger overprisede vekstaksjer. Årsaken til at vekstaksjene presterer bedre, er i følge dem grunnet i at investorer er alt for optimistiske i sitt syn på fortsatt framtidig vekst. I tillegg styrer institusjonelle investorer muligens unna verdiaksjer av karrieremessige grunner. Malkiel (2003) mener derimot at årsakene til dette er at modellene som er brukt i disse studiene ikke fanger godt nok opp all risiko, slik at den høyere avkastningen kan forklares med en risikopremie, og at trefaktormodellen av Fama og French er bedre egnet til å forklare dette.

Det kan også være slik at størrelsen av et firma har utslag på avkastningen aksjen oppnår.

"Since 1926, small-company stocks in the United States have produced annual rates of return over 1 percentage point larger than the returns from large stocks." (Keim (1983) in Malkiel, 2003, p. 68)

Hvis man bruker KVM som forklaringsmodell gir dette små firma meravkastning ettersom selskapene gir høyere avkastning enn betaverdien skulle tilsi.

"Average returns on small (low ME) stocks are too high given their β estimates, and average returns on large stocks are too low." (Fama & French, 1992, p. 427)

Det kan være flere forklaringer på dette fenomenet. På den ene siden mener Bodie et al. (2009) at institusjonelle investorer bruker mindre ressurser på informasjonsinnhentning om små firma ettersom disse aksjene ofte er mindre likvide enn i større firma. Både mangelen på informasjon og dårlig likviditet kan være med å forklare den høyere avkastningen. På den andre siden mener Fama og French at det er feil med modellen, og at størrelse og forholdet mellom bokført EK / markedsbasert EK gir bedre estimater på risiko.

Til slutt viser det seg at økonomiske kriser kan gi en god pekepinn på at markedene ikke opptrer rasjonelt og effisient. Eksempler er børskrakket i 1987 hvor børsene falt med en tredjedel, dot.com boblen i begynnelsen av årtusenet og finanskrisen i 2008. Særlig etter finanskrisen har markedseffisienshypotesen fått stor kritikk. Enkelte av kritikerne går så langt som å si at hypotesen var skyld i krisen:

"They were blinded by an irrational faith in a discredited EMH and failed to see the bubble in asset prices and to give due warning of its collapse. The irony is that the strong implication of this hypothesis is that nobody, no practitioner, no academic and no regulator had the ability to foresee the collapse of this most recent bubble" (Brown, 2011, p. 93).

2.3 Oppsummering

I kapittelet om markedseffisiens har vi gjennomgått de sentrale sidene ved markedseffisienshypotesen. Først tok vi for oss utgangspunktet for hypotesen der man hevdet at aksjekursene kun beveger seg som følge av ny informasjon i markedet slik at forventningsverdien for morgendagens pris er lik dagens. Denne modellen kalles random walk, utledet av fair game modellen.

Videre diskuterte vi de ulike formene for markedseffisiens som begrenser seg til svak, halvsterk og sterk markedseffisiens med etterfølgende tester som er utført på de respektive formene. Til slutt presenterte vi kritikk av hypotesen på bakgrunn av at enkelte studier hevder å oppnå unormal avkastning som bryter med fundamentene i hypotesen.

3. Arbitrasje

3.1 Hva er arbitrasje?

Sentralt i økonomisk teori er det faktum at like goder må være priset likt.

“Therefore, any two portfolios that provide the same pattern of wealth must be worth the same amount.” (Varian, 1987, p. 59)

Dette innebærer at både like aktiva og prosjekter som gir samme konstantstrøm skal ha samme markedsverdi, noe som kalles ”The Law of One Price”. I et effisient marked skal det ikke være muligheter for å oppnå en gratis gevinst. Man skal ikke kunne oppnå avkastning utover risikofri rente uten å påta seg risiko. I motsetning til dette handler arbitrasje veldig forenklet om å utnytte prisforskjellene mellom to eller flere markeder, der formålet er risikofri avkastning.

“Basically, the no arbitrage condition must rule out "free lunches"-configurations of prices such that an individual can get something for nothing” (Varian, 1987, p. 59).

I følge Reverre (2001) har arbitrasje ytterligere en karakteristikk som skiller arbitrasje fra vanlig handel, nemlig det faktum at man tar motstående posisjoner i flere aktiva samtidig i motsetning til en investeringsstrategi hvor man vanligvis bare tar en posisjon. Følgelig er det slik at man må ta hensyn til kostnadene ved å innta en arbitrasjeposisjon, slik at en mulighet bare er lønnsom dersom gevinsten av feilprisingen er større enn summen av kostnadene knyttet til å innta posisjonen.

Vi skiller i følge Dubil (2004) mellom forskjellige typer arbitrasje; ren arbitrasje, relativ arbitrasje (risk/near) og rene spekulative strategier.

3.1.1 Ren arbitrasje

“A pure arbitrage opportunity (PAO) is a zero-cost trading strategy that offers the possibility of a gain with no possibility of a loss” (Bondarenko, 2003, p. 875).

Det er denne typen strategi Varian (1987) kaller en gratis lunsj. I følge Damodaran (2003) er man avhengig av to like aktiva for å få en ren arbitrasjemulighet, noe som er vanskelig å finne i dagens markeder. Det er naturligvis sjelden man finner slike muligheter, da et velfungerende marked stort sett vil være effisient og ikke la prisforskjeller i relativt like aktiva oppstå, noe som utelukker arbitrasjemuligheter. Dersom det allikevel skulle oppstå en

mulighet vil det grunnet de mange markedsaktørene raskt opprettes en ny likevekt som gjør at arbitrasjemuligheten forsvinner.

Som eksempel på rene arbitrasjemuligheter kan man dra paralleller til sportsbetting. De siste årene har det vært en stor vekst i denne industrien med store markeder på internett og tilsvarende mange aktører. Dette markedet tilbyr spill hvor oddsen er i kontinuerlig endring, derfor oppstår det noen ganger arbitrasjemuligheter. Man kan se for seg et spill med to utfall, og på grunn av feilprising har de flittige vært i stand til å oppnå risikofri gevinst uansett utfall i spillet. Et annet eksempel kan være en aksje som er listet på flere børser og har ulik pris. Man kan da kjøpe den billig på en børs og selge den dyrt på en annen børs uten risiko, forutsatt at det er likvide aksjer. Arbitrasjemuligheter i feilprisede valutakryss er også en mulighet. Naturligvis kan det ikke være slik at det ligger rene arbitrasjemuligheter billedlig rundt hvert hjørne. Hadde det vært slik hadde alle vært milliardærer nærmest uten innsats. I følge Dubil (2004) fokuserer derfor de fleste arbitrasjører på relativ arbitrasje.

3.1.2 Relativ arbitrasje

I følge Damodaran (2003) skiller relativ arbitrasje seg fra ren arbitrasje med at arbitrasjøren inntar posisjoner i veldig like aktiva som er priset forskjellig, eller i like aktiva som er relativt feilpriset. Dette medfører at han kvitter seg med markedsrisiko, men lar sekundærrisiko være udekket. Det er dermed ikke en gratis lunsj som i eksempelet med ren arbitrasje, ettersom arbitrasjøren fremdeles er eksponert for sekundærrisiko. Dette skyldes at det aldri er noen garanti for at posisjonene skal konvergere. Eksempler på relativ arbitrasje kan være å innta motsatte posisjoner i to langsiktige obligasjoner med ulik løpetid. Markedet vil forvente at rentesatsene på obligasjonene skal konvergere, mens man som arbitrasjører spekulerer i at de ikke konvergerer. I tillegg finnes det en rekke muligheter når firma endrer kapitalstruktur, som ved f.eks. fusjoner, tilbakekjøp av aksjer o.l.

3.1.3 Spekulativ arbitrasje

Det finnes i tillegg arbitrasjestrategier som er mer spekulative. Eksempler på dette er pairs trading som vi skriver mer om i neste kapittel.

3.1.4 Aktører

I følge Schleifer (1997) fungerer ikke modellen med kapitalfri arbitrasje i praksis. Ettersom det er små marginer som det handles på trenger man store volum for å oppnå en akseptabel kroneavkastning, noe som innebærer at aktørene som driver med arbitrasje trenger store kapitalbeløp både for å utføre handler og til å dekke tap. Siden det er de færreste som har

tilgang på så store kapitalbeløp betyr dette at det er få og høyt spesialiserte investorer som bruker sin kunnskap sammen med egen og/eller andres kapital til arbitrasjehandel.

Et annet interessant fenomen oppstår på grunnlag av bruken av fremmedkapital:

Investors may rationally allocate money based on past returns of arbitrageurs” (A. Shleifer, Vishny, Robert W., 1997, p. 37).

Dette forklarer at de fremmedkapitalfinansierte aktørene er avhengige av historiske gode resultater for å få tilgang på kapitalen de behøver. Shleifer (1997) viser videre at dette kan føre til tilfeller hvor de ikke fullt ut utnytter mulighetene i markedet. For eksempel kan feilprisingen som aktørene har spilt mot bli enda verre og føre til at de må lukke posisjonen som følge av press fra investorene. En annen grunn kan være at de er forsiktigere i sine handler for å opprettholde stabile prestasjoner, noe han mener kan begrense effisiensen i markedet. Arbitrasjørene oppfyller dermed ikke sin rolle i markedet med å tilføre likviditet når det trengs mest.

3.2 Risiko

3.2.1 Effektueringsrisiko

Effektueringsrisiko er knyttet til både inngåelse av posisjoner og likvidering av disse. Det er naturlig å snakke om sikringsstrategier knyttet til inngåelse av posisjoner. Vi skiller mellom statiske og dynamiske sikringsstrategier. Ettersom vi ikke skal bruke opsjoner ser vi det ikke som naturlig å definere dynamisk sikring. Ved en statisk sikringsstrategi forutsetter man at det finnes en statisk sammenheng mellom to aktiva i forhold til deres respektive størrelser på et sluttidspunkt. Man aksepterer at dette også gjelder for alle perioder før dette, og man holder derfor en konstant andel over hele leveperioden (Reverre, 2001). Arbitrasjehandel avhenger derfor av at man har proporsjonelle posisjoner på kjøps- og salgssiden. Hvis man har en long posisjon må man også ha en fornuftig størrelse short. I følge Reverre (2001) er det i et risikoperspektiv langt viktigere at man får balansert posisjonene, i motsetning til å innta store nominelle posisjoner som man ofte trenger for å utnytte arbitrasjemulighetene.

På samme måte som at det er viktig å ha riktige proporsjoner på kjøps- og salgssiden er det av stor betydning at transaksjonene skjer samtidig, timing er derfor avgjørende. I situasjoner hvor mange aktører ser en arbitrasjemulighet og forsøker å innta like posisjoner, vil en form for risiko være om man har mulighet til å innta posisjoner til de prisene man ønsker.

Vanskeligheter med å oppnå ønskede priser kan skyldes stor etterspørsel i markedet. Det kan

også være tilfeller med lite likviditet i markedene hvor det kan oppstå problemer med å gå inn i, lukke posisjoner, eller at det kun er mulighet for å lukke den ene siden av posisjonen. Om det er dårlig likviditet i markedet vil det derfor ut fra et risikoperspektiv være bedre å nøye seg med mindre posisjoner, kontra å ha for stor andel på den ene siden i håp om bedre likviditet. Vidyamurthy (2004) mener derfor at arbitrasjører bør fokusere på sin kjernekompetanse, og at outsourcing av handel til spesialiserte meglere kan være med å dempe denne type risiko. Videre mener han at man skal forsøke å oppnå så stor spread som mulig når man inngår en posisjon, samt så liten spread som mulig når man lukker en posisjon for å oppnå høyest mulig avkastning. Dersom megleren har stordriftsfordeler kan man dra nytte av dette for å oppnå beste spread.

3.2.2 Konvergeringsrisiko og likviditet

Som nevnt inntar arbitrasjører motstående posisjoner i aktiva i den tro at de skal konvergere til en forventet likevekt. Det er naturlig at forventningene ikke alltid stemmer, og posisjonen derfor ikke gir noe profitt. I følge Reverre (2001) er suksessfulle arbitrasjører flinke til å fastsette mål på hvor lang tid de ønsker å holde posisjonene åpne basert på historiske handler, samt å stole på egen magefølelse. Han sier at en av de viktigste egenskapene til arbitrasjører er evnen til å være disiplinerte og til å ta tap.

Årsaken til dette er at ettersom man ofte må låne penger for å få en stor nok handel, vil mange arbitrasjehandlere foregå på margin eller med sikkerhet i andre aktiva, da kapital er en knapp ressurs. Dersom man da har en posisjon i en forventet lønnsom arbitrasjemulighet og posisjonene fortsetter å divergere i stedet for konvergere så er det en kostnad involvert. Man kan få en margin-call fra långivere og dermed bli tvunget til å likvidere posisjonen med et tap, dette på det verst tenkelige tidspunkt ettersom posisjonen er på sitt mest lønnsomme og det optimale kanskje hadde vært å ta en enda større andel i posisjonen.

3.2.3 Short-squeeze

Short-squeeze er en viktig risiko å ta i betraktning da det ikke er mulig å sikre seg mot denne. I følge Reverre (2001) kan en short-squeeze skje når en aksje uventet får en kraftig stigning. Dersom det er investorer som forvalter egne penger har de som regel høyere risikoaversjon enn institusjonelle aktører. Dersom de førstnevnte har store salgsposisjoner og aksjen plutselig stiger vil de ønske å lukke sine posisjoner ved å kjøpe tilbake aksjene for å begrense tapene, som igjen fører til en flom av kjøpsordrer som fører til at prisene stiger.

Fusjoner kan være et annet tilfelle. Ved en fusjon vil som regel aksjeprisene til begge selskapene stige, og hvis det i tillegg kommer gode nyheter fortsetter prisen å stige. Dette gir insentiver til de som eier aksjene til å selge for å sikre seg profitt. I tillegg vil mange aktører være på utkikk etter arbitrasjemuligheter som fører til at etterspørselen etter aksjene stiger, noe som dermed gjør dem til en knapp ressurs og får prisen til å stige ytterligere.

3.2.4 Kredittrisiko

Dersom man har gjort en handel og motparten ikke kan oppfylle sine forpliktelser, er dette en risiko. Det kan være at en utsteder av obligasjoner går konkurs eller at man har inngått en futureskontrakt for levering av råvarer. For å sikre seg mot denne typen risiko kan man kjøpe forsikringer.

3.2.5 Long Term Capital Management

For å avrunde kapittelet om risiko er det naturlig å se på et eksempel som viser at det er langt fra risikofritt å drive med arbitrasjehandel. Long Term Capital Management (LTCM) var et stort hedgefond som på 1990-tallet samlet noen av de beste hodene i finans for å utnytte arbitrasjemuligheter i markeder over hele verden. I følge Damodaran (2003) hadde de som konkurransefortrinn at de kunne låne penger til en lav kostnad og dermed fikk de tilgang til en stor kapitalbase, noe som gjorde at de i flere år kunne slå de andre traderne på Wall Street.

På grunn av sin suksess fikk de tilgang til så store mengder kapital at de måtte begynne å se seg om etter mer spekulative arbitrasjemuligheter. Fondet argumenterte med at selv om posisjonene kunne være mer risikable, hadde de så mange posisjoner slik de at de var diversifiserte. Et tap på en posisjon ville jevnes ut av gevinst på en annen.

I følge Shleifer (2000) maktet ikke Russland å følge opp sine gjeldsforpliktelser i 1998, og devaluerte rubelen. I tillegg til at de la restriksjoner på russiske banker slik at de ikke fikk betale sine forpliktelser til vestlige kreditorer. Størsteparten av gjelden var holdt av hedgefond og investorer lokket av den høye avkastningen på obligasjonene. Siden fondene forutså at Russland ville misligholde den innenlandske gjelden og devaluere valutaen, tok hedgefondene salgsposisjoner i utenlandsopsjonene i den tro at landet også skulle misligholde disse. I tillegg solgte de terminkontrakter på russiske rubler slik at dersom landet devaluerte valutaen skulle de tjene penger på dette. Det viste seg at ingen av disse sikringsstrategiene fungerte ettersom Russland ikke misligholdt utenlandsgjelden og la

restriksjoner på bankene sine. Fondene måtte derfor ta store tap og markedssituasjonen gjorde at LTCM måtte reddes av den amerikanske sentralbanken.

Ettersom det ligger i arbitrasjørerers natur å forvalte store kapitalmengder bruker som nevnt de fleste fond lånt kapital. LTCM var intet unntak. De brukte en ekstrem grad av lånefinansiering i tillegg til en stor kapitalbase. I løpet av en måned tok de et tap på 3-4 milliarder dollar. I følge Damodaran (2003) er nettopp størrelse et tve-egget sverd, det gir kanskje stordriftsfordeler, men gjør det vanskelig å lukke posisjoner. I tillegg gjør bruken av lånefinansiering at selv små prisendringer kan gjøre store endringer i egenkapitalandelen. Til slutt viste historien om LTCM at selv de beste hodene i bransjen ikke kan sikre seg mot uforutsette hendelser i markedene.

3.3 Implikasjoner av EMH og arbitrasje

Markedseffisienshypotesen har vært dominerende siden 1970-tallet med uttallige artikler som både støtter og kritiserer den. Det har vist seg vanskelig å definitivt bevise eller motbevise hypotesen. Vi noterer at Bodie et al. (2009) mener markedene er meget effisiente, men at det kan være gevinster å hente for de spesielt kreative, intelligente eller flittige. Det er sentralt i økonomisk teori at to like goder skal være likt priset, dette er kjent som "The Law of One Price". I følge Gatev et al. (2006) kan man også forstå arbitrasjehandel som en test av denne hypotesen. For vår problemstilling som omfatter statistisk arbitrasje vil det ikke være mulig å hente noen gevinst dersom markedene var perfekt effisiente. Det blir derfor spennende å se i hvilken grad det kan være mulig å oppnå meravkastning også justert for transaksjonskostnadene.

3.4 Oppsummering

I kapittelet om arbitrasje tok vi først for oss hva det er, for så å stadfeste at EHM ikke støtter at det finnes gratis godbiter i markedene. Videre drøftet vi ulike typer arbitrasje som grener seg fra tilnærmet risikofri ren arbitrasje til relativ arbitrasje som medfører sekundærrisiko. Så forklarte vi hvilke aktører som i størst grad bedriver arbitrasje og hvilken tilnærming og utfordringer dette medfører. Til slutt diskuterte vi risiko og hvilke fatale konsekvenser dette kan få ved høy belåning.

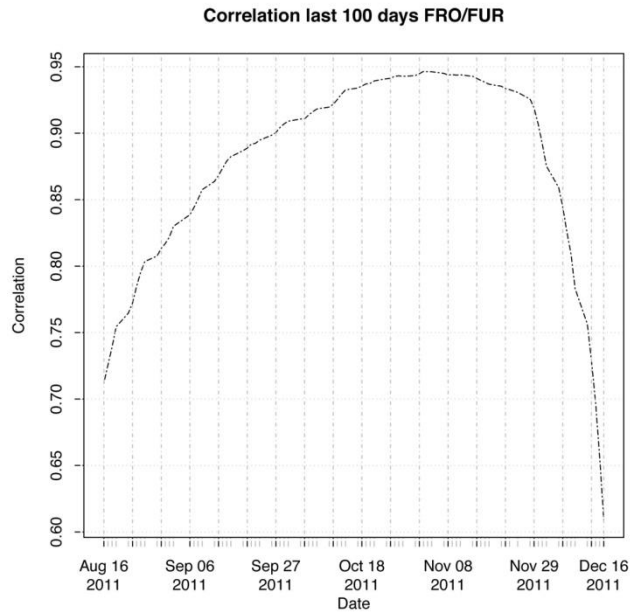
4. Pairs trading

4.1 Hva er pairs trading?

Investeringsmiljøer har lenge vært interessert i kvantitative metoder for å forbedre sine muligheter til investeringer og spekulasjon. En populær strategi for kortsiktig spekulasjon kalles pairs trading, også kjent som statistisk arbitrasje. Strategien har vært brukt de siste 30 årene, mest på Wall Street i regi av hedgefond og investeringsbanker. Det hele startet midt på 80-tallet da en gruppe med fysikere, matematikere og dataprogrammerere gikk sammen for å finne arbitrasjemuligheter i aksjemarkedet i regi av investeringsbanken Morgan Stanley. Tanken bak var å finne sofistikerte statistiske metoder som tok intuisjonen ut av spekulasjonen, og erstattet dette med mekanisk disiplin og handleregler. En av styrkene til strategien er utvilsomt graden av nøytralitet i forhold til svingninger i markedet og har siden oppdagelsen blitt en populær strategi som i dag også brukes av private aksjetradere (Gatev et al., 2006).

Konseptet bak pairs trading er ganske enkelt. Det hele handler om å finne to aksjer som over tid beveger seg likt. Det finnes ulike tilnæringer til dette, der en av disse er å måle korrelasjonen. Dersom to aksjer beveger seg likt impliserer dette at det er høy korrelasjon mellom aksjene (se figur 2). Poenget med dette er at man antar at to aksjer som historisk har en høy grad av korrelasjon også vil ha det i fremtiden (Whistler, 2004). En annen måte er å bruke kointegrasjon. Kointegrasjon har den fordelen at det impliserer en langsiktig likevekt, mens korrelasjon forutsetter en mer kortsiktig tidshorisont. Kointegrasjon vil bli nærmere beskrevet i kapittel 4.1.2.

Normalt finner man slike par innenfor en bestemt sektor fordi man antar at informasjon som påvirker aksjepriser er mer relatert til hele sektorer enn til kun et spesifikt selskap (Reverre, 2001). Når spreaden mellom de to aksjene blir unormalt stor selger man den overprisede og kjøper den som er underpriset. Tanken da er at historien vil gjenta seg og at spreaden mellom aksjene vil konvergere til et normalnivå hvor man trer ut av posisjonene med en profitt tilsvarende bevegelsene i aksjene. Spreaden som vi refererer til i denne oppgave er avstanden mellom prisen i to aksjer, altså aksjepris A minus aksjepris B.

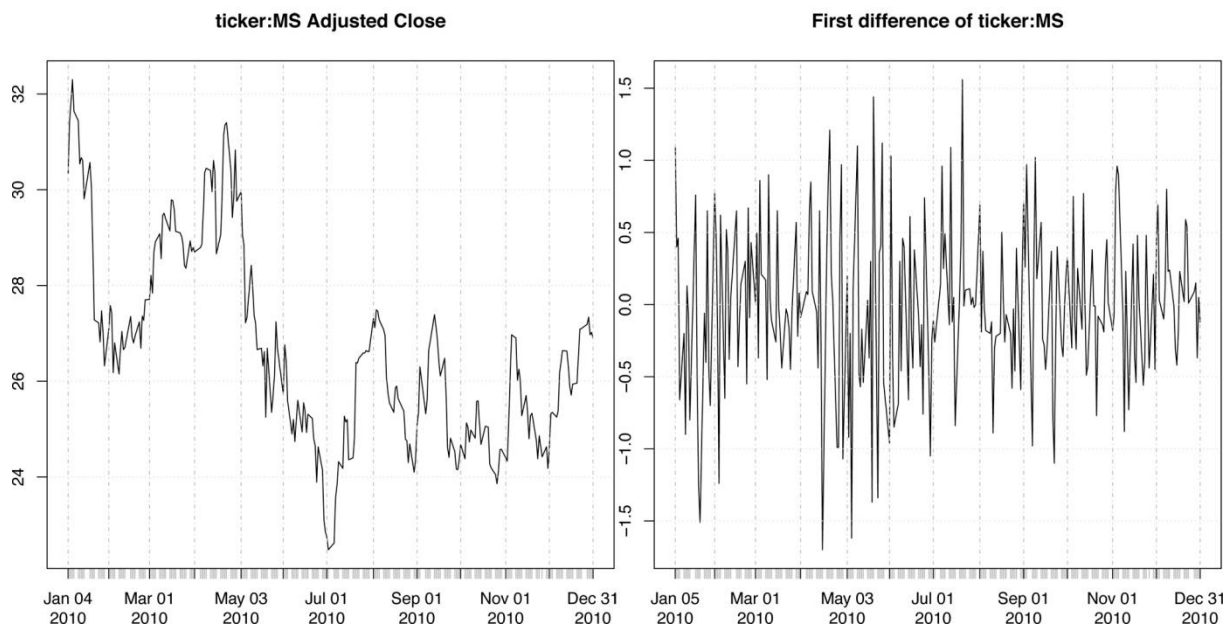


Figur 2 Korrelasjon mellom FRO og FUR

4.1.1 Stasjonære og ikke-stasjonære variabler.

De to neste kapitlene omhandler analyse av tidsserier med flere forklaringsvariabler som aksjekurser er et typisk eksempel på.

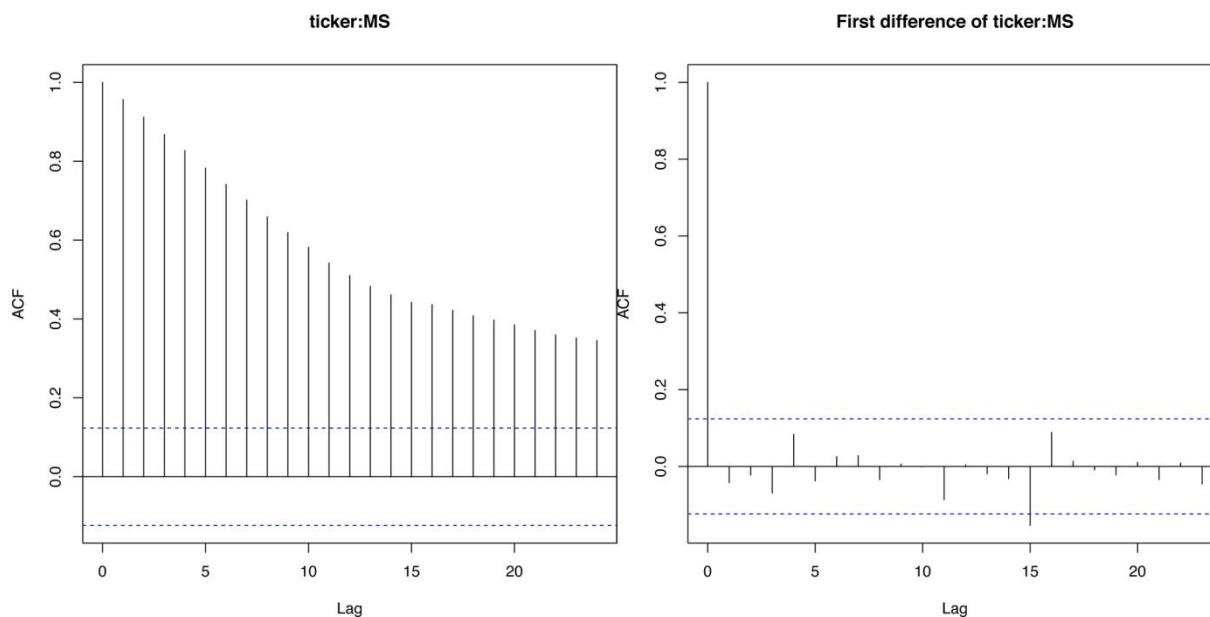
En stokastisk prosess kan i følge Lütkepohl og Krätzig (2004) defineres som stasjonær dersom dens første og andre moment ikke avhenger av tiden. Dette betyr at forventningsverdien til prosessen vil være gjennomsnittet av prosessen uavhengig av hvor på tidsaksen man befinner seg. I tillegg er den andre forutsetningen til en stasjonær prosess at den har en endelig varians og kovarians, som også er uavhengig av tiden. I tillegg må både forventningen, variansen og kovariansen være endelige tall, i motsetning til en ikke-stasjonær prosess som i prinsippet kan ha uavhengig forventning og varians.



Figur 3 Tidsserier

Figur 3 viser to tidsserier. Den til venstre viser en klar trend og vi kan dermed si at denne ikke er stasjonær. For å sjekke om en tidsserie er stasjonær kan man benytte seg av en Augmented Dickey Fuller (ADF) test. Den førstnevnte tidsserien fikk en testverdi på 1.47, og vi kan dermed ikke forkaste nullhypotesen om at tidsserien er ikke-stasjonær på noen signifikansnivå. På høyre side har vi en tidsserie som ser ut som den beveger seg rundt et fast gjennomsnitt. Den fikk i ADF testen en testverdi på -12.8 og vi kan dermed forkaste nullhypotesen på alle signifikansnivå. Tidsserien er dermed stasjonær.

Videre kan man også sjekke hvordan autokorrelasjonsfunksjonen ser ut for å si noe om en tidsserie er stasjonær. I figur 4 ser vi at til venstre avtar autokorrelasjonen sakte, men til høyre avtar den raskt og skifter fortegn flere ganger. Det sistnevnte indikerer en stasjonær tidsserie.



Figur 4 Autokorrelasjon

Det er vanlig å definere aksjepriser som ikke-stasjonære.

”The log levels of asset prices are usually treated as $I(1)$ with drift. Indeed, the random walk model of stock prices is a special case of an $I(1)$ process” (Zivot & Wang, 2006, p. 120).

Det at en prosess er integrert i første orden $I(1)$ betyr i følge Lütkepohl og Krätzig (2004) at den ikke-stasjonære prosessen kan bli stasjonær ved å se på den første forskjellen ($\Delta Y = Y_t - Y_{t-1}$). I følge Gatev et al. (2006) er risikoen for konkurs en av grunnene til at individuelle aksjekurser ikke kan ses på som stasjonære. I tillegg skriver Vidyamurthy (2004) at det er en vanlig forutsetning at den naturlige logaritmen av en aksjekurs er en random walk, ergo ikke stasjonær. Dette betyr at aksjepriser forutsettes å være ikke-stasjonære.

4.1.2 Kointegrasjon

Engle og Granger (1987) oppdaget at selv om to tidsserier er ikke-stasjonære så er det mulig i noen tilfeller å finne en lineær kombinasjon av de to som er stasjonær. Denne oppdagelsen fikk så navnet kointegrasjon. Kointegrasjon impliserer at på lang sikt opprettholdes en likevekt med en stasjonær og begrenset varians i motsetning til ikke-stasjonære tidsserier med uendelig varians. Det er verdt å merke seg at de to økonomene fikk Nobelprisen i økonomi for denne oppdagelsen.

I følge Vidyamurthy (2004) er ideen bak kointegrasjon i en formell sammenheng slik: La Y_t og X_t være to ikke-stasjonære tidsserier. Hvis en gitt verdi av Y , gir at $Y_t - \gamma X_t$ er stasjonær,

da er de to seriene kointegrerte. I ettertid har metoden blitt brukt i en rekke makroøkonomiske sammenhenger; konsum og inntekt, kort- og langsiktig rente, pengeetterspørsel og BNP m.m. Hovedresonnementet bak kointegrasjon er en langsiktig likevekt som er gjennomsnittet av den lineære kombinasjonen. Poenget er da at hvis det er et avvik fra det langsiktige gjennomsnittet så vil en eller begge tidsseriene justere seg selv for å gjenopprette den langsiktige likevekten (Vidyamurthy, 2004). Fra dette poenget oppsto blant annet handelsstrategien pairs trading.

4.1.3 Noen avklaringer

Pairs trading søker å finne relativ feilprising mellom to aksjer. Dette gjøres som regel på to måter, som er kalt statistisk arbitrasje og risiko arbitrasje. Statistisk arbitrasje har som formål å finne feilprising igjennom å analysere en tidsserie av prisinformasjon, mens risikoarbitrasje handler om strategier knyttet til fusjon av to selskaper (Gatev et al., 2006).

I følge Avellaneda & Lee (2010) omfatter statistisk arbitrasje en rekke strategier og investeringsprogrammer. Likhetene dem imellom er som følger:

1. Signalene til å handle er systematiske eller regelbaserte i motsetning til å bli drevet av fundamentale forhold.
2. Ulike posisjoner medfører at de sammen er markedsnøytrale med en beta lik null.
3. Mekanismen for å oppnå meravkastning er statistisk.

Metoden som beskrives i denne oppgaven er opptatt av at to aksjer konvergerer mot en statistisk likevekt. Men Whistler (2004) beskriver også varianter som spekulerer i forventning om at to aksjer skal divergere fra en likevekt basert på fundamentale forhold, kalt divergence trading. Han påpeker at dette kan være en mer komplisert måte å handle på. Dette fordi man aldri kan være sikker på når en topp er nådd og parene begynner å konvergere, noe som medfører større usikkerhet. På den andre siden kan man på en enkel måte kan skaffe seg tilgang til store mengder data og på denne måten regne ut gjennomsnittlig verdi. Derfor mener han at dette er en enklere og bedre metode, ettersom det er lettere å vite når aksjene er på gjennomsnittet, enn i ekstremiteten.

4.1.3.1 Hvorfor metoden hevdes å fungere

Hovedmannen bak strategien pairs trading, Nunzio Tartaglia mener psykologi gjør at metoden fungerer fordi mennesker har en tendens til å handle aksjer som går opp og ikke de som går ned. Dette kan tolkes som at strategien prøver å utnytte overreaksjonene til udisiplinerte investorer (Gatev et al., 2006). Grøtte (2002) forklarer at de fleste aktørene legger sine kjøps- og salgsordre i et selskap på bakgrunn av tro om oppgang eller nedgang i en aksje, ikke på spreaden mellom aksjer. På bakgrunn av dette mener han at det hele tiden vil oppstå muligheter for pairs trading.

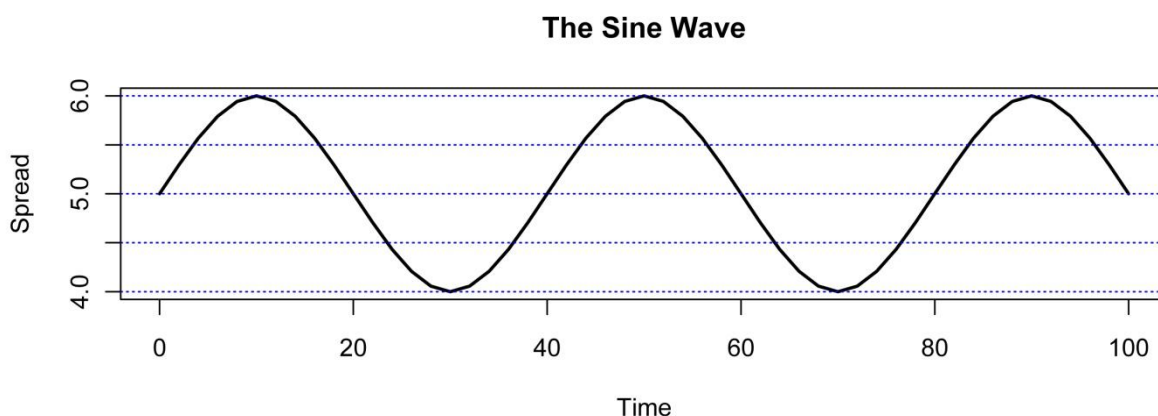
4.2 Markedsnøytral strategi

En markedsnøytral strategi innebærer å holde både en kjøpsposisjon og en salgsposisjon med lik eksponering av markedsrisiko hele tiden. Dette blir gjort ved å sammenføre de to vektete betaene fra både kjøps- og salgsposisjonen. Resultatet blir da at porteføljen ikke lenger er eksponert mot markedsrisiko og korrelerer derfor ikke med markedet, som igjen betyr at strategien er markedsnøytral. Man kan derfor fastslå at avkastningen til en markedsnøytral strategi er uavhengig av markedsavkastningen.

4.2.1 Pairs trading med enkel spread som mål

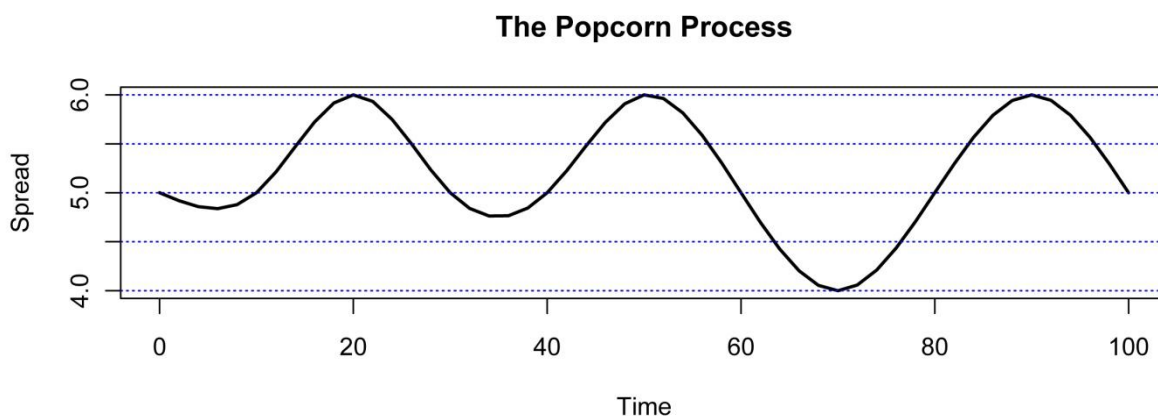
En av de første strategiene som ble brukt i pairs trading var å se på spreaden mellom to aksjer som beveget seg over tid i samme retning. Ved å identifisere bevegelsene i spreaden så kunne man lage en handleregel hvor man skulle kjøpe en aksje og selge en annen. For eksempel kunne to aksjer ha en spread som varierte mellom \$-2 og \$6. Handleregelen kunne da være å legge både kjøps- og salgsordre når spreaden var på \$4, for så å stenge posisjonen når den gikk tilbake til \$0. Denne strategien var ekstremt profitabel på 80- tallet da det fantes tusenvis av slike par og spekulere i. Pole (2007) mener at mye av grunnen til at dette var mulig, skyldtes at kun de store aktørene hadde innsikt, resurser og kapasitet til å utføre strategien. De hadde tilgang til datamaskiner som kunne behandle store mengder data, og på den tiden kostet maskinene flere titalls tusen dollar. I tillegg var det kun institusjonelle aktører som fikk daglige oppdateringer av aksjepriser. Dette var en barriere som forhindret andre fra å utnytte strategien.

Til å begynne med brukte man som sagt spreaden mellom to aksjer som beveget seg likt og forutsatte at den ville systematisk variere, fra vesentlig over gjennomsnittet til vesentlig under gjennomsnittet. Dette blir gjengitt i en sinuskurve som vist i figur 5. I starten var dette den teoretiske modellen for spread analyse.



Figur 5 The Sine Wave (Pole, 2007, p. 19)

Handleregelen var at man skulle åpne posisjonene på toppnivået, og deretter lukke dem på bunnivået av spreaden. Det viste seg derimot at med denne modellen mistet man mange muligheter, ettersom flere topper ofte fulgte hverandre i stedet for de svingningene man ser i en sinuskurve. Det medførte at man med denne metoden bandt kapitalen over lengre tidsperioder som forårsaket høyere risiko, enn ved å bruke popcornprosessen som beskrevet under.



Figur 6 The Popcorn Process (Pole, 2007, p. 19)

I figur 6 ser man popcornprosessen som ga ny innsikt til metoden. Her var tanken at spreaden ofte reverserer seg tilbake til gjennomsnittet slik at om man stoppet posisjonene der så ville flere muligheter åpne seg. En annen positiv ting med metoden er periodene uten eksponering. I dag har de aller fleste strategiene knyttet til pairs trading sin bakgrunn i forståelsen av popcorn prosessen (Pole, 2007).

Selv om det viktige her er å forstå grafene kan sinusfunksjonen skrives slik: $Y_t = \sin(t)$, mens popcornfunksjonen kan skrives: $Y_t = I_t \sin(t)$ der I_t er en indikator funksjon som ved verdiene 1 eller -1 signaliserer en topp eller bunn.

4.2.2 Pairs trading ved bruk av standardavvik

Selv om pairs trading er fellesbetegnelsen på long/short posisjoner i to selskaper som beveger seg likt finnes det utallige måter å gjøre dette på. Beregninger av korrelasjon, kointegrasjon, spread, gjennomsnitt og standardavvik er noe som hver enkelt aktør definerer ut i fra egne preferanser. Noen aktører tenker langsiktig med å finne feilprising ut i fra en lang tidsperiode på flere måneder og år, mens andre ser på relativ feilprising innenfor et par minutter.

Grøtte (2002) presenterer en metode av pairs trading som også betegnes som spread trading. Metoden er svært vanlig blant dem som trader profesjonelt. Fordelene med metoden skal være at den er nøytral i forhold til markedet og har relativt lav risiko, da sjansen for store tap er liten, samtidig som store gevinstmuligheter heller ikke er særlig sannsynlig. Argumentet for metoden er altså mange små gevinster over tid. Som drøftet tidligere handler metoden om å innta posisjonen når det statistisk og historisk er stor sjanse for at de to aksjene konvergerer etter å ha beveget seg bort fra likevekten en kort periode. Som annen profitabel aksjehandel er prinsippene om å følge trenden, kutte tapene og holde på vinnerne noe som også gjelder denne metoden (Grøtte, 2002).

Metoden til Grøtte (2002) handler om å normalisere verdien på spreaden mellom to aksjer som er beskrevet av Reverre (2001). Dette medfører at verdien i handleregele handler om feilprisingen av spreaden uttrykt i form av standardavvik som uttrykker deviasjoner fra normalen, og posisjonene inntas ved statistisk signifikante nivåer. I metoden til Grøtte gjøres dette når prisen har beveget seg minimum to standardavvik fra normalverdi. Formelen uttrykkes slik.

$$\Delta_G = \frac{\Delta - MA_{500} \Delta}{SD_{500}(\Delta)} \quad (6)$$

(Grøtte, 2002, p. 511)

Der Δ er forskjellen mellom de to aksjene i nominell valuta. $MA_{500}(\Delta)$ er glidende gjennomsnitt de siste 500 minutter på forskjellen mellom aksjene. $SD_{500}(\Delta)$ er standardavviket de siste 500 minutter på forskjellen og Δ_G uttrykker feilprisingen i spread uttrykket ved standardavvik.

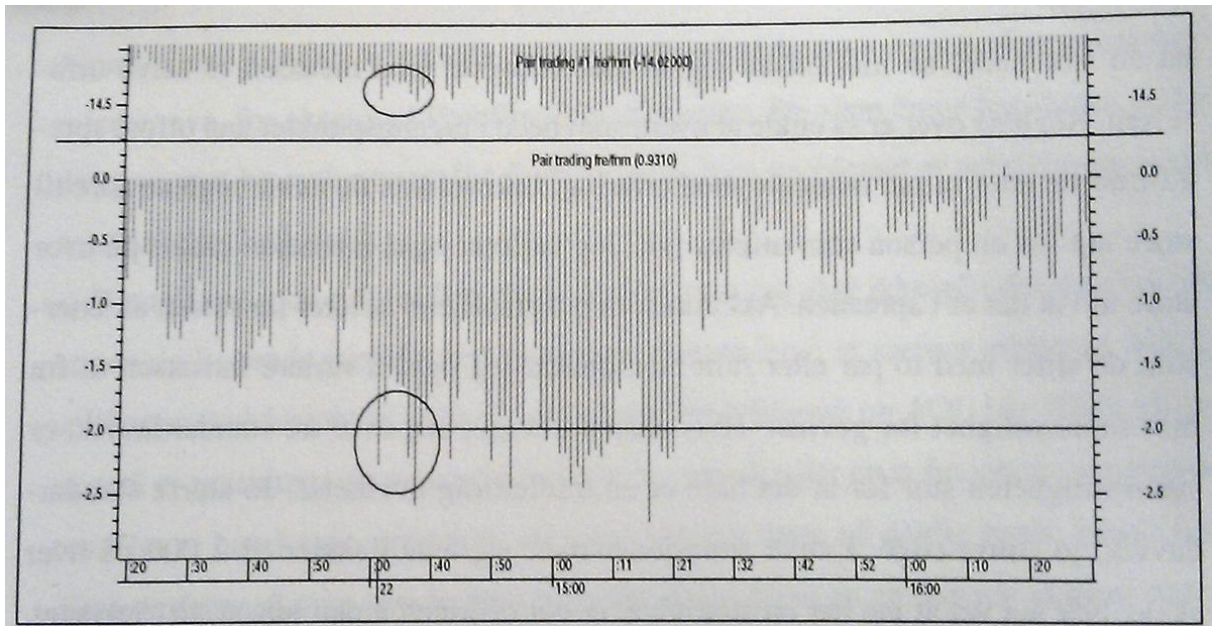
Formelen som er vist ovenfor har sitt utgangspunkt i formelen til Reverre, som er helt lik bortsett fra at den bruker 30 dager istedenfor 500 minutter. Formelen ser derfor slik ut:

$$\Delta_{norm} = \frac{\Delta - MA_{30} \Delta}{\sigma_{30}(\Delta)} \quad (7)$$

(Reverre, 2001, p. 476)

4.2.3 Standardavvik eller spread ratio

Metoden som beskrives ovenfor tilsier at spreaden og avviket er en tilfeldig variabel som er normalfordelt med snitt rundt null og standardavvik rundt en. Ved å ta denne forutsetningen så kan man anta distribusjonen til standardavviket, noe som ikke er like innlysende ved spread ratio. For eksempel kan man anta at det er ca. 15 % sannsynlighet for at standardavviket vil være enten over 1 eller under -1, samt at det er ca. 2,3 % sannsynlighet for at det er enten over 2 eller under -2. Dette gjør at man kan være rimelig sikker på at avvik utover dette er ekstremnivå eller midlertidige feilprising som snart vil medføre at aksjene vil bevege seg inn igjen mot likevekten. I metoden til Grøtte (2002) skal man innta posisjon når spreaden avviker med et standardavvik på 2 i minutt-chartet som illustrert i figur 7.



Figur 7 Chart (Grøtte, 2002, p. 513)

Pole (2007) poengterer at det ved bruk av standardavvik oppstår en feil siden alle observasjoner inkludert ekstremverdien blir kvadrert. Dette medfører at ekstremverdiene får proporsjonelt større vekt enn det de burde ha. Han argumenterer derfor for å fjerne ekstremverdier eller mer komplekst å justere dem ned. En annen feil med standardavvik er at finansiell data ikke er normalfordelt noe som kan gi feilkalkulasjoner som igjen medfører undervurdering av risiko. Hvis standardavvik gir slike feil, hvordan har det seg at aktører velger dette foran empirisk distribusjon, altså at de kun ser på spreaden? Pole (2007) forklarer dette med at ekstremverdiene som oppstår lettere kan hanskles med hvis man ser på det analytisk igjennom standardavvik i motsetning til kun å se på spread. Samtidig er det slik at standardavvik gir et tilfredsstillende resultat som derfor anses som godt nok.

4.2.4 Metode ved bruk av beta.

La oss si at aksje P og Q er i den samme sektoren og i tillegg er ganske like selskaper som for eksempel Exxon Mobile og Conoco Phillips. Man forventer da at avkastningen til de to aksjene vil følge hverandre etter å ha kontrollert for beta. Følgelig hvis P_t og Q_t uttrykker en gitt historisk tidsserie av priser så kan vi modellere dette som

$$\ln(P_t/P_{t0}) = \alpha(t - t_0) + \beta \ln(Q_t/Q_{t0}) + X_t \quad (8)$$

Dette betyr at etter å ha kontrollert for beta så vil long-short porteføljen svinge i nærheten av en statistisk likevekt. Modellen foreslår en investeringsstrategi der vi går long en dollar i

aksje P og short β dollar i aksje Q hvis X_t er liten og motsatt, gå short P og long Q hvis X_t er stor. X_t uttrykker avstanden fra likevekten (cointegration residual). Porteføljen forventes da å oppnå en positiv avkastning når verdien konvergerer mot likevekten (Avellaneda & Lee, 2010).

4.3 Pairs trading i praksis

4.3.1 Strategiens profitabilitet

Potensiell avkastning ved pairs trading avhenger selvsagt av hvilke kriterier en forvalter bruker. Gatev et al. (2006) slår i sin artikkel fast at pairs trading er profitabel. Studien er basert på data fra 1962 – 2002 ved bruk av sluttpriser for hver dag. Handelsregelen de har brukt er den samme som traderne de intervjuet brukte. Dette innebærer å åpne posisjoner i et par når prisene gir et avvik på to historiske standardavvik. For nærmere beskrivelse av metoden og beregning av avkastning så henvises det til artikkelen. Studiet viste at de fem beste parene ga en gjennomsnittlig avkastning på 1,31 % per måned, mens de 20 beste en avkastning på 1,44 % per måned, hvor avkastningen ble beregnet ut fra antall par som faktisk handlet i perioden. Fra 1999 til 2002 ga de 20 beste parene en årlig meravkastning på 10.4 %, en signifikant størrelse. Et interessant funn i studien viser at diversifiseringseffekten ved å øke antall par gir en høyere minimumsavkastning samtidig som maksimumsavkastningen holder seg relativt stabil. I perioden på 474 måneder, ga de 20 beste parene 71 måneder med negativ avkastning, mens de 5 beste ga 124 måneder med negativ avkastning, som igjen antyder mer stabilitet ved diversifisering. Denne artikkelen, samt Pole (2007) viser at pairs trading har hatt fallende profitabilitet siden sin storhetstid på 80-tallet. Dette forklares med den økende konkurransen blant de som praktiserer strategien, da det er blitt flere som ønsker sin del av avkastningen. I en artikkel av Do & Faff (2009) forsøker de å replisere studiet til Gatev et al. De klarer i ganske god grad å replisere funnene i perioden de sistnevnte studerte. Men de utvider i tillegg tidshorisonten til å gjelde perioden 2003 – 2008, hvor gjennomsnittsavkastningen faller til magre 0,06 % i perioden. Dette er i følge dem et resultat som ikke er statistisk signifikant annerledes fra 0 %.

4.3.2 Fundamentale forhold

Selv om utgangspunktet for denne oppgaven omhandler statistisk arbitrasje eller pairs trading er det også hensiktsmessig å ha en forståelse av de fundamentale driverne rundt kursbevegelsene i markedet. Årsaken til dette er at de statistiske signalene som oppstår i bunn og grunn skyldes fundamentale forhold. Poenget her er altså i følge Whistler (2004) å studere noen få fundamentale forhold slik at man kan unngå feil i de tilfellene statistikken er feilslått. Man blir altså en mer effektiv spekulant ved å dobbelsjekke statistikken mot fundamentale vurderinger. Han har følgende eksempler på fundamentale vurderinger som bør sjekkes:

Aksjene som handles bør være likvide aksjer. Begrunnelsen for dette er ganske enkelt at man lettere får kjøpt og solgt aksjene når man selv ønsker det. Handler man mindre likvide aksjer så kan man risikere og ikke få kjøpt/solgt alle postene til den pris og det tidspunktet man ønsker.

Vekst – forholdstall er viktig for å forstå fremgangen til et hvert selskap. Har man for eksempel en salgsvekst på 80 % mens inntektsveksten kun er på 10 %, så indikerer dette at noe er galt. Ulike veksttall mellom to selskaper kan derfor indikere at de vil divergere fra hverandre. En god regel er å se på tallene i forhold til gjennomsnittet innenfor samme sektor som selskapet.

Pris-fortjeneste (PE) er et forholdstall som for mange kan virke som irrelevant siden det ikke sier noe om fremtiden. PE finnes med å ta nåværende pris av aksjen og dele på nåværende fortjeneste per aksje (EPS). Noen mener at PE bør være rundt 1, mens andre mener den bør ligge mellom 20 til 40 for å indikere at den handles med en premie til den nåværende verdivurderingen. En nyere versjon av PE er PEG-forholdstallet som fremkommer ved å dele PE på den årlige EPS veksten. Dette forholdstallet sies å ha bedre evne til og prediktere fremtidig prisutvikling (Whistler, 2004). For eksempel så vil en PEG på 3.0 indikere at aksjen er overkjøpt og at investorer dermed forventer en tilbakegang i prisen. Tallene varierer allikevel mye fra en sektor til en annen, så man må derfor gjøre sammenligninger deretter. Andre viktige vurderinger knyttes til gjeldssituasjonen i selskaper, samt likvide midler til å betjene denne gjelden. Poenget er i følge Whistler (2004) at man ikke trenger å ha et totalt overblikk over den økonomiske situasjonen i selskapene, men at man bør ha en ide om hvor de ligger i terrenget.

4.3.3 Money management og risiko ved pairs trading

Selv om pairs trading hevdes å kunne gi meravkastning er det viktig å forstå risiko som er forbundet med strategien. En måte å håndtere risiko på, er å forstå hva som kan gå galt, og ha en plan for hva som skal gjøres i enhver situasjon som kan oppstå. Man kan beskytte kapitalen gjennom god money management. Whistler (2004) mener at selv om pairs trading består av hedging så medfører strategien mer risiko enn ved å investere i åpne posisjoner. Han begrunner dette med at man holder to posisjoner samtidig, og dersom posisjonene går feil vei, så har man dobbelt så stor eksponering mot tap. Et annet poeng er at transaksjonskostnadene er dobbelt så høye som ved vanlig handel. Han mener derfor at forståelsen av både tekniske og fundamentale forhold er avgjørende for profitabel utførelse av pairs trading. Videre poengterer han at når et par divergerer så må man være opptatt av å forstå de fundamentale forklaringene for at dette skjer. Man må derfor skille mellom årsaker som vil føre til en varig endring og de som mest sannsynlig kun er midlertidige og fører til at parene igjen vil konvergere.

Et annet sentralt poeng som Whistler (2004) påpeker er forståelsen av at menneskers følelser i investering omhandler frykt og grådighet. Han viser til menneskers svakhet når det kommer til å kutte tap og la vinnere løpe, og å holde seg konsistent til en strategi.

Han kommer med konkrete tips om å sette stopp ved 8.5 % på hver side av posisjonene. Maksimalt tap vil da være 18 %. Man kan være mer konservativ ved å gå så lavt som til 2,5 % - 5 % for å kutte tap. Et poeng som forklarer dette bedre er og aldri la en posisjon kunne skade den totale egenkapitalen med mer en 2,5 %. Man kan derfor si at jo større andel av egenkapitalen som eksponeres i en posisjon, jo lavere "stop loss" bør settes. En annen måte å justere risiko er ved å justere nivå av standardavvik. Ved å ha en handleregel som setter posisjoner på 3 standardavvik istedenfor 2, vil dette medføre et mindre antall handler, men som igjen potensielt gir bedre odds de gangene 3 standardavvik oppstår. Whistler argumenterer også med at pairs trading er mye mer avansert enn vanlig retningsbestemt aksjehandel. Hans begrunnelse for dette er at han mener man må gjøre de samme tekniske og fundamentale vurderingene ved å investere i en aksje, bare at her må man gjøre dette med to. For eksempel mener han at en stop-loss ikke nødvendigvis bør legges mekanisk, men at man ved bruk av teknisk analyse bør se på tekniske triggere som relevante plasseringer av stopp.

Metoden til Whistler (2004) er altså mer omfattende og stiller store krav til undersøkelser ved utførelse av pairs trading. Fundamentale forhold må heller ikke overses da statistiske

signaler ikke fanger opp om et selskap for eksempel er i økonomiske problemer. Han avslutter sitt resonnement med å fastslå at man kun vil handle profitabelt ved pairs trading om man legger en større innsats i arbeidet enn sine konkurrenter.

4.4 Oppsummering

Kapitlet om pairs trading starter først med å forklare hva pairs trading er. Deretter forklarer vi hvordan man kan definere om en tidsserie er stasjonær eller ikke-stasjonær. Videre forklares det hvordan man ved hjelp av kointegrasjon kan finne en lineær sammenheng mellom to ikke-stasjonære tidsserier. For så å forklare hvorfor metoden lar seg gjøre.

Underkapitlet om markedsnøytral strategi forklarer hvorfor strategien gir en beta rundt null som impliserer en markedsnøytral strategi. Videre beskrives hvordan strategien først ble praktisert med bruk av enkel spread, mens den senere har utviklet seg med bruk av spread ratio og avvik i standardavvik. I underkapittel 4.3 beskriver vi at strategiens profitabilitet basert på tidligere studier er dalende. Videre diskuteres en metode for å implementere fundamentale forhold som en del av strategien. Til slutt diskuteres money management og risiko ved pairs trading.

5. Metode

5.1 Analyseverktøy

For å kunne gjøre våre analyser har vi vært avhengige av et godt verktøy. Vi besluttet å benytte oss av statistikkprogrammet R², som er et gratis open-source program. Fordelen med R er at man kan tilpasse det slik at det passer til den problemstillingen man arbeider med ved hjelp av tilleggspakker. Innen fagfeltet finans finnes det en rekke utviklingsprosjekter hvor programmere tar i bruk de til enhver tid nyeste metoder innenfor fagfeltet og lager pakker som kan brukes til analyse av disse.

Ulempen med R er for det første at det er en ganske høy brukerterskel for å forstå programmet om man ikke er kjent med programmering. Vi har brukt mange timer på å forstå hvordan selv enkle operasjoner skal gjøres. I tillegg har det vært delvis tynn dokumentasjon på enkelte av funksjonene i R, ettersom disse fremdeles er under utvikling.

For å teste par for kointegrasjon har vi tatt i bruk pakken 'urca'³ hvor man finner Johansen testen. Vi har brukt pakken 'PairTrading'⁴ for å teste om spreaden mellom aksjene er stasjonær ved hjelp av ADF og PP testen.

For analysen i handleperioden har vi brukt pakkene i prosjektet 'Trade Analytics'⁵, i tillegg til pakken 'PerformanceAnalytics'⁶ for utregning av deskriptiv statistikk.

5.2 Utvelgelse av aksjer, sektor og tidsperioder

Vår metode startet med å definere hvilke aksjer vi skulle bruke i studien. På forhånd definerte vi at børsene New York Stock Exchange og NASDAQ Stock Exchange var godt egnet på bakgrunn av at disse børsene gir en stor bredde av industrier med store likvide selskaper. De fleste studier om temaet pairs trading er utført i disse markedene som igjen gjør at vi kan sammenlikne vår studie mot disse. Av hensyn til arbeidsmengde bestemte vi oss for å begrense studien til de fem hovedsektorene; financials, technology, consumer goods, conglomerates og basic materials. Videre definerte vi at kun selskaper med en markedsverdi over \$100 millioner, samt handelsvolum på minst 100 000 aksjer per dag skulle tas i betraktning. Dersom det var selskaper som manglet prisdata på enkelte dager ekskluderte vi disse. Med å gjøre dette valget ekskluderer vi mindre selskapet med liten

² <http://www.r-project.org/>

³ <http://cran.r-project.org/web/packages/urca/>

⁴ <http://cran.r-project.org/web/packages/PairTrading/index.html>

⁵ http://r-forge.r-project.org/R/?group_id=316

⁶ <http://cran.r-project.org/web/packages/PerformanceAnalytics/index.html>

omsetning som igjen er essensielt for å komme seg inn og ut av posisjoner når man måtte ønske uten å risikere stor avstand mellom kjøper og selger i ordreboken. Til slutt ekskluderte vi også aksjer som manglet prisdata for hele perioden, ettersom det var vanskelig å implementere en regel i R som tilsa at vi skulle gå ut av posisjoner når aksjene ble tatt av børs.

Når disse valgene var gjort ble vi sittende igjen med i overkant av 1500 aksjer fordelt på 92 undersektorer. Vi hentet daglige data for disse aksjene fra Yahoo Finance.

Valg av periode for testing av par og handelsperiode har sin bakgrunn i studier av Gatev et al. (2006) og Do & Faff (2009). Begge disse studiene har benyttet seg av 12 måneders periode for å finne par, med påfølgende seks måneders handelsperiode for disse parene. Studien til Do & Faff (2009) har de ferskeste data og omfatter handelsperiode ut 2008. Vi ønsket derfor å se om trenden fra deres studie har fortsatt, slik at vår handelsperiode er de første seks månedene av 2011. Parene vi handler i denne perioden ble testet for kointegrasjon i perioden 4.1.2010 til 31.12.2010.

5.3 Utvelgelse av par

Vi har kun testet par innen hver undersektor. Dette fordi det skal være størst sjans for at det skal være like fundamentale forhold i bedriftene som gjør at de er nære økonomiske substitutter. Do & Faff (2011) kommer i sin artikkel med en kritikk av metodikken til Gatev et al (2006) for utvelgelse av par på tvers av sektorer, fordi de finner at risikoen for at par ikke skal konvergere til en likevekt har økt etter 2002. Ved bare å velge par innenfor en gitt undersektor forsøker vi dermed å redusere denne risikoen. En viktig faktor er at vi med denne fremgangsmåten for utvelgelse får en håndterlig mengde par i utvalget i forhold til størrelsen på oppgaven.

For å teste for kointegrerte par har vi har brukt Johansen metoden på alle mulige kombinasjoner av par innenfor en gitt sektor. Vi valgte her å bruke eigenvalue metoden. Testen har en null-hypotese ($r=0$) som angir ingen kointegrasjon, mens alternativhypotesen angir at det er et kointegrasjonsforhold.

Dersom vi fikk en testverdi som var høyere enn kritisk verdi på 1 %, 5 % eller 10 % signifikansnivå var dette par vi ville se nærmere på. Etter å ha testet for kointegrasjon med Johansen metoden gjorde vi derfor flere tester; ADF og PP test hvor fremgangsmåten blir beskrevet under.

I neste steg omgjorde vi aksjeprisen til den naturlige logaritmen av aksjeprisene. Etter vi hadde gjort aksjekursene om til logaritmiske priser gjorde vi en lineær regresjon på de to.

$$lm \ln(y) \sim \ln(x) \quad (9)$$

På bakgrunn av resultatet fra regresjonen fikk vi en hedge ratio (regresjonskoeffisient 1) og premium (regresjonskoeffisient 2). Ved hjelp av disse kunne vi lage en spread mellom aksjene.

$$spread = \ln(y) - (hedge\ ratio * \ln(x) + premium) \quad (10)$$

Vi testet så om spreaden mellom aksjene var kointegrerte. Først med en Phillips-Perron (PP) test, for så i tillegg å teste med en Augmented Dickey-Fuller (ADF) test.

Ettersom vi ikke ønsket å bruke tid på å forklare forskjeller mellom testene og derfor rettferdiggjøre hvorfor vi hypotetisk sett skulle velge et par som gikk gjennom en av de tre testene på et gitt signifikansnivå, men ikke var signifikant i en annen test besluttet vi å dele aksjene inn i to grupper.

Alle aksjeparene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe, som vi ved hjelp av god fantasi kalte 5 %.

De parene som ble signifikante på minimum 10 % nivå i alle testene har vi gruppert sammen i 10 % gruppen. De parene som ble signifikante i Johansen testen, men ikke signifikante på minimum 10 % nivå i både PP og ADF ble forkastet. Alle parene som ikke ble statistisk signifikante ved hjelp av Johansen metoden ble forkastet uten øvrige tester.

Vi testet totalt 1537 par for kointegrasjon og endte til slutt opp med 161 par i gruppen 5 %. I gruppen 10 % endte vi også opp med 161 par. Parene sorteres som f.eks. BVN.HMY, der den førstnevnte er aksje 1 og den sistnevnte aksje 2.

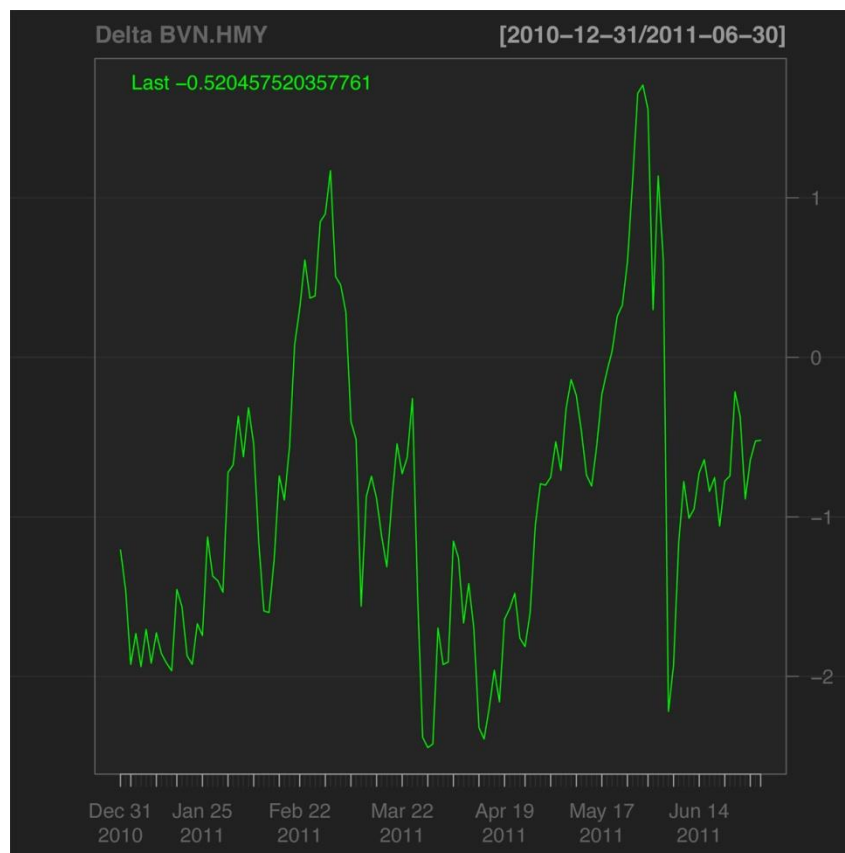
5.4 Handleregul

Vi har kalt vår handelsindikator delta. Den regnes ut på følgende måte:

$$\Delta_{norm} = \frac{\Delta - MA_{50}(\Delta)}{\sigma_{50}(\Delta)} \quad (11)$$

Der Δ er forskjellen/spreaden mellom de to aksjene (aksje 1 – aksje 2). $MA_{50}(\Delta)$ er glidende gjennomsnitt de siste 50 dager på forskjellen/spread mellom aksjene. $SD_{50}(\Delta)$ er standardavviket de siste 50 dager på forskjellen/spreaden og Δ_{norm} uttrykker feilprisingen i forskjellen/spreaden uttrykket ved standardavvik, som vi har valgt å kalle delta. I kalkuleringen av delta bruker vi prisen fra siste handel per dag, ikke justert for dividende og splits (Close).

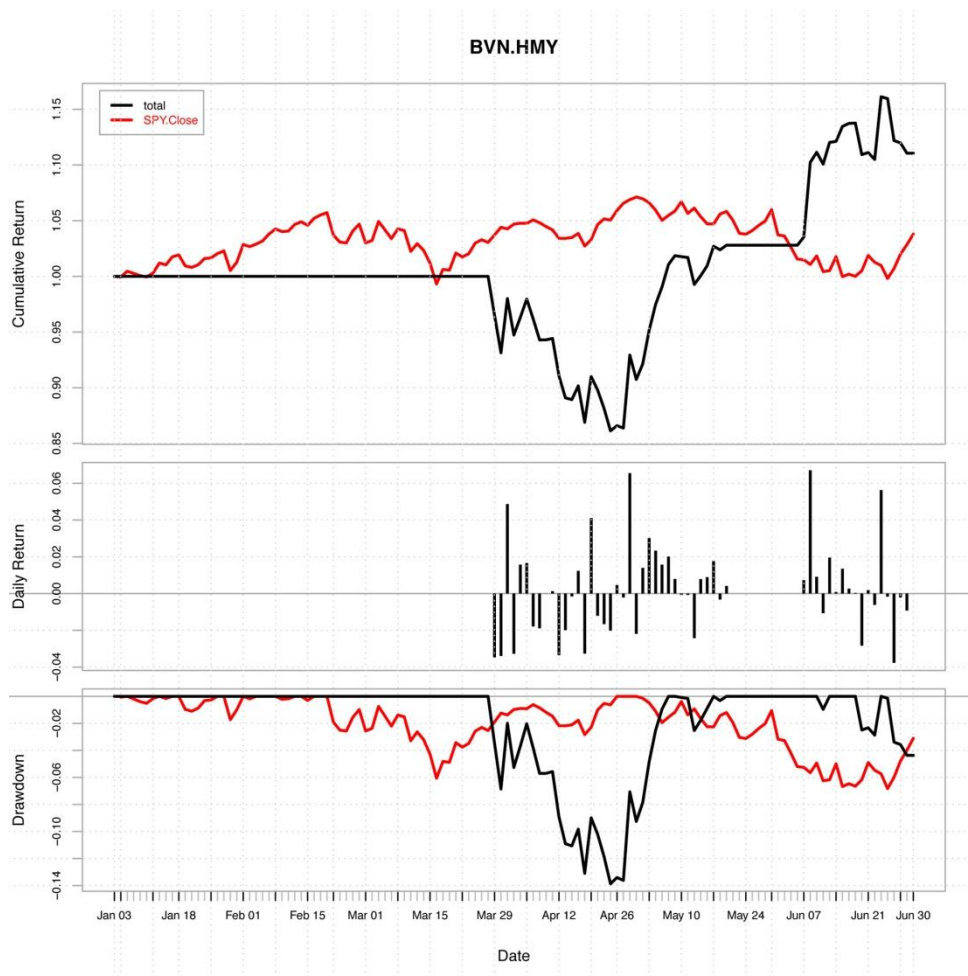
Handleregulene er som følger. Vi inngår handel på 2 eller - 2 standardavvik, i tråd med Reverre (2001). Det vil si salg av den overprisede aksjen, samt kjøp av den underprisede aksjen med like stor sum på begge sider. Dersom indikatoren viser 2 kjøper vi aksje 1 og selger aksje 2, og motsatt dersom indikatoren viser minus 2. Posisjonene lukkes ved konvergering til standardavvik på 0 eller på siste dag av handelsperioden. Etersom det er risiko knyttet til handel forutsetter vi videre at vi lukker posisjonene dersom paret divergerer ut over 3 eller -3 standardavvik. Delta kan krysse 2 og -2 flere ganger etter vi har åpnet en posisjon uten at vi har løst den inn. Vi har derfor lagt inn en regel om at vi bare åpner posisjonen første gang delta krysser indikatorverdiene. For eksempel ser vi i figur 8 at delta krysser - 2 på nytt 12. april etter vi allerede har åpnet en posisjon 28. mars, som enda er åpen. I dette tilfellet åpner vi ikke en ny posisjon selv om indikatoren skulle tilsi det.



Figur 8 Delta

For å illustrere metoden skal vi derfor gå igjennom et eksempel fra studien.

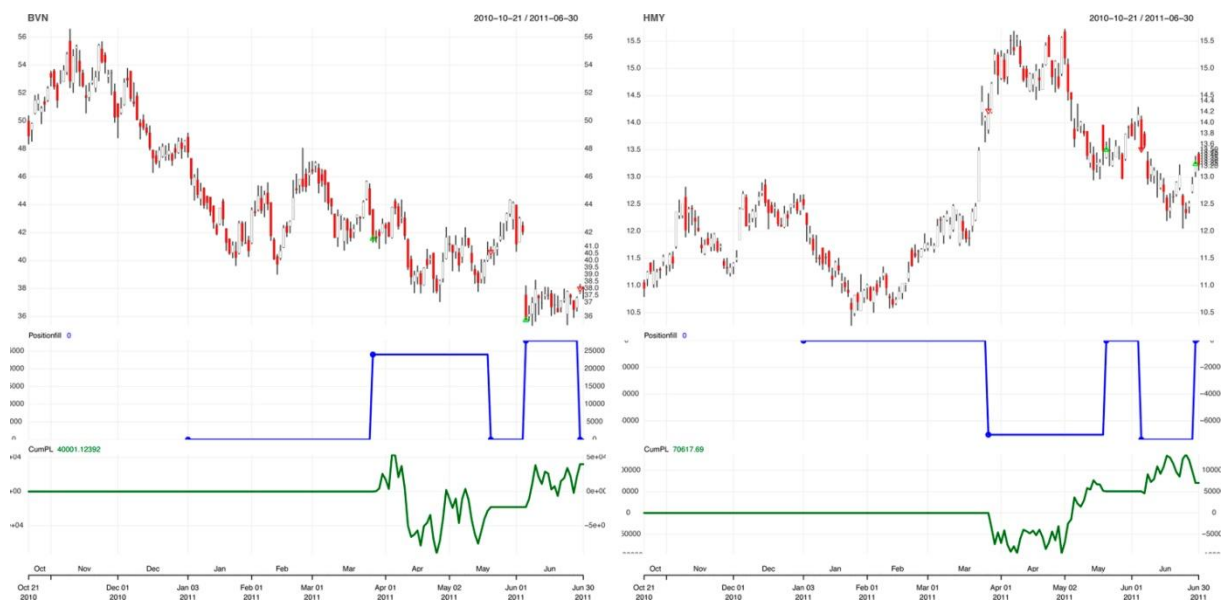
I dette eksempelet har vi tatt for oss aksjeparet BVN (aksje 1) og HMY (aksje 2). Figur 8 viser oss bevegelsen av feilprisingen i spread uttrykt i indikatoren delta. Som handlereglene over forklarer har vi kjøpt BVN og solgt HMY når delta lik -2 og lukket posisjonene ved delta lik 0. Som man da kan lese av diagrammet fikk vi tre signaler om å innta posisjon. Den første posisjonen ble innløst ved delta lik null, ved det andre signalet skjedde det ingen handel ettersom vi hadde en åpen posisjon, mens den siste posisjonen ble innløst på slutten av handelsperioden.



Figur 9 Sammendrag BVN.HMY

Dette resulterte i en periodeavkastning på 11,06 % som man kan lese av figur 9. I tillegg sammenliknes resultatet mot referanseindeksen S&P 500. Diagrammet Daily return viser den daglige avkastningen hver gang vi har hatt åpne posisjoner.

I figur 10 ser man kursutviklingen for BVN og HMY i den aktuelle perioden på 6 måneder. I dette tilfellet ser man at det nederste diagrammet (grønn linje) viser positiv avkastning på begge aksjene basert på handlene som ble gjort. Videre viser det midterste diagrammet (blå linje) position fill, og viser som tidligere at vi har kjøpt BVN, mens vi solgte HMY i begge åpningene av posisjoner. Alle kjøp og salg av aksjer skjer til prisen fra siste handel per dag, ikke justert for dividende og splits (Close).



Figur 10 Oversikt aksjekurser

5.5 Beregninger i analysen

Avkastning

I analysen har vi beregnet avkastning og sammenlignet med referanseindeksen som i vårt tilfelle er S&P 500 eller sektorindeksen til den respektive sektor. Den daglige avkastningen på long og short posisjonene ble kalkulert som verdivektede avkastninger på denne måten som uttrykker daglig avkastning:

$$r_{P,t} = \frac{\sum_{i \in P} w_i t r_{i,t}}{\sum_{i \in P} w_i t} \quad (12)$$

Vi antar at hver posisjon består av 1 dollar, altså 1 dollar på en long posisjon og 1 dollar på short posisjon.

$$r_{P,k} = \frac{\sum_{i \in P} \sum_{t \in k} c_{i,t}}{n} \quad (13)$$

Gatev et al. (2006) samt Do & Faff (2009) har to mål på avkastningen. Den første kaller de ”return on committed capital”. Ved utregning av avkastningen deler de den samlede daglige avkastningen på antall par i porteføljen. Den andre måten kaller de ”fully invested return”. Ved denne metoden beregner man avkastningen ved å ta den samlede daglige avkastningen delt på antall par som faktisk handler i perioden. Den førstnevnte metoden er en klart mer konservativ tilnærming til å beregne avkastningen. De mener at man ved denne metoden tar

hensyn til alternativkostnaden ved å binde kapital til en strategi, selv om strategien ikke gjør handler.

Dersom man har god tilgang på kapital og er fleksibel i finansieringen kan det derimot være mer korrekt å benytte seg av den faktiske sysselsatte kapitalen i beregningen av avkastningen.

Vi har valgt den mest konservative måten å beregne avkastningen på, nemlig ”return on committed capital”.

Periodeavkastningen beregnes ved å beregne den kumulative daglige avkastningen.

$$w_{i,t} = w_{i,t-1} (1 + r_{i,t-1}) = (1 + r_1) \dots (1 + r_{i,t-1}) \quad (14)$$

Den årlige avkastningen beregner vi som en simple return av periodeavkastningen. I følge Bodie et al. (2009) er det vanlig å bruke simple returns på investeringer med en tidshorisont kortere enn ett år.

$$\overline{R}_a * scale \quad (15)$$

(Peter Carl, 2012)

Videre beregnet vi høyeste og laveste avkastning per posisjon, fordeling av antallet gevinster versus tap og gjennomsnittlig holdingperiode per handel.

Risikofri rente

6 måneders risikofri rente (R_f) har vi satt til 0,2 %, hentet fra U.S Department of the Treasury (n.d). Programmet ville ikke la oss bruke flere desimaler, men forskjellene ville vært minimale i utregninger.

Sharpe Ratio (Årlig)

For å måle meravkastning utover risikofri rente per enhet totalrisiko så bruker vi Sharpe ratio. Forholdet mellom avkastning og risiko sier noe om hvor attraktiv en portefølje er. Dette avgjøres av forholdet mellom risikopremien og standardavviket til avkastningen utover risikofritt aktivum. Måling av Sharpe ratio er svært vanlig for å vurdere resultater til forvaltere (Bodie et al., 2009).

$$\text{Sharpe ratio} = \frac{\sqrt[n]{\text{prod}(1 + R_a)^{\text{scale}}} - 1}{\sqrt{\text{scale}}\sqrt{\sigma}} \quad (16)$$

$n = 125$ som er antall handelsdager i perioden og $\text{scale} = 252$ som er antall handelsdager i året. R_a er justert for risikofri rente.

(Peter Carl, 2012)

Information Ratio

Et alternativ til Sharpe ratio som prestasjonsmål er Information Ratio (IR). Dette prestasjonsmålet måler meravkastningen (α) en forvalter oppnår utover referanseindeksen i forhold til en tracking error. Denne defineres som forskjellen i standardavvik mellom porteføljen og referanseindeksen. Fordelen med IR er at den er lett å tyde. Dersom IR er større enn null, har forvalteren gjort en god jobb og motsatt dersom IR er under null. Markedsporteføljen vil bestandig ha en IR lik null.

$$IR = \frac{\alpha_p}{\sigma(e_h)} \quad (17)$$

(Bodie et al. (2009)

Value at Risk (VaR)

Value at Risk (VaR) måler en enhets eksponering mot markedsrisiko. VaR gir et statistisk mål på mulige tap i en portefølje basert på ”normal” markedsutvikling. Tap større en VaR skjer kun ved en spesifisert liten sannsynlighet. Kvartilen q av en distribusjon, er verdiene som ligger under q % i distribusjonen. Den vanlige estimeringen som vi også har benyttet oss av er 5 % VaR, som betyr at 95 % av avkastningene vil overgå VaR, og 5 % vil være mindre. Man kan derfor si at VaR er det minste tapet av de 5 % dårligste avkastningene.

$$\text{VaR}(.05 \sim N) \quad (18)$$

(Bodie et al., 2009)

Expected shortfall

Et mer realistisk syn på nedsiderisiko fokuserer på det forventede tapet gitt at man befinner seg i det verst tenkelige scenario. Denne verdien kalles expected shortfall (ES) som vektet de verste avkastningene mest, som igjen gir et verre utslag for potensielle tap enn VaR.

$$ES = \frac{1}{.05} \text{esp}(\mu)N[-\sigma - F(.95)] \quad (19)$$

(Bodie et al., 2009, p. 166)

Max drawdown

Max drawdown er ifølge Chan (2009) definert som den prosentvise forskjellen mellom det høyeste punktet i avkastningskurven og det laveste punktet som inntreffer på et senere tidspunkt.

Beta

Beta er et mål på volatiliteten i porteføljen sammenlignet med volatiliteten til referanseindeksen. Med vår markedsnøytrale strategi kan vi forvente en beta rundt 0, da den er uavhengig av endringer i referanseindeksen.

$$\beta_{a, b} = \frac{CoV_{a, b}}{\sigma_b} = \frac{\sum((R_a - \mu_a)(R_b - \mu_b))}{\sum(R_b - \mu_b)^2} \quad (20)$$

(Peter Carl, 2012)

Standardavvik (Årlig)

$$\text{Årlig } \sigma = \sqrt{\sigma^2} = \sqrt{\text{observert } \sigma^2} * \sqrt{n} \quad (21)$$

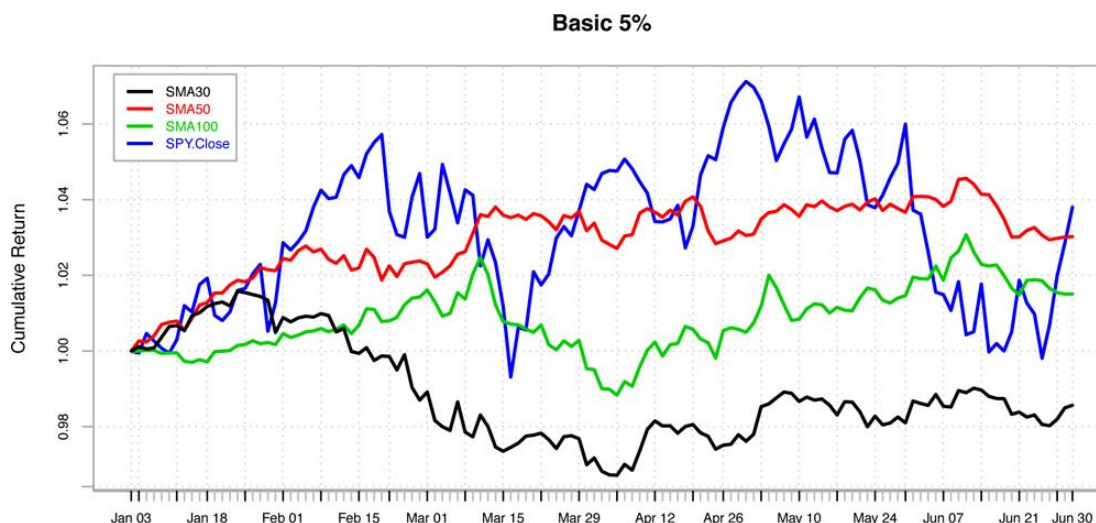
(Peter Carl, 2012)

Vi har brukt $n = 252$ dager, som tilsvarer antall handelsdager i et år.

5.6 Diskusjon, valg av kriterier

5.6.1 Indikator, valg av lengde på glidende gjennomsnitt

I studien har vi brukt et glidende gjennomsnitt på 50 dager. Grunnlaget for dette valget baserer seg på vår forståelse av hvordan feilprising vist ved standardavvik reagerer på antall dager med glidende gjennomsnitt. Mekanismen er slik at dersom det er et høyt antall dager som blir brukt til glidende gjennomsnitt, resulterer dette i færre handler med lengre holdingperiode. En lang holdingperiode kan være en risikofaktor ettersom Do & Faff (2011) har kommet fram til at den største risikoen ved handelsstrategien er konvergeringsrisikoen. I tillegg vil en trader se seg nødt til å binde kapitalen over et lengre tidsrom. På den andre siden vil for kort holdingperiode føre til at strategien vil bli veldig handelsintensiv, som naturlig nok medfører høyere transaksjonskostnader.



Figur 11 Chart, diverse SMA.

	Basic 5% SMA 30	Basic 5% SMA 50	Basic 5% SMA 100
Avkastning	- 2.89 %	6.1 %	3.0 %
Standardavvik	4.58 %	3.75 %	4.5 %
Antall handler	776	584	400
Holdingperiode	11.3 dager	18.7 dager	24.6 dager

Tabell 1 Resultater, forskjellige SMA.

Vi ser i tabell 1 at ved å bruke 30 dagers glidende gjennomsnitt (SMA) til beregning av indikatoren delta får vi nesten dobbelt så mange handler som med 100 dagers SMA. I tillegg

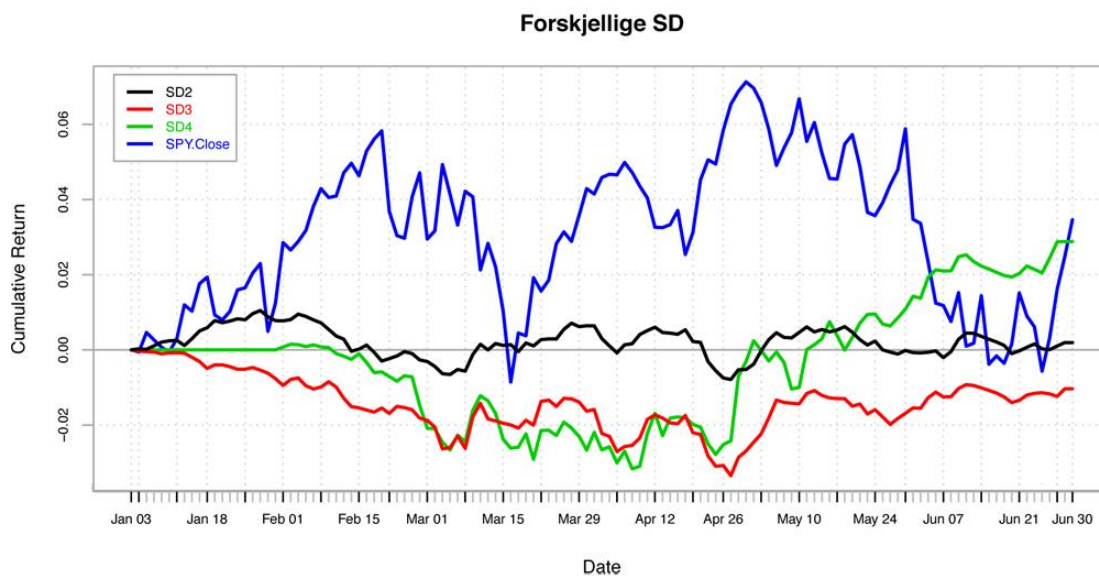
blir holdingperioden mer enn halvert ved bruk av 30 dagers SMA i forhold til 100 dagers SMA. Det finnes ingen fasitsvar på hva som er korrekt å bruke til beregning av indikator. Den enkelte trader må selv vurdere hva de vektlegger mest. Det kan være at for noen er det viktig å ha kort holdingperiode og dermed kortere perioder med kapitalbinding, de vil da bruke et relativt kort SMA i indikatoren. For andre er det kanskje viktigere å ha en mindre handelsintensiv strategi og da vil det være mer hensiktsmessig med en lengre tidshorison i indikatoren.

Som en supplerende kommentar har vi lagt ved en graf for kumulativ avkastning for gruppen 5 % i sektoren Basic Material målt mot referanseindeksen S&P 500 i figur 11. Som man kan se fra grafen og også lese ut av tabellen har valget av tidshorison i beregning av glidende gjennomsnitt, innvirkning på hvilken avkastning strategien oppnår. Man kan allikevel ikke trekke noen konklusjoner om en tidshorison er bedre enn den andre, ettersom avkastningen avhenger av hvilken tidsperiode man ser på og ikke minst hvilke par man benytter seg av.

5.6.2 Valg av handleregel

Vi har sett at man kan velge forskjellig lengde på glidende gjennomsnitt i indikatoren, og at forskjellige lengder gir forskjellige resultater. På samme vis har naturlig nok valg av handelssignal en stor innvirkning på prestasjonene til strategien. Gatev et al. (2006) viser i sin studie at de valgte to standardavvik, selv om handleregelen ikke bestandig ville gi stor nok profitt til å dekke transaksjonskostnadene. Dessuten poengterer de at det var fristende å prøve potensielt mer profitable handleregler, men at de på grunn av fare for data-snooping bias ikke ønsket å optimalisere strategien.

Vi ønsker i dette kapittelet å se hva som skjer dersom vi øker handleregelen til handel på henholdsvis 3 og 4 standardavvik, hvor vi også løser inn posisjonen dersom den fortsetter å divergere til henholdsvis 4 og 5 standardavvik. Begrunnelsen for dette er på samme måte som i forrige underkapittel; å belyse at valget av handleregel er viktig dersom man ønsker å forsøke seg på denne typen kvantitative investeringsstrategier. Vi bruker i eksempelet porteføljen for alle parene i gruppen 5 %.



Figur 12 Forskjellige SD

	SD2	SD3	SD4
Antall handler	1772	540	76
Gjennomsnitt holdingperiode	19,60	0,07	2,32
Periodeavkastning	0,0021	-0,0101	0,0292
Standardavvik	0,0245	0,0351	0,0557

Tabell 2 Forskjellige SD

Ved å øke nivået i deviasjoner fra normalen der hvor man åpner en posisjon ser vi at antall handler blir dramatisk redusert. Som en naturlig konsekvens av færre posisjoner blir også standardavviket høyere. Videre ser vi at den gjennomsnittlige holdingperioden blir langt kortere. Årsaken til at holdingperioden på SD3 er lavere enn SD4 kan være at posisjonene i dette eksempelet fortsetter å divergere og dermed blir raskt løst inn på stop-loss nivået. Vi kan som tidligere ikke si noe om hvilket alternativ som er det beste, men ser at vi oppnår forskjellig avkastning ettersom hvilken handleregel som legges til grunn. Periodeavkastningen er i dette eksempelet for øvrig beregnet ut fra antall par som faktisk handlet i perioden.

5.6.3 Kointegrasjon

Ettersom Gatev et al. (2006) hadde en litt annerledes metode for etablering av par enn den metoden vi brukte, manglet vi retningslinjer på hvilket signifikansnivå som er vanlig å benytte seg av når man sjekker aksjepar for kointegrasjon. Intuitivt tror man tro at jo høyere signifikansnivå vi hadde benyttet, desto større sikkerhet ville det ha vært for at aksjeparene

var kointegrerte. Dermed hadde det vært størst mulig sjanse for at de har mean-reverting egenskaper ved seg, noe som vi er avhengige av for at strategien skal fungere. Siden vi som sagt ikke hadde retningslinjer på dette, valgte vi å sortere parene i grupper hvor par som var signifikante på 1 % og 5 % nivå var i en gruppe, mens den andre gruppen besto utelukkende av par signifikante på 10 % nivå.

Gatev et. al (2006) bruker en seks måneders periode til handel etter parene er testet. Disse periodene er i følge artikkelforfatterne valgt tilfeldig. Man kan da stille spørsmål ved om man kanskje skulle benyttet seg av enten kortere, eller lengre periode for handel. Et argument for å bruke en lengre tidshorisont til handel er at kointegrasjon impliserer en langsiktig likevekt. På den andre siden vil man med kortere handelsperioder sikre seg hyppigere tester av om parene fremdeles er kointegrerte.

5.6.4 Risikomål

I tillegg til de risikomålene vi har implementert i vår studie, bruker Gatev et al. (2006) også en utvidet Fama & French tre-faktor modell for å undersøke porteføljens systematiske risikoeksponering. I tillegg til modellens faktorer; markedspremie, SMB (small minus big) og HML (high BM ratio) utvider de denne med en short-term reversal faktor og en momentum faktor. Argumentet for å inkludere de to ekstra faktorene er at pairs trading handler på den relative styrken til aksjene i et par.

Hvis man selger short-term vinnere og kjøper short-term tapere forventer de dermed at avkastningen vil være positiv for reversal faktoren. Samtidig vil det være slik at dersom man selger medium-term vinnere og kjøper medium-term tapere vil avkastningen være negativt korrelert med momentum faktoren.

I sin artikkel finner Gatev et al. (2006) at bare en ubetydelig del av avkastningen kan tillegges de fem faktorene i modellen. På bakgrunn av resultatene fra artikkelen, samt omfanget på oppgaven vår valgte vi ikke å inkludere denne modellen i vår studie, men nøyer oss med å henvise til artikkelen.

5.6.6 Ordretype

Den vanligste måten å handle aksjer på børsen er å legge inn en markedsordre. Dette betyr at man legger en kjøps- eller salgsordre til ønsket pris. Når både kjøpere og selgere vil kjøpe og selge til samme pris gjennomføres handelen. Dersom aksjeprisutviklingen går i feil retning,

risikerer man i verste fall å tape hele investeringen om man holder en kjøpsposisjon eller en salgsposisjon.

For å redusere mulige tap kan man benytte en stop-loss ordre. Man legger da inn ordre på et kjøp, samt at man skal selge til markedspris dersom aksjeprisen faller med en gitt proSENTSATS. Stop-loss er en måte å sikre seg mot store tap i en portefølje.

I studien har vi utelukkende benyttet oss av markedsordre. Dessuten har vi benyttet oss av en slags stop-loss ved at posisjonen løses inn dersom parene divergerer ytterligere og passerer tre standardavvik fra normalen. En måte å forbedre dette på hadde vært ved å legge stop-loss ordre på f.eks. 10 % på hver av aksjene i et par. Samtidig måtte det ha vært automatikk i at dersom en av sidene i en posisjon ble lukket, skulle den andre også bli lukket, ettersom man da ville eksponeres for markedsrisiko. En enda bedre metode ville vært å kunne stoppe en posisjon ved et gitt prosentvis tap av investeringen i et par.

5.7 Mulige feilkilder

5.7.1 Survivorship bias

Det finnes flere databaser for å tilegne seg historiske prisdata til aksjer. Eksempler er Thompson Reuters, CRSP, Yahoo Finance og Google. De to førstnevnte er kostbare alternativer og i følge Chan (2009) fri for survivorship bias ettersom man får tilgang til prisdata til aksjen selv etter den er blitt tatt av børs. De to sistnevnte er gratis å bruke, men ulempen er at databasene har survivorship bias ettersom historiske prisdata for aksjer tatt av børs fjernes.

Survivorship bias kan gi store feilkilder. For eksempel dersom man tester en strategi på å kjøpe billige aksjer, og velger å kjøpe de 20 billigste aksjene på en gitt indeks. Dersom databasen bare inkluderer de aksjene som har overlevd i perioden, vil man i testen få en avkastning som ikke gjenspeiler det man faktisk ville oppnådd dersom man handlet aksjene. Dette skyldes at man ikke tar hensyn til tapene på de selskapene som enten gikk konkurs eller ble tatt av børs, og som skulle vært i det aktuelle utvalget.

Chan (2009) viser til at en metode for å minske påvirkningen av survivorship bias er å bruke ferske data, slik at man ikke mangler for mange aksjer i utvalget. Vi mener at vi har gjort dette i vår studie. I tillegg har vi et bredt utvalg slik at påvirkningen fra enkelte par ikke bør gi for store utslag på funnene våre.

5.7.2 Data snooping bias

Problemer med data snooping kan forekomme dersom man overoptimaliserer ett sett data slik at man klarer å produsere de resultatene man ønsker i datasettet. For eksempel kan man bruke altfor mange parametre i modellen, slik at den tilpasses historiske tilfeldigheter som ikke vil gjenta seg i framtiden.

Dette må, i følge Chan (2009), man være særlig oppmerksom på når man analyserer finansielle data på grunn av få relevante data. Om man tester en handlestrategi vil man i noen tilfeller ha data 70 – 80 år tilbake, mens kanskje bare de fem til ti siste år er relevante for modellen man bruker. Forklaringen kan være at regulatoriske endringer eller makroøkonomiske begivenheter gjør at man ikke kan sammenlikne data mellom periodene.

Han gir flere anbefalinger, bl.a. at man bruker modeller med få parametre og kun lineær regresjon til modellene. I tillegg mener han at optimaliseringen av strategien skjer i et eget datasett, og at resultatene testes på et nytt datasett av nyere dato.

Vi mener vi har minimert risikoen for data snooping bias ettersom vi har brukt en veldig enkel handlestrategi, samtidig som vi overhodet ikke har gjort noen optimalisering av strategien. Videre bygger vi for en stor del vår metodikk på tidligere empiriske studier.

5.7.3 Look ahead bias

Look ahead bias inntreffer dersom man lager en modell hvor et handelssignal er at man skal kjøpe en aksje på den laveste prisen en dag. Naturlig nok vil man bare i ettertid kunne fastslå hva som var den laveste prisen den dagen.

I følge Chan (2009) er løsningen for å unngå look ahead bias å introdusere en forsinkelse slik at handelen først inntreffer etter at et signal har blitt aktivert. Likeledes mener han at man ikke trenger å introdusere forsinkelser dersom man bare inngår posisjoner ved periodens slutt. Vi anser handelsdagen som en periode, og har derfor ikke lagt inn forsinkelse i signalet. I tillegg er programmet vi bruker bygget for å modellere hvordan handel foregår i praksis. De skriver i dokumentasjonen at det er bygget inn en forsinkelse fra handelssignalet utløses til handelen utføres, uten at det spesifiseres nærmere hvor lang denne er.

5.8 Oppsummering

I metodekapitlet har vi utredet alle aspektene med hvordan vi kommer frem til resultatene i studien, samt begrunnet valgene vi har gjort.

Først startet vi med å beskrive analyseverktøyet R som vi har brukt i alle beregninger i studien. Videre forklarte vi hvordan vi har valgt ut aksjer, sektorer og tidsrommet for studien. Ved utvelgelse av par har vi tatt utgangspunkt i undersektorer og testet aksjene parvis for grad av kointegrasjon. Vi endte deretter opp med 2 grupper på følgende måte: Alle aksjeparene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe som vi ga navnet 5 % gruppen. De parene som ble signifikante på minimum 10 % nivå i alle testene ble gruppert i den andre gruppen som fikk navnet 10 % gruppen. Øvrige par ble forkastet. Videre beskriver vi handelsstrategien som baserer seg på misprising av standardavvik basert på spread. Deretter viste vi et praktisk eksempel fra et av parene i studien.

I underkapittel 5.5 presenterer vi de ulike beregningsmetodene vi har brukt til å analysere resultatene fra strategien. Disse spenner seg fra beregning av avkastning og prestasjonsmål til win/loss ratio og gjennomsnittlig holdingperiode.

Så diskuterer vi valg av kriterier i handelsstrategien. Til slutt diskuteres ulike feilkilder som kan oppstå ved å bruke historiske data som grunnlag i kvantitativ forskning.

6. Analyse

Vårt utgangspunkt er å teste hvorvidt tendensene fra tidligere studier på investeringsstrategien pairs trading vedvarer. Studier av Gatev et al. (2006) og Do & Faff (2009) viser at strategien hadde en eventyrlig avkastning på 80-tallet, mens den i fortsettelsen ga dalende avkastning, før den til slutt ikke gav noen avkastning som var signifikant annerledes enn null i perioden 2003 – 2008.

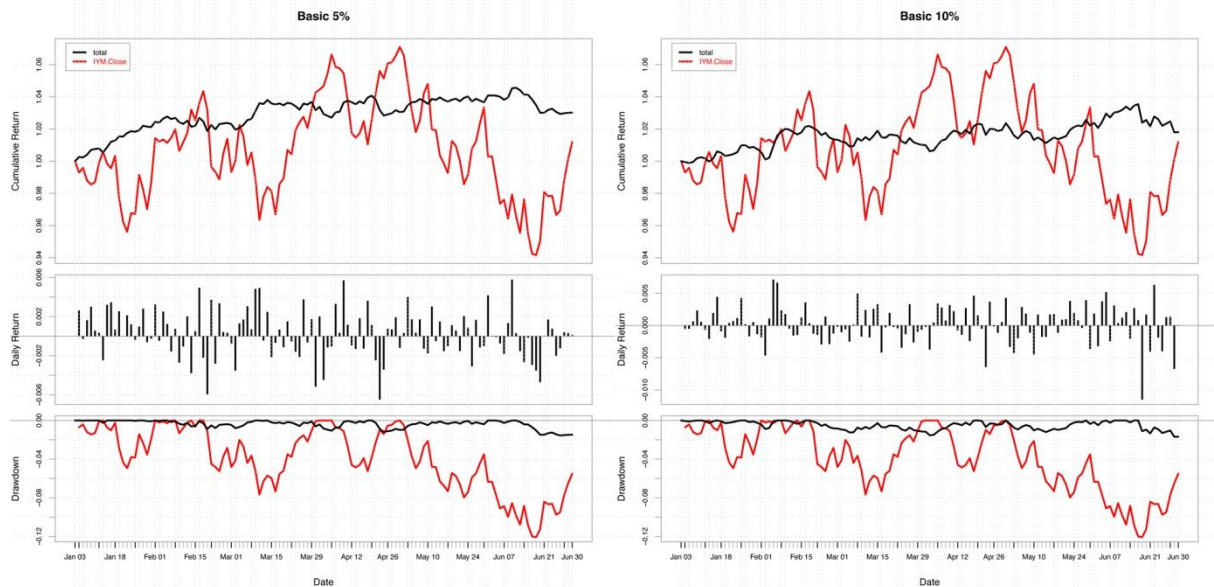
Pairs trading som investeringsstrategi er ikke ulik andre spekulative investeringsstrategier. Eksempler på slike er momentum- og reversal strategier, eller at markedet ga unormalt høy avkastning i januar måned. For mange av disse strategiene har tidlige studier funnet at de ga høy avkastning, for så å bli motbevist av senere studier.

Vi vil i vår studie forsøke å sammenlikne våre resultater opp mot studiene til Gatev et al. (2006) og Do & Faff (2009). Gatev et al. (2006) har også som en del av sin studie gitt kriterier om at parene må være selskaper i samme sektor, basert på Standard & Poors industrikategorier: Utilities, Transportation, Financial og Industrial. Dette er en veldig grov inndeling av sektorer, og dermed ikke direkte sammenlignbar med vår studie. Vi kan allikevel sammenligne porteføljen med alle aksjene mot de tidligere nevnte studier. Vi har ikke grunnlag til å sammenlikne de andre sektorene fordi vi ikke har med par fra transport- og utilitiessektoren i vår studie, mens industrisektoren er alt for bred i forhold til de sektorer vi har i vår studie.

Vi har ikke noen grunn til å tro at trenden i avkastning vist gjennom studien til Do & Faff (2009) skal ha forandret seg nevneverdig de senere år. Våre forventninger er derfor å finne sammenlignbare resultater med denne studien.

6.1 Sektorer

6.1.1 Basic Materials



Figur 13 Graf, Basic Materials

	Basic Materials 5 %	Basic Materials 10 %	
Antall par	51	61	
Antall posisjoner	146	175	
Antall handler	584	700	
Gj.snitt holdingperiode	18,7	23	
Største gevinst	42,86 %	56,04 %	
Største tap	-40,44 %	-47,59 %	
Gevinst %	51,37 %	56 %	
Tap %	48,63 %	44 %	
W/L ratio	1,056	1,273	
	Basic Materials 5 %	Basic Materials 10 %	IYM
Periodeavkastning	0,0302	0,0180	0,0119
Årlig avkastning	0,0610	0,0364	0,0243
Sharpe-ratio (Årlig)	1,5759	0,7682	0,1100
Information Ratio	0,2732	0,1513	-
VaR	-0,0035	-0,0042	-0,0206
Conditional VaR (ES)	-0,0049	-0,0060	-0,0269
Standardavvik	0,0376	0,0450	0,2045
Max Drawdown	0,0155	0,0169	0,1207
Beta	-0,0432	-0,0375	1,0000

Tabell 3 Basic Materials

Sektoren Basic materials inneholder selskaper som driver med leting, utvinning og foredling av råvarer. Eksempler er olje, metall og kjemiske produkter. Vi startet her med 356 aksjer, fordelt på 17 undersektorer. Denne sektoren ble den nest største i gruppen 5 % med 51 par, mens den ble den største sektoren i gruppen 10 % med 61 par. Som referanseindeks til sektoren brukte vi indeksen Dow Jones US Basic Materials gjennom iShares ETF ticker: IYM som repliserer den nevnte indeks.

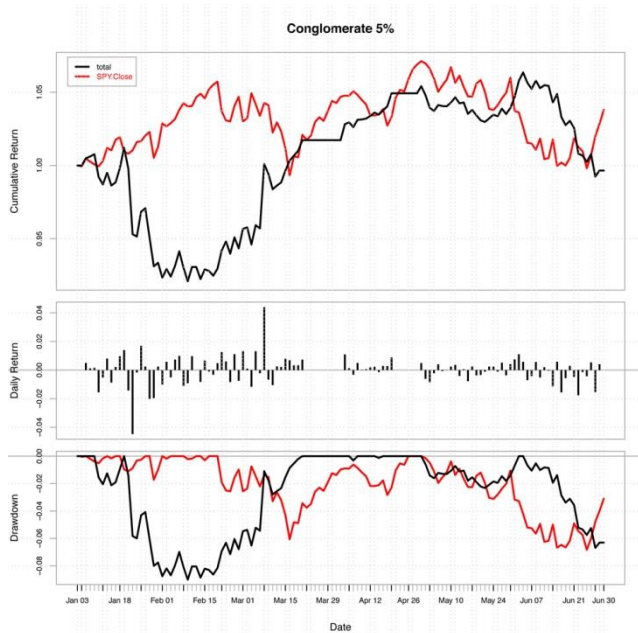
I gruppen 5 % gikk vi inn i totalt 146 posisjoner, mens handelen i 10 % gruppen resulterte i 175 posisjoner. Den førstnevnte gruppen hadde en gjennomsnittlig holdingperiode på 18,7 handelsdager, mens sistnevnte var noe høyere med en holdingperiode på 23 handelsdager.

For porteføljen i gruppen 5 % oppnådde vi i denne sektoren en periodeavkastning på 3,02 %, noe som tilsvarer en årlig avkastning på 6,1 %. I porteføljen for gruppen 10 % ble resultatet litt magrere med en periodeavkastning på 1,8 % som tilsvarer en årlig avkastning på 3,64 %. Avkastningen for begge gruppene er derimot bra når man sammenlikner med resultatene til indeksen som bare oppnådde en periodeavkastning på 1,19 %, tilsvarende årlig avkastning på 2,43 %.

Strategien oppnådde som forventet en beta nær null. Dersom man legger kapitalverdimodellen til grunn har vi dermed en portefølje som er uavhengig av volatiliteten i markedet. Når det gjelder porteføljenes risiko uttrykt gjennom standardavvik er de langt lavere enn indeksens standardavvik på 0,2045. Porteføljene i gruppen 5 % og 10 % fikk et standardavvik på henholdsvis 0,0376 og 0,045. Value at Risk (VaR) uttrykker hva det minste forventede tapet vil være over et gitt tidsintervall. Ettersom vi valgte 5 % som signifikansnivå og bruker daglige data, uttrykker VaR at vi i gjennomsnitt kan forvente et tap på minimum 2,06 % en gang hver 20 dag dersom vi holder markedsporteføljen. Ettersom porteføljene i vår studie holder både kjøps- og salgsposisjoner som skal eliminere markedsrisiko er det derfor ikke overraskende at VaR er lavere. I porteføljen til gruppen 5 % ble VaR 0,35 %, mens 10 % gruppens VaR ble 0,42 %.

Som en konsekvens av lav risiko og bra avkastning oppnår vi naturlig nok gode prestasjonsmål. Sektorindeksen som vi måler oss mot oppnådde en Sharpe ratio på 0,11. Vi ser derimot at for gruppen 5 % ble Sharpe ratio 1,58, og noe lavere for gruppen 10 % med 0,77. På samme måte viser også Information Ratio at begge porteføljene presterte bra med en IR på hhv. 0,27 og 0,15.

6.1.2 Conglomerates



Figur 14 Graf, Conglomerates

Conglomerates 5 %	
Antall par	2
Antall posisjoner	6
Antall handler	24
Gj.snitt holdingperiode	26,5
Største gevinst	6,40 %
Største tap	-5,39 %
Gevinst %	50 %
Tap %	50 %
W/L ratio	1

	Conglomerates 5 %	S&P 500
Periodeavkastning	-0,0035	0,0380
Årlig avkastning	-0,0070	0,0772
Sharpe-ratio (Årlig)	-0,0617	0,5926
Information Ratio	-0,4473	-
VaR	-0,0151	-0,0164
Conditional VaR (ES)	-0,0212	-0,0190
Standardavvik	0,1429	0,1273
Max Drawdown	0,0900	0,0683
Beta	-0,0798	1,0000

Tabell 4 Conglomerates

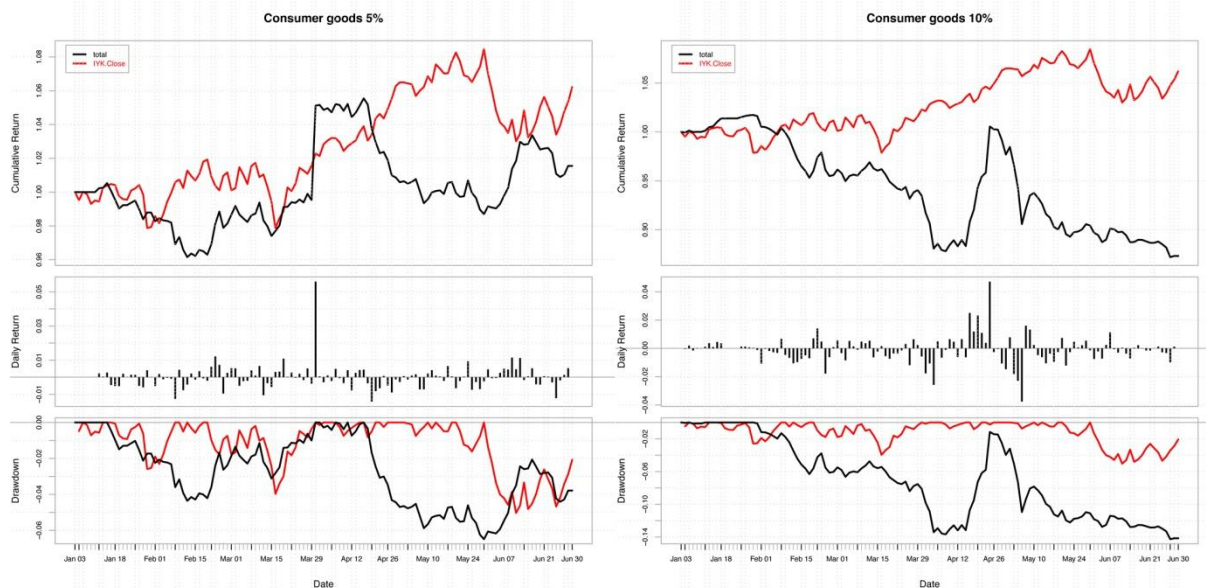
Sektoren Conglomerates består av selskaper som driver med flere ulike forretningsområder. På grunn av manglende referanseindeks til denne sektoren har vi sammenlignet resultatene i tabellen med S&P 500.

Vi testet 10 aksjer som resulterte i 2 par som falt inn under 5 % gruppen. Det ble kun tatt 6 posisjoner i perioden med like mange tap og gevinster, med en gjennomsnittlig holdingperiode på 26,5 dager. Den største gevinsten var 6,4 %, mens det største tapet ble 5,39 %. Periodeavkastningen ble -0,35 % som tilsvarer en årlig avkastning på -0,7 %. Dette er langt lavere enn avkastningen til referanseindeksen som oppnådde en årlig avkastning på 7,72 % . Prestasjonsmålene ble følgelig dårlige med en Sharpe ratio på -0,0617.

Få handler gjenspeiles ved stor volatilitet som man kan se av avkastningsgrafene, samt maksimal drawdown på 9 % som er forskjellen på høyeste og etterfølgende laveste punkt i avkastningskurven. Naturlig nok får vi derfor også i denne porteføljen det høyeste standardavviket i studien. Dette bekreftes også med en høy VaR på 1,51 %, bare marginalt lavere enn indeksens VaR.

Et viktig resultat fra denne sektoren er at beta for porteføljen ligger rundt null som indikerer at avkastningen er uavhengig i forhold til markedet selv ved et lite antall par i porteføljen.

6.1.3 Consumer goods



Figur 15 Graf, Consumer goods

	Consumer goods 5 %	Consumer goods 10 %	
Antall par	6	7	
Antall posisjoner	15	18	
Antall handler	60	72	
Gj.snitt holdingperiode	28,6	26,2	
Største gevinst	29,37 %	9,69 %	
Største tap	-26,09 %	-32,44 %	
Gevinst %	53,33 %	27,78 %	
Tap %	46,67 %	72,22 %	
W/L ratio	1,143	0,385	
	Consumer goods 5 %	Consumer goods 10 %	IYK
Periodeavkastning	0,0155	-0,1267	0,0622
Årlig avkastning	0,0313	-0,2553	0,1264
Sharpe-ratio (Årlig)	0,2648	-1,7319	1,1831
Information Ratio	-0,6565	-1,9566	-
VaR	-0,0081	-0,0142	-0,0109
Conditional VaR (ES)	-0,0110	-0,0221	-0,0145
Standardavvik	0,1114	0,1485	0,1053
Max Drawdown	0,0649	0,1427	0,0503
Beta	-0,0394	-0,0583	1

Tabell 5 Consumer goods

Sektoren Consumer goods inneholder selskaper innen bransjene klær og sko, tobakk, leker samt motoriserte kjøretøy. Som referanseindeks til sektoren har vi benyttet oss av iShares ETF ticker: IYK som repliserer indeksen Dow Jones US Consumer Goods.

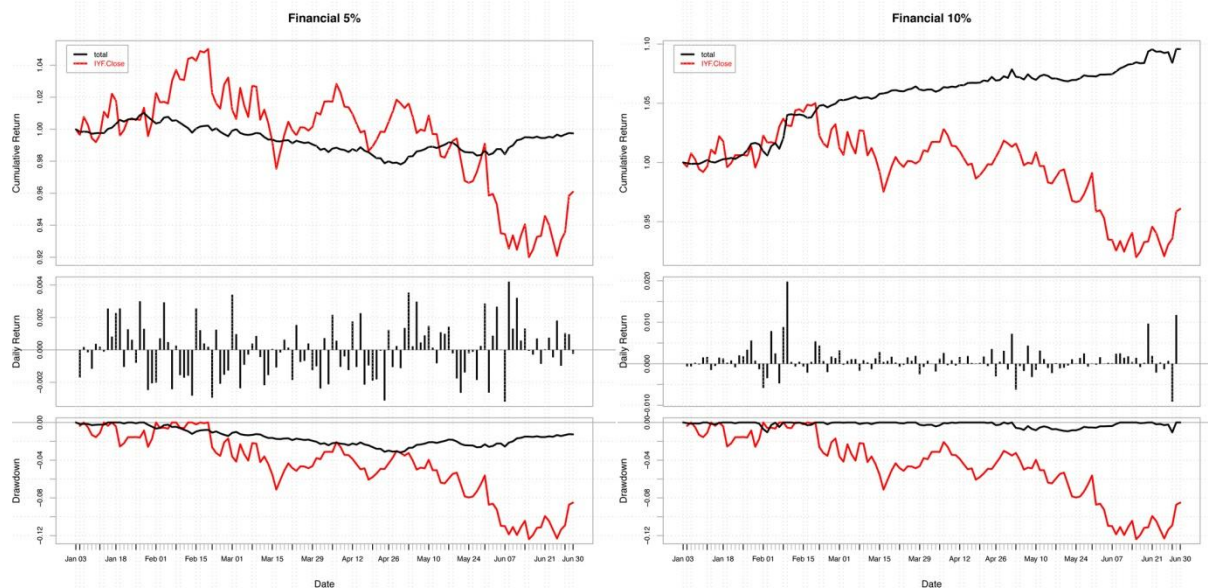
Samlet testet vi 25 aksjer innenfor 4 undersektorer. I 5 % gruppen fant vi 6 par som resulterte i 15 posisjoner i løpet av perioden. For 10 % gruppen fant vi 7 par som resulterte i 18 posisjoner. Gjennomsnittlig holdingperiode per par var henholdsvis 28,6 og 26,2 handelsdager for de to porteføljene. En stor forskjell mellom gruppene er antall gevinster der 5 % gruppen oppnår positiv avkastning i 53,33 % av posisjonene, mens 10 % gruppen kun oppnår positiv avkastning i 27,78 % av posisjonene.

Ser vi på grafene viser den at 5 % gruppen tenderer svakt i positiv retning mens 10 % gruppen tenderer i negativ retning. Periodeavkastningen ble 1,55 % for 5 % gruppen og – 12,67 % for 10 % gruppen hvor sistnevnte representerer den svakeste prestasjonen i studien. Begge gruppene presterer langt dårligere enn referanseindeksen som oppnår en periodeavkastning på 6,22 %. Naturlig nok blir dermed prestasjonsmålene til porteføljene lave. Sett i forhold til referanseindeksen viser det seg at indeksen både ga bedre avkastning samt lavere risiko uttrykt i standardavvik enn de to gruppene.

Standardavviket i 5 % gruppen er 0,1114, mens for 10 % gruppen ble det noe høyere med 0,1485 slik at førstnevnte presterte klart best av de to. Videre er VaR 0,0081 for gruppen 5 %, mens den er 0,0142 for gruppen 10 %. Max drawdown bekrefter også at det i denne sektoreren var stor volatilitet med henholdsvis 6.49 % og 14.27 % hvor sistnevnte er det høyeste i studien.

Selv om begge gruppene viser ulike prestasjoner er det viktig å påpeke at begge gruppene indikerer relativt større risiko sammenlignet med sektorer hvor vi har flere par i porteføljen. Det kan derfor tyde på at det er en diversifiseringsgevinst ved å holde flere par i en portefølje. Dette er i tråd med studien til Gatev et al. (2006) som også finner at standardavviket til en portefølje faller når man øker antall par i porteføljen.

6.1.4 Financials



Figur 16 Graf, Financials

	Financials 5 %	Financials 10 %	
Antall par	53	49	
Antall posisjoner	139	137	
Antall handler	556	548	
Gj.snitt holdingperiode	30,6	32,2	
Største gevinst	28,11 %	184,62 %	
Største tap	-21,60 %	-22,96 %	
Gevinst %	56,12 %	65,69 %	
Tap %	43,88 %	34,31 %	
W/L ratio	1,279	1,915	
	Financials 5 %	Financials 10 %	IYF
Periodeavkastning	-0,0026	0,0958	-0,0392
Årlig avkastning	-0,0053	0,1931	-0,0796
Sharpe-ratio (Årlig)	-0,2763	3,7643	-0,5266
Information Ratio	0,5210	1,8228	-
VaR	-0,0025	-0,0029	-0,0180
Conditional VaR (ES)	-0,0028	-0,0051	-0,0231
Standardavvik	0,0256	0,0508	0,1546
Max Drawdown	0,0318	0,0105	0,1238
Beta	-0,0023	-0,0150	1

Tabell 6 Financials

Financials er en sektor som består av selskaper innenfor bransjene bank, finans og forsikring.

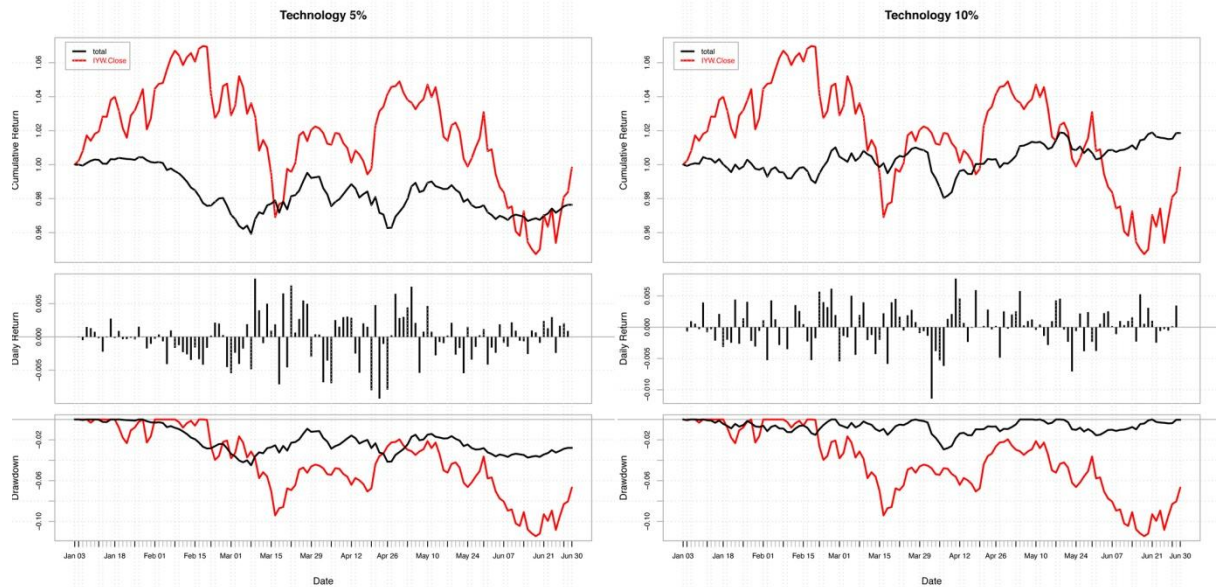
Samlet testet vi 384 aksjer fordelt på 30 undersektorer. I 5 % gruppen fant vi 53 par som resulterte i 139 posisjoner i løpet av perioden. For 10 % gruppen fant vi 49 par som resulterte i 137 posisjoner. Gjennomsnittlig holdingperiode per par var henholdsvis 30,6 og 32,2 dager. Som referanseindeks har vi brukt iShares ETF ticker: IYF som repliserer indeksen Dow Jones US Financial Sector.

Periodeavkastningen ble -0,26 % for 5 % gruppen og 9,58 % for 10 % gruppen hvor sistnevnte presterte best i studien. Standardavviket er meget lavt for begge gruppene på 0,0256 og 0,0508. Man ser av grafen at 5 % gruppen tenderer svakt negativt etter toppen i starten av perioden. 10 % gruppen viser derimot en sterk positiv trend igjennom hele perioden, dog med noen marginale drawdowns underveis. Denne kombinasjonen av en kontinuerlig positiv trend kombinert med svært lav volatilitet er et prakt eksemp lar på hvordan en god handelsstrategi skal fungere. Win/loss ratio er positiv for begge gruppene selv om 10 % gruppen skiller seg ut med gevinst i hele 65.69 % av posisjonene.

Ettersom referanseindeksen presterte dårligere enn begge gruppene ble også prestasjonsmålene til begge gruppene bedre enn indeksens. Det er mest interessant å se på 10 % gruppens prestasjonsmål ettersom denne presterte best i studien. Vi oppnådde her en Sharpe ratio på 3,76 og en Information Ratio på 1,82.

Et funn er at 10 % gruppen som har prestert klart best i studien, også er den gruppen med høyest gjennomsnittlig holdingperiode per posisjon. Forklaringen på dette kan være så enkel som at kursbevegelsene i denne gruppen passer handelsstrategien godt, som igjen medfører at færre posisjonene blir løst inn ved 3 standardavvik i forhold til de andre gruppene, samtidig som at vinnerposisjoner holdes lenger. I utgangspunktet er det naturlig å tro at gruppen med 5 % er mer kointegrerte som igjen betyr en bedre reversaleffekt gjenspeilet ved en lavere holdingperiode. Problemet som oppstår ved hurtig reversal er man mister noe av avkastningen, fordi posisjonen også lukkes raskere på bakgrunn av at indikatoren og handlere gelen består av standardavvik og glidende gjennomsnitt. Man kan derfor argumentere for at 5 % gruppen kutter gevinstposisjonene for tidlig, mens 10 % gruppen holder dem lenger. Dette er en refleksjon over hvorfor 10 % har lenger holding periode og bedre avkastning enn 5 % gruppen, men om dette er korrekt er vanskelig å bevise.

6.1.5 Technology



Figur 17 Graf, Technology

	Technology 5 %	Technology 10 %
Antall par	49	44
Antall posisjoner	137	124
Antall handler	548	496
Gj.snitt holdingperiode	24,2	26,5
Største gevinst	32,89 %	26,33 %
Største tap	-47,79 %	-47,01 %
Gevinst %	52,55 %	55,65 %
Tap %	47,45 %	44,35 %
W/L ratio	1,108	1,255

	Technology 5 %	Technology 10 %	IYW
Periodeavkastning	-0,0236	0,0186	-0,0017
Årlig avkastning	-0,0476	0,0376	-0,0034
Sharpe-ratio (Årlig)	-0,9469	0,7237	-0,0330
Information Ratio	-0,1917	0,3276	-
VaR	-0,0054	-0,0053	-0,0186
Conditional VaR (ES)	-0,0074	-0,0066	-0,0228
Standardavvik	0,0522	0,0495	0,1589
Max Drawdown	0,0449	0,0294	0,1145
Beta	-0,0086	0,0263	1,0000

Tabell 7 Technology

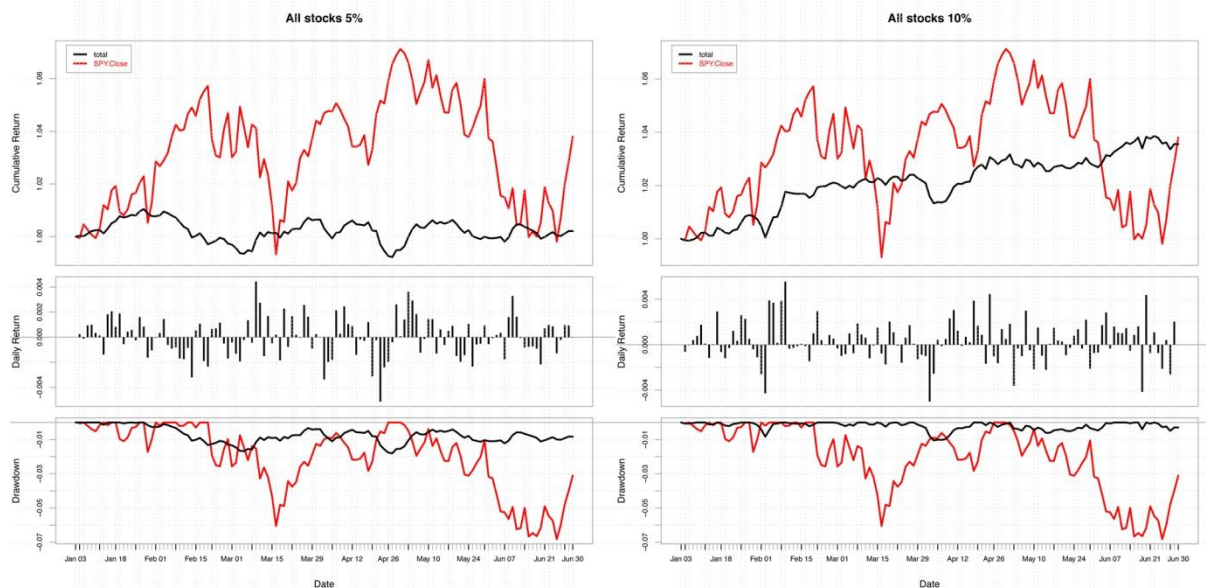
Technology består av teknologibedrifter som blant annet arbeider med programvare, IT og kommunikasjon. Som referanseindeks har vi brukt iShares ETF ticker: IYW som repliserer indeksen Dow Jones US Technology.

Totalt sett testet vi 464 aksjer som er fordelt på 30 undersektorer. I 5 % og 10 % gruppen fant vi henholdsvis 49 og 44 aksjepar som medførte totalt 137 og 124 posisjoner i perioden. 10 % gruppen presterte klart best med positiv avkastning på 1,86 %, mens 5 % gruppen måtte ta et tap på -2,36 %. Prestasjonsmålene Sharpe- og Information ratio gjenspeiler dette med svakt positiv for 10 % gruppen og svakt negativ for 5 % gruppen. Man kan se av tabellen at 5 % gruppen inneholdt det paret som ga høyest avkastning på 32,89 % mens 10 % gruppen viser seg mer konsistent med en høyere win/loss ratio. Verdt å merke seg inneholdt begge gruppene parene med høyest tap i hele studien med henholdsvis -47,79 % og -47,01 %.

Strategien hadde som forventet lavere risiko enn referanseindeksen. Vi ser at standardavviket til 5 % gruppen er 0,0522, mens det bare er marginalt lavere for 10 % gruppen med 0,0495. Dette er bare en drøy tredjedel av standardavviket til referanseindeksen. VaR var på henholdsvis 0,0054 og 0,0053 for 5 % og 10 % gruppen. Maksimal drawdown for 5 % gruppen ble på hele 4,49 %, mens 10 % gruppen er lavere med 2,94 %. Alt i alt ser det ut til at begge gruppene har relativt lav risiko sammenliknet med referanseindeksen. Et godt eksempel på dette er å se på grafen med indeksen og 10 % gruppen fra midten av perioden. Her ser vi at 10 % gruppene viser en stabil positiv trend med liten volatilitet, mens indeksen viser klart høyere volatilitet.

Videre ser man at begge gruppene viser en beta rundt null, som antyder at avkastningen i stor grad er uavhengig av markedsutviklingen.

6.1.6 Alle parene



Figur 18 Graf, alle parene

	Alle 5 %	Alle 10 %	Alle 10 % - Financials 10 %	
Antall par	161	161	112	
Antall posisjoner	443	454	317	
Antall handler	1772	1816	1268	
Gj.snitt holdingperiode	19,6	23,4	23,3	
Største gevinst	42,86 %	184,62 %	56,04 %	
Største tap	-47,79 %	-47,59 %	-47,59 %	
Gevinst %	53,27 %	57,71 %	54,26 %	
Tap %	46,73 %	42,29 %	45,74 %	
W/L ratio	1,140	1,365	1,186	
	Alle 5 %	Alle 10 %	Alle 10 % - Financials 10 %	S&P 500
Periodeavkastning	0,0021	0,0356	0,0092	0,0380
Årlig avkastning	0,0042	0,0717	0,0186	0,0772
Sharpe-ratio (Årlig)	0,0966	2,4849	0,5253	0,5926
Information Ratio	-0,5035	0,0221	-0,3971	-
VaR	-0,0023	-0,0025	-0,0027	-0,0164
Conditional VaR (ES)	-0,0031	-0,0035	-0,0042	-0,0190
Standardavvik	0,0245	0,0281	0,0320	0,1273
Max Drawdown	0,0180	0,0105	0,0167	0,0683
Beta	-0,0393	-0,0245	-0,0215	1,0000

Tabell 8 Alle parene

Deskriptiv statistikk

Det viser seg at pairs trading er en veldig handelsintensiv investeringsstrategi. Samlet sett fant vi i gruppen 5 % 161 par som ifølge våre tester egnet seg for handel. I perioden gikk vi inn i til sammen 443 posisjoner. Ettersom hver posisjon består av fire handler, gjorde vi dermed 1772 handler. Den gjennomsnittlige holdingperioden for hver posisjon var 19,6 handelsdager. Fordelingen av antall gevinster og tap var rimelig jevnt fordelt med en win/loss ratio på 1,14. Samtidig var største gevinst 42,86 %, mens største tap ble -47,79 %.

Det var ikke store forskjeller i antall posisjoner inngått mellom gruppen 5 % og gruppen 10 %. I den sistnevnte gruppen åpnet vi 454 posisjoner, som tilsvarer 1816 handler. Det kan derimot virke som at den 10 % gruppen har dårligere mean-reverting egenskaper, ettersom den gjennomsnittlige holdingperioden er 23,4 handelsdager. Dette representerer en økning på nesten fire handelsdager i denne gruppen i forhold til gruppen 5 %. Videre synes det som at forholdet mellom antall gevinster og tap har økt i porteføljen til 10 % gruppen som følge av at win/loss ratio er økt til 1,365. Samtidig er det største tapet på en posisjon 47,59 %, omtrent det samme som i 5 % gruppen. Den største gevinsten er på hele 184,6 %.

Den opplagte forskjellen denne studien sammenliknet med Gatev et al. (2006) er hvor lenge posisjoner er åpne og gjennomsnittlig antall åpning av posisjoner per par. I deres studie fant de i gjennomsnitt at den gjennomsnittlige holdingperioden på et par var 3,97 måneder, mer i tråd med en "medium term" investeringsstrategi. I vår studie er den gjennomsnittlige holdingperioden på rundt en måned i gjennomsnitt, som må kunne anses som en kort tidshorisont. Videre har vi i studien samlet 322 par i begge gruppene, med totalt 897 posisjoner åpnet i perioden. Vi kan derfor si at et par gjennomsnittlig gjorde 2,79 "roundtrips", noe som er markant høyere enn 1,62 "roundtrips" som Gatev et al. (2006) finner for alle parene i hans portefølje.

Bakgrunnen for forskjellene kan være basert i handelsindikatoren. Gatev et al. (2006) bruker i sin studie en handelsindikator basert på historisk standardavvik på spreaden som blir estimert i oppbyggingsfasen av hvert enkelt par. Ved at vi bruker glidende gjennomsnitt på 50 dager resulterer nok dette i at vår indikator gir flere handelssignaler.

Avkastning og prestasjonsmål

I porteføljen med alle aksjene i gruppen 5 % oppnådde vi en periodeavkastning på 0,21 %, noe som tilsvarer en årlig avkastning på 0,42 %. Ettersom vi oppnådde så lav avkastning, gjenspeiles dette også i prestasjonsmålene Sharpe ratio og Information ratio, som begge er betydelig lavere enn de samme forholdstallene til referanseindeksen.

Vi oppnådde overraskende resultater i porteføljen med alle parene i gruppen 10 %.

Periodeavkastningen ble 3,56 %, godt hjulpet av bra resultater fra de finansielle aksjene.

Dette tilsvarer en årlig avkastning på 7,17 %. Med tanke på at referanseindeksen oppnådde en årlig avkastning på 7,72 % var dette et veldig bra resultat.

Ettersom vi oppnådde en betydelig avkastning med lav risiko i forhold til referanseindeksen ble naturlig nok prestasjonsmålene bra. Sharpe ratio på 2,48 på porteføljen i motsetning til 0,59 på referanseindeksen. Information ratio var også positiv, noe som indikerer at vi oppnådde meravkastning.

For å sammenlikne med tidligere studier, ser vi at Gatev et al. (2006) finner en gjennomsnittlig månedlig avkastning på 0,614 % for alle parene i sin studie, dersom man legger den mest konservative måten å regne ut avkastningen på til grunn. Dette tilsvarer en gjennomsnittlig årlig avkastning på 7,37 % i perioden 1963 til 2002. Do & Faff (2009) finner på sin side en gjennomsnittlig månedlig avkastning på 0,07 % i perioden 2003 til 2008, som i følge dem ikke er signifikant annerledes enn null. Dette tilsvarer en årlig avkastning på 0,84 %.

Vi ser fra våre resultater at avkastningen i gruppen 5 % er mest i tråd med Do & Faff (2009), mens resultatene i gruppen 10 % er sammenliknbar med funnene til Gatev et al. (2006).

Forskjellen skyldes i hovedsak avkastningen til porteføljen med parene i sektoren Financials i gruppen 10 % som oppnådde en årlig avkastning på 19,31 %. I vår studie var dette den porteføljen som gjorde det aller best. Ettersom vi bare har en periode i vår studie, er det vanskelig å trekke noen konklusjoner om hvorvidt en slik avkastning er representativ, eller om det dreier seg om tilfeldigheter. Porteføljen skiller seg så klart fra de øvrige funnene våre slik at det kan være nyttig å se hvilke resultater man oppnår uten denne.

Dersom vi ser bort fra denne porteføljen vil de øvrige parene i gruppen 10 % gi en årlig avkastning på 1,86 % og en Sharpe ratio på 0,5253. Dette er altså langt lavere enn hva Gatev et al. (2006) sine estimater, men fremdeles drøye ett prosentpoeng høyere enn Do & Faff

(2009). Samtidig er det viktig å huske på at avkastningen vil variere fra periode til periode. Selv i studien til Do & Faff (2009) varierer den månedlige avkastningen fra et minimum på -2,72 %, til et maksimum på 3,19 %. Dette kan tyde på at vi ikke har noe grunnlag for å konkludere med at funnene i vår studie skiller seg fra deres.

Risikoparametre

Det kan virke som at vi oppnådde en markedsnøytral strategi i porteføljen 5 % ettersom beta ikke er langt fra null. Når vi ser på porteføljens risikoparametre finner vi et veldig lavt årlig standardavvik på 0,0245, mens VaR og Conditional VaR er hhv. -0,0023 og -0,0031. VaR impliserer at gjennomsnittlig vil vi en gang hver 20 dag tape minimum 0,23 % av porteføljen. Det største daglige tapet i perioden for denne porteføljen ble 0,51 %.

Maksimal drawdown er 0,018, noe som også er lavere enn referanseindeksen. Strategien innebærer dermed mindre volatilitet i forhold til å holde markedsporteføljen, noe vi også forventet.

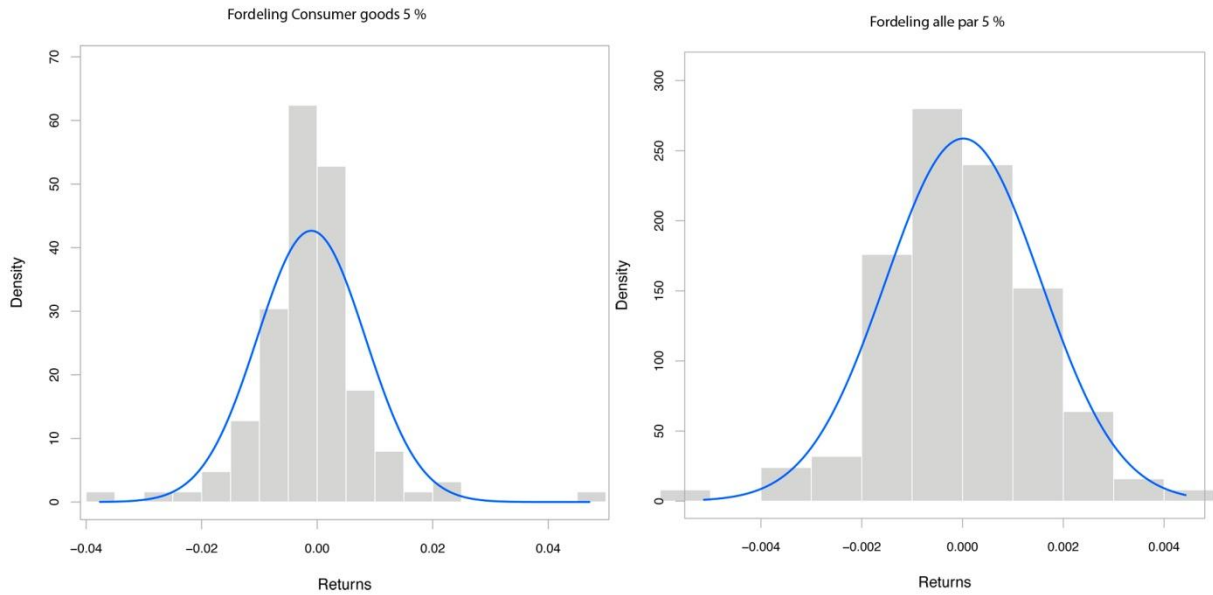
Også i porteføljen 10 % virker det som vi var markedsnøytrale med en beta nær null. Samtidig var det årlige standardavviket på 0,0281, mens VaR og Conditional var hhv. 0,0025 og 0,0035. Maksimal drawdown ble i denne porteføljen 0,0105.

Gatev et al. (2006) finner i sin studie at den daglige Value at Risk (5 %) i porteføljen for alle parene er -0,00202. Dette stemmer godt overens med vår studie.

Fordelingen til avkastningen

Et interessant funn fra vår studie er fordelingen til avkastningen. I Gatev et al.'s (2006) studie finner de en skewnesskoeffisient på 3,42 og kurtosiskoeffisient på 25,25 på fordelingen til alle parene ved handel samme dag som handelssignalet utløses. I Do & Faff (2009) sin studie finner de tilsvarende en skewnesskoeffisient på 0,41 og kurtosiskoeffisient på 6,93. Fordelingen til avkastningen i de to studiene har altså en høy kurtosiskoeffisient.

Vi finner like tendenser i vår studie der vi har få aksjer i porteføljen. For eksempel har fordelingen til avkastningen til porteføljen Consumer goods i gruppen 5 % en positiv skewness på 0,61 og kurtosis på 6,67.



Figur 19 Fordeling

Ser vi derimot på fordelingen til porteføljene som inneholder alle parene i en gruppe virker fordelingen til avkastningen å nærme seg normalfordelingen. Avkastningen til porteføljen med alle parene innenfor 5 % gruppen har en skewnesskoeffisient på $-0,073$ og en kurtosiskoeffisient på $0,42$, noe som er nært normalfordelt. Dette gjør at man ikke trenger å gjøre vurderinger hvorvidt utregninger vi har gjort som forutsetter normalfordeling er representative. Som et eksempel nevner Gatev et al. (2006) at Sharpe ratio kan gi et feilaktig bilde dersom man har en distribusjon med et negativt tredjemoment.

Han finner videre at selv om man øker antall par i en portefølje, har dette egenskapen at minimumsavkastningen øker mens maksimumsavkastningen holder seg relativt stabil. Dette representert med en økende skewnesskoeffisient. Basert på data fra vår studie finner vi at det motsatte er tilfelle, nemlig en avtagende skewnesskoeffisient, slik at vi kan ikke konkludere på samme måte som Gatev et al. (2006).

6.2 Transaksjonskostnader

Når man tester investeringsstrategier vil naturligvis transaksjonskostnader ha stor innvirkning på resultatet. Dette gjelder særlig handelsintensive strategier slik som den vi har testet.

Det er ikke tatt høyde for transaksjonskostnader ved utregning av avkastning i vår studie. Vi ønsker derfor å diskutere hvorvidt strategien er robust for transaksjonskostnader. Gatev et al. (2006) hadde i sin studie to handleregler. For den første gjorde de handler samme dag som handlesignalet indikerte at de burde inngå en posisjon, mens for den andre ventet de med handelen til dagen etter når handlesignalet indikerte åpning.

De tar en konservativ tilnærming for å estimere transaksjonskostnadene. Her ser de for seg en ekstrem tilnærming ved at når man åpner posisjonene er den relative vinneren ask pris og taperen er bid pris. Ved å vente en dag er det lik sannsynlighet for at prisene skal være henholdsvis bid eller ask, og dette vil da redusere avkastningen med halvparten av spreaden til kjøps- og salgsposisjon. Når parene konvergerer forutsetter de at vinneren er bid, mens taperen er ask. Ved igjen å vente en dag vil man redusere avkastningen med halvparten av bid-ask spreaden til aksjene. Dette impliserer at den reduserte avkastningen med å vente en dag transaksjonskostnadene til et par.

De fant i studien at periodeavkastningen falt med 324 basispoeng dersom man ventet en dag. Ut fra dette mente de at forskjellene i avkastningen gjenspeilet gjennomsnittlig bid-ask spread som igjen kan forklares som transaksjonskostnader.

Siden de i gjennomsnitt inngikk to posisjoner per par i en seks måneders periode, impliserer dette ifølge dem en gjennomsnittlig transaksjonskostnad per posisjon på 162 basispoeng, eller en estimert bid-ask spread på 81 basispoeng. Dette var høyere enn estimer funnet i tidligere studier av Peterson and Fialkowski (1994) i Gatev et al. (2006) som fant en spread på 37 basispoeng. Av denne grunn mente Gatev et al. (2006) at deres konservative estimering i verste fall er for høy. Ved å se på studiene vi tidligere refererte til kan det virke som at transaksjonskostnadene for en institusjonell aktør ligger et sted mellom 74 og 162 basispoeng per par.

Metoden Gatev et al. (2006) brukte for å vurdere hvorvidt strategien var robust for transaksjonskostnader var dermed å ta den gjennomsnittlige seks måneders fully invested avkastningen og trekke transaksjonskostnadene fra denne. I studiet fant de en avkastning

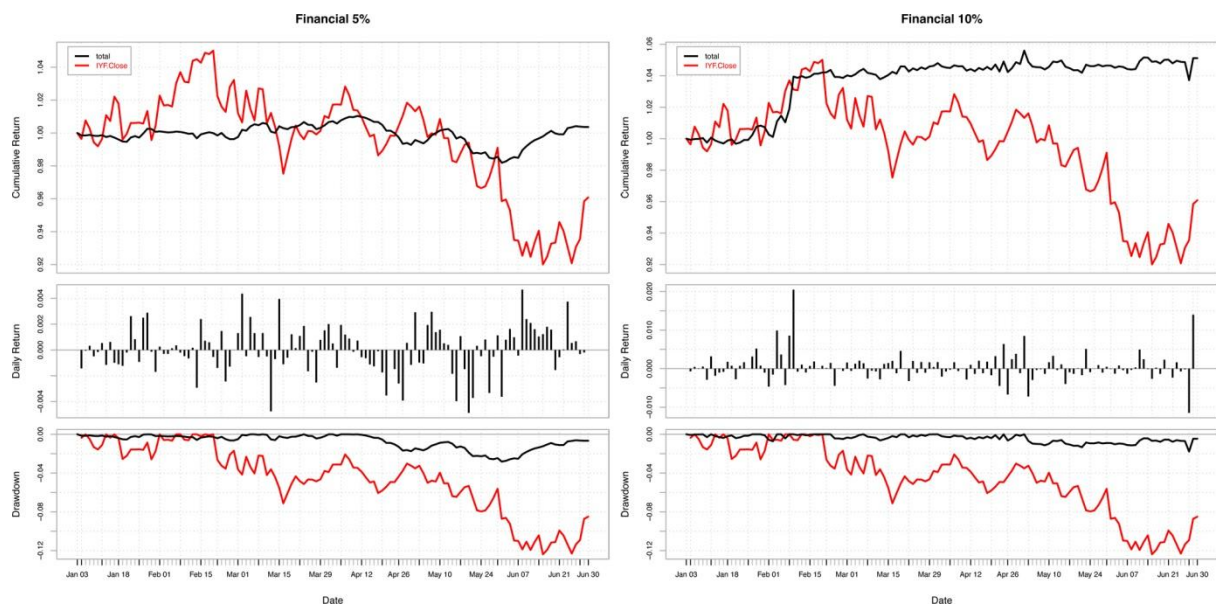
som varierte fra 437 til 549 basispoeng som de trakk fra de 324 basispoengene de fant at avkastningen falt med. På bakgrunn av at de fremdeles fikk et overskudd på 113 til 225 basispoeng konkluderte de med at deres strategi var robust for transaksjonskostnader.

I vår studie inngikk vi gjennomsnittlig 2,8 posisjoner i de seks månedene vi handlet. Dersom vi legger Gatev et al. (2006) til grunn med 162 basispoeng i transaksjonskostnad per posisjon, vil den gjennomsnittlige transaksjonskostnaden vår være 453,6 basispoeng eller 4,536 %. Etter å ha gått gjennom alle parene viser det seg at det bare finnes ett par som ikke handlet i perioden. Av denne grunn vil dermed avkastningen vi kom fram til, være svært lik begge målene på avkastningen til Gatev et al. (2006). Ettersom alle parene i porteføljen er vektet likt, kan vi forstå periodeavkastningen som gjennomsnittlig avkastning per par i perioden. For at avkastningen vi fant i studien vår skal være robust for transaksjonskostnader, må dermed periodeavkastningen være minimum 4,536 %. Vi finner i vår studie bare en portefølje som presterer bedre enn dette, nemlig porteføljen til gruppen 10 % i sektoren Financials. Det kan derfor virke som at vår strategi totalt sett ikke er robust for transaksjonskostnader.

Dermed blir spørsmålet hvordan man kan gjøre strategien robust for transaksjonskostnader, noe vi vil diskutere i neste kapittel.

6.2.1 Sensitivitetsanalyse transaksjonskostnader

Intuitivt kan man tenke seg at dersom en gjør færre handler vil transaksjonskostnadene også bli lavere. Vi har i kapittel 4.6.1 diskutert valg av lengde på glidende gjennomsnitt i indikatoren. Ved å velge et lengre tidsrom for beregning av glidende gjennomsnitt, ser vi at holdingperioden per par blir lengre, som igjen reduserer hyppigheten av handler. En annen måte vil være å vente med å inngå posisjoner til aksjene divergerer til et større antall standardavvik enn det vi bruker. Man vil da teoretisk sett få en større gevinst per par som igjen vil dekke inn transaksjonskostnadene.



Figur 20 Graf, Financials SMA150

	Financials 5% SMA150	Financials 10% SMA150
Periodeavkastning	0,0037	0,0512
Antall posisjoner	82	71
Antall par	53	49
Roundtrips	1,55	1,45
Holdingsperiode	40,0	42,8

Tabell 9 Financials SMA150

I studien fant vi at gjennomsnittlige transaksjonskostnader per par var 453,6 basispoeng, dette med utgangspunkt i estimeringen av transaksjonskostnader i studien til Gatev et al. (2006). Vi har forsøkt å øke lengden på glidende gjennomsnitt for å se hvilken effekt dette vil få for transaksjonskostnadene. I eksempelet vist i tabell 9, har vi brukt 150 dagers glidende gjennomsnitt, dette tilsvarer i overkant av 7 måneder med handelsdager. Som forventet øker holdingsperioden per par markant. Vi får her en gjennomsnittlig holdingsperiode påhenholdsvis 40 og 42,8 handelsdager for gruppene 5 % og 10 %, mot tidligere hhv. 30,6 og 32,2 dager. Dette tilsvarer en økning i holdingsperioden på rundt 33 % for begge gruppene.

Som en konsekvens av lengre holdingsperiode, faller også gjennomsnittlig antall roundtrips per par i perioden. Ved å bruke 50 dagers glidende gjennomsnitt i indikatoren hadde vi for denne sektoren i gjennomsnitt 2,6 og 2,9 roundtrips per par i perioden for gruppen 5 % og 10 %, ikke ulikt gjennomsnittet vi fant for alle parene. Som vi ser i tabell 9 har antallet

roundtrips falt markant til henholdsvis 1,55 og 1,45 i perioden ved å øke glidende gjennomsnitt til 150 dager. Det impliserer at transaksjonskostnadene faller fra 453,6 basispoeng til 243 basispoeng, altså en reduksjon på 44 %. Dette er stemmer bra overens med antall handler vi gjør, ettersom vi gjør 41 % færre handler i 5 % gruppen mot 48 % færre handler i 10 % gruppen.

Dersom holdingperioden øker vil ceteris paribus, avkastningen øke ettersom transaksjonskostnadene reduseres. Dessverre er det ikke så enkelt, ettersom vi får en annen avkastning når vi endrer indikatoren. For gruppen 10 % får vi en periodeavkastning på 5,12 %. Trekker vi fra transaksjonskostnadene får vi en nettoavkastning på 2,69 %. Hvis vi ser på eksempelet med 50 dagers glidende gjennomsnitt har vi en periodeavkastning på 9,58 %. Med transaksjonskostnader estimert til 453,6 basispoeng, tilsvarer dette en nettoavkastning på 5,04 % som er bedre enn i tilfellet med 150 dagers glidende gjennomsnitt.

Allikevel må vi si at forskjellen mellom de to transaksjonskostnadene på 210,6 basispoeng er stor. Hvis man tenker seg at man har en langsiktig investeringshorisont vil en årlig besparelse i transaksjonskostnader på 421,2 basispoeng kunne utgjøre forskjellen mellom tap og gevinst.

6.3 Oppsummering

Vi vil i dette kapittelet gjøre en oppsummering av de viktigste funnene i vår studie.

Antall par i porteføljene har variert ettersom hvilken sektor de tilhører. Porteføljen til sektoren Conglomerates inneholdt bare to par, men vi oppnådde allikevel en beta nær null. Dette kan tyde på at pairs trading er en markedsnøytral strategi uavhengig av antall par i porteføljen. Videre har vi funnet indikasjoner på at det er en diversifiseringsgevinst ved å øke antall par i en portefølje, ettersom både porteføljens standardavvik og VaR reduseres når man introduserer flere par i porteføljen. Investeringsstrategien har for øvrig lav risiko sammenliknet med å holde referanseindeksen målt i både standardavvik og VaR.

Ved å bruke den handelsindikatoren vi presenterte i kapittel 4.4 resulterer dette i en veldig handelsintensiv strategi. Vi ser at holdingperioden ligger rundt 20 handelsdager. For gruppen 10 % fant vi at gjennomsnittlig holdingperiode var lengre enn for 5 % gruppen. Den korte holdingperioden fører naturlig nok til flere handler, noe som øker transaksjonskostnadene. Ved å øke lengden på glidende gjennomsnitt i indikatoren fra 50 dager til 150 dager ble transaksjonskostnadene redusert med 44 %.

For porteføljen med alle parene i gruppen 5 % oppnådde vi i studiet nær nullavkastning, noe som stemmer overens med resultater fra studien til Do & Faff (2009), mens i gruppen 10 % ble periodeavkastningen 3,56 %.

7. Avslutning

7.1 Oppsummering og konklusjon

Teori om markedseffisiens er essensielt for å forstå hvorfor det er vanskelig å oppnå unormal avkastning i aksjemarkedet. Vår tilnærming ble derfor ikke å stille spørsmålstegn med hypotesen, men å se på om avkastningen til handelsstrategien er sammenfallende med tidligere forskning.

Gatev et al. (2006) hevder i sin studie at mulighetene ved pairs trading skyldes en kompensasjon til arbitrasjører for å håndheve "The Law of One Price". Arbitrasje er derfor ikke ment til å slå markedet, men å oppnå positiv avkastning uavhengig av markedsrisiko ved å gjenopprette likevekter.

Vårt utgangspunkt var å finne aksjer som er kointegrerte. Vi sorterte aksjeparane i to grupper på følgende måte: Alle aksjeparane som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe som vi ga navnet 5 % gruppen. De parene som ble signifikante på minimum 10 % nivå i alle testene ble gruppert i den andre gruppen som fikk navnet 10 % gruppen. I begge gruppene fikk vi 161 aksjepar.

Vi spekulerte derfor i at aksjepar som divergerer fra likevekten igjen vil konvergere, ved å selge den overprisede og å kjøpe den underprisede med like store posisjoner på begge sider. Handleregelen vår er basert på mispricing uttrykt i standardavvik på spreaden. Glidende gjennomsnitt var 50 dager og posisjonene åpnet på 2 standardavvik. Vi lukket posisjonene ved konvergering til likevekten og på slutten av handelsperioden, eller hvis aksjene divergerte utover 3 standardavvik.

Problemstillingen vår er som følger:

"Gir handelsstrategien pairs trading meravkastning kombinert med lav risiko eller er avkastningen avtagende jamfør tidligere forskning?"

I 5 % gruppen fikk vi en årlig avkastning på 0,42 %. Dette resultatet samsvarer med Do & Faff (2009) som oppnådde nær nullavkastning i sin studie. 10 % gruppen presterte langt bedre med en årlig avkastning tilsvarende 7,17 %. Dette resultatet er mer i samsvar med Gatev et al. (2006). Siden vi bare har en periode i studien er det vanskelig å gi en forklaring på hvorfor 10 % gruppen presterer såpass mye bedre enn 5 % gruppen. Dette kan skyldes tilfeldigheter, ettersom Do & Faff (2009) fant stor variasjon i den månedlige avkastningen fra -2,72 %, til 3,19 %. Dette kan tyde på at vi ikke har noe grunnlag for å konkludere med at funnene i vår studie skiller seg fra deres. Et fellestrekk ved hele studien er at strategien er markedsnøytral, med langt lavere risiko uttrykt i standardavvik og VaR enn referanseindeksen. Dette er også i samsvar med studiene til Gatev et al. (2006) og Do & Faff (2009).

Resultatene våre er totalt sett ikke lønnsomme justert for transaksjonskostnader. Ved å øke glidende gjennomsnitt til 150 dager så reduserer vi transaksjonskostnadene med 44 %. Har man derfor en lang investeringshorisont kan dette utgjøre forskjellen mellom tap og gevinst. Ser vi derimot på 10 % gruppen innenfor finanssektoren så er strategien vært attraktiv på bakgrunn av positiv avkastning samt svært lav risiko. Det bør derfor gjøres videre forskning på om dette er tilfeldigheter eller ikke.

I forhold til andre studier så har vi stort sett valgt å se på Do & Faff (2009), samt Gatev et al. (2006). Begge disse studiene er omfattende på bakgrunn av antall aksjer og aksjepar. I et slikt perspektiv kan man si noe om pairs trading ved store datasett, men kan man på bakgrunn av dette generalisere pairs trading som helhet? Dette kan man trolig ikke, da en trader som oftest vil spesialisere seg innenfor en sektor eller et antall aksjepar som vedkommende føler seg komfortabel med.

Som konklusjon kan det tyde på at strategien totalt sett gir nullavkastning som er sammenfallende med Do & Faff (2009). Studien til Gatev et al. (2006) er noe eldre, men også denne viser en avtagende lønnsomhet. Vi henviser derfor til videre forskning i kapittel 7.3.

7.2 Kritikk av studiet

I en oppgave som denne vil det naturlige spørsmålet være i hvilken grad strategien kan overføres til virkelighetens verden. I vårt tilfelle er dette mulig. Det finnes tilleggspakker til R som gjør det mulig å koble programmet opp mot en internettbasert megler, og handlere ville da automatisk blitt utført etter hvert som handelssignalene oppstår. Naturligvis kan man også benytte seg av andre løsninger som gir samme effekt.

Når det er sagt er det viktig å klargjøre at vi har brukt sluttkurser, som vil si den siste handelen som blir gjennomført i sluttauksjonen hver dag. I den sammenheng kan man selvsagt argumentere for at handelssignal basert på intradagpriser er en mer realistisk tilnærming. Dette er derfor noe som må tas i betraktning når man skal vurdere resultatene i studien. Vårt utgangspunkt er allikevel at utslagene ved bruk av intradagpriser vil være minimale med tanke på den gjennomsnittlige holdingperioden til parene i studiet. Hvis man tenker seg handelsstrategier som er beregnet kun på daytrading eller opp til noen dager, så vil bruk av intradagpriser naturlig nok være mer essensielt.

Som nevnt i kapittel 4.3.3 viser Whistler (2004) til metoder hvor man gjør undersøkelser vedrørende fundamentale forhold i selskaper, og at man på bakgrunn av funnene i undersøkelsene kan luke ut par som ikke egner seg til handel. På bakgrunn av det store utvalget av par i vår studie, var det ikke hensiktsmessig for oss å undersøke fundamentale forhold mellom selskapene i parene vi fant. Det er grunn til å tro at de som trader i den virkelige verden gjør slike undersøkelser. I følge Grøtte (2002) følger han i underkant av 10 par som han handler med. Bredt utvalg av par gjør dette at vår studie kanskje ikke er sammenlignbar med hvordan tradere i den virkelige verden opptrer, men på den andre siden gjør det at vi kan sammenligne vår studie med tidligere studier.

7.3 Forslag til videre forskning

I studien har vi testet strategien over de seks første måneder i 2011. Det ville vært meget interessant å se hvordan resultatene ville blitt om vi hadde testet strategien over en lengre tidsperiode. I den sammenheng måtte man da brukt den samme metoden for hver nye handelsperiode. Dette betyr altså testing av par for kointegrasjon i 12 måneder forut for hver seks måneders handelsperiode. For og ytterligere få flere perioder kan man suksessivt flytte seg en måned framover med testing og handel.

Et sentralt poeng som leseren bør være klar over er at metoden vår ikke inneholder stop-loss ved prosentvis tap, men ved divergering utover ± 3 standardavvik. Selv om dette stopper en del tapsposisjoner åpner det allikevel mulighet for betydelige tap. Det ville derfor være meget interessant å teste ut hvordan resultatene ville blitt ved å stoppe tapsposisjoner ved ulike nivå for stop-loss, som ved 10 %, 15 % og 20 %.

En annen metode er å kjøre en backtest av strategien i de samme 12 måneder som man finner parene, for så kun å bruke de parene som presterer hensiktsmessig i handelsperioden. Målet ville da være å se om resultatene bedrer seg eller ikke.

I sakens natur er det kun fantasien som begrenser mulighetene for videre forskning. Alle parametere kan endres og optimaliseres for å finne den beste metoden og tilhørende personlige preferanser.

8. Litteraturliste

- Avellaneda, M., & Lee, J. H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7), 761-782.
- Björk, T. (1998). *Arbitrage theory in continuous time*. Oxford: Oxford University Press.
- Bodie, Z., Kane, A., & Marcus, A. J. (2009). *Investments* (8th ed.). Boston, Mass.: McGraw-Hill.
- Bondarenko, O. (2003). Statistical Arbitrage and Securities Prices. *Review of Financial Studies*, 16(3), 875-919.
- Brown, S. J. (2011). The efficient markets hypothesis: The demise of the demon of chance? *Accounting & Finance*, 51(1), 79-95. doi: 10.1111/j.1467-629X.2010.00366.x
- Bruce, N. L. (1990). Fads, Martingales, and Market Efficiency: National Bureau of Economic Research, Inc.
- Chan, E. P. (2009). *Quantitative trading: how to build your own algorithmic trading business*. Hoboken, N.J.: Wiley.
- Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of Financial Economics*, 87(2), 249-268.
- Damodaran, A. (2003). *Investment philosophies: successful strategies and the investors who made them work*. Hoboken, N.J.: Wiley.
- Do, B., & Faff, R. (2009). Does simple pairs trading still work? *Financial Analysts Journal*, 66(4), 1-13.
- Do, B. H., & Faff, R. W. (2011). Are Pairs Trading Profits Robust to Trading Costs? *SSRN eLibrary*. doi: 10.2139/ssrn.1707125
- Dubil, R. (2004). *An arbitrage guide to financial markets*. Chichester: Wiley Finance.
- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.
- Gatev, E., Rouwenhorst, K. G., & Goetzmann, W. N. (2006). *Pairs trading: performance of a relative value arbitrage rule* (Vol. 7032). Cambridge, Mass.: NBER.
- Grøtthe, O. (2002). *Aksjekjøp og daytrading: metode, psykologi, risiko og strategier*. [Lysaker]: Hegnar media.
- Keim, D. B. (1983). Size-related anomalies and stock return seasonality. Further empirical evidence. *Journal of Financial Economics*, 12(1), 13-32.
- Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, 49(5), 1541-1578.
- Lütkepohl, H., & Krätzig, M. (2004). *Applied time series econometrics*. Cambridge: Cambridge University Press.
- Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *The Journal of Economic Perspectives*, 17(1), 59-82.
- Peter Carl, B. G. P., Kris Boudt, Eric Zivot. (2012). PerformanceAnalytics: Econometric tools for performance and risk analysis.
- Pole, A. (2007). *Statistical arbitrage: algorithmic trading insights and techniques*. Hoboken, N.J.: Wiley.
- Reverre, S. (2001). *The complete arbitrage deskbook*. New York: McGraw-Hill.
- Ross, S. A., Westerfield, R. W., Jordan, B., Roberts, G. (2007). *Fundamentals of Corporate Finance* (Sixth Canadian ed.): McGraw-Hill Ryerson.
- Setterberg, H. (2007). *Swedish post-earnings announcement drift and momentum return*. Center for Financial Analysis and Managerial Economics in Accounting. Stockholm School of Economics.

- Shleifer, A. (2000). *Inefficient markets: an introduction to behavioral finance*. Oxford: Oxford University Press.
- Shleifer, A., Vishny, Robert W. (1997). The Limits of Arbitrage. *The Journal of Finance*, 52(1), 35-55.
- U.S Department of the Treasury (n.d). Resource Center. Daily Treasury Yield Curve Rates. Retrieved 22.04.2012 from <http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2011>
- Varian, H. R. (1987). The Arbitrage Principle in Financial Economics. *The Journal of Economic Perspectives*, 1(2), 55-72.
- Vidyamurthy, G. (2004). *Pairs trading: quantitative methods and analysis*. Hoboken, N.J.: Wiley.
- Whistler, M. (2004). *Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies*. Hoboken, N.J.: Wiley.
- Zivot, E., & Wang, J. (2006). *Modeling financial time series with S-PLUS*. New York: Springer.

Vedlegg 1 Script for å få data inn i R. Eksempel sektoren Agricultural i Basic Materials.

```
mon <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/MON.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
mon$V1 <- mon$V3 <- mon$V4 <- mon$V5 <- mon$V7 <- NULL
colnames(mon) <-c("Dato", "MON")
pot <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/POT.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
pot$V1 <- pot$V3 <- pot$V4 <- pot$V5 <- pot$V7 <- NULL
colnames(pot) <-c("Dato", "POT")
syt <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/SYT.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
syt$V1 <- syt$V3 <- syt$V4 <- syt$V5 <- syt$V7 <- NULL
colnames(syt) <-c("Dato", "SYT")
mos <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/MOS.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
mos$V1 <- mos$V3 <- mos$V4 <- mos$V5 <- mos$V7 <- NULL
colnames(mos) <-c("Dato", "MOS")
agu <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/AGU.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
agu$V1 <- agu$V3 <- agu$V4 <- agu$V5 <- agu$V7 <- NULL
colnames(agu) <-c("Dato", "AGU")
cf <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/CF.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
cf$V1 <- cf$V3 <- cf$V4 <- cf$V5 <- cf$V7 <- NULL
colnames(cf) <-c("Dato", "CF")
smg <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/SMG.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
smg$V1 <- smg$V3 <- smg$V4 <- smg$V5 <- smg$V7 <- NULL
colnames(smg) <-c("Dato", "SMG")
cmp <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/CMP.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
cmp$V1 <- cmp$V3 <- cmp$V4 <- cmp$V5 <- cmp$V7 <- NULL
colnames(cmp) <-c("Dato", "CMP")
tnh <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/TNH.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
tnh$V1 <- tnh$V3 <- tnh$V4 <- tnh$V5 <- tnh$V7 <- NULL
colnames(tnh) <-c("Dato", "TNH")
avd <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/AVD.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
avd$V1 <- avd$V3 <- avd$V4 <- avd$V5 <- avd$V7 <- NULL
colnames(avd) <-c("Dato", "AVD")
#Har hentet script fra http://www.stockhistoricaldata.com/ og det til å hente data fra Yahoo
Finance.
agricultural <- data.frame
agricultural <- merge(mon,pot,by="Dato")
rm(mon,pot)
agricultural <- merge(agricultural,syt,by="Dato")
rm(syt)
agricultural <- merge(agricultural,mos,by="Dato")
rm(mos)
agricultural <- merge(agricultural,agu,by="Dato")
rm(agu)
agricultural <- merge(agricultural,cf,by="Dato")
rm(cf)
agricultural <- merge(agricultural,smg,by="Dato")
rm(smg)
agricultural <- merge(agricultural,cmp,by="Dato")
rm(cmp)
agricultural <- merge(agricultural,tnh,by="Dato")
rm(tnh)
agricultural <- merge(agricultural,avd,by="Dato")
rm(avd)
agricultural <-na.omit(agricultural)
write.csv(agricultural, file= "agricultural.csv")
```

Vedlegg 2 Johansen test av en hel sektor. Eksempel Property and Casualty Insurance.

```
library(urca)
result <- list()
j <- 0
startindex <- which(pcinsurance$Dato>"20100100")[1]
sluttindex <- which(pcinsurance$Dato>"20110100")[1]-1
for(i1 in 2:46){
  for (i2 in (i1+1):47){
    j <- j+1
    a1 <- pcinsurance[startindex:sluttindex,i1]
    a2 <- pcinsurance[startindex:sluttindex,i2]
    a3 <- cbind(a1,a2)
    colnames(a3) <- c(colnames(pcinsurance)[i1], colnames(pcinsurance)[i2])
    result[[j]] <- ca.jo(a3, type = c("eigen"), ecdet = c("none"), K = 10,
spec=c("transitory"), season = NULL, dumvar = NULL)
  }
}
for (j in 1:1035){
  print(summary(result[[j]]))
}
```

Vedlegg 3 Eksempel output Johansen test.

```
#####
# Johansen-Procedure #
#####

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.081569696 0.005161012

Values of teststatistic and critical values of test:

      test 10pct  5pct  1pct
r <= 1 |  1.25  6.50  8.18 11.65
r = 0  | 20.59 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)

      HTH.11  AMSF.11
HTH.11 1.000000 1.00000000
AMSF.11 1.303256 -0.09938318

Weights W:
(This is the loading matrix)

      HTH.11  AMSF.11
HTH.d -0.002493368 -0.009520242
AMSF.d -0.123693611 -0.008397578
```

Vedlegg 4 ADF og PP-test av spreaden mellom to aksjer. Eksempel Aksjeparet HTH.AMSF.

```
Date <- strptime(pcinsurance$Dato,format="%Y%m%d")
hth.amsf <- data.frame
hth.amsf <- pcinsurance[c("Dato","HTH","AMSF")]
hth.amsf <-cbind(hth.amsf,Date)
hth.amsf$Dato <- NULL
hth.amsf <-hth.amsf[ ,c(3,1:2)]
hth.amsf.xts <- xts(hth.amsf[,-1], order.by= hth.amsf[,1])
rm(hth.amsf)
library(PairTrading)
price.pair <- hth.amsf.xts[,1:2]["2010-01-01::2010-12-31"]
reg <- EstimateParameters(price.pair, method = lm)
str(reg)
plot(reg$spread)
IsStationary(reg$spread, 0.05)
```

Vedlegg 5 Eksempel output ADF og PP-test

```
List of 3
 $ spread      :An 'xts' object from 2010-01-04 to 2010-12-31 containing:
  Data: num [1:252, 1] 0.0599 0.0287 0.0179 0.0126 0.0242 ...
- attr(*, "dimnames")=List of 2
 ..$ : NULL
 ..$ : chr "AMSF"
Indexed by objects of class: [POSIXct,POSIXt] TZ:
xts Attributes:
List of 2
 ..$ tclass: chr [1:2] "POSIXct" "POSIXt"
 ..$ tzone : chr ""
 $ hedge.ratio: num -0.416
 $ premium   : num 3.86
> plot(reg$spread)
> IsStationary(reg$spread, 0.05)
PP.test adf.test
TRUE FALSE
```

Vedlegg 6 Script for testing av handelsstrategi. Eksempel fra alle 5 %.

```
#by Garrett See
#code borrowed heavily from existing quantstrat demos

#This is a simple pairs trading example intended to illustrate how you can extend
#existing quantstrat functionality. Also, it uses addPosLimits to specify
#levels and position limits, and shows how to pass a custom order sizing function to osFUN

#Note that it would be easier to build a spread first and treat it as a single instrument
#instead of dealing with a portfolio of stocks.

## given 2 stocks, calculate their ratio. If the ratio falls below it's
# 2 stdev band, then when it crosses back above it, buy stock 1 and sell stock 2.
# If the ratio rises above it's 2 stdev band, then when it crosses back below
# it, sell stock 1 and buy stock 2. If the ratio crosses it's moving average,
# then flatten any open positions.

# The Qty of Stock A that it buys (sells) = MaxPos / lvl
# The Qty of Stock B that it sells (buys) = MaxPos * Ratio / lvl
pdf(file="PairTrade-alle50.pdf", onefile=TRUE)
try(rm(list=.blotter[["portfolios"]], pos=.blotter), silent=TRUE)
try(rm(list=.blotter[["accounts"]], pos=.blotter), silent=TRUE)
try(rm(list=.strategy[["order_books"]], pos=.strategy), silent=TRUE)
try(rm("portfolios","accounts", pos=.blotter), silent=TRUE)
try(rm("order_books", pos=.strategy), silent=TRUE)
ls(envir=.blotter)
ls(envir=.strategy)
try(rm("initDate", "endDate", "startDate", "initEq", "SD", "N", "symb1", "symb2", "M",
      "portfoliol.st", "account.st", "pairStrat", "out1", "ret1"), silent=TRUE)

require(quantstrat)
initDate = '2011-01-03'
endDate = '2011-06-30'
#StartTrade = '2011-01-03'
startDate = '2010-10-21'
#xstartDate <- index(CEO[which(index(CEO)==StartTrade)-100])
initEq = 1000000
#SD = 2
#N = 100

#MaxPos = 100000 #max position in stockA;
#max position in stock B will be max * ratio, i.e. no hard position limit in Stock B
lvl = 1 #how many times to fade; Each order's qty will = MaxPos/lvl

#symb1<-'FII'
#symb2<-'CLMS'
#portfoliol.st <- 'pair1'
#account.st <- 'pairs'

symb1 <- c('GFI','BVN','NUE','SCHN','CEO','CEO','SNP','OXY','ECA','ECA',
  'ECA','SU','APC','APC','SWN','SWN','SWN','CHK','STR','NFX','NFX','PXP','CRK',
  'FST','SM','PVA','HES','VLO','TSO','IOC','IOC','EROC','CLMT','HAL','BHI','PDS','SIAL',
  'SXT','WPZ','WPZ','ENB','ENB','EP','KMR','DO','BBG','BBG','BBG','MWE','HERO','GIFI',
  'CBE','GE','ALV','KO','SNE','UA','HBI','CPO',
  'IVZ','FII','MS','BBD','BBD','UBS','CS','C','TD','BPO','CXW','HXM','RYN','DEI','NNN',
  'FSP','BDN','FPO','PSA','ARE','USB','NTRS','FITB','FMER','ONB','MBFI','MBFI',
  'VLY','VLY','NPBC','TRST','SIVB','UMPQ','ESS','CPT','BRE','HME','SIM','SPG','VNO',
  'KIM','REG','HPT','BRK-A','BRK-A','ACE','AIG','RNR','RNR','THG','VR','VR','CB',
  'CPWR','BMC','LFT','ADP','VSAT','EFII','ZBRA','NICE','LOGI','BMI','MTSC','CPHD',
  'IIVI','RSTI','MTD','TRMB','WAT','BHE','USMO','LEAP','CEL','SBAC','PTNR','NIHD','NIHD',
  'USM','VIV','PCS','PT','VIP','TU','MBT','TKC','TKC','CHU','AMX','DCM',
  'DCM','VOD','CHL','CHL','NTT','SHEN','DLB','SNDA','VRSN','CRM','SYMC','YHOO')
symb2 <- c('HMY','HMY','HSC','HAYN','ECA','FST','ECA','CNQ','EPD','STR',
  'XTEX','FST','KWK','ATPG','STR','CRK','PVA','GDP','GMXR','GPOR','BEXP','KWK','PVA',
  'GPOR','PQ','GMXR','CLMT','DK','CVI','CLMT','XTXI','DK','XTXI','WFT','NGS','TESO',
  'NEU','NEU','PPO','KWR','MMP','WES','CPNO','NGLS','VQ','MWE','BBEP','PSE','EVEF',
  'VQ','LNG','DHR','GY','TRW','HANS','HAR','OXM','ZQK','FLO',
  'AMP','CLMS','HGT','BMA','BFR','LYG','BCS','KEY','KEY','CXW','CIM','CIM','FUR','FUR',
  'FUR','CUZ','FUR','FUR','ARE','BMR','PVTB','PVTB','CHFC','PVTB','PVTB','FMBI','WTFC',
  'SUSQ','WBS','WFD','WBS','EWBC','HAFc','BRE','AEC','SUI','MAA','PHH','UBA','FRT','REG',
  'EPR','IRC','THG','PTP','THG','EIG','THG','ERIE','ERIE','MRH','TWGP','ACE',
  'TRAK','TRAK','NTCT','BLKB','ELMG','SSYS','SSYS','SSYS','SSYS','ALOG','BMI','BMI','BMI',
```

```

'CPHD', 'BMI', 'BMI', 'BMI', 'PLXS', 'UTSI', 'UTSI', 'UTSI', 'UTSI', 'UTSI', 'UTSI', 'USMO', 'UTSI', 'UTSI
',
'UTSI', 'UTSI', 'NIHD', 'NIHD', 'UTSI', 'UTSI', 'VIP', 'UTSI', 'UTSI', 'UTSI', 'VIP', 'UTSI', 'UTSI', 'MBT
',
'NZT', 'CNSL', 'CTS', 'ARBA', 'SNDA', 'SNDA', 'SNDA', 'UNTD')
M <-length(symb1)
if (M != length(symb2)) stop('symb1 and symb2 must have equal length\n')
portfoliol.st<-paste(symb1,symb2,sep=".")
account.st<-paste(symb1,symb2,sep=" ")
.blotter[["portfolios"]]<-paste("portfolio",portfoliol.st,sep=".")
.blotter[["accounts"]]<-paste("account",account.st,sep=".")
.strategy[["order_books"]]<-paste("order_book",portfoliol.st,sep=".")
pairStrat<-list()
ret1<-list()
out1<-list()

calcRatio <- function(x) { #returns the ratio of close prices for 2 symbols
x1 <- get(x[1])
x2 <- get(x[2])
rat <- Cl(x1) / Cl(x2)
colnames(rat) <- 'Ratio'
rat
}
getRatio <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Ratio <- last(portf$Ratio[toDate])
print(timestamp)
as.numeric(Ratio)
}
calcPrice <- function(x) { #returns the price of stock A
x1 <- get(x[1])
price <- Cl(x1)
colnames(price) <- 'Price'
price
}
getxPrice <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Price <- last(portf$Price[toDate])
as.numeric(Price)
}
calcDelta <- function(x) { #returns the ratio of close prices for 2 symbols
x1 <- get(x[1])
x2 <- get(x[2])
spread <- Cl(x1) - Cl(x2)
mean50 <-SMA(spread,n=50)
sd50 <- runSD(spread,n=50,sample=TRUE,cumulative=FALSE)
delta <- (spread-mean50)/sd50
colnames(delta) <- 'Delta'
delta <- as.xts(delta)
delta
}
getDelta <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Delta <- last(portf$Delta[toDate])
as.numeric(Delta)
}
calcFinish <- function(x) {
finish <- index(mktdata)=="2011-06-29"
finish <- as.numeric(finish)
names(finish) <- index(mktdata)
finish <- as.xts(finish)
colnames(finish) <- c("Finish")
finish
}
osSpreadMaxPos <- function (data, timestamp, orderqty, ordertype, orderside, portfolio,
symbol, ruletype, ..., orderprice)
{
portf <- getPortfolio(portfolio)
legside <- portf$pair[symbol] #"long" if symbol=symb1[i], "short" if symbol=symb2[i]
cat("legside=",legside,"\n")
if (legside != "long" && legside != "short") stop('pair must contain "long" and "short"')
ratio <- getRatio(portfolio, timestamp)
}

```

```

price <- getxPrice(portfolio, timestamp)
MaxPos <- initEq/price
if (orderqty=='all') qty <- MaxPos
if (mktdata[timestamp,"cross.p2"]==1) qty = -qty
# addPosLimit(portfolio=portfolio, timestamp=timestamp, symbol=symbol,
# maxpos=MaxPos, longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)
pos <- getPosQty(portfolio, symbol, timestamp)
cat("pos=",pos,"qty=",qty,"\n")
if((pos> 0.5) | (pos< -0.5)) qty<-0
# PosLimit <- getPosLimit(portfolio, symbol, timestamp)
# qty <- orderqty
if (legside == "short") {#symbol is 2nd leg
## Comment out next line to use equal ordersizes for each stock.
# addPosLimit(portfolio=portfolio, timestamp=timestamp,
# symbol=symbol, maxpos=MaxPos*ratio, longlevels=lvls,
# minpos=-MaxPos*ratio, shortlevels=lvls)
# TODO: is it okay that MaxPos and lvls come from .GlobalEnv ?
qty <- -qty*ratio #switch orderqty for Stock B
}
if (qty > 0) orderside = "long"
if (qty < 0) orderside = "short"
# orderqty <- osMaxPos(data=data,timestamp=timestamp,orderqty=qty,
# ordertype=ordertype,orderside=orderside,portfolio=portfolio,
# symbol=symbol,ruletype=ruletype, ...)
orderqty <- round(qty,0)
cat("orderqty=",orderqty,"\n")

#Add the order here instead of in the ruleSignal function
if (!is.null(orderqty) & !orderqty == 0 & !is.null(orderprice)) {
addOrder(portfolio = portfolio, symbol = symbol,
timestamp = timestamp, qty = orderqty, price = as.numeric(orderprice),
ordertype = ordertype, side = orderside, status = "open", ... = ...)
}
return(0) #so that ruleSignal function doesn't also try to place an order
}

for (i in 1:M) {
getSymbols(c(symb1[i], symb2[i]), from=startDate, to=endDate, adjust=TRUE)

currency("USD")
stock(symb1[i], currency="USD", multiplier=1)
stock(symb2[i], currency="USD", multiplier=1)

#Initialize Portfolio, Account, and Orders
initPortf(name=portfoliol.st[i], c(symb1[i],symb2[i]), initDate=initDate)
initAcct(account.st[i], portfolios=portfoliol.st[i], initDate=initDate, initEq=initEq)
initOrders(portfolio=portfoliol.st[i],initDate=initDate)
mktdata<-get(symb1[i])
#create a slot in portfolio for symb1[i] and symb2[i] to make them available to osFUN
pair <- c('long','short')
names(pair) <- c(symb1[i],symb2[i])
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$pair <- pair

# Create initial position limits and levels by symbol
# allow 3 entries for long and short.
#addPosLimit(portfolio=portfoliol.st[i], timestamp=initDate, symbol=symb1[i],
#maxpos=MaxPos, #longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)
#addPosLimit(portfolio=portfoliol.st[i], timestamp=initDate, symbol=symb2[i],
#maxpos=MaxPos, #longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)

# Create a strategy object
pairStrat[[i]] <- strategy(paste('pairStrat',symb1[i],symb2[i],sep="."))

Ratio <- calcRatio(c(symb1[i],symb2[i]))
#let's go ahead and put this in a slot in portfolio
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Ratio <- Ratio
#and make a function to get the most recent Ratio

#Get the price from one stock
Price <- calcPrice(c(symb1[i]))
#let's go ahead and put this in a slot in portfolio
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Price <- Price

```



```

Delta <- calcDelta(c(symb1[i],symb2[i]))
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Delta <- Delta
#and make the most recent Ratio

Finish <- calcFinish(c(symb1[i],symb2[i]))
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Finish <- Finish
#and make a function to get the most recent Ratio

# Create an indicator - BBands on the Ratio
pairStrat[[i]] <- add.indicator(strategy = pairStrat[[i]], name = "calcDelta",
arguments = list(x=c(symb1[i],symb2[i])))
pairStrat[[i]] <- add.indicator(strategy = pairStrat[[i]], name = "calcFinish",
arguments = list(x=c(symb1[i],symb2[i])))

applyIndicators(strategy=pairStrat[[i]],mktdata=get(symb1[i])) #for debugging

# Create signals - buy when crossing lower band from below, sell when crossing upper band from
above, flatten when crossing mavg from above or from below
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=-2, relationship="lte",
cross=TRUE), label="cross.m2")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=-3, relationship="lte", cross=TRUE),
label="cross.m3")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=2, relationship="gte",
cross=TRUE), label="cross.p2")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=3, relationship="gte",
cross=TRUE), label="cross.p3")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=0,relationship="gte",
cross=TRUE), label="cross.mid.fa")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=0, relationship="lte",
cross=TRUE), label="cross.mid.fb")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Finish", threshold=0, relationship="gt",
cross=FALSE), label="sigfinish")

# Create entry and exit rules for longs and for shorts. Both symbols will get the same
buy/sell signals, but osMaxPos will reverse those for the second symbol.
# orderqty's are bigger than PosLimits allow. osMaxPos will adjust the orderqty down to 1/3
the max allowed. (1/3 is because we are using 3 levels in PosLimit)
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.m2", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL, osFUN='osSpreadMaxPos'), type='enter' )
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.m3", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.p2", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL, osFUN='osSpreadMaxPos'), type='enter')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.p3", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.mid.fb", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.mid.fa", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="sigfinish", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')

#applySignals(strategy=pairStrat[[i]],
mktdata=applyIndicators(strategy=pairStrat[[i]],mktdata=get(symb1[i]))) #for debugging

out1[[i]]<-try(applyStrategy(strategy=pairStrat[[i]], portfolios=portfoliol.st[i]))

updatePortf(Portfolio=portfoliol.st[i],
Dates=paste("::",as.Date(Sys.time()),sep=''))

```

```

updateAcct(account.st[i],Dates=paste(startDate,endDate,sep="::"))
updateEndEq(account.st[i],Dates=paste(startDate,endDate,sep="::"))
getEndEq(account.st[i],Sys.time())

chart.Posn(Portfolio=portfoliol.st[i],Symbol=symb1[i])
chart.Posn(Portfolio=portfoliol.st[i],Symbol=symb2[i])
chartSeries(Delta,name=paste("Delta",portfoliol.st[i],sep=" "))

ret1[[i]] <- PortfReturns(account.st[i])
ret1[[i]]$total <- rowSums(ret1[[i]])
#ret1

if("package:PerformanceAnalytics" %in% search() ||
    require("PerformanceAnalytics",quietly=TRUE)) {
  getSymbols("SPY", from='2011-01-03')
  SPY.ret <- Return.calculate(SPY$SPY.Close)
  tmp <- merge(SPY.ret,ret1[[i]]$total,all=FALSE)
  charts.PerformanceSummary(cbind(tmp[,2],tmp[,1]),main=portfoliol.st[i],geometric=FALSE,
    wealth.index=TRUE)
}
}
ExcessReturn <- ret1[[1]]$total
for (i in 2:M) {
  ExcessReturn <- ExcessReturn+ret1[[i]]$total
  ExcessReturn <- ExcessReturn["2011-01-01::2011-06-30"]
}
ExcessReturn <- ExcessReturn/M

#zzPort<-getPortfolio(portfoliol.st[i])
#str(zzPort)
#zzAcct<-getAccount(account.st[i])
#str(zzAcct)
#str(pairStrat[[i]])
require(sandwich)
require(lmtest)
ExcessReturn.lm<-lm(as.vector(ExcessReturn) ~1)
ExcessReturn.lm.coefstest<-coefstest(ExcessReturn.lm,df=3,
  vcov=NeweyWest(ExcessReturn.lm,lag=4,prewhite=FALSE))
ExcessReturn.lm.coefstest

portefolje <-
  blotter:::.getByPortf(.blotter[.blotter[["accounts"]][1]],Attribute="Realized.PL",
  Dates=index(mktdata))
for (i in 2:M) {
  portefolje <-
  cbind(portefolje,blotter:::.getByPortf(.blotter[.blotter[["accounts"]][i]],Attribute="Realized.PL",
  Dates=index(mktdata)))
}

portefoljel <- (initEq+portefolje - initEq)/initEq
portefoljel <- portefoljel["2011-01-01::2011-06-30"]
x <- table(portefoljel,exclude=0)
x <- as.data.frame(x)
x <- na.omit(x)
posisjoner <-sum(x$Freq) #antall lukninger av posisjoner, ergo antall posisjoner åpnet.
y <- x$portefoljel
y <- as.numeric(levels(y)[y])
stortap <- min(y) #største tap
storgev <- max(y) #største gevinst
snitt <- mean(y) #gjennomsnittlig avkastning
zmin<- table(y[y<0])
zmin <-as.data.frame(zmin)
zpos<- table(y[y>0])
zpos <- as.data.frame(zpos)
zb <- sum(zmin$Freq)
zc <- sum(zpos$Freq)

win <- (zc/posisjoner)*100 #prosent gevinst
tap <- (zb/posisjoner)*100 #prosent tap
wlratio <- win/tap #win/loss ratio
handler <- posisjoner*4

portefolje3 <-

```

```

blotter:::.getByPortf(.blotter[.blotter[["accounts"]][1]],Attribute="Short.Value",
Dates=index(mktdata))
for (i in 2:M) {
  portefolje3 <-
  cbind(portefolje3,blotter:::.getByPortf(.blotter[.blotter[["accounts"]][2]],Attribute="Short.Value",
Dates=index(mktdata)))
}

portefolje3 <- portefolje3["2011-01-01::2011-06-30"]
portefolje3 <- table(portefolje3,exclude=0)
portefolje3 <- as.data.frame(portefolje3)
holding <- sum(portefolje3$Freq)/posisjoner

res <- cbind(M,posisjoner, handler, holding, stortap, storgev, win, tap, wlratio)
colnames(res) <- c("Antall Par", "Antall posisjoner", "Antall handler", "Gj.snitt
holdingperiode", "Største tap", "Største gevinst", "Gevinst %", "Tap %", "W/L ratio")

getSymbols('SPY', from='2011-01-03', to='2011-06-30')
getSymbols('^FVX', from='2011-01-03', to='2011-06-30')
Rf <- FVX$FVX.Close
Rb <- SPY$SPY.Close
Rb <-Return.calculate(Rb, method="compound")
return.cumulative <- Return.cumulative(ExcessReturn, geometric=FALSE)
return.ann <- Return.annualized(ExcessReturn, geometric=FALSE, scale=252)
sharpe <- SharpeRatio.annualized(ExcessReturn, geometric=FALSE, Rf=0.0018/252)
var <- VaR(ExcessReturn, method=c("historical"),portfolio_method=c("single"))
es <- ES(ExcessReturn, method=c("historical"),portfolio_method=c("single"))#conditionalVAR
sdev <- sd.annualized(ExcessReturn, scale=252)
maxdraw <- maxDrawdown(ExcessReturn, weights=NULL, geometric = FALSE, invert=TRUE)
table.CAPM(ExcessReturn, Rb, scale = 252, Rf = 0.0018/252, digits = 4)
capm <- CAPM.beta(ExcessReturn, Rb, Rf=0.0018/252)
charts.PerformanceSummary(cbind(ExcessReturn,Rb), geometric=FALSE, wealth.index=TRUE,
main="All stocks 5%")
table.Stats(ExcessReturn)
info <- InformationRatio(ExcessReturn, Rb, scale=252)

res2 <- cbind(return.cumulative, return.ann, sharpe, var, es, sdev,maxdraw, capm, info)
colnames(res2) <- c("Periode avkastning", "Årlig avkastning", "Sharpe-ratio (Årlig)", "VaR",
"Conditional VaR", "Standardavvik (Årlig)", "Max Drawdown", "Beta", "Information Ratio")
rownames(res2) <- NULL

return.spy <- Return.cumulative(Rb, geometric=FALSE)
ann.spy <- Return.annualized(Rb, geometric=FALSE, scale=252)
sharpe.spy <- SharpeRatio.annualized(Rb, geometric=FALSE, Rf=0.0018/252)
var.spy <- VaR(Rb, method=c("historical"),portfolio_method=c("single"))
es.spy <- ES(Rb, method=c("historical"),portfolio_method=c("single"))#conditionalVAR
StdDev(Rb)
sdev.spy <- sd.annualized(Rb, scale=252)
maxdraw.spy <- maxDrawdown(Rb, weights=NULL, geometric = FALSE, invert=TRUE)
capm.spy <- CAPM.beta(Rb, Rb, Rf=0.0018/252)

res3 <- ExcessReturn.lm.coefest

res4 <- cbind(return.spy, ann.spy, sharpe.spy, var.spy, es.spy, sdev.spy,maxdraw.spy,
capm.spy)
colnames(res4) <- c("Periode avkastning", "Årlig avkastning", "Sharpe-ratio (Årlig)", "VaR",
"Conditional VaR", "Standardavvik (Årlig)", "Max Drawdown", "Beta")
rownames(res3) <- NULL

con <- file("tradingstats.csv","w")
write.table(res, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res2, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res3, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res4, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
close(con)

dev.off(which = dev.cur())

```

Vedlegg 7 Finne antall par som faktisk handler.

```
port <- portefoljel
port <- na.omit(port)
colSums(port)
colSums(port) != 0
port <- port[, colSums(port) != 0]
antallpar <- ncol(port)

port1 <- port[,1]
for (i in 2:antallpar) {
  port1 <- port1+port[,i]
}
port <- port/antallpar
fullyinvested <- Return.cumulative(port, geometric=FALSE)
fullyinvested
```

Abstract

Pairs trading is a speculative investment strategy existing since the 1980s, originally exclusive to investment banks and hedge funds because of their economies of scale. The rapid development of ICT tools and internet based brokers has recently made the strategy a viable option for small scale investors.

The idea behind the strategy is to find two securities that move similarly over time. We can thus assume that they are in equilibrium. When the securities diverge from the equilibrium we go long in the relative underpriced security and go short in the overpriced. The idea is that the securities will eventually converge to the equilibrium and this will yield us a profit.

We have tried to follow the methodology of earlier studies of Gatev, Rouwenhorst & Goetzmann (2006) and Do & Faff (2009) to some degree. To start with we chose roughly 1500 stocks from NYSE and NASDAQ. We use a co-integration approach to find pairs and have two groups of securities. One group that was significant at 5 % level or better and one group at 10 % level. The indicator to trade is calculated using the mispricing expressed in standard deviations using a 50 day simple moving average.

The 5 % group made an annual return of 0,42 %, which is in line with the findings of Do & Faff (2009), the 10 % group did however perform better with an annual return of 7,17 %. Both portfolios had much lower risk metrics than the market portfolio, and the beta value indicated that they were market neutral.

Forord

Denne masteroppgaven er skrevet som en avsluttende studie i spesialiseringen Finansiering og Investering ved Handelshøgskolen i Bodø. Tidsrommet for studien strekker seg over hele vårsemesteret 2012 og oppgaven har et omfang på 30 studiepoeng. Temaet til oppgaven kunne velges fritt innenfor spesialiseringens grenser, med den hensikt at vi skal lære å praktisere teoretisk kunnskap.

Vi vil takke vår veileder, professor Frode Sættem ved Norges Handelshøyskole for god hjelp og oppmuntrende tilbakemeldinger underveis i arbeidet med oppgaven. Vi ønsker også å rette en stor takk til førsteamanuensis Berner Larsen ved Universitetet i Nordland for uvurderlig hjelp med programmering i R og statistiske utfordringer.

Avslutningsvis vil vi nevne at eventuelle feil og mangler i oppgaven er undertegnede alene ansvarlige for.

Handelshøgskolen i Bodø, 23.5.2012

Andreas Mikkelsen

Petter Størksen

Sammendrag

Pairs trading er en spekulativ handlestrategi som har eksistert siden 1980-tallet. Opprinnelig var det investeringsbanker og hedgefond på Wall Street som benyttet seg av strategien. Men utviklingen innen IT-løsninger og nettbaserte meglere har gjort at også privatpersoner kan benytte seg av denne. Utgangspunktet for pairs trading er å finne to aksjer som over tid beveger seg likt. Dette gjør at man antar at det finnes en likevekt mellom aksjene. Når aksjene avviker fra denne likevekten kjøper man den aksjen som relativt sett er underpriset, og selger den aksjen som er overpriset. Tanken er da at når aksjene går tilbake til likevekten, løser man inn posisjonene og innkasserer en hyggelig gevinst. For å definere likevekten kan man benytte seg av ulike mål. Eksempler på slike mål er korrelasjon eller kointegrasjon. Korrelasjon er et velkjent begrep, men ulempen er at den bare indikerer en kortsiktig likevekt. Kointegrasjon forutsetter derimot en langsiktig likevekt.

Vi har i vår studie forsøkt å følge metodikken til tidligere studier av Gatev et al. (2006) samt Do & Faff (2009) for å kunne sammenlikne vår studie med resultatene fra deres studier. Begge studiene har benyttet seg av kointegrasjon for å finne aksjer som beveger seg likt, og av den grunn besluttet vi å gjøre det samme.

Vi startet med et utvalg på i overkant av 1500 aksjer fordelt på fem supersektorer fra børsene New York Stock Exchange og NASDAQ. Vi testet bare aksjer innenfor undersektorer mot hverandre for kointegrasjon, fordi at det skulle være størst mulig sjanse for at parene var kointegrerte basert på at de var økonomiske substitutter. Vår framgangsmåte var først å teste alle mulige kombinasjoner av par innen en undersektor med en Johansen kointegrasjonstest. Deretter ble de parene som var statistisk signifikante testet på nytt for kointegrasjon med en Augmented Dickey Fuller (ADF) test og Phillips-Perron (PP) test. De parene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg signifikante på 5 % nivå i ADF og PP testene ble plassert i en gruppe med benevnelsen 5 %. De parene som var signifikante på minimum 10 % nivå i alle testene ble sortert i en gruppe vi kalte 10 %. De øvrige parene ble forkastet.

Videre har vi benyttet oss av en handelsindikator basert på 50 dagers glidende gjennomsnitt. Indikatoren uttrykker feilprisingen i standardavvik av spreaden mellom aksjene. Dersom aksjene divergerte to standardavvik fra likevekten åpnet vi en posisjon. For å gjøre det så realistisk som mulig benyttet vi en form for risikostyring ved å lukke posisjonene dersom

parene fortsetter å divergere. Dette ble gjort ved å programmere indikatoren til å løse inn posisjonene dersom parene divergerte ut over tre standardavvik fra likevekten. Vi lukket også posisjonene dersom parene konvergente tilbake til likevekten, eller når handelsperioden var over. Etter å ha testet par for kointegrasjon satt vi igjen med 161 par i gruppen 5 % i tillegg til 161 par i gruppen 10 %. Vi testet så handelsstrategien over en tidsperiode på seks måneder, fra 3. januar 2011 til 30. juni 2011. Vi lagde porteføljer basert på sektorer, og i tillegg en portefølje med alle parene i en gruppe.

Vi fant at handelsstrategien basert på et 50 dagers glidende gjennomsnitt er veldig handelsintensivt. I gruppen 5 % gjennomførte vi 1772 handler over en seks måneders periode, og tilsvarende 1816 handler for 10 % gruppen. Samtidig viste det seg at parene hadde relativt kort gjennomsnittlig holdingperiode. Holdingperioden for gruppen 5 % var ca. 20 handelsdager. For gruppen 10 % var den noe høyere med i overkant av 23 handelsdager. Dette impliserer at basert på handelsindikatoren brukt i vår studie oppnår man en handelsstrategi med kort tidshorison.

For begge gruppene oppnådde vi en markedsnøytral strategi ettersom beta var nær null uavhengig av hvor mange par vi hadde i porteføljene. Videre viste det seg at vi oppnådde en diversifiseringsgevinst ettersom risiko uttrykt i både standardavvik og Value at Risk ble lavere når flere par ble introdusert i en portefølje. Strategien var gjennomgående mindre volatil sammenliknet med å holde markedsporteføljen.

Strategien oppnådde en årlig avkastning på 0,42 % for gruppen 5 %. Dette var som forventet ettersom Do & Faff (2009) finner en avtagende avkastning for strategien; en avkastning ikke signifikant annerledes enn null i perioden 2003 - 2008. For gruppen 10 % ble den årlige avkastningen 7,17 %, nesten like bra som S&P 500 oppnådde i samme periode med lavere risiko. Vi kan dessverre ikke dra noen konklusjoner på hva grunnen til den gode avkastningen kan være ettersom vi bare har en handelsperiode i vår studie.

Som nevnt er pairs trading en intensiv handelsstrategi, noe som naturlig nok medfører høye transaksjonskostnader. Det virker ikke som at strategien er robust for transaksjonskostnadene som ble estimert til å være i gjennomsnitt 453,6 basispoeng per par i perioden. Vi finner at ved å øke lengden på glidende gjennomsnitt til 150 dager i indikatoren reduserer dette transaksjonskostnadene for strategien med 44 %. Det må kunne sies å være en substansiell reduksjon og dermed ha potensial til å være forskjellen mellom tap og gevinst på lang sikt.

Innholdsfortegnelse

Abstract	II
Forord	III
Sammendrag	IV
Innholdsfortegnelse	VI
Figurliste	IX
Tabelliste	IX
Vedleggsliste	IX
1. Innledning.....	1
1.1 Aktualisering.....	1
1.2 Bakgrunn for oppgaven.....	1
1.3 Avgrensninger	2
1.4 Problemstilling.....	2
1.5 Oppgavens oppbygning.....	3
2. Markedseffisiens	4
2.1 Markedseffisienshypotesen.....	4
2.1.1 Svak form for markedseffisiens.....	6
2.1.2 Semi sterk form for markedseffisiens	7
2.1.3 Sterk form for markedseffisiens.....	8
2.2 Kritikk av hypotesen	8
2.3 Oppsummering.....	10
3. Arbitrasje	11
3.1 Hva er arbitrasje?	11
3.1.1 Ren arbitrasje	11
3.1.2 Relativ arbitrasje	12
3.1.3 Spekulativ arbitrasje.....	12
3.1.4 Aktører	12
3.2 Risiko	13
3.2.1 Effektueringsrisiko.....	13
3.2.2 Konvergeringsrisiko og likviditet	14
3.2.3 Short-squeeze.....	14
3.2.4 Kredittrisiko	15
3.2.5 Long Term Capital Management.....	15

3.3 Implikasjoner av EMH og arbitrasje	16
3.4 Oppsummering.....	16
4. Pairs trading	17
4.1 Hva er pairs trading?	17
4.1.1 Stasjonære og ikke-stasjonære variabler.....	18
4.1.2 Kointegrasjon	20
4.1.3 Noen avklaringer	21
4.1.3.1 Hvorfor metoden hevdes å fungere.....	22
4.2 Markedsnøytral strategi.....	22
4.2.1 Pairs trading med enkel spread som mål.....	22
4.2.2 Pairs trading ved bruk av standardavvik.....	24
4.2.3 Standardavvik eller spread ratio.....	25
4.2.4 Metode ved bruk av beta.....	26
4.3 Pairs trading i praksis	27
4.3.1 Strategiens profitabilitet	27
4.3.2 Fundamentale forhold.....	28
4.3.3 Money management og risiko ved pairs trading	29
4.4 Oppsummering.....	30
5. Metode	31
5.1 Analyseverktøy.....	31
5.2 Utvelgelse av aksjer, sektor og tidsperioder	31
5.3 Utvelgelse av par	32
5.4 Handleregler	34
5.5 Beregninger i analysen	37
Avkastning	37
Risikofri rente	38
Sharpe Ratio (Årlig)	38
Information Ratio	39
Value at Risk (VaR)	39
Expected shortfall.....	40
Max drawdown	40
Beta	40
Standardavvik (Årlig)	40
5.6 Diskusjon, valg av kriterier	41

5.6.1 Indikator, valg av lengde på glidende gjennomsnitt	41
5.6.2 Valg av handleregel	42
5.6.3 Kointegrasjon	43
5.6.4 Risikomål	44
5.6.6 Ordretype	44
5.7 Mulige feilkilder.....	45
5.7.1 Survivorship bias	45
5.7.2 Data snooping bias	46
5.7.3 Look ahead bias.....	46
5.8 Oppsummering.....	47
6. Analyse	48
6.1 Sektorer	49
6.1.1 Basic Materials	49
6.1.2 Conglomerates	51
6.1.3 Consumer goods.....	53
6.1.4 Financials	55
6.1.5 Technology	57
6.1.6 Alle parene	59
6.2 Transaksjonskostnader.....	64
6.2.1 Sensitivitetsanalyse transaksjonskostnader.....	65
6.3 Oppsummering.....	67
7. Avslutning.....	68
7.1 Oppsummering og konklusjon	68
7.2 Kritikk av studiet.....	70
7.3 Forslag til videre forskning	71
8. Litteraturliste.....	72

Figurliste

Figur 1 Oppgavens oppbygning.....	3
Figur 2 Korrelasjon mellom FRO og FUR	18
Figur 3 Tidsserier	19
Figur 4 Autokorrelasjon.....	20
Figur 5 The Sine Wave (Pole, 2007, p. 19).....	23
Figur 6 The Popcorn Process (Pole, 2007, p. 19).....	23
Figur 7 Chart (Grøtte, 2002, p. 513)	26
Figur 8 Delta	35
Figur 9 Sammendrag BVN.HMY.....	36
Figur 10 Oversikt aksjekurser	37
Figur 11 Chart, diverse SMA.....	41
Figur 12 Forskjellige SD	43
Figur 13 Graf, Basic Materials	49
Figur 14 Graf, Conglomerates	51
Figur 15 Graf, Consumer goods.....	53
Figur 16 Graf, Financials	55
Figur 17 Graf, Technology	57
Figur 18 Graf, alle parene.....	59
Figur 19 Fordeling.....	63
Figur 20 Graf, Financials SMA150.....	66

Tabelliste

Tabell 1 Resultater, forskjellige SMA.....	41
Tabell 2 Forskjellige SD.....	43
Tabell 3 Basic Materials.....	49
Tabell 4 Conglomerates.....	51
Tabell 5 Consumer goods	53
Tabell 6 Financials	55
Tabell 7 Technology.....	57
Tabell 8 Alle parene.....	59
Tabell 9 Financials SMA150.....	66

Vedleggsliste

Vedlegg 1 Script for å få data inn i R. Eksempel sektoren Agricultural i Basic Materials.....	A
Vedlegg 2 Johansentest av en hel sektor. Eksempel Property and Casualty Insurance.	B
Vedlegg 3 Eksempel output Johansentest.	B
Vedlegg 4 ADF og PP-test av spredningen mellom to aksjer. Eksempel Aksjeparet HTH.AMSF.....	C
Vedlegg 5 Eksempel output ADF og PP-test.....	C
Vedlegg 6 Script for testing av handelsstrategi. Eksempel fra alle 5 %.....	D
Vedlegg 7 Finne antall par som faktisk handler.	J

1. Innledning

Innledningsvis vil vi gå igjennom bakgrunnen for valg av tema og problemstilling. Vi vil deretter forklare hvorfor dette er et dagsaktuelt tema. Til slutt gjør vi noen avgrensninger for studien, samt går gjennom oppgavens oppbygning.

1.1 Aktualisering

Innen akademia har det vært stor interesse for forskning på aksjemarkedet. Studier av Keim (1983), Lehmann (1990), Lakonishok (1994) og Gatev et al. (2006) omhandler alle ulike fenomener innenfor temaet. I norske aviser¹ har Setterberg (2007) denne våren fått oppmerksomhet med sin studie som påviser at det finnes en momentumeffekt på Stockholmsbørsen.

Blant praktikere har det også blitt brukt store ressurser på utvikling av metoder for å oppnå høy avkastning på deres investeringer. Tradisjonelt har dette vært forbeholdt hedgefond og institusjonelle investorer grunnet deres stordriftsfordeler. Utviklingen innen IT-løsninger og nettbaserte meglere har imidlertid ført til at også privatpersoner har tilgang til markedet til en lav kostnad uten å gå via mellommenn.

Fenomenet pairs trading har blitt benyttet siden 1980-tallet av investeringsbanken Morgan Stanley med stor suksess.

1.2 Bakgrunn for oppgaven

Konsistent positiv avkastning i aksjemarkedet med lavest mulig risiko er svært vanskelig å oppnå. Årsaken til dette skriver seg fra en rekke forhold, men først og fremst kommer det av at det stilles store krav til riktig kunnskap, dedikasjon og erfaring. Et av kjennetegnene på dem som lykkes er deres evne til å tenke i sannsynligheter og hele tiden være trygg på at metoden de til enhver tid bruker har vist seg å være historisk profitabel. Ved å teste strategier på historiske data kan man derfor få kunnskap om hvordan man bør bevege seg i markedet for å redusere risiko og oppnå målet om konsekvent positiv avkastning over tid. Pairs trading er en strategi som over lang tid har vist seg som en fornuftig tilnærming til aksjemarkedet på bakgrunn av eliminert systematisk risiko samt stabil positiv avkastning. Målet i seg selv er derfor ikke å slå markedet til en hver tid, men å unngå den store variansen som man kan oppleve med passiv forvaltning.

¹ <http://www.dn.no/forsiden/borsMarked/article2379758.ece> (Retrieved 13.05.2012)

På bakgrunn av denne tankegangen ble vi nysgjerrige på om pairs trading innehar kvaliteter som lav risiko og positiv avkastning, eller om strategien i årenes løp har mistet disse egenskapene. Et annet poeng er at tilnærmingen til statistisk arbitrasje og algoritmehandel i stor grad er to sider av samme sak, der strategien kan automatiseres og all handel kan utføres mekanisk. Et siste poeng som øker betydningen av studien er at temaet relativt sett er lite utforsket i akademia når man sammenligner med mer kjente og tradisjonelle tilnærminger til aksjemarkedet innen teknisk og fundamental analyse. Det vil derfor være meget interessant å sammenligne nye funn med eldre studier. Summen av dette er bakgrunnen for at vi ønsket å gjennomføre denne studien.

1.3 Avgrensninger

Etttersom vi har hatt begrenset med tid til å gjennomføre studien så vi oss nødt til å avgrense den noe. På grunn av arbeidsmengden hadde vi bare mulighet til å teste par i fem sektorer. Vi mener likevel at vi oppnådde et bredt utvalg ettersom vi hadde prisdata fra i overkant av 1500 aksjer som utgangspunkt for å finne par. Det viste seg at det mest tidkrevende var å teste mulige par for kointegrasjon, da vi ble nødt til å gjøre dette manuelt. Vi har brukt de samme tidsrammer for oppbygging av par og påfølgende handelsperiode som tidligere studier. Dette førte til at vi bare fikk testet strategien over seks måneder.

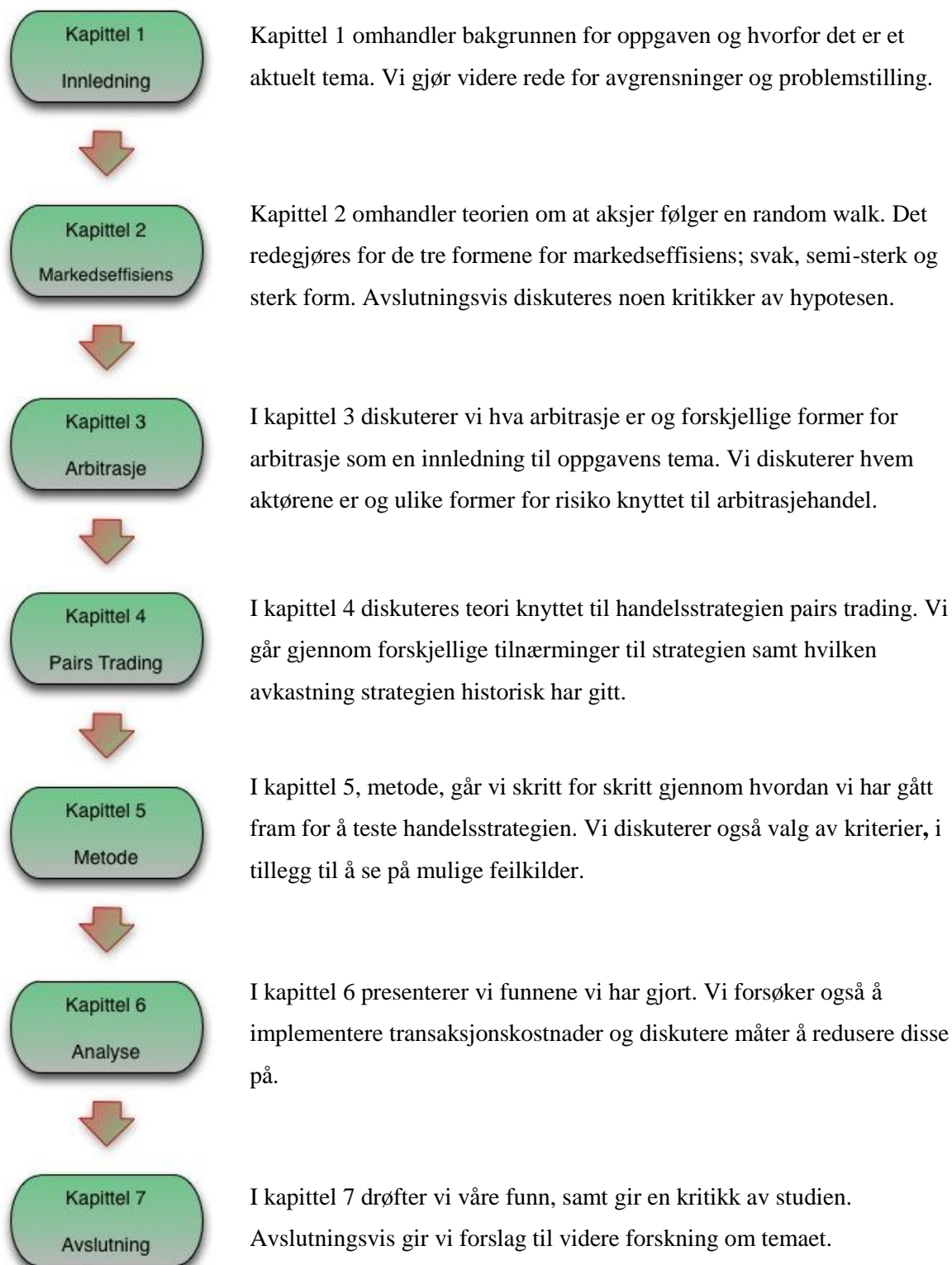
1.4 Problemstilling

Bakgrunnen for oppgaven er at vi ønsket å se på om "vår" pairs trading strategi oppnådde positiv avkastning kombinert med lav risiko justert for transaksjonskostnader. Det var videre naturlig å sammenligne våre resultater med tidligere studier. Testene gjorde vi på historiske data ved hjelp av statistikkprogrammet R.

Problemstillingen er derfor som følger:

"Gir handelsstrategien pairs trading meravkastning kombinert med lav risiko eller er avkastningen avtagende jamfør tidligere forskning?"

1.5 Oppgavens oppbygning



Figur 1 Oppgavens oppbygning

2. Markedseffisiens

2.1 Markedseffisienshypotesen

Etter at IT-teknologi fikk større utbredelse på 1950-tallet ble store mengder data tilgjengelig for forskere til bruk i empiriske studier. Dette inkluderte også studier av prisutviklingen på aksjer. Etter at dataene ble gjennomgått antydte flere studier at prisutviklingen på aksjer fulgte en random-walk modell. Dette stemte godt overens med matematikeren Bacheliers teori om hvordan priser utvikler seg fra 1900-tallet.

"His "fundamental principle" for the behavior of prices was that speculation should be a "fair game"; in particular, the expected profits to the speculator should be zero. With the benefit of the modern theory of stochastic processes, we know now that the process implied by this fundamental principle is a martingale" (Fama, 1970, p. 389).

I følge Björk (1998, p. 34) er egenskapene til en martingale slik at morgendagens forventede verdi er lik dagens observerte verdi, det er ingen drift – trinnvis utvikling i prisen. Fama (1970) beskriver fair game modellen formelt på denne måten, og vi vil kort kommentere den for å få fram noen poenger:

$$E(\tilde{p}_{j,t+1}|\Phi_t) = [1 + E(\tilde{r}_{j,t+1}|\Phi_t)]p_{jt}, \quad (1)$$

E er forventningsverdien. Prisen på aktivum j på tidspunkt t er angitt som p_{jt} . $r_{j,t+1}$ er periodeavkastningen til aktivum j. Symbolet Φ_t angir at all tilgjengelig informasjon på tidspunkt t er bakt inn i prisen. Tødlene over p og r angir at disse er tilfeldige variabler på tidspunkt t.

$$x_{j,t+1} = p_{j,t+1} - E(p_{j,t+1}|\Phi_t) \quad (2)$$

$$E(\tilde{x}_{j,t+1}|\Phi_t) = 0 \quad (3)$$

Likning (2) og (3) angir derfor at prosessen $x_{j,t+1}$ er et fair game med hensyn til informasjonen (Φ) tilgjengelig på tidspunkt t. Ettersom den forventete avkastningen er betinget informasjonen tilgjengelig på tidspunktet t, gjør dette at informasjonen allerede er bakt inn i prisen på tidspunkt t. Dette gjør at man ikke kan oppnå meravkastning utover forventet avkastning, med handelsstrategier basert på Φ .

Fama (1970) mente at antakelsen om at den observerte prisen på et aktivum fullt ut reflekterer all tilgjengelig informasjon, impliserte at suksessive endringer i prisen er

uavhengige. I tillegg var det en vanlig forutsetning at endringene hadde identisk fordeling. Dette førte til en utvidelse av fair game modellen og resulterte i random walk modellen i likning (4).

$$f(r_{j,t+1}|\Phi_t) = f(r_{j,t+1}) \quad (4)$$

Hvis man videre forutsetter at forventningsverdien til et aktivum er konstant over tid får vi likning (5).

$$E(r_{j,t+1}|\Phi_t) = E(r_{j,t+1}) \quad (5)$$

Man kan forstå av uttrykket at gjennomsnittet av distribusjonen til $r_{j,t+1}$ er uavhengig av informasjonen tilgjengelig på tidspunkt t . I tillegg er hele distribusjonen uavhengig av informasjonen (Φ) tilgjengelig på tidspunkt t .

Fama (1970) mener at man bør se på random walk modellen som en utvidelse av fair game modellen. Forskjellen mellom de to er at den førstnevnte tar hensyn til den økonomiske virkeligheten, slik som investorers preferanser og prosesser som gir ny informasjon bidrar til likevekter hvor avkastningsfordelinger repeterer seg over tid. Vi anser det som tilstrekkelig å bruke random walk og fair game som synonymer. Dette vi vil gjøre i resten av oppgaven.

Basert på egen og tidligere forskning publiserte økonomen Fama (1970) teorien om effisiente markeder eller markedseffisienshypotesen (EMH).

I følge markedseffisienshypotesen er alle aktiva korrekt priset i et velfungerende marked. Det sentrale er at prisene fullt ut reflekterer tilgjengelig informasjon. Den eneste grunnen til at man får høyere avkastning på enkelte aktivaklasser kontra andre forklares med høyere risiko og dermed en risikopremie for å holde aktiva med høyere risiko (Fama, 1970).

Ettersom markedseffisienshypotesen er nettopp en hypotese er det derfor ikke overraskende at den som andre økonomiske teorier forsøker å gi et tilstrekkelig, men dog forenklet bilde av virkeligheten. Følgende forutsetninger for hypotesen er gjeldende:

- (i) *there are no transactions costs in trading securities,*
- (ii) *all available information is costlessly available to all market participants, and*
- (iii) *all agree on the implications of current information for the current price and distributions of future prices of each security. In such a market, the current price*

of a security obviously "fully reflects" all available information (Fama, 1970, p. 387).

Selv om dette er ganske urealistiske forutsetninger for hvordan et marked fungerer mener Fama (1970) at dette er tilfredsstillende forutsetninger. Han argumenterer for at selv om det er høye transaksjonskostnader, noe som fører til færre handler og mindre likviditet, vil dette ikke være likestilt med å si at prisene ikke avspeiler all tilgjengelig informasjon. Med hensyn til forutsetningen om at alle markedsaktørene har tilgang til all markedsinformasjon, mener han at det er tilstrekkelig dersom majoriteten av aktørene opptrer rasjonelt.

Vi vil videre gå inn på de tre forskjellige formene for markedseffisiens som beskrevet av Fama (1970) før vi avslutter med noen kritikker av hypotesen.

2.1.1 Svak form for markedseffisiens

Den svake formen for markedseffisiens fokuserer på historiske data som pris og volum. Det vil si at man ikke vil kunne finne under- eller overprisede aksjer basert på historiske data fordi prisen på en aksje allerede reflekterer all tilgjengelig informasjon.

Å få tilgang til denne type data påfører ikke en investor store kostnader i verken ressursbruk eller penger. I følge Bodie et al. (2009) forutsetter man derfor at alle investorer har tilgang til dette, og dersom det dermed fantes muligheter for å utnytte mønster i aksjeprisingen ville denne umiddelbart fått en ny likevekt som følge av tilbud og etterspørsel. I et effisient marked selges alle aktiva til en rettferdig pris.

"Investors get exactly what they pay for when they buy securities, and firms receive exactly what their stocks and bonds are worth when they sell them" (Ross, 2007, p. 357).

I følge Bodie et al. (2009) finnes det derfor ingen måte å forutsi hvilken retning aksjene som er listet på verdens børser vil bevege seg basert på historisk utvikling. De kaller dette for en "random walk". Med dette mener de at pris på forskjellige aktiva kun beveger seg som følge av ny informasjon i markedet. Denne informasjonen er i seg selv uforutsigbar, slik at aksjeprisene umiddelbart stiger ved gode nyheter, og tilsvarende synker ved dårlige nyheter, uavhengig av historisk utvikling.

"The current price of a security "fully reflects" available information was assumed to imply that successive price changes (or more usually, successive one-period returns) are independent" (Fama, 1970, p. 386).

Siden historiske data ikke har noen betydning for framtidig pris vil dette si at man ikke vil kunne oppnå høyere avkastning på å benytte seg av teknisk analyse ifølge hypotesen.

Allikevel bruker fondsforvaltere store summer hvert år på aktiv forvaltning, hvor de leter etter f.eks. underprisede aksjer. De er villige til å bruke betydelige beløp på dette ettersom de bare trenger å få en marginal økning i avkastningen for å øke pengeavkastningen betydelig. I følge EMH skulle det ikke være noen grunn for forvaltere å bruke penger på dette siden aksjene allerede er riktig priset. Det argumenteres for at denne praksisen fører til høyere konkurranse, som i seg selv er med på å gjøre markedene mer effisiente. Fama mener selv at dette ikke motbeviser EMH, men at markedene derimot er tilstrekkelig effisiente.

Spørsmålet er derfor om spekulanter kan finne trender i historiske data som kan gi dem unormal høy avkastning. En måte å identifisere trender i aksjepriser er ved å se på seriekorrelasjonen, som angir tendenser for om avkastningen i aksjepriser kan relateres til tidligere avkastninger. Studier på dette finner en liten sammenheng på ukesbasis, men ikke klart nok til å hevde at det er grunnlag for en god strategi (Bodie et al. 2009).

Andre studier viser at aksjer som enten stiger mye eller faller mye har en tendens til å reversere i etterkant. Et eksempel på dette er en studie som tok for seg de 35 beste og de 35 dårligste aksjene over en fem års periode. Resultatet viste at over de neste tre årene presterte de 35 "dårlige" aksjene gjennomsnittlig 25 % bedre enn de 35 beste. Dette funnet konkluderer med at aksjer overreagerer på relevante nyheter for så å reversere når overreaksjonen er oppdaget (Bodie et al. 2009).

2.1.2 Semi sterk form for markedseffisiens

I den semi sterke formen for markedseffisiens forutsetter en også at all tilgjengelig informasjon om et firma er bakt inn i tilhørende aksjepris. I tillegg til informasjonen man benytter seg av i svak form tas det også hensyn til fundamentale firmaspesifikke data. Dette kan være regnskapspraksis, sammensetning av ledergruppe og immaterielle eiendeler m.m. Siden dette er informasjon som er offentlig tilgjengelig vil dette reflekteres i aksjeprisen, i følge hypotesen.

Fama (1970) konkluderer med at mange typer ny informasjon i markedet underbygger hypotesen. For og støtte hypotesen ser han på prisutviklingen etter at ny informasjon som f.eks. aksjesplitter, kvartalsrapporter og annet har blitt offentliggjort, for så å se hvordan markedet reagerer ved hjelp av event-studier.

“The available semi-strong form evidence on the effect of various sorts of public announcements on common stock returns is all consistent with the efficient markets model ” (Fama, 1970, p. 409).

Overraskende nok viser det seg at enkle fundamentale mål som PE ratio eller børsverdi synes å forutsette unormal risikojustert avkastning. En annen strategi som å kjøpe aksjer etter at de har lagt fram et positivt resultat virker å gi unormal risikojustert avkastning. Dette er i stor kontrast til semi sterk form for EHM som hevder at dette ikke er mulig om markedet er effisient (Bodie et al. 2009).

2.1.3 Sterk form for markedseffisiens

Her forutsettes det at investorer har tilgang til all informasjon som i de to andre formene. I tillegg har de også tilgang til innsideinformasjon om selskapene, og dette skal dermed reflekteres i markedsprisene. Dette er den mest ekstreme formen for markedseffisiens. Det er derfor i følge Bodie et al. (2009) laget reguleringer i markedene som skal forhindre innsidere i å oppnå ekstraordinær avkastning ved handel med innsideinformasjon. De mener derfor at man ikke anser markedet for å ha en sterk form for markedseffisiens.

Selv om man skulle tro at innsidere hadde en fordel med handel i eget selskap, viser en studie av Seyhun i Bodie et al. (2009) at man ikke oppnår unormal avkastning fratrukket transaksjonskostnader ved å følge offentlig kjente innsidekjøp.

I følge Fama (1970) kan sterk form for EHM best sees på som en referanse for å fastslå avvik fra markedseffisiens. Eksempel på slike avvik er i følge Fama (1970) spesialisten på en børs, samt innsidere. Spesialisten har monopol ved å kjenne hele ordreboken, mens innsidere ofte har monopol på informasjon om et selskap. Disse tilfellene er i følge Fama (1970) avvik som kan generere unormalt høy avkastning. Fama (1970) konkluderer videre med at alle andre (unntatt spesialisten og innsidere) må ta til takke med et sterkt effisient marked så lenge det motsatte ikke er bevist.

2.2 Kritikk av hypotesen

Markedseffisienshypotesen har mottatt kritikk, og da særlig fra forskere innen fagfeltet atferdsfinans. Bakgrunnen for kritikken er at man har observert hendelser med unormal avkastning som ikke kan forklares med bakgrunn i modeller, slik som prissettingsmodeller som kapitalverdimodellen og dividendemodellen. Studier har også funnet tilfeller hvor det har vært mulig å forutsi hvilken retning prisene utvikler seg fra en dag til neste, ved hjelp av seriekorrelasjon. Vi vil gjøre rede for noen av de viktigste kritikkene av hypotesen.

Lehmann (1990) viser i sin studie at det er grader av momentum i aksjeprisene på kort sikt. Han mener at markedet overreagerer på nyheter, slik at aksjeprisene stiger for mye ved gode nyheter og tilsvarende faller for mye ved dårlige nyheter. Forklaringen på dette mener han er flokkmentalitet (fads), og at det derfor finnes arbitrasjemuligheter som følge av mean reversal i påfølgende periode. Dersom markedseffisienshypotesen holdt skulle det ikke være muligheter for dette siden alle aksjene er riktig priset. Også andre forskere har i sine studier påvist at det går an å finne pristrender på kort sikt:

"In an efficient market, return predictability from past information should be short-lived and minimal. Given the evidence that such predictability does exist in the short run, understanding its time variation and its relation to other financial market attributes, such as liquidity, are of fundamental importance" (Chordia, Roll, & Subrahmanyam, 2008, p. 266).

Det har også blitt sett på unormal avkastning i enkelte måneder og dager. I følge Malkiel (2003) ble det observert at måneden januar ga unormal høy avkastning for en stund, i tillegg til at mandager også skulle gi meravkastning. Problemet med så banale handlingsregler er at når de blir kjent vil alle investorer prøve å oppnå denne avkastningen slik at handlingsregelen dreper seg selv. Man kan derfor si at for å opprettholde et effisient marked, må markedsaktører jakte etter godbitene slik at de forsvinner like raskt som de oppstår. Dette betegnes som effisiensparadokset.

Lakonishok, et al. (1994) finner i sin forskning at såkalte verdiaksjer gjør det bedre enn vekstaksjer. De viser at ved å benytte seg av motsatt strategi enn den som majoriteten av investorene bruker, er det mulig å oppnå ekstra avkastning, ettersom man kjøper underprisede verdiaksjer og selger overprisede vekstaksjer. Årsaken til at vekstaksjene presterer bedre, er i følge dem grunnet i at investorer er alt for optimistiske i sitt syn på fortsatt framtidig vekst. I tillegg styrer institusjonelle investorer muligens unna verdiaksjer av karrieremessige grunner. Malkiel (2003) mener derimot at årsakene til dette er at modellene som er brukt i disse studiene ikke fanger godt nok opp all risiko, slik at den høyere avkastningen kan forklares med en risikopremie, og at trefaktormodellen av Fama og French er bedre egnet til å forklare dette.

Det kan også være slik at størrelsen av et firma har utslag på avkastningen aksjen oppnår.

"Since 1926, small-company stocks in the United States have produced annual rates of return over 1 percentage point larger than the returns from large stocks." (Keim (1983) in Malkiel, 2003, p. 68)

Hvis man bruker KVM som forklaringsmodell gir dette små firma meravkastning ettersom selskapene gir høyere avkastning enn betaverdien skulle tilsi.

"Average returns on small (low ME) stocks are too high given their β estimates, and average returns on large stocks are too low." (Fama & French, 1992, p. 427)

Det kan være flere forklaringer på dette fenomenet. På den ene siden mener Bodie et al. (2009) at institusjonelle investorer bruker mindre ressurser på informasjonsinnhentning om små firma ettersom disse aksjene ofte er mindre likvide enn i større firma. Både mangelen på informasjon og dårlig likviditet kan være med å forklare den høyere avkastningen. På den andre siden mener Fama og French at det er feil med modellen, og at størrelse og forholdet mellom bokført EK / markedsbasert EK gir bedre estimater på risiko.

Til slutt viser det seg at økonomiske kriser kan gi en god pekepinn på at markedene ikke opptrer rasjonelt og effisient. Eksempler er børskrakket i 1987 hvor børsene falt med en tredjedel, dot.com boblen i begynnelsen av årtusenet og finanskrisen i 2008. Særlig etter finanskrisen har markedseffisienshypotesen fått stor kritikk. Enkelte av kritikerne går så langt som å si at hypotesen var skyld i krisen:

"They were blinded by an irrational faith in a discredited EMH and failed to see the bubble in asset prices and to give due warning of its collapse. The irony is that the strong implication of this hypothesis is that nobody, no practitioner, no academic and no regulator had the ability to foresee the collapse of this most recent bubble" (Brown, 2011, p. 93).

2.3 Oppsummering

I kapittelet om markedseffisiens har vi gjennomgått de sentrale sidene ved markedseffisienshypotesen. Først tok vi for oss utgangspunktet for hypotesen der man hevdet at aksjekursene kun beveger seg som følge av ny informasjon i markedet slik at forventningsverdien for morgendagens pris er lik dagens. Denne modellen kalles random walk, utledet av fair game modellen.

Videre diskuterte vi de ulike formene for markedseffisiens som begrenser seg til svak, halvsterk og sterk markedseffisiens med etterfølgende tester som er utført på de respektive formene. Til slutt presenterte vi kritikk av hypotesen på bakgrunn av at enkelte studier hevder å oppnå unormal avkastning som bryter med fundamentene i hypotesen.

3. Arbitrasje

3.1 Hva er arbitrasje?

Sentralt i økonomisk teori er det faktum at like goder må være priset likt.

“Therefore, any two portfolios that provide the same pattern of wealth must be worth the same amount.” (Varian, 1987, p. 59)

Dette innebærer at både like aktiva og prosjekter som gir samme konstantstrøm skal ha samme markedsverdi, noe som kalles ”The Law of One Price”. I et effisient marked skal det ikke være muligheter for å oppnå en gratis gevinst. Man skal ikke kunne oppnå avkastning utover risikofri rente uten å påta seg risiko. I motsetning til dette handler arbitrasje veldig forenklet om å utnytte prisforskjellene mellom to eller flere markeder, der formålet er risikofri avkastning.

“Basically, the no arbitrage condition must rule out "free lunches"-configurations of prices such that an individual can get something for nothing” (Varian, 1987, p. 59).

I følge Reverre (2001) har arbitrasje ytterligere en karakteristikk som skiller arbitrasje fra vanlig handel, nemlig det faktum at man tar motstående posisjoner i flere aktiva samtidig i motsetning til en investeringsstrategi hvor man vanligvis bare tar en posisjon. Følgelig er det slik at man må ta hensyn til kostnadene ved å innta en arbitrasjeposisjon, slik at en mulighet bare er lønnsom dersom gevinsten av feilprisingen er større enn summen av kostnadene knyttet til å innta posisjonen.

Vi skiller i følge Dubil (2004) mellom forskjellige typer arbitrasje; ren arbitrasje, relativ arbitrasje (risk/near) og rene spekulative strategier.

3.1.1 Ren arbitrasje

“A pure arbitrage opportunity (PAO) is a zero-cost trading strategy that offers the possibility of a gain with no possibility of a loss” (Bondarenko, 2003, p. 875).

Det er denne typen strategi Varian (1987) kaller en gratis lunsj. I følge Damodaran (2003) er man avhengig av to like aktiva for å få en ren arbitrasjemulighet, noe som er vanskelig å finne i dagens markeder. Det er naturligvis sjelden man finner slike muligheter, da et velfungerende marked stort sett vil være effisient og ikke la prisforskjeller i relativt like aktiva oppstå, noe som utelukker arbitrasjemuligheter. Dersom det allikevel skulle oppstå en

mulighet vil det grunnet de mange markedsaktørene raskt opprettes en ny likevekt som gjør at arbitrasjemuligheten forsvinner.

Som eksempel på rene arbitrasjemuligheter kan man dra paralleller til sportsbetting. De siste årene har det vært en stor vekst i denne industrien med store markeder på internett og tilsvarende mange aktører. Dette markedet tilbyr spill hvor oddsen er i kontinuerlig endring, derfor oppstår det noen ganger arbitrasjemuligheter. Man kan se for seg et spill med to utfall, og på grunn av feilprising har de flittige vært i stand til å oppnå risikofri gevinst uansett utfall i spillet. Et annet eksempel kan være en aksje som er listet på flere børser og har ulik pris. Man kan da kjøpe den billig på en børs og selge den dyrt på en annen børs uten risiko, forutsatt at det er likvide aksjer. Arbitrasjemuligheter i feilprisede valutakryss er også en mulighet. Naturligvis kan det ikke være slik at det ligger rene arbitrasjemuligheter billedlig rundt hvert hjørne. Hadde det vært slik hadde alle vært milliardærer nærmest uten innsats. I følge Dubil (2004) fokuserer derfor de fleste arbitrasjører på relativ arbitrasje.

3.1.2 Relativ arbitrasje

I følge Damodaran (2003) skiller relativ arbitrasje seg fra ren arbitrasje med at arbitrasjøren inntar posisjoner i veldig like aktiva som er priset forskjellig, eller i like aktiva som er relativt feilpriset. Dette medfører at han kvitter seg med markedsrisiko, men lar sekundærrisiko være udekket. Det er dermed ikke en gratis lunsj som i eksempelet med ren arbitrasje, ettersom arbitrasjøren fremdeles er eksponert for sekundærrisiko. Dette skyldes at det aldri er noen garanti for at posisjonene skal konvergere. Eksempler på relativ arbitrasje kan være å innta motsatte posisjoner i to langsiktige obligasjoner med ulik løpetid. Markedet vil forvente at rentesatsene på obligasjonene skal konvergere, mens man som arbitrasjør spekulerer i at de ikke konvergerer. I tillegg finnes det en rekke muligheter når firma endrer kapitalstruktur, som ved f.eks. fusjoner, tilbakekjøp av aksjer o.l.

3.1.3 Spekulativ arbitrasje

Det finnes i tillegg arbitrasjestrategier som er mer spekulative. Eksempler på dette er pairs trading som vi skriver mer om i neste kapittel.

3.1.4 Aktører

I følge Schleifer (1997) fungerer ikke modellen med kapitalfri arbitrasje i praksis. Ettersom det er små marginer som det handles på trenger man store volum for å oppnå en akseptabel kroneavkastning, noe som innebærer at aktørene som driver med arbitrasje trenger store kapitalbeløp både for å utføre handler og til å dekke tap. Siden det er de færreste som har

tilgang på så store kapitalbeløp betyr dette at det er få og høyt spesialiserte investorer som bruker sin kunnskap sammen med egen og/eller andres kapital til arbitrasjehandel.

Et annet interessant fenomen oppstår på grunnlag av bruken av fremmedkapital:

Investors may rationally allocate money based on past returns of arbitrageurs” (A. Shleifer, Vishny, Robert W., 1997, p. 37).

Dette forklarer at de fremmedkapitalfinansierte aktørene er avhengige av historiske gode resultater for å få tilgang på kapitalen de behøver. Shleifer (1997) viser videre at dette kan føre til tilfeller hvor de ikke fullt ut utnytter mulighetene i markedet. For eksempel kan feilprisingen som aktørene har spilt mot bli enda verre og føre til at de må lukke posisjonen som følge av press fra investorene. En annen grunn kan være at de er forsiktigere i sine handler for å opprettholde stabile prestasjoner, noe han mener kan begrense effisiensen i markedet. Arbitrasjørene oppfyller dermed ikke sin rolle i markedet med å tilføre likviditet når det trengs mest.

3.2 Risiko

3.2.1 Effektueringsrisiko

Effektueringsrisiko er knyttet til både inngåelse av posisjoner og likvidering av disse. Det er naturlig å snakke om sikringsstrategier knyttet til inngåelse av posisjoner. Vi skiller mellom statiske og dynamiske sikringsstrategier. Ettersom vi ikke skal bruke opsjoner ser vi det ikke som naturlig å definere dynamisk sikring. Ved en statisk sikringsstrategi forutsetter man at det finnes en statisk sammenheng mellom to aktiva i forhold til deres respektive størrelser på et sluttidspunkt. Man aksepterer at dette også gjelder for alle perioder før dette, og man holder derfor en konstant andel over hele leveperioden (Reverre, 2001). Arbitrasjehandel avhenger derfor av at man har proporsjonelle posisjoner på kjøps- og salgssiden. Hvis man har en long posisjon må man også ha en fornuftig størrelse short. I følge Reverre (2001) er det i et risikoperspektiv langt viktigere at man får balansert posisjonene, i motsetning til å innta store nominelle posisjoner som man ofte trenger for å utnytte arbitrasjemulighetene.

På samme måte som at det er viktig å ha riktige proporsjoner på kjøps- og salgssiden er det av stor betydning at transaksjonene skjer samtidig, timing er derfor avgjørende. I situasjoner hvor mange aktører ser en arbitrasjemulighet og forsøker å innta like posisjoner, vil en form for risiko være om man har mulighet til å innta posisjoner til de prisene man ønsker.

Vanskeligheter med å oppnå ønskede priser kan skyldes stor etterspørsel i markedet. Det kan

også være tilfeller med lite likviditet i markedene hvor det kan oppstå problemer med å gå inn i, lukke posisjoner, eller at det kun er mulighet for å lukke den ene siden av posisjonen. Om det er dårlig likviditet i markedet vil det derfor ut fra et risikoperspektiv være bedre å nøye seg med mindre posisjoner, kontra å ha for stor andel på den ene siden i håp om bedre likviditet. Vidyamurthy (2004) mener derfor at arbitrasjører bør fokusere på sin kjernekunnskap, og at outsourcing av handel til spesialiserte meglere kan være med å dempe denne type risiko. Videre mener han at man skal forsøke å oppnå så stor spread som mulig når man inngår en posisjon, samt så liten spread som mulig når man lukker en posisjon for å oppnå høyest mulig avkastning. Dersom megleren har stordriftsfordeler kan man dra nytte av dette for å oppnå beste spread.

3.2.2 Konvergeringsrisiko og likviditet

Som nevnt inntar arbitrasjører motstående posisjoner i aktiva i den tro at de skal konvergere til en forventet likevekt. Det er naturlig at forventningene ikke alltid stemmer, og posisjonen derfor ikke gir noe profitt. I følge Reverre (2001) er suksessfulle arbitrasjører flinke til å fastsette mål på hvor lang tid de ønsker å holde posisjonene åpne basert på historiske handler, samt å stole på egen magefølelse. Han sier at en av de viktigste egenskapene til arbitrasjører er evnen til å være disiplinerte og til å ta tap.

Årsaken til dette er at ettersom man ofte må låne penger for å få en stor nok handel, vil mange arbitrasjehandlere foregå på margin eller med sikkerhet i andre aktiva, da kapital er en knapp ressurs. Dersom man da har en posisjon i en forventet lønnsom arbitrasjemulighet og posisjonene fortsetter å divergere i stedet for konvergere så er det en kostnad involvert. Man kan få en margin-call fra långivere og dermed bli tvunget til å likvidere posisjonen med et tap, dette på det verst tenkelige tidspunkt ettersom posisjonen er på sitt mest lønnsomme og det optimale kanskje hadde vært å ta en enda større andel i posisjonen.

3.2.3 Short-squeeze

Short-squeeze er en viktig risiko å ta i betraktning da det ikke er mulig å sikre seg mot denne. I følge Reverre (2001) kan en short-squeeze skje når en aksje uventet får en kraftig stigning. Dersom det er investorer som forvalter egne penger har de som regel høyere risikoaversjon enn institusjonelle aktører. Dersom de førstnevnte har store salgsposisjoner og aksjen plutselig stiger vil de ønske å lukke sine posisjoner ved å kjøpe tilbake aksjene for å begrense tapene, som igjen fører til en flom av kjøpsordrer som fører til at prisene stiger.

Fusjoner kan være et annet tilfelle. Ved en fusjon vil som regel aksjeprisene til begge selskapene stige, og hvis det i tillegg kommer gode nyheter fortsetter prisen å stige. Dette gir insentiver til de som eier aksjene til å selge for å sikre seg profitt. I tillegg vil mange aktører være på utkikk etter arbitrasjemuligheter som fører til at etterspørselen etter aksjene stiger, noe som dermed gjør dem til en knapp ressurs og får prisen til å stige ytterligere.

3.2.4 Kredittrisiko

Dersom man har gjort en handel og motparten ikke kan oppfylle sine forpliktelser, er dette en risiko. Det kan være at en utsteder av obligasjoner går konkurs eller at man har inngått en futureskontrakt for levering av råvarer. For å sikre seg mot denne typen risiko kan man kjøpe forsikringer.

3.2.5 Long Term Capital Management

For å avrunde kapittelet om risiko er det naturlig å se på et eksempel som viser at det er langt fra risikofritt å drive med arbitrasjehandel. Long Term Capital Management (LTCM) var et stort hedgefond som på 1990-tallet samlet noen av de beste hodene i finans for å utnytte arbitrasjemuligheter i markeder over hele verden. I følge Damodaran (2003) hadde de som konkurransefortrinn at de kunne låne penger til en lav kostnad og dermed fikk de tilgang til en stor kapitalbase, noe som gjorde at de i flere år kunne slå de andre traderne på Wall Street.

På grunn av sin suksess fikk de tilgang til så store mengder kapital at de måtte begynne å se seg om etter mer spekulative arbitrasjemuligheter. Fondet argumenterte med at selv om posisjonene kunne være mer risikable, hadde de så mange posisjoner slik de at de var diversifiserte. Et tap på en posisjon ville jevnes ut av gevinst på en annen.

I følge Shleifer (2000) maktet ikke Russland å følge opp sine gjeldsforpliktelser i 1998, og devaluerte rubelen. I tillegg til at de la restriksjoner på russiske banker slik at de ikke fikk betale sine forpliktelser til vestlige kreditorer. Størsteparten av gjelden var holdt av hedgefond og investorer lokket av den høye avkastningen på obligasjonene. Siden fondene forutså at Russland ville misligholde den innenlandske gjelden og devaluere valutaen, tok hedgefondene salgsposisjoner i utenlandsopsjonene i den tro at landet også skulle misligholde disse. I tillegg solgte de terminkontrakter på russiske rubler slik at dersom landet devaluerte valutaen skulle de tjene penger på dette. Det viste seg at ingen av disse sikringsstrategiene fungerte ettersom Russland ikke misligholdt utenlandsgjelden og la

restriksjoner på bankene sine. Fondene måtte derfor ta store tap og markedssituasjonen gjorde at LTCM måtte reddes av den amerikanske sentralbanken.

Ettersom det ligger i arbitrasjørerers natur å forvalte store kapitalmengder bruker som nevnt de fleste fond lånt kapital. LTCM var intet unntak. De brukte en ekstrem grad av lånefinansiering i tillegg til en stor kapitalbase. I løpet av en måned tok de et tap på 3-4 milliarder dollar. I følge Damodaran (2003) er nettopp størrelse et tve-egget sverd, det gir kanskje stordriftsfordeler, men gjør det vanskelig å lukke posisjoner. I tillegg gjør bruken av lånefinansiering at selv små prisendringer kan gjøre store endringer i egenkapitalandelen. Til slutt viste historien om LTCM at selv de beste hodene i bransjen ikke kan sikre seg mot uforutsette hendelser i markedene.

3.3 Implikasjoner av EMH og arbitrasje

Markedseffisienshypotesen har vært dominerende siden 1970-tallet med uttallige artikler som både støtter og kritiserer den. Det har vist seg vanskelig å definitivt bevise eller motbevise hypotesen. Vi noterer at Bodie et al. (2009) mener markedene er meget effisiente, men at det kan være gevinster å hente for de spesielt kreative, intelligente eller flittige. Det er sentralt i økonomisk teori at to like goder skal være likt priset, dette er kjent som "The Law of One Price". I følge Gatev et al. (2006) kan man også forstå arbitrasjehandel som en test av denne hypotesen. For vår problemstilling som omfatter statistisk arbitrasje vil det ikke være mulig å hente noen gevinst dersom markedene var perfekt effisiente. Det blir derfor spennende å se i hvilken grad det kan være mulig å oppnå meravkastning også justert for transaksjonskostnadene.

3.4 Oppsummering

I kapittelet om arbitrasje tok vi først for oss hva det er, for så å stadfeste at EHM ikke støtter at det finnes gratis godbiter i markedene. Videre drøftet vi ulike typer arbitrasje som grener seg fra tilnærmet risikofri ren arbitrasje til relativ arbitrasje som medfører sekundærrisiko. Så forklarte vi hvilke aktører som i størst grad bedriver arbitrasje og hvilken tilnærming og utfordringer dette medfører. Til slutt diskuterte vi risiko og hvilke fatale konsekvenser dette kan få ved høy belåning.

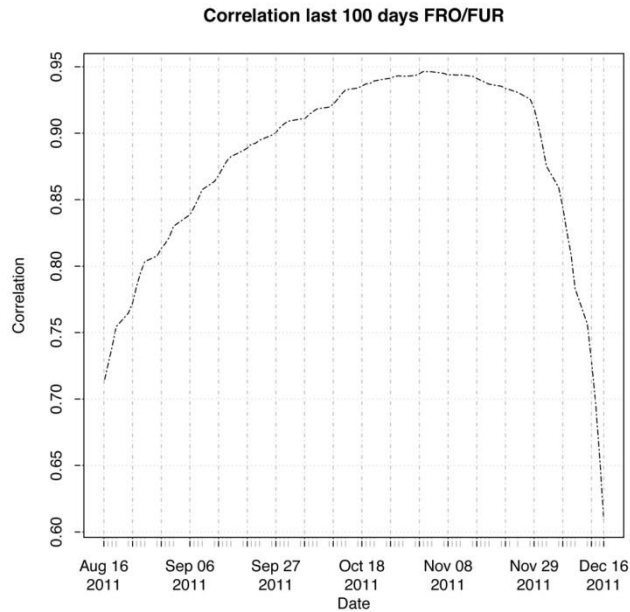
4. Pairs trading

4.1 Hva er pairs trading?

Investeringsmiljøer har lenge vært interessert i kvantitative metoder for å forbedre sine muligheter til investeringer og spekulasjon. En populær strategi for kortsiktig spekulasjon kalles pairs trading, også kjent som statistisk arbitrasje. Strategien har vært brukt de siste 30 årene, mest på Wall Street i regi av hedgefond og investeringsbanker. Det hele startet midt på 80-tallet da en gruppe med fysikere, matematikere og dataprogrammerere gikk sammen for å finne arbitrasjemuligheter i aksjemarkedet i regi av investeringsbanken Morgan Stanley. Tanken bak var å finne sofistikerte statistiske metoder som tok intuisjonen ut av spekulasjonen, og erstattet dette med mekanisk disiplin og handleregler. En av styrkene til strategien er utvilsomt graden av nøytralitet i forhold til svingninger i markedet og har siden oppdagelsen blitt en populær strategi som i dag også brukes av private aksjetradere (Gatev et al., 2006).

Konseptet bak pairs trading er ganske enkelt. Det hele handler om å finne to aksjer som over tid beveger seg likt. Det finnes ulike tilnæringer til dette, der en av disse er å måle korrelasjonen. Dersom to aksjer beveger seg likt impliserer dette at det er høy korrelasjon mellom aksjene (se figur 2). Poenget med dette er at man antar at to aksjer som historisk har en høy grad av korrelasjon også vil ha det i fremtiden (Whistler, 2004). En annen måte er å bruke kointegrasjon. Kointegrasjon har den fordel at det impliserer en langsiktig likevekt, mens korrelasjon forutsetter en mer kortsiktig tidshorisont. Kointegrasjon vil bli nærmere beskrevet i kapittel 4.1.2.

Normalt finner man slike par innenfor en bestemt sektor fordi man antar at informasjon som påvirker aksjepriser er mer relatert til hele sektorer enn til kun et spesifikt selskap (Reverre, 2001). Når spreaden mellom de to aksjene blir unormalt stor selger man den overprisede og kjøper den som er underpriset. Tanken da er at historien vil gjenta seg og at spreaden mellom aksjene vil konvergere til et normalnivå hvor man trer ut av posisjonene med en profitt tilsvarende bevegelsene i aksjene. Spreaden som vi refererer til i denne oppgave er avstanden mellom prisen i to aksjer, altså aksjepris A minus aksjepris B.

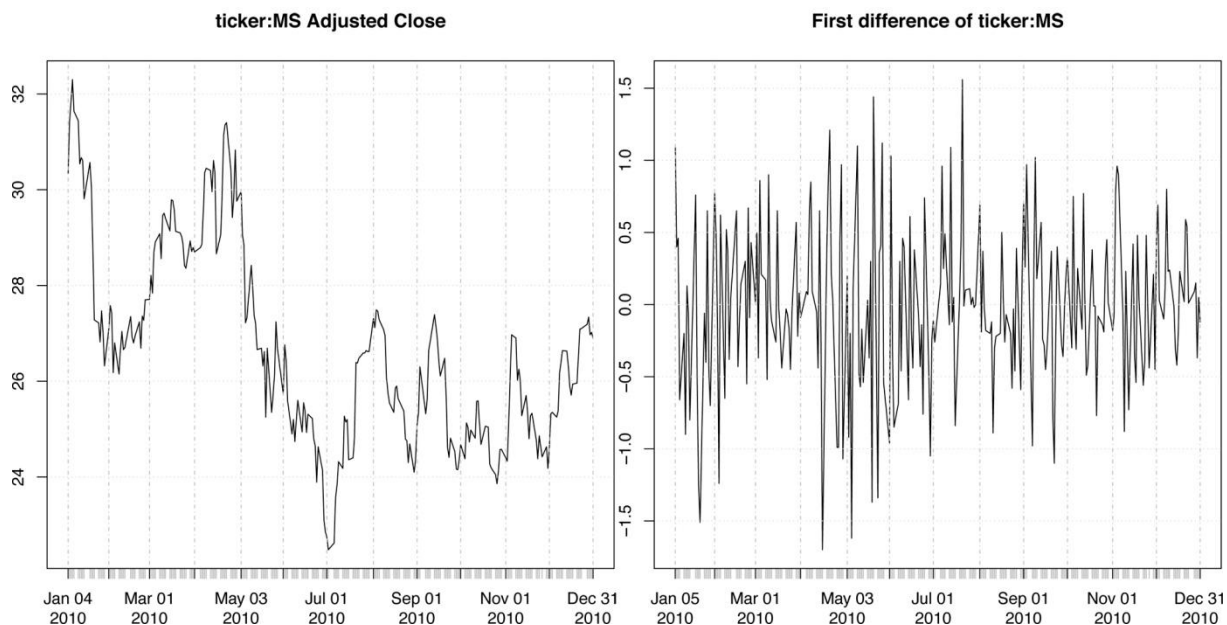


Figur 2 Korrelasjon mellom FRO og FUR

4.1.1 Stasjonære og ikke-stasjonære variabler.

De to neste kapitlene omhandler analyse av tidsserier med flere forklaringsvariabler som aksjekurser er et typisk eksempel på.

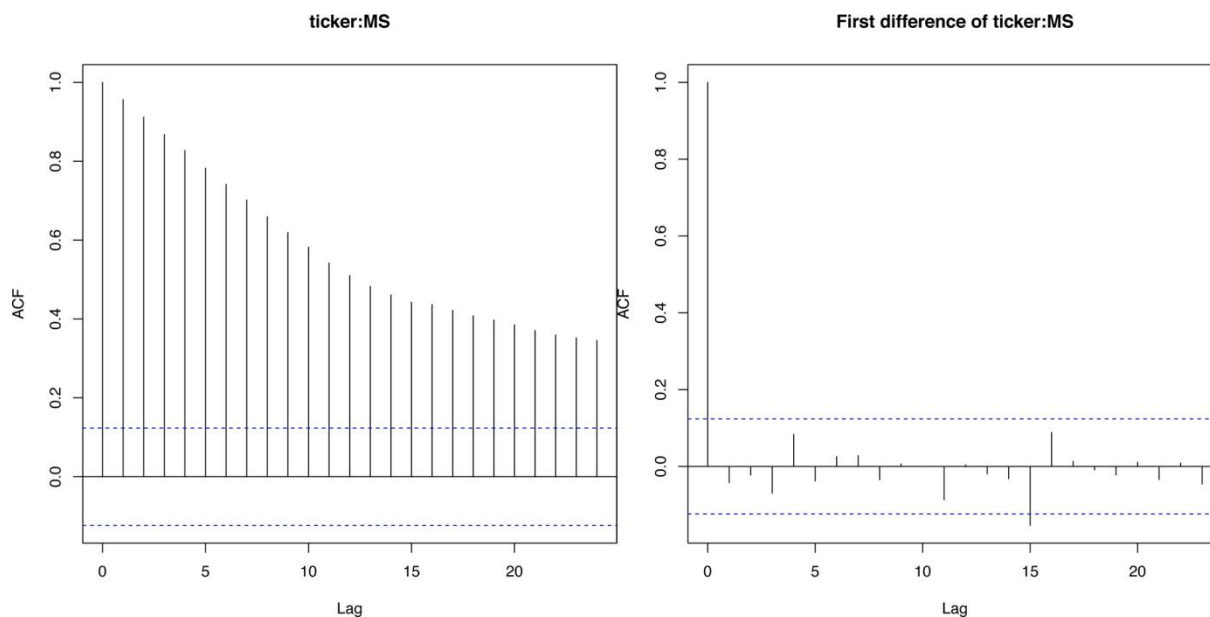
En stokastisk prosess kan i følge Lütkepohl og Krätzig (2004) defineres som stasjonær dersom dens første og andre moment ikke avhenger av tiden. Dette betyr at forventningsverdien til prosessen vil være gjennomsnittet av prosessen uavhengig av hvor på tidsaksen man befinner seg. I tillegg er den andre forutsetningen til en stasjonær prosess at den har en endelig varians og kovarians, som også er uavhengig av tiden. I tillegg må både forventningen, variansen og kovariansen være endelige tall, i motsetning til en ikke-stasjonær prosess som i prinsippet kan ha uavhengig forventning og varians.



Figur 3 Tidsserier

Figur 3 viser to tidsserier. Den til venstre viser en klar trend og vi kan dermed si at denne ikke er stasjonær. For å sjekke om en tidsserie er stasjonær kan man benytte seg av en Augmented Dickey Fuller (ADF) test. Den førstnevnte tidsserien fikk en testverdi på 1.47, og vi kan dermed ikke forkaste nullhypotesen om at tidsserien er ikke-stasjonær på noen signifikansnivå. På høyre side har vi en tidsserie som ser ut som den beveger seg rundt et fast gjennomsnitt. Den fikk i ADF testen en testverdi på -12.8 og vi kan dermed forkaste nullhypotesen på alle signifikansnivå. Tidsserien er dermed stasjonær.

Videre kan man også sjekke hvordan autokorrelasjonsfunksjonen ser ut for å si noe om en tidsserie er stasjonær. I figur 4 ser vi at til venstre avtar autokorrelasjonen sakte, men til høyre avtar den raskt og skifter fortegn flere ganger. Det sistnevnte indikerer en stasjonær tidsserie.



Figur 4 Autokorrelasjon

Det er vanlig å definere aksjepriser som ikke-stasjonære.

”The log levels of asset prices are usually treated as $I(1)$ with drift. Indeed, the random walk model of stock prices is a special case of an $I(1)$ process” (Zivot & Wang, 2006, p. 120).

Det at en prosess er integrert i første orden $I(1)$ betyr i følge Lütkepohl og Krätzig (2004) at den ikke-stasjonære prosessen kan bli stasjonær ved å se på den første forskjellen ($\Delta Y = Y_t - Y_{t-1}$). I følge Gatev et al. (2006) er risikoen for konkurs en av grunnene til at individuelle aksjekurser ikke kan ses på som stasjonære. I tillegg skriver Vidyamurthy (2004) at det er en vanlig forutsetning at den naturlige logaritmen av en aksjekurs er en random walk, ergo ikke stasjonær. Dette betyr at aksjepriser forutsettes å være ikke-stasjonære.

4.1.2 Kointegrasjon

Engle og Granger (1987) oppdaget at selv om to tidsserier er ikke-stasjonære så er det mulig i noen tilfeller å finne en lineær kombinasjon av de to som er stasjonær. Denne oppdagelsen fikk så navnet kointegrasjon. Kointegrasjon impliserer at på lang sikt opprettholdes en likevekt med en stasjonær og begrenset varians i motsetning til ikke-stasjonære tidsserier med uendelig varians. Det er verdt å merke seg at de to økonomene fikk Nobelprisen i økonomi for denne oppdagelsen.

I følge Vidyamurthy (2004) er ideen bak kointegrasjon i en formell sammenheng slik: La Y_t og X_t være to ikke-stasjonære tidsserier. Hvis en gitt verdi av Y , gir at $Y_t - \gamma X_t$ er stasjonær,

da er de to seriene kointegrerte. I ettertid har metoden blitt brukt i en rekke makroøkonomiske sammenhenger; konsum og inntekt, kort- og langsiktig rente, pengeetterspørsel og BNP m.m. Hovedresonnementet bak kointegrasjon er en langsiktig likevekt som er gjennomsnittet av den lineære kombinasjonen. Poenget er da at hvis det er et avvik fra det langsiktige gjennomsnittet så vil en eller begge tidsseriene justere seg selv for å gjenopprette den langsiktige likevekten (Vidyamurthy, 2004). Fra dette poenget oppsto blant annet handelsstrategien pairs trading.

4.1.3 Noen avklaringer

Pairs trading søker å finne relativ feilprising mellom to aksjer. Dette gjøres som regel på to måter, som er kalt statistisk arbitrasje og risiko arbitrasje. Statistisk arbitrasje har som formål å finne feilprising igjennom å analysere en tidsserie av prisinformasjon, mens risikoarbitrasje handler om strategier knyttet til fusjon av to selskaper (Gatev et al., 2006).

I følge Avellaneda & Lee (2010) omfatter statistisk arbitrasje en rekke strategier og investeringsprogrammer. Likhetene dem imellom er som følger:

1. Signalene til å handle er systematiske eller regelbaserte i motsetning til å bli drevet av fundamentale forhold.
2. Ulike posisjoner medfører at de sammen er markedsnøytrale med en beta lik null.
3. Mekanismen for å oppnå meravkastning er statistisk.

Metoden som beskrives i denne oppgaven er opptatt av at to aksjer konvergerer mot en statistisk likevekt. Men Whistler (2004) beskriver også varianter som spekulerer i forventning om at to aksjer skal divergere fra en likevekt basert på fundamentale forhold, kalt divergence trading. Han påpeker at dette kan være en mer komplisert måte å handle på. Dette fordi man aldri kan være sikker på når en topp er nådd og parene begynner å konvergere, noe som medfører større usikkerhet. På den andre siden kan man på en enkel måte kan skaffe seg tilgang til store mengder data og på denne måten regne ut gjennomsnittlig verdi. Derfor mener han at dette er en enklere og bedre metode, ettersom det er lettere å vite når aksjene er på gjennomsnittet, enn i ekstremiteten.

4.1.3.1 Hvorfor metoden hevdes å fungere

Hovedmannen bak strategien pairs trading, Nunzio Tartaglia mener psykologi gjør at metoden fungerer fordi mennesker har en tendens til å handle aksjer som går opp og ikke de som går ned. Dette kan tolkes som at strategien prøver å utnytte overreaksjonene til udisiplinerte investorer (Gatev et al., 2006). Grøtte (2002) forklarer at de fleste aktørene legger sine kjøps- og salgsordre i et selskap på bakgrunn av tro om oppgang eller nedgang i en aksje, ikke på spreaden mellom aksjer. På bakgrunn av dette mener han at det hele tiden vil oppstå muligheter for pairs trading.

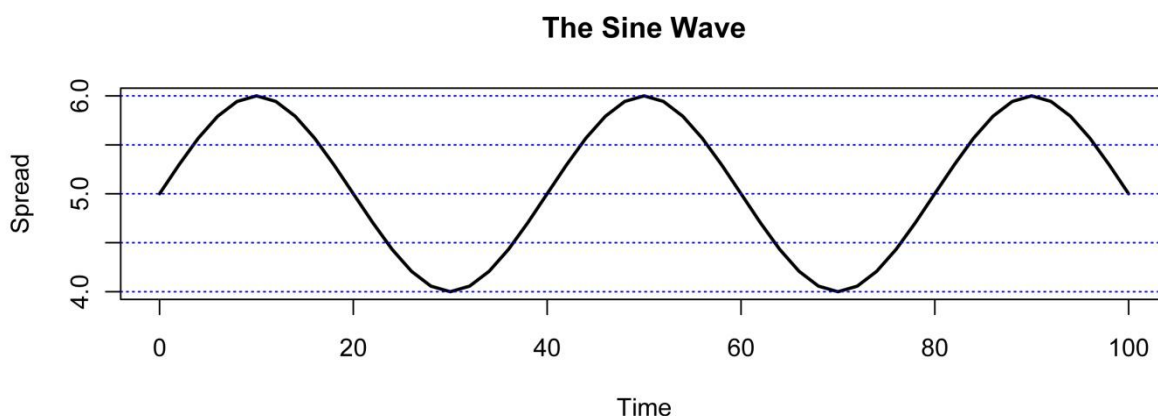
4.2 Markedsnøytral strategi

En markedsnøytral strategi innebærer å holde både en kjøpsposisjon og en salgsposisjon med lik eksponering av markedsrisiko hele tiden. Dette blir gjort ved å sammenføre de to vektete betaene fra både kjøps- og salgsposisjonen. Resultatet blir da at porteføljen ikke lenger er eksponert mot markedsrisiko og korrelerer derfor ikke med markedet, som igjen betyr at strategien er markedsnøytral. Man kan derfor fastslå at avkastningen til en markedsnøytral strategi er uavhengig av markedsavkastningen.

4.2.1 Pairs trading med enkel spread som mål

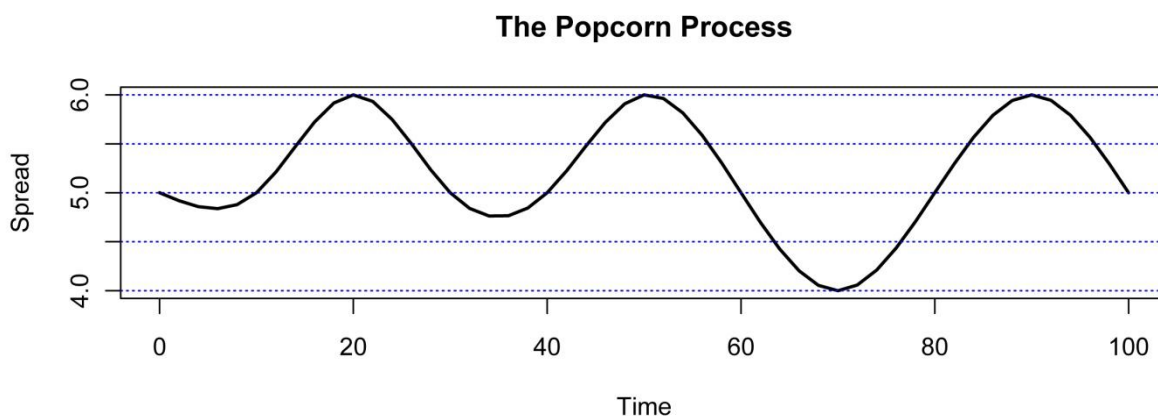
En av de første strategiene som ble brukt i pairs trading var å se på spreaden mellom to aksjer som beveget seg over tid i samme retning. Ved å identifisere bevegelsene i spreaden så kunne man lage en handleregel hvor man skulle kjøpe en aksje og selge en annen. For eksempel kunne to aksjer ha en spread som varierte mellom \$-2 og \$6. Handleregelen kunne da være å legge både kjøps- og salgsordre når spreaden var på \$4, for så å stenge posisjonen når den gikk tilbake til \$0. Denne strategien var ekstremt profitabel på 80- tallet da det fantes tusenvis av slike par og spekulere i. Pole (2007) mener at mye av grunnen til at dette var mulig, skyldtes at kun de store aktørene hadde innsikt, resurser og kapasitet til å utføre strategien. De hadde tilgang til datamaskiner som kunne behandle store mengder data, og på den tiden kostet maskinene flere titalls tusen dollar. I tillegg var det kun institusjonelle aktører som fikk daglige oppdateringer av aksjepriser. Dette var en barriere som forhindret andre fra å utnytte strategien.

Til å begynne med brukte man som sagt spreaden mellom to aksjer som beveget seg likt og forutsatte at den ville systematisk variere, fra vesentlig over gjennomsnittet til vesentlig under gjennomsnittet. Dette blir gjengitt i en sinuskurve som vist i figur 5. I starten var dette den teoretiske modellen for spread analyse.



Figur 5 The Sine Wave (Pole, 2007, p. 19)

Handleregelen var at man skulle åpne posisjonene på toppnivået, og deretter lukke dem på bunnivået av spreaden. Det viste seg derimot at med denne modellen mistet man mange muligheter, ettersom flere topper ofte fulgte hverandre i stedet for de svingningene man ser i en sinuskurve. Det medførte at man med denne metoden bandt kapitalen over lengre tidsperioder som forårsaket høyere risiko, enn ved å bruke popcornprosessen som beskrevet under.



Figur 6 The Popcorn Process (Pole, 2007, p. 19)

I figur 6 ser man popcornprosessen som ga ny innsikt til metoden. Her var tanken at spreaden ofte reverserer seg tilbake til gjennomsnittet slik at om man stoppet posisjonene der så ville flere muligheter åpne seg. En annen positiv ting med metoden er periodene uten eksponering. I dag har de aller fleste strategiene knyttet til pairs trading sin bakgrunn i forståelsen av popcorn prosessen (Pole, 2007).

Selv om det viktige her er å forstå grafene kan sinusfunksjonen skrives slik: $Y_t = \sin(t)$, mens popcornfunksjonen kan skrives: $Y_t = I_t \sin(t)$ der I_t er en indikator funksjon som ved verdiene 1 eller -1 signaliserer en topp eller bunn.

4.2.2 Pairs trading ved bruk av standardavvik

Selv om pairs trading er fellesbetegnelsen på long/short posisjoner i to selskaper som beveger seg likt finnes det utallige måter å gjøre dette på. Beregninger av korrelasjon, kointegrasjon, spread, gjennomsnitt og standardavvik er noe som hver enkelt aktør definerer ut i fra egne preferanser. Noen aktører tenker langsiktig med å finne feilprising ut i fra en lang tidsperiode på flere måneder og år, mens andre ser på relativ feilprising innenfor et par minutter.

Grøtte (2002) presenterer en metode av pairs trading som også betegnes som spread trading. Metoden er svært vanlig blant dem som trader profesjonelt. Fordelene med metoden skal være at den er nøytral i forhold til markedet og har relativt lav risiko, da sjansen for store tap er liten, samtidig som store gevinstmuligheter heller ikke er særlig sannsynlig. Argumentet for metoden er altså mange små gevinster over tid. Som drøftet tidligere handler metoden om å innta posisjonen når det statistisk og historisk er stor sjanse for at de to aksjene konvergerer etter å ha beveget seg bort fra likevekten en kort periode. Som annen profitabel aksjehandel er prinsippene om å følge trenden, kutte tapene og holde på vinnerne noe som også gjelder denne metoden (Grøtte, 2002).

Metoden til Grøtte (2002) handler om å normalisere verdien på spreaden mellom to aksjer som er beskrevet av Reverre (2001). Dette medfører at verdien i handleregelen handler om feilprisingen av spreaden uttrykt i form av standardavvik som uttrykker deviasjoner fra normalen, og posisjonene inntas ved statistisk signifikante nivåer. I metoden til Grøtte gjøres dette når prisen har beveget seg minimum to standardavvik fra normalverdi. Formelen uttrykkes slik.

$$\Delta_G = \frac{\Delta - MA_{500} \Delta}{SD_{500}(\Delta)} \quad (6)$$

(Grøtte, 2002, p. 511)

Der Δ er forskjellen mellom de to aksjene i nominell valuta. $MA_{500}(\Delta)$ er glidende gjennomsnitt de siste 500 minutter på forskjellen mellom aksjene. $SD_{500}(\Delta)$ er standardavviket de siste 500 minutter på forskjellen og Δ_G uttrykker feilprisingen i spread uttrykket ved standardavvik.

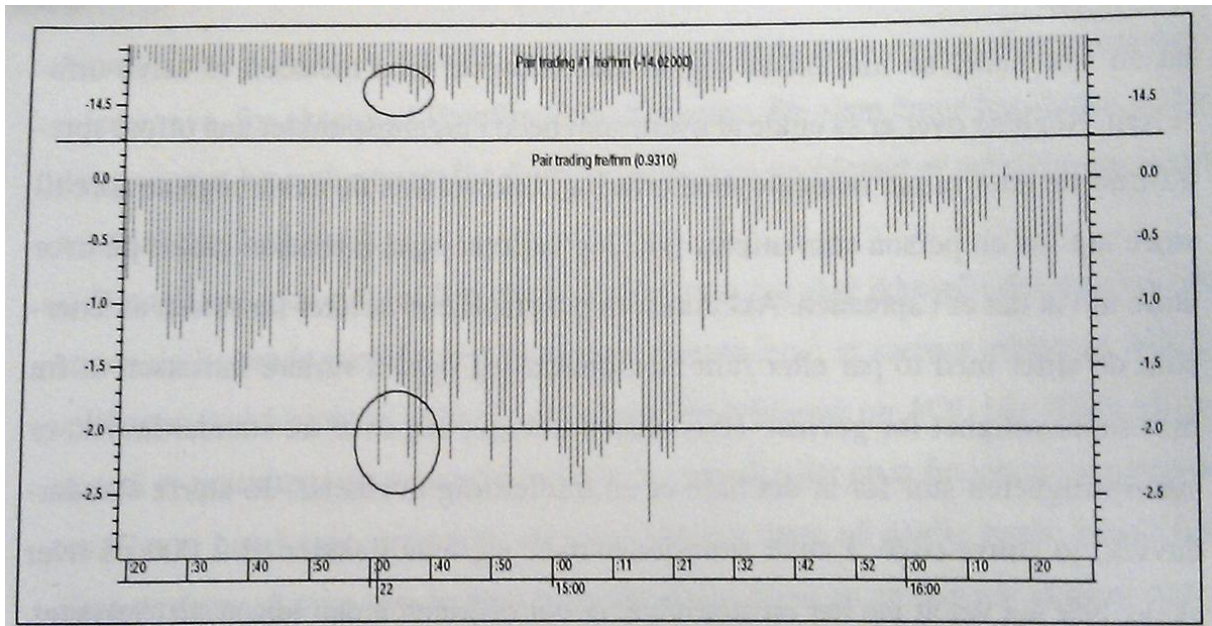
Formelen som er vist ovenfor har sitt utgangspunkt i formelen til Reverre, som er helt lik bortsett fra at den bruker 30 dager istedenfor 500 minutter. Formelen ser derfor slik ut:

$$\Delta_{norm} = \frac{\Delta - MA_{30} \Delta}{\sigma_{30}(\Delta)} \quad (7)$$

(Reverre, 2001, p. 476)

4.2.3 Standardavvik eller spread ratio

Metoden som beskrives ovenfor tilsier at spreaden og avviket er en tilfeldig variabel som er normalfordelt med snitt rundt null og standardavvik rundt en. Ved å ta denne forutsetningen så kan man anta distribusjonen til standardavviket, noe som ikke er like innlysende ved spread ratio. For eksempel kan man anta at det er ca. 15 % sannsynlighet for at standardavviket vil være enten over 1 eller under -1, samt at det er ca. 2,3 % sannsynlighet for at det er enten over 2 eller under -2. Dette gjør at man kan være rimelig sikker på at avvik utover dette er ekstremnivå eller midlertidige feilprising som snart vil medføre at aksjene vil bevege seg inn igjen mot likevekten. I metoden til Grøtte (2002) skal man innta posisjon når spreaden avviker med et standardavvik på 2 i minutt-chartet som illustrert i figur 7.



Figur 7 Chart (Grøtte, 2002, p. 513)

Pole (2007) poengterer at det ved bruk av standardavvik oppstår en feil siden alle observasjoner inkludert ekstremverdien blir kvadrert. Dette medfører at ekstremverdiene får proporsjonelt større vekt enn det de burde ha. Han argumenterer derfor for å fjerne ekstremverdier eller mer komplekst å justere dem ned. En annen feil med standardavvik er at finansiell data ikke er normalfordelt noe som kan gi feilkalkulasjoner som igjen medfører undervurdering av risiko. Hvis standardavvik gir slike feil, hvordan har det seg at aktører velger dette foran empirisk distribusjon, altså at de kun ser på spreaden? Pole (2007) forklarer dette med at ekstremverdiene som oppstår lettere kan hanskles med hvis man ser på det analytisk igjennom standardavvik i motsetning til kun å se på spread. Samtidig er det slik at standardavvik gir et tilfredsstillende resultat som derfor anses som godt nok.

4.2.4 Metode ved bruk av beta.

La oss si at aksje P og Q er i den samme sektoren og i tillegg er ganske like selskaper som for eksempel Exxon Mobile og Conoco Phillips. Man forventer da at avkastningen til de to aksjene vil følge hverandre etter å ha kontrollert for beta. Følgelig hvis P_t og Q_t uttrykker en gitt historisk tidsserie av priser så kan vi modellere dette som

$$\ln(P_t/P_{t0}) = \alpha(t - t_0) + \beta \ln(Q_t/Q_{t0}) + X_t \quad (8)$$

Dette betyr at etter å ha kontrollert for beta så vil long-short porteføljen svinge i nærheten av en statistisk likevekt. Modellen foreslår en investeringsstrategi der vi går long en dollar i

aksje P og short β dollar i aksje Q hvis X_t er liten og motsatt, gå short P og long Q hvis X_t er stor. X_t uttrykker avstanden fra likevekten (cointegration residual). Porteføljen forventes da å oppnå en positiv avkastning når verdien konvergerer mot likevekten (Avellaneda & Lee, 2010).

4.3 Pairs trading i praksis

4.3.1 Strategiens profitabilitet

Potensiell avkastning ved pairs trading avhenger selvsagt av hvilke kriterier en forvalter bruker. Gatev et al. (2006) slår i sin artikkel fast at pairs trading er profitabel. Studien er basert på data fra 1962 – 2002 ved bruk av sluttpriser for hver dag. Handelsregelen de har brukt er den samme som traderne de intervjuet brukte. Dette innebærer å åpne posisjoner i et par når prisene gir et avvik på to historiske standardavvik. For nærmere beskrivelse av metoden og beregning av avkastning så henvises det til artikkelen. Studiet viste at de fem beste parene ga en gjennomsnittlig avkastning på 1,31 % per måned, mens de 20 beste en avkastning på 1,44 % per måned, hvor avkastningen ble beregnet ut fra antall par som faktisk handlet i perioden. Fra 1999 til 2002 ga de 20 beste parene en årlig meravkastning på 10.4 %, en signifikant størrelse. Et interessant funn i studien viser at diversifiseringseffekten ved å øke antall par gir en høyere minimumsavkastning samtidig som maksimumsavkastningen holder seg relativt stabil. I perioden på 474 måneder, ga de 20 beste parene 71 måneder med negativ avkastning, mens de 5 beste ga 124 måneder med negativ avkastning, som igjen antyder mer stabilitet ved diversifisering. Denne artikkelen, samt Pole (2007) viser at pairs trading har hatt fallende profitabilitet siden sin storhetstid på 80-tallet. Dette forklares med den økende konkurransen blant de som praktiserer strategien, da det er blitt flere som ønsker sin del av avkastningen. I en artikkel av Do & Faff (2009) forsøker de å replisere studiet til Gatev et al. De klarer i ganske god grad å replisere funnene i perioden de sistnevnte studerte. Men de utvider i tillegg tidshorisonten til å gjelde perioden 2003 – 2008, hvor gjennomsnittsavkastningen faller til magre 0,06 % i perioden. Dette er i følge dem et resultat som ikke er statistisk signifikant annerledes fra 0 %.

4.3.2 Fundamentale forhold

Selv om utgangspunktet for denne oppgaven omhandler statistisk arbitrasje eller pairs trading er det også hensiktsmessig å ha en forståelse av de fundamentale driverne rundt kursbevegelsene i markedet. Årsaken til dette er at de statistiske signalene som oppstår i bunn og grunn skyldes fundamentale forhold. Poenget her er altså i følge Whistler (2004) å studere noen få fundamentale forhold slik at man kan unngå feil i de tilfellene statistikken er feilslått. Man blir altså en mer effektiv spekulant ved å dobbelsjekke statistikken mot fundamentale vurderinger. Han har følgende eksempler på fundamentale vurderinger som bør sjekkes:

Aksjene som handles bør være likvide aksjer. Begrunnelsen for dette er ganske enkelt at man lettere får kjøpt og solgt aksjene når man selv ønsker det. Handler man mindre likvide aksjer så kan man risikere og ikke få kjøpt/solgt alle postene til den pris og det tidspunktet man ønsker.

Vekst – forholdstall er viktig for å forstå fremgangen til et hvert selskap. Har man for eksempel en salgsvekst på 80 % mens inntektsveksten kun er på 10 %, så indikerer dette at noe er galt. Ulike veksttall mellom to selskaper kan derfor indikere at de vil divergere fra hverandre. En god regel er å se på tallene i forhold til gjennomsnittet innenfor samme sektor som selskapet.

Pris-fortjeneste (PE) er et forholdstall som for mange kan virke som irrelevant siden det ikke sier noe om fremtiden. PE finnes med å ta nåværende pris av aksjen og dele på nåværende fortjeneste per aksje (EPS). Noen mener at PE bør være rundt 1, mens andre mener den bør ligge mellom 20 til 40 for å indikere at den handles med en premie til den nåværende verdivurderingen. En nyere versjon av PE er PEG-forholdstallet som fremkommer ved å dele PE på den årlige EPS veksten. Dette forholdstallet sies å ha bedre evne til og prediktere fremtidig prisutvikling (Whistler, 2004). For eksempel så vil en PEG på 3.0 indikere at aksjen er overkjøpt og at investorer dermed forventer en tilbakegang i prisen. Tallene varierer allikevel mye fra en sektor til en annen, så man må derfor gjøre sammenligninger deretter. Andre viktige vurderinger knyttes til gjeldssituasjonen i selskaper, samt likvide midler til å betjene denne gjelden. Poenget er i følge Whistler (2004) at man ikke trenger å ha et totalt overblikk over den økonomiske situasjonen i selskapene, men at man bør ha en ide om hvor de ligger i terrenget.

4.3.3 Money management og risiko ved pairs trading

Selv om pairs trading hevdes å kunne gi meravkastning er det viktig å forstå risiko som er forbundet med strategien. En måte å håndtere risiko på, er å forstå hva som kan gå galt, og ha en plan for hva som skal gjøres i enhver situasjon som kan oppstå. Man kan beskytte kapitalen gjennom god money management. Whistler (2004) mener at selv om pairs trading består av hedging så medfører strategien mer risiko enn ved å investere i åpne posisjoner. Han begrunner dette med at man holder to posisjoner samtidig, og dersom posisjonene går feil vei, så har man dobbelt så stor eksponering mot tap. Et annet poeng er at transaksjonskostnadene er dobbelt så høye som ved vanlig handel. Han mener derfor at forståelsen av både tekniske og fundamentale forhold er avgjørende for profitabel utførelse av pairs trading. Videre poengterer han at når et par divergerer så må man være opptatt av å forstå de fundamentale forklaringene for at dette skjer. Man må derfor skille mellom årsaker som vil føre til en varig endring og de som mest sannsynlig kun er midlertidige og fører til at parene igjen vil konvergere.

Et annet sentralt poeng som Whistler (2004) påpeker er forståelsen av at menneskers følelser i investering omhandler frykt og grådighet. Han viser til menneskers svakhet når det kommer til å kutte tap og la vinnere løpe, og å holde seg konsistent til en strategi.

Han kommer med konkrete tips om å sette stopp ved 8.5 % på hver side av posisjonene. Maksimalt tap vil da være 18 %. Man kan være mer konservativ ved å gå så lavt som til 2,5 % - 5 % for å kutte tap. Et poeng som forklarer dette bedre er og aldri la en posisjon kunne skade den totale egenkapitalen med mer en 2,5 %. Man kan derfor si at jo større andel av egenkapitalen som eksponeres i en posisjon, jo lavere "stop loss" bør settes. En annen måte å justere risiko er ved å justere nivå av standardavvik. Ved å ha en handleregel som setter posisjoner på 3 standardavvik istedenfor 2, vil dette medføre et mindre antall handler, men som igjen potensielt gir bedre odds de gangene 3 standardavvik oppstår. Whistler argumenterer også med at pairs trading er mye mer avansert enn vanlig retningsbestemt aksjehandel. Hans begrunnelse for dette er at han mener man må gjøre de samme tekniske og fundamentale vurderingene ved å investere i en aksje, bare at her må man gjøre dette med to. For eksempel mener han at en stop-loss ikke nødvendigvis bør legges mekanisk, men at man ved bruk av teknisk analyse bør se på tekniske triggere som relevante plasseringer av stopp.

Metoden til Whistler (2004) er altså mer omfattende og stiller store krav til undersøkelser ved utførelse av pairs trading. Fundamentale forhold må heller ikke overses da statistiske

signaler ikke fanger opp om et selskap for eksempel er i økonomiske problemer. Han avslutter sitt resonnement med å fastslå at man kun vil handle profitabelt ved pairs trading om man legger en større innsats i arbeidet enn sine konkurrenter.

4.4 Oppsummering

Kapitlet om pairs trading starter først med å forklare hva pairs trading er. Deretter forklarer vi hvordan man kan definere om en tidsserie er stasjonær eller ikke-stasjonær. Videre forklares det hvordan man ved hjelp av kointegrasjon kan finne en lineær sammenheng mellom to ikke-stasjonære tidsserier. For så å forklare hvorfor metoden lar seg gjøre.

Underkapitlet om markedsnøytral strategi forklarer hvorfor strategien gir en beta rundt null som impliserer en markedsnøytral strategi. Videre beskrives hvordan strategien først ble praktisert med bruk av enkel spread, mens den senere har utviklet seg med bruk av spread ratio og avvik i standardavvik. I underkapittel 4.3 beskriver vi at strategiens profitabilitet basert på tidligere studier er dalende. Videre diskuteres en metode for å implementere fundamentale forhold som en del av strategien. Til slutt diskuteres money management og risiko ved pairs trading.

5. Metode

5.1 Analyseverktøy

For å kunne gjøre våre analyser har vi vært avhengige av et godt verktøy. Vi besluttet å benytte oss av statistikkprogrammet R², som er et gratis open-source program. Fordelen med R er at man kan tilpasse det slik at det passer til den problemstillingen man arbeider med ved hjelp av tilleggspakker. Innen fagfeltet finans finnes det en rekke utviklingsprosjekter hvor programmere tar i bruk de til enhver tid nyeste metoder innenfor fagfeltet og lager pakker som kan brukes til analyse av disse.

Ulempen med R er for det første at det er en ganske høy brukerterskel for å forstå programmet om man ikke er kjent med programmering. Vi har brukt mange timer på å forstå hvordan selv enkle operasjoner skal gjøres. I tillegg har det vært delvis tynn dokumentasjon på enkelte av funksjonene i R, ettersom disse fremdeles er under utvikling.

For å teste par for kointegrasjon har vi tatt i bruk pakken 'urca'³ hvor man finner Johansen testen. Vi har brukt pakken 'PairTrading'⁴ for å teste om spreaden mellom aksjene er stasjonær ved hjelp av ADF og PP testen.

For analysen i handleperioden har vi brukt pakkene i prosjektet 'Trade Analytics'⁵, i tillegg til pakken 'PerformanceAnalytics'⁶ for utregning av deskriptiv statistikk.

5.2 Utvelgelse av aksjer, sektor og tidsperioder

Vår metode startet med å definere hvilke aksjer vi skulle bruke i studien. På forhånd definerte vi at børsene New York Stock Exchange og NASDAQ Stock Exchange var godt egnet på bakgrunn av at disse børsene gir en stor bredde av industrier med store likvide selskaper. De fleste studier om temaet pairs trading er utført i disse markedene som igjen gjør at vi kan sammenlikne vår studie mot disse. Av hensyn til arbeidsmengde bestemte vi oss for å begrense studien til de fem hovedsektorene; financials, technology, consumer goods, conglomerates og basic materials. Videre definerte vi at kun selskaper med en markedsverdi over \$100 millioner, samt handelsvolum på minst 100 000 aksjer per dag skulle tas i betraktning. Dersom det var selskaper som manglet prisdata på enkelte dager ekskluderte vi disse. Med å gjøre dette valget ekskluderer vi mindre selskapet med liten

² <http://www.r-project.org/>

³ <http://cran.r-project.org/web/packages/urca/>

⁴ <http://cran.r-project.org/web/packages/PairTrading/index.html>

⁵ http://r-forge.r-project.org/R/?group_id=316

⁶ <http://cran.r-project.org/web/packages/PerformanceAnalytics/index.html>

omsetning som igjen er essensielt for å komme seg inn og ut av posisjoner når man måtte ønske uten å risikere stor avstand mellom kjøper og selger i ordreboken. Til slutt ekskluderte vi også aksjer som manglet prisdata for hele perioden, ettersom det var vanskelig å implementere en regel i R som tilsa at vi skulle gå ut av posisjoner når aksjene ble tatt av børs.

Når disse valgene var gjort ble vi sittende igjen med i overkant av 1500 aksjer fordelt på 92 undersektorer. Vi hentet daglige data for disse aksjene fra Yahoo Finance.

Valg av periode for testing av par og handelsperiode har sin bakgrunn i studier av Gatev et al. (2006) og Do & Faff (2009). Begge disse studiene har benyttet seg av 12 måneders periode for å finne par, med påfølgende seks måneders handelsperiode for disse parene. Studien til Do & Faff (2009) har de ferskeste data og omfatter handelsperiode ut 2008. Vi ønsket derfor å se om trenden fra deres studie har fortsatt, slik at vår handelsperiode er de første seks månedene av 2011. Parene vi handler i denne perioden ble testet for kointegrasjon i perioden 4.1.2010 til 31.12.2010.

5.3 Utvelgelse av par

Vi har kun testet par innen hver undersektor. Dette fordi det skal være størst sjans for at det skal være like fundamentale forhold i bedriftene som gjør at de er nære økonomiske substitutter. Do & Faff (2011) kommer i sin artikkel med en kritikk av metodikken til Gatev et al (2006) for utvelgelse av par på tvers av sektorer, fordi de finner at risikoen for at par ikke skal konvergere til en likevekt har økt etter 2002. Ved bare å velge par innenfor en gitt undersektor forsøker vi dermed å redusere denne risikoen. En viktig faktor er at vi med denne fremgangsmåten for utvelgelse får en håndterlig mengde par i utvalget i forhold til størrelsen på oppgaven.

For å teste for kointegrerte par har vi har brukt Johansen metoden på alle mulige kombinasjoner av par innenfor en gitt sektor. Vi valgte her å bruke eigenvalue metoden. Testen har en null-hypotese ($r=0$) som angir ingen kointegrasjon, mens alternativhypotesen angir at det er et kointegrasjonsforhold.

Dersom vi fikk en testverdi som var høyere enn kritisk verdi på 1 %, 5 % eller 10 % signifikansnivå var dette par vi ville se nærmere på. Etter å ha testet for kointegrasjon med Johansen metoden gjorde vi derfor flere tester; ADF og PP test hvor fremgangsmåten blir beskrevet under.

I neste steg omgjorde vi aksjeprisen til den naturlige logaritmen av aksjeprisene. Etter vi hadde gjort aksjekursene om til logaritmiske priser gjorde vi en lineær regresjon på de to.

$$lm \ln(y) \sim \ln(x) \quad (9)$$

På bakgrunn av resultatet fra regresjonen fikk vi en hedge ratio (regresjonskoeffisient 1) og premium (regresjonskoeffisient 2). Ved hjelp av disse kunne vi lage en spread mellom aksjene.

$$spread = \ln(y) - (hedge\ ratio * \ln(x) + premium) \quad (10)$$

Vi testet så om spreaden mellom aksjene var kointegrerte. Først med en Phillips-Perron (PP) test, for så i tillegg å teste med en Augmented Dickey-Fuller (ADF) test.

Ettersom vi ikke ønsket å bruke tid på å forklare forskjeller mellom testene og derfor rettferdiggjøre hvorfor vi hypotetisk sett skulle velge et par som gikk gjennom en av de tre testene på et gitt signifikansnivå, men ikke var signifikant i en annen test besluttet vi å dele aksjene inn i to grupper.

Alle aksjeparene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe, som vi ved hjelp av god fantasi kalte 5 %.

De parene som ble signifikante på minimum 10 % nivå i alle testene har vi gruppert sammen i 10 % gruppen. De parene som ble signifikante i Johansen testen, men ikke signifikante på minimum 10 % nivå i både PP og ADF ble forkastet. Alle parene som ikke ble statistisk signifikante ved hjelp av Johansen metoden ble forkastet uten øvrige tester.

Vi testet totalt 1537 par for kointegrasjon og endte til slutt opp med 161 par i gruppen 5 %. I gruppen 10 % endte vi også opp med 161 par. Parene sorteres som f.eks. BVN.HMY, der den førstnevnte er aksje 1 og den sistnevnte aksje 2.

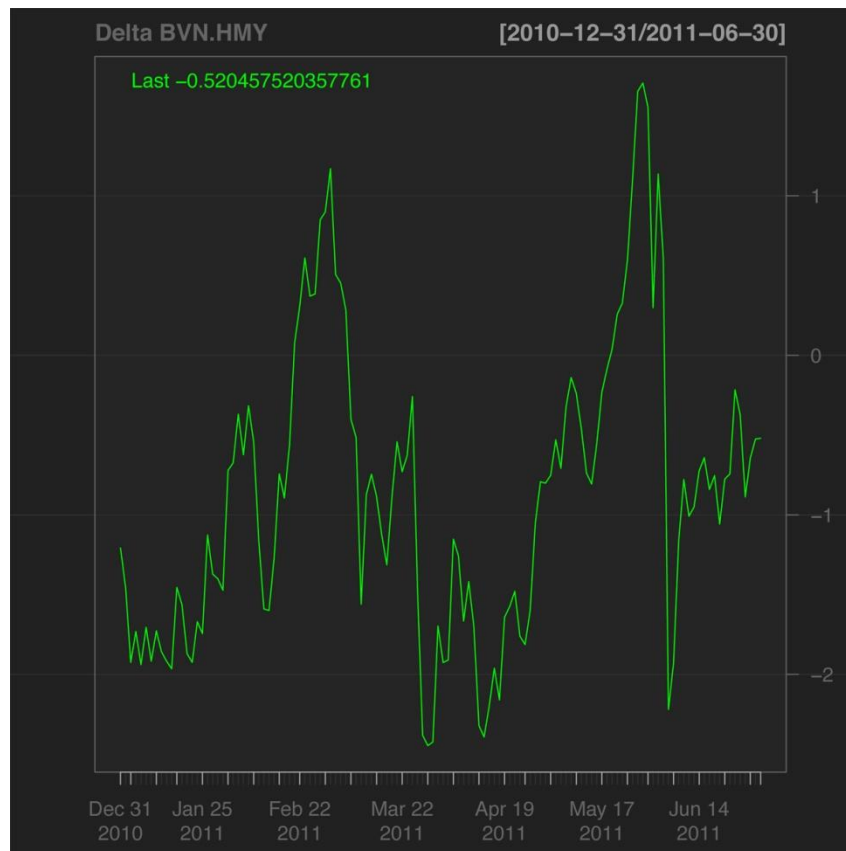
5.4 Handleregul

Vi har kalt vår handelsindikator delta. Den regnes ut på følgende måte:

$$\Delta_{norm} = \frac{\Delta - MA_{50}(\Delta)}{\sigma_{50}(\Delta)} \quad (11)$$

Der Δ er forskjellen/spreaden mellom de to aksjene (aksje 1 – aksje 2). $MA_{50}(\Delta)$ er glidende gjennomsnitt de siste 50 dager på forskjellen/spread mellom aksjene. $SD_{50}(\Delta)$ er standardavviket de siste 50 dager på forskjellen/spreaden og Δ_{norm} uttrykker feilprisingen i forskjellen/spreaden uttrykket ved standardavvik, som vi har valgt å kalle delta. I kalkuleringen av delta bruker vi prisen fra siste handel per dag, ikke justert for dividende og splits (Close).

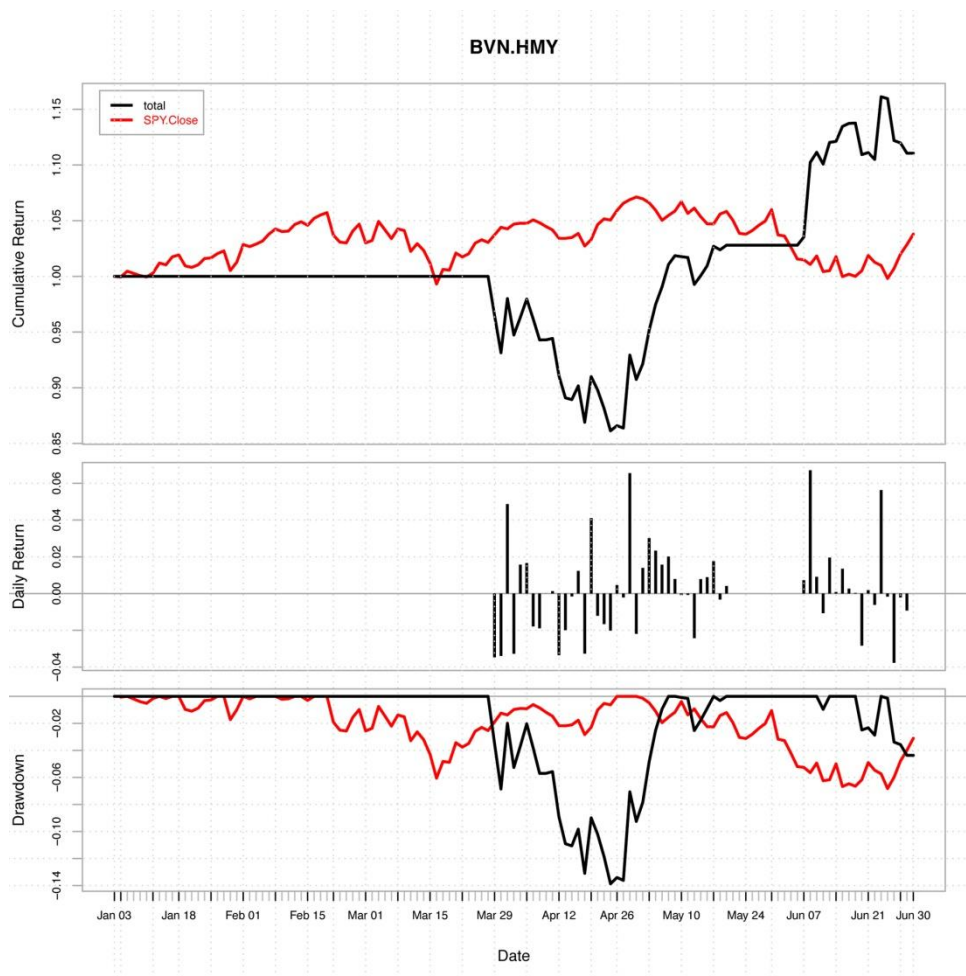
Handleregulene er som følger. Vi inngår handel på 2 eller - 2 standardavvik, i tråd med Reverre (2001). Det vil si salg av den overprisede aksjen, samt kjøp av den underprisede aksjen med like stor sum på begge sider. Dersom indikatoren viser 2 kjøper vi aksje 1 og selger aksje 2, og motsatt dersom indikatoren viser minus 2. Posisjonene lukkes ved konvergering til standardavvik på 0 eller på siste dag av handelsperioden. Etersom det er risiko knyttet til handel forutsetter vi videre at vi lukker posisjonene dersom paret divergerer ut over 3 eller -3 standardavvik. Delta kan krysse 2 og -2 flere ganger etter vi har åpnet en posisjon uten at vi har løst den inn. Vi har derfor lagt inn en regel om at vi bare åpner posisjonen første gang delta krysser indikatorverdiene. For eksempel ser vi i figur 8 at delta krysser - 2 på nytt 12. april etter vi allerede har åpnet en posisjon 28. mars, som enda er åpen. I dette tilfellet åpner vi ikke en ny posisjon selv om indikatoren skulle tilsi det.



Figur 8 Delta

For å illustrere metoden skal vi derfor gå igjennom et eksempel fra studien.

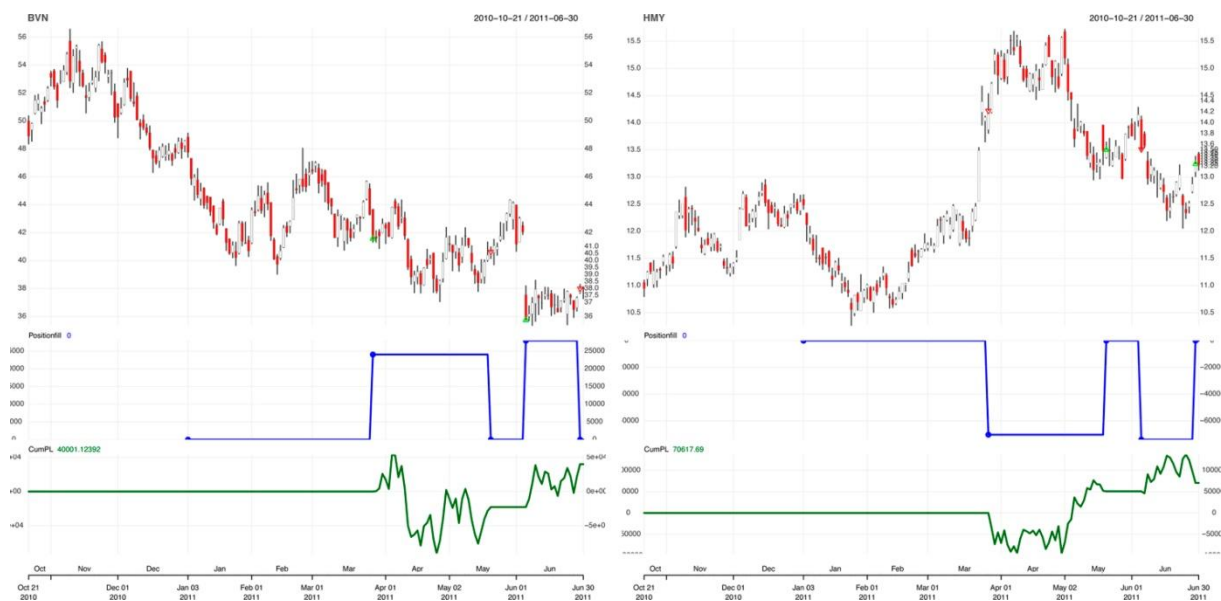
I dette eksempelet har vi tatt for oss aksjeparet BVN (aksje 1) og HMY (aksje 2). Figur 8 viser oss bevegelsen av feilprisingen i spread uttrykt i indikatoren delta. Som handlereglene over forklarer har vi kjøpt BVN og solgt HMY når delta lik -2 og lukket posisjonene ved delta lik 0. Som man da kan lese av diagrammet fikk vi tre signaler om å innta posisjon. Den første posisjonen ble innløst ved delta lik null, ved det andre signalet skjedde det ingen handel ettersom vi hadde en åpen posisjon, mens den siste posisjonen ble innløst på slutten av handelsperioden.



Figur 9 Sammendrag BVN.HMY

Dette resulterte i en periodeavkastning på 11,06 % som man kan lese av figur 9. I tillegg sammenliknes resultatet mot referanseindeksen S&P 500. Diagrammet Daily return viser den daglige avkastningen hver gang vi har hatt åpne posisjoner.

I figur 10 ser man kursutviklingen for BVN og HMY i den aktuelle perioden på 6 måneder. I dette tilfellet ser man at det nederste diagrammet (grønn linje) viser positiv avkastning på begge aksjene basert på handlene som ble gjort. Videre viser det midterste diagrammet (blå linje) position fill, og viser som tidligere at vi har kjøpt BVN, mens vi solgte HMY i begge åpningene av posisjoner. Alle kjøp og salg av aksjer skjer til prisen fra siste handel per dag, ikke justert for dividende og splits (Close).



Figur 10 Oversikt aksjekurser

5.5 Beregninger i analysen

Avkastning

I analysen har vi beregnet avkastning og sammenlignet med referanseindeksen som i vårt tilfelle er S&P 500 eller sektorindeksen til den respektive sektor. Den daglige avkastningen på long og short posisjonene ble kalkulert som verdivektede avkastninger på denne måten som uttrykker daglig avkastning:

$$r_{P,t} = \frac{\sum_{i \in P} w_i t r_{i,t}}{\sum_{i \in P} w_i t} \quad (12)$$

Vi antar at hver posisjon består av 1 dollar, altså 1 dollar på en long posisjon og 1 dollar på short posisjon.

$$r_{P,k} = \frac{\sum_{i \in P} \sum_{t \in k} c_{i,t}}{n} \quad (13)$$

Gatev et al. (2006) samt Do & Faff (2009) har to mål på avkastningen. Den første kaller de ”return on committed capital”. Ved utregning av avkastningen deler de den samlede daglige avkastningen på antall par i porteføljen. Den andre måten kaller de ”fully invested return”. Ved denne metoden beregner man avkastningen ved å ta den samlede daglige avkastningen delt på antall par som faktisk handler i perioden. Den førstnevnte metoden er en klart mer konservativ tilnærming til å beregne avkastningen. De mener at man ved denne metoden tar

hensyn til alternativkostnaden ved å binde kapital til en strategi, selv om strategien ikke gjør handler.

Dersom man har god tilgang på kapital og er fleksibel i finansieringen kan det derimot være mer korrekt å benytte seg av den faktiske sysselsatte kapitalen i beregningen av avkastningen.

Vi har valgt den mest konservative måten å beregne avkastningen på, nemlig "return on committed capital".

Periodeavkastningen beregnes ved å beregne den kumulative daglige avkastningen.

$$w_{i,t} = w_{i,t-1} (1 + r_{i,t-1}) = (1 + r_1) \dots (1 + r_{i,t-1}) \quad (14)$$

Den årlige avkastningen beregner vi som en simple return av periodeavkastningen. I følge Bodie et al. (2009) er det vanlig å bruke simple returns på investeringer med en tidshorisont kortere enn ett år.

$$\overline{R}_a * scale \quad (15)$$

(Peter Carl, 2012)

Videre beregnet vi høyeste og laveste avkastning per posisjon, fordeling av antallet gevinster versus tap og gjennomsnittlig holdingperiode per handel.

Risikofri rente

6 måneders risikofri rente (R_f) har vi satt til 0,2 %, hentet fra U.S Department of the Treasury (n.d). Programmet ville ikke la oss bruke flere desimaler, men forskjellene ville vært minimale i utregninger.

Sharpe Ratio (Årlig)

For å måle meravkastning utover risikofri rente per enhet totalrisiko så bruker vi Sharpe ratio. Forholdet mellom avkastning og risiko sier noe om hvor attraktiv en portefølje er. Dette avgjøres av forholdet mellom risikopremien og standardavviket til avkastningen utover risikofritt aktivum. Måling av Sharpe ratio er svært vanlig for å vurdere resultater til forvaltere (Bodie et al., 2009).

$$\text{Sharpe ratio} = \frac{\sqrt[n]{\text{prod}(1 + R_a)^{\text{scale}}} - 1}{\sqrt{\text{scale}}\sqrt{\sigma}} \quad (16)$$

$n = 125$ som er antall handelsdager i perioden og $\text{scale} = 252$ som er antall handelsdager i året. R_a er justert for risikofri rente.

(Peter Carl, 2012)

Information Ratio

Et alternativ til Sharpe ratio som prestasjonsmål er Information Ratio (IR). Dette prestasjonsmålet måler meravkastningen (α) en forvalter oppnår utover referanseindeksen i forhold til en tracking error. Denne defineres som forskjellen i standardavvik mellom porteføljen og referanseindeksen. Fordelen med IR er at den er lett å tyde. Dersom IR er større enn null, har forvalteren gjort en god jobb og motsatt dersom IR er under null. Markedsporteføljen vil bestandig ha en IR lik null.

$$IR = \frac{\alpha_p}{\sigma(e_h)} \quad (17)$$

(Bodie et al. (2009)

Value at Risk (VaR)

Value at Risk (VaR) måler en enhets eksponering mot markedsrisiko. VaR gir et statistisk mål på mulige tap i en portefølje basert på ”normal” markedsutvikling. Tap større en VaR skjer kun ved en spesifisert liten sannsynlighet. Kvartilen q av en distribusjon, er verdiene som ligger under q % i distribusjonen. Den vanlige estimeringen som vi også har benyttet oss av er 5 % VaR, som betyr at 95 % av avkastningene vil overgå VaR, og 5 % vil være mindre. Man kan derfor si at VaR er det minste tapet av de 5 % dårligste avkastningene.

$$\text{VaR}(.05 \sim N) \quad (18)$$

(Bodie et al., 2009)

Expected shortfall

Et mer realistisk syn på nedsiderisiko fokuserer på det forventede tapet gitt at man befinner seg i det verst tenkelige scenario. Denne verdien kalles expected shortfall (ES) som vekter de verste avkastningene mest, som igjen gir et verre utslag for potensielle tap enn VaR.

$$ES = \frac{1}{.05} \text{esp}(\mu)N[-\sigma - F(.95)] \quad (19)$$

(Bodie et al., 2009, p. 166)

Max drawdown

Max drawdown er ifølge Chan (2009) definert som den prosentvise forskjellen mellom det høyeste punktet i avkastningskurven og det laveste punktet som inntreffer på et senere tidspunkt.

Beta

Beta er et mål på volatiliteten i porteføljen sammenlignet med volatiliteten til referanseindeksen. Med vår markedsnøytrale strategi kan vi forvente en beta rundt 0, da den er uavhengig av endringer i referanseindeksen.

$$\beta_{a, b} = \frac{CoV_{a, b}}{\sigma_b} = \frac{\sum((R_a - \mu_a)(R_b - \mu_b))}{\sum(R_b - \mu_b)^2} \quad (20)$$

(Peter Carl, 2012)

Standardavvik (Årlig)

$$\text{Årlig } \sigma = \sqrt{\sigma^2} = \sqrt{\text{observert } \sigma^2} * \sqrt{n} \quad (21)$$

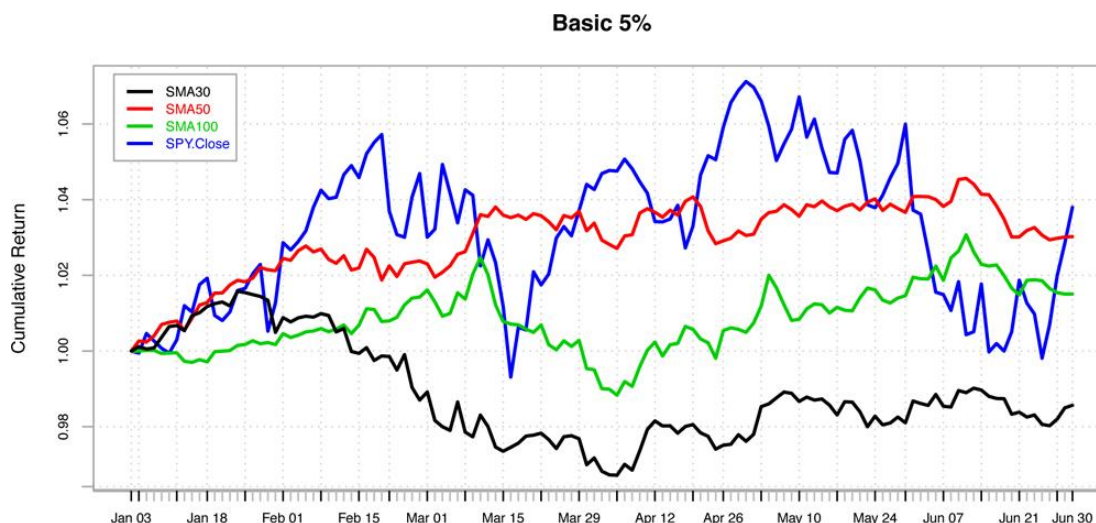
(Peter Carl, 2012)

Vi har brukt $n = 252$ dager, som tilsvarer antall handelsdager i et år.

5.6 Diskusjon, valg av kriterier

5.6.1 Indikator, valg av lengde på glidende gjennomsnitt

I studien har vi brukt et glidende gjennomsnitt på 50 dager. Grunnlaget for dette valget baserer seg på vår forståelse av hvordan feilprising vist ved standardavvik reagerer på antall dager med glidende gjennomsnitt. Mekanismen er slik at dersom det er et høyt antall dager som blir brukt til glidende gjennomsnitt, resulterer dette i færre handler med lengre holdingperiode. En lang holdingperiode kan være en risikofaktor ettersom Do & Faff (2011) har kommet fram til at den største risikoen ved handelsstrategien er konvergeringsrisikoen. I tillegg vil en trader se seg nødt til å binde kapitalen over et lengre tidsrom. På den andre siden vil for kort holdingperiode føre til at strategien vil bli veldig handelsintensiv, som naturlig nok medfører høyere transaksjonskostnader.



Figur 11 Chart, diverse SMA.

	Basic 5% SMA 30	Basic 5% SMA 50	Basic 5% SMA 100
Avkastning	- 2.89 %	6.1 %	3.0 %
Standardavvik	4.58 %	3.75 %	4.5 %
Antall handler	776	584	400
Holdingperiode	11.3 dager	18.7 dager	24.6 dager

Tabell 1 Resultater, forskjellige SMA.

Vi ser i tabell 1 at ved å bruke 30 dagers glidende gjennomsnitt (SMA) til beregning av indikatoren delta får vi nesten dobbelt så mange handler som med 100 dagers SMA. I tillegg

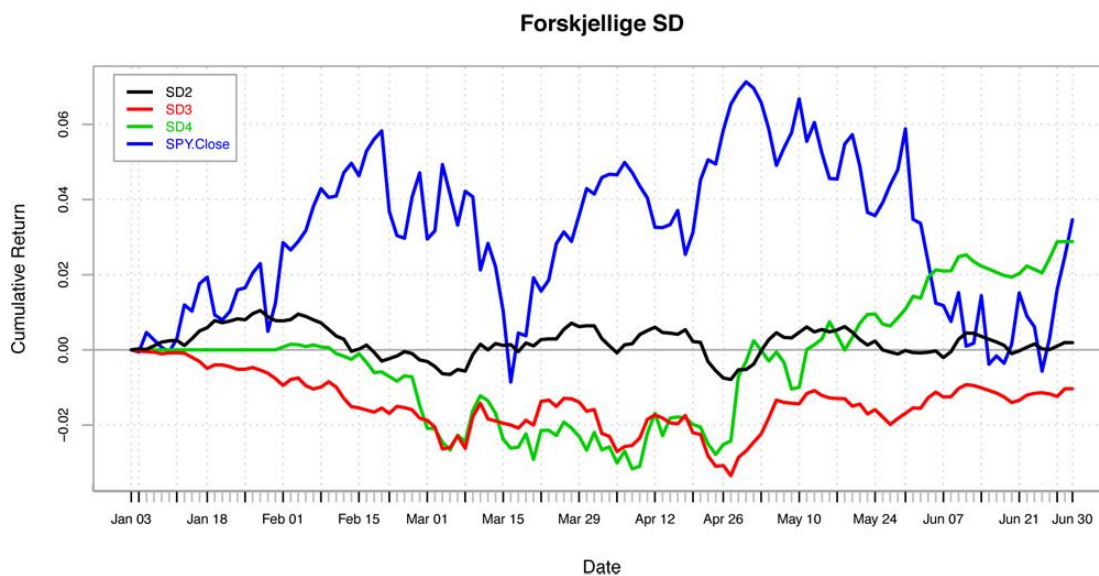
blir holdingperioden mer enn halvert ved bruk av 30 dagers SMA i forhold til 100 dagers SMA. Det finnes ingen fasitsvar på hva som er korrekt å bruke til beregning av indikator. Den enkelte trader må selv vurdere hva de vektlegger mest. Det kan være at for noen er det viktig å ha kort holdingperiode og dermed kortere perioder med kapitalbinding, de vil da bruke et relativt kort SMA i indikatoren. For andre er det kanskje viktigere å ha en mindre handelsintensiv strategi og da vil det være mer hensiktsmessig med en lengre tidshorisont i indikatoren.

Som en supplerende kommentar har vi lagt ved en graf for kumulativ avkastning for gruppen 5 % i sektoren Basic Material målt mot referanseindeksen S&P 500 i figur 11. Som man kan se fra grafen og også lese ut av tabellen har valget av tidshorisont i beregning av glidende gjennomsnitt, innvirkning på hvilken avkastning strategien oppnår. Man kan allikevel ikke trekke noen konklusjoner om en tidshorisont er bedre enn den andre, ettersom avkastningen avhenger av hvilken tidsperiode man ser på og ikke minst hvilke par man benytter seg av.

5.6.2 Valg av handleregel

Vi har sett at man kan velge forskjellig lengde på glidende gjennomsnitt i indikatoren, og at forskjellige lengder gir forskjellige resultater. På samme vis har naturlig nok valg av handelssignal en stor innvirkning på prestasjonene til strategien. Gatev et al. (2006) viser i sin studie at de valgte to standardavvik, selv om handleregelen ikke bestandig ville gi stor nok profitt til å dekke transaksjonskostnadene. Dessuten poengterer de at det var fristende å prøve potensielt mer profitable handleregler, men at de på grunn av fare for data-snooping bias ikke ønsket å optimalisere strategien.

Vi ønsker i dette kapittelet å se hva som skjer dersom vi øker handleregelen til handel på henholdsvis 3 og 4 standardavvik, hvor vi også løser inn posisjonen dersom den fortsetter å divergere til henholdsvis 4 og 5 standardavvik. Begrunnelsen for dette er på samme måte som i forrige underkapittel; å belyse at valget av handleregel er viktig dersom man ønsker å forsøke seg på denne typen kvantitative investeringsstrategier. Vi bruker i eksempelet porteføljen for alle parene i gruppen 5 %.



Figur 12 Forskjellige SD

	SD2	SD3	SD4
Antall handler	1772	540	76
Gjennomsnitt holdingperiode	19,60	0,07	2,32
Periodeavkastning	0,0021	-0,0101	0,0292
Standardavvik	0,0245	0,0351	0,0557

Tabell 2 Forskjellige SD

Ved å øke nivået i deviasjoner fra normalen der hvor man åpner en posisjon ser vi at antall handler blir dramatisk redusert. Som en naturlig konsekvens av færre posisjoner blir også standardavviket høyere. Videre ser vi at den gjennomsnittlige holdingperioden blir langt kortere. Årsaken til at holdingperioden på SD3 er lavere enn SD4 kan være at posisjonene i dette eksempelet fortsetter å divergere og dermed blir raskt løst inn på stop-loss nivået. Vi kan som tidligere ikke si noe om hvilket alternativ som er det beste, men ser at vi oppnår forskjellig avkastning ettersom hvilken handleregel som legges til grunn. Periodeavkastningen er i dette eksempelet for øvrig beregnet ut fra antall par som faktisk handlet i perioden.

5.6.3 Kointegrasjon

Ettersom Gatev et al. (2006) hadde en litt annerledes metode for etablering av par enn den metoden vi brukte, manglet vi retningslinjer på hvilket signifikansnivå som er vanlig å benytte seg av når man sjekker aksjepar for kointegrasjon. Intuitivt tror man tro at jo høyere signifikansnivå vi hadde benyttet, desto større sikkerhet ville det ha vært for at aksjeparene

var kointegrerte. Dermed hadde det vært størst mulig sjanse for at de har mean-reverting egenskaper ved seg, noe som vi er avhengige av for at strategien skal fungere. Siden vi som sagt ikke hadde retningslinjer på dette, valgte vi å sortere parene i grupper hvor par som var signifikante på 1 % og 5 % nivå var i en gruppe, mens den andre gruppen besto utelukkende av par signifikante på 10 % nivå.

Gatev et. al (2006) bruker en seks måneders periode til handel etter parene er testet. Disse periodene er i følge artikkelforfatterne valgt tilfeldig. Man kan da stille spørsmål ved om man kanskje skulle benyttet seg av enten kortere, eller lengre periode for handel. Et argument for å bruke en lengre tidshorisont til handel er at kointegrasjon impliserer en langsiktig likevekt. På den andre siden vil man med kortere handelsperioder sikre seg hyppigere tester av om parene fremdeles er kointegrerte.

5.6.4 Risikomål

I tillegg til de risikomålene vi har implementert i vår studie, bruker Gatev et al. (2006) også en utvidet Fama & French tre-faktor modell for å undersøke porteføljens systematiske risikoeksponering. I tillegg til modellens faktorer; markedspremie, SMB (small minus big) og HML (high BM ratio) utvider de denne med en short-term reversal faktor og en momentum faktor. Argumentet for å inkludere de to ekstra faktorene er at pairs trading handler på den relative styrken til aksjene i et par.

Hvis man selger short-term vinnere og kjøper short-term tapere forventer de dermed at avkastningen vil være positiv for reversal faktoren. Samtidig vil det være slik at dersom man selger medium-term vinnere og kjøper medium-term tapere vil avkastningen være negativt korrelert med momentum faktoren.

I sin artikkel finner Gatev et al. (2006) at bare en ubetydelig del av avkastningen kan tillegges de fem faktorene i modellen. På bakgrunn av resultatene fra artikkelen, samt omfanget på oppgaven vår valgte vi ikke å inkludere denne modellen i vår studie, men nøyer oss med å henvise til artikkelen.

5.6.6 Ordretype

Den vanligste måten å handle aksjer på børsen er å legge inn en markedsordre. Dette betyr at man legger en kjøps- eller salgsordre til ønsket pris. Når både kjøpere og selgere vil kjøpe og selge til samme pris gjennomføres handelen. Dersom aksjeprisutviklingen går i feil retning,

risikerer man i verste fall å tape hele investeringen om man holder en kjøpsposisjon eller en salgsposisjon.

For å redusere mulige tap kan man benytte en stop-loss ordre. Man legger da inn ordre på et kjøp, samt at man skal selge til markedspris dersom aksjeprisen faller med en gitt proSENTSATS. Stop-loss er en måte å sikre seg mot store tap i en portefølje.

I studien har vi utelukkende benyttet oss av markedsordre. Dessuten har vi benyttet oss av en slags stop-loss ved at posisjonen løses inn dersom parene divergerer ytterligere og passerer tre standardavvik fra normalen. En måte å forbedre dette på hadde vært ved å legge stop-loss ordre på f.eks. 10 % på hver av aksjene i et par. Samtidig måtte det ha vært automatikk i at dersom en av sidene i en posisjon ble lukket, skulle den andre også bli lukket, ettersom man da ville eksponeres for markedsrisiko. En enda bedre metode ville vært å kunne stoppe en posisjon ved et gitt prosentvis tap av investeringen i et par.

5.7 Mulige feilkilder

5.7.1 Survivorship bias

Det finnes flere databaser for å tilegne seg historiske prisdata til aksjer. Eksempler er Thompson Reuters, CRSP, Yahoo Finance og Google. De to førstnevnte er kostbare alternativer og i følge Chan (2009) fri for survivorship bias ettersom man får tilgang til prisdata til aksjen selv etter den er blitt tatt av børs. De to sistnevnte er gratis å bruke, men ulempen er at databasene har survivorship bias ettersom historiske prisdata for aksjer tatt av børs fjernes.

Survivorship bias kan gi store feilkilder. For eksempel dersom man tester en strategi på å kjøpe billige aksjer, og velger å kjøpe de 20 billigste aksjene på en gitt indeks. Dersom databasen bare inkluderer de aksjene som har overlevd i perioden, vil man i testen få en avkastning som ikke gjenspeiler det man faktisk ville oppnådd dersom man handlet aksjene. Dette skyldes at man ikke tar hensyn til tapene på de selskapene som enten gikk konkurs eller ble tatt av børs, og som skulle vært i det aktuelle utvalget.

Chan (2009) viser til at en metode for å minske påvirkningen av survivorship bias er å bruke ferske data, slik at man ikke mangler for mange aksjer i utvalget. Vi mener at vi har gjort dette i vår studie. I tillegg har vi et bredt utvalg slik at påvirkningen fra enkelte par ikke bør gi for store utslag på funnene våre.

5.7.2 Data snooping bias

Problemer med data snooping kan forekomme dersom man overoptimaliserer ett sett data slik at man klarer å produsere de resultatene man ønsker i datasettet. For eksempel kan man bruke altfor mange parametre i modellen, slik at den tilpasses historiske tilfeldigheter som ikke vil gjenta seg i framtiden.

Dette må, i følge Chan (2009), man være særlig oppmerksom på når man analyserer finansielle data på grunn av få relevante data. Om man tester en handlestrategi vil man i noen tilfeller ha data 70 – 80 år tilbake, mens kanskje bare de fem til ti siste år er relevante for modellen man bruker. Forklaringen kan være at regulatoriske endringer eller makroøkonomiske begivenheter gjør at man ikke kan sammenlikne data mellom periodene.

Han gir flere anbefalinger, bl.a. at man bruker modeller med få parametre og kun lineær regresjon til modellene. I tillegg mener han at optimaliseringen av strategien skjer i et eget datasett, og at resultatene testes på et nytt datasett av nyere dato.

Vi mener vi har minimert risikoen for data snooping bias ettersom vi har brukt en veldig enkel handlestrategi, samtidig som vi overhodet ikke har gjort noen optimalisering av strategien. Videre bygger vi for en stor del vår metodikk på tidligere empiriske studier.

5.7.3 Look ahead bias

Look ahead bias inntreffer dersom man lager en modell hvor et handelssignal er at man skal kjøpe en aksje på den laveste prisen en dag. Naturlig nok vil man bare i ettertid kunne fastslå hva som var den laveste prisen den dagen.

I følge Chan (2009) er løsningen for å unngå look ahead bias å introdusere en forsinkelse slik at handelen først inntreffer etter at et signal har blitt aktivert. Likeledes mener han at man ikke trenger å introdusere forsinkelser dersom man bare inngår posisjoner ved periodens slutt. Vi anser handelsdagen som en periode, og har derfor ikke lagt inn forsinkelse i signalet. I tillegg er programmet vi bruker bygget for å modellere hvordan handel foregår i praksis. De skriver i dokumentasjonen at det er bygget inn en forsinkelse fra handelssignalet utløses til handelen utføres, uten at det spesifiseres nærmere hvor lang denne er.

5.8 Oppsummering

I metodekapitlet har vi utredet alle aspektene med hvordan vi kommer frem til resultatene i studien, samt begrunnet valgene vi har gjort.

Først startet vi med å beskrive analyseverktøyet R som vi har brukt i alle beregninger i studien. Videre forklarte vi hvordan vi har valgt ut aksjer, sektorer og tidsrommet for studien. Ved utvelgelse av par har vi tatt utgangspunkt i undersektorer og testet aksjene parvis for grad av kointegrasjon. Vi endte deretter opp med 2 grupper på følgende måte: Alle aksjeparene som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe som vi ga navnet 5 % gruppen. De parene som ble signifikante på minimum 10 % nivå i alle testene ble gruppert i den andre gruppen som fikk navnet 10 % gruppen. Øvrige par ble forkastet. Videre beskriver vi handelsstrategien som baserer seg på misprising av standardavvik basert på spread. Deretter viste vi et praktisk eksempel fra et av parene i studien.

I underkapittel 5.5 presenterer vi de ulike beregningsmetodene vi har brukt til å analysere resultatene fra strategien. Disse spenner seg fra beregning av avkastning og prestasjonsmål til win/loss ratio og gjennomsnittlig holdingperiode.

Så diskuterer vi valg av kriterier i handelsstrategien. Til slutt diskuteres ulike feilkilder som kan oppstå ved å bruke historiske data som grunnlag i kvantitativ forskning.

6. Analyse

Vårt utgangspunkt er å teste hvorvidt tendensene fra tidligere studier på investeringsstrategien pairs trading vedvarer. Studier av Gatev et al. (2006) og Do & Faff (2009) viser at strategien hadde en eventyrlig avkastning på 80-tallet, mens den i fortsettelsen ga dalende avkastning, før den til slutt ikke gav noen avkastning som var signifikant annerledes enn null i perioden 2003 – 2008.

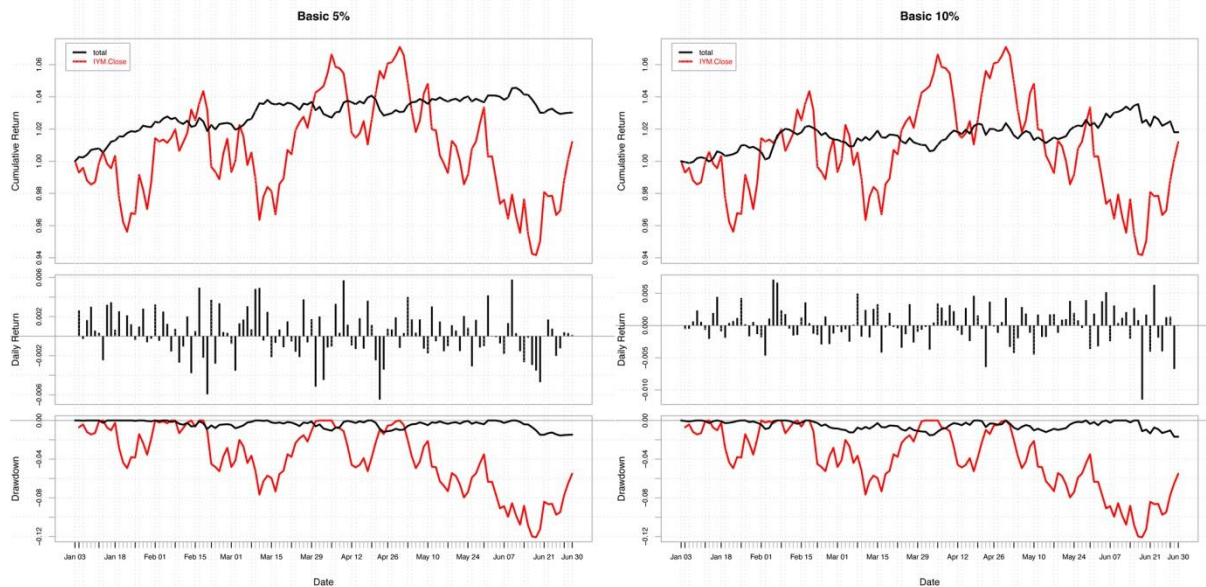
Pairs trading som investeringsstrategi er ikke ulik andre spekulative investeringsstrategier. Eksempler på slike er momentum- og reversal strategier, eller at markedet ga unormalt høy avkastning i januar måned. For mange av disse strategiene har tidlige studier funnet at de ga høy avkastning, for så å bli motbevist av senere studier.

Vi vil i vår studie forsøke å sammenlikne våre resultater opp mot studiene til Gatev et al. (2006) og Do & Faff (2009). Gatev et al. (2006) har også som en del av sin studie gitt kriterier om at parene må være selskaper i samme sektor, basert på Standard & Poors industrikategorier: Utilities, Transportation, Financial og Industrial. Dette er en veldig grov inndeling av sektorer, og dermed ikke direkte sammenlignbar med vår studie. Vi kan allikevel sammenligne porteføljen med alle aksjene mot de tidligere nevnte studier. Vi har ikke grunnlag til å sammenlikne de andre sektorene fordi vi ikke har med par fra transport- og utilitessektoren i vår studie, mens industrisektoren er alt for bred i forhold til de sektorer vi har i vår studie.

Vi har ikke noen grunn til å tro at trenden i avkastning vist gjennom studien til Do & Faff (2009) skal ha forandret seg nevneverdig de senere år. Våre forventninger er derfor å finne sammenlignbare resultater med denne studien.

6.1 Sektorer

6.1.1 Basic Materials



Figur 13 Graf, Basic Materials

	Basic Materials 5 %	Basic Materials 10 %	
Antall par	51	61	
Antall posisjoner	146	175	
Antall handler	584	700	
Gj.snitt holdingperiode	18,7	23	
Største gevinst	42,86 %	56,04 %	
Største tap	-40,44 %	-47,59 %	
Gevinst %	51,37 %	56 %	
Tap %	48,63 %	44 %	
W/L ratio	1,056	1,273	
	Basic Materials 5 %	Basic Materials 10 %	IYM
Periodeavkastning	0,0302	0,0180	0,0119
Årlig avkastning	0,0610	0,0364	0,0243
Sharpe-ratio (Årlig)	1,5759	0,7682	0,1100
Information Ratio	0,2732	0,1513	-
VaR	-0,0035	-0,0042	-0,0206
Conditional VaR (ES)	-0,0049	-0,0060	-0,0269
Standardavvik	0,0376	0,0450	0,2045
Max Drawdown	0,0155	0,0169	0,1207
Beta	-0,0432	-0,0375	1,0000

Tabell 3 Basic Materials

Sektoren Basic materials inneholder selskaper som driver med leting, utvinning og foredling av råvarer. Eksempler er olje, metall og kjemiske produkter. Vi startet her med 356 aksjer, fordelt på 17 undersektorer. Denne sektoren ble den nest største i gruppen 5 % med 51 par, mens den ble den største sektoren i gruppen 10 % med 61 par. Som referanseindeks til sektoren brukte vi indeksen Dow Jones US Basic Materials gjennom iShares ETF ticker: IYM som repliserer den nevnte indeks.

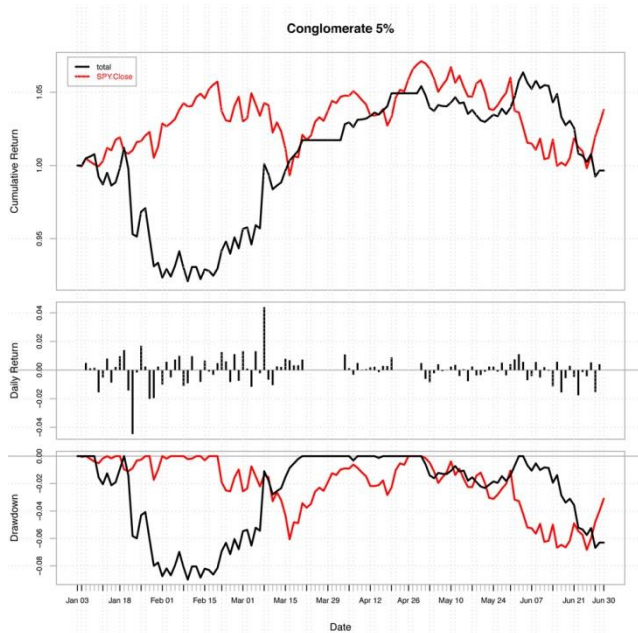
I gruppen 5 % gikk vi inn i totalt 146 posisjoner, mens handelen i 10 % gruppen resulterte i 175 posisjoner. Den førstnevnte gruppen hadde en gjennomsnittlig holdingperiode på 18,7 handelsdager, mens sistnevnte var noe høyere med en holdingperiode på 23 handelsdager.

For porteføljen i gruppen 5 % oppnådde vi i denne sektoren en periodeavkastning på 3,02 %, noe som tilsvarer en årlig avkastning på 6,1 %. I porteføljen for gruppen 10 % ble resultatet litt magrere med en periodeavkastning på 1,8 % som tilsvarer en årlig avkastning på 3,64 %. Avkastningen for begge gruppene er derimot bra når man sammenlikner med resultatene til indeksen som bare oppnådde en periodeavkastning på 1,19 %, tilsvarende årlig avkastning på 2,43 %.

Strategien oppnådde som forventet en beta nær null. Dersom man legger kapitalverdimodellen til grunn har vi dermed en portefølje som er uavhengig av volatiliteten i markedet. Når det gjelder porteføljenes risiko uttrykt gjennom standardavvik er de langt lavere enn indeksens standardavvik på 0,2045. Porteføljene i gruppen 5 % og 10 % fikk et standardavvik på henholdsvis 0,0376 og 0,045. Value at Risk (VaR) uttrykker hva det minste forventede tapet vil være over et gitt tidsintervall. Ettersom vi valgte 5 % som signifikansnivå og bruker daglige data, uttrykker VaR at vi i gjennomsnitt kan forvente et tap på minimum 2,06 % en gang hver 20 dag dersom vi holder markedsporteføljen. Ettersom porteføljene i vår studie holder både kjøps- og salgsposisjoner som skal eliminere markedsrisiko er det derfor ikke overraskende at VaR er lavere. I porteføljen til gruppen 5 % ble VaR 0,35 %, mens 10 % gruppens VaR ble 0,42 %.

Som en konsekvens av lav risiko og bra avkastning oppnår vi naturlig nok gode prestasjonsmål. Sektorindeksen som vi måler oss mot oppnådde en Sharpe ratio på 0,11. Vi ser derimot at for gruppen 5 % ble Sharpe ratio 1,58, og noe lavere for gruppen 10 % med 0,77. På samme måte viser også Information Ratio at begge porteføljene presterte bra med en IR på hhv. 0,27 og 0,15.

6.1.2 Conglomerates



Figur 14 Graf, Conglomerates

Conglomerates 5 %	
Antall par	2
Antall posisjoner	6
Antall handler	24
Gj.snitt holdingperiode	26,5
Største gevinst	6,40 %
Største tap	-5,39 %
Gevinst %	50 %
Tap %	50 %
W/L ratio	1

	Conglomerates 5 %	S&P 500
Periodeavkastning	-0,0035	0,0380
Årlig avkastning	-0,0070	0,0772
Sharpe-ratio (Årlig)	-0,0617	0,5926
Information Ratio	-0,4473	-
VaR	-0,0151	-0,0164
Conditional VaR (ES)	-0,0212	-0,0190
Standardavvik	0,1429	0,1273
Max Drawdown	0,0900	0,0683
Beta	-0,0798	1,0000

Tabell 4 Conglomerates

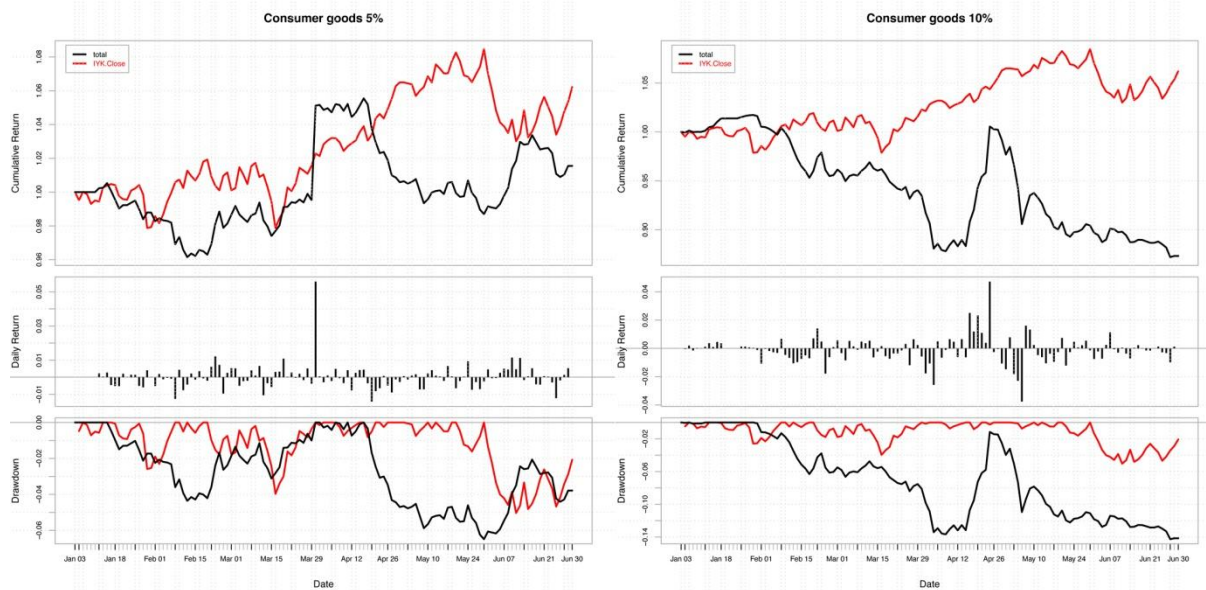
Sektoren Conglomerates består av selskaper som driver med flere ulike forretningsområder. På grunn av manglende referanseindeks til denne sektoren har vi sammenlignet resultatene i tabellen med S&P 500.

Vi testet 10 aksjer som resulterte i 2 par som falt inn under 5 % gruppen. Det ble kun tatt 6 posisjoner i perioden med like mange tap og gevinster, med en gjennomsnittlig holdingperiode på 26,5 dager. Den største gevinsten var 6,4 %, mens det største tapet ble 5,39 %. Periodeavkastningen ble -0,35 % som tilsvarer en årlig avkastning på -0,7 %. Dette er langt lavere enn avkastningen til referanseindeksen som oppnådde en årlig avkastning på 7,72 % . Prestasjonsmålene ble følgelig dårlige med en Sharpe ratio på -0,0617.

Få handler gjenspeiles ved stor volatilitet som man kan se av avkastningsgrafene, samt maksimal drawdown på 9 % som er forskjellen på høyeste og etterfølgende laveste punkt i avkastningskurven. Naturlig nok får vi derfor også i denne porteføljen det høyeste standardavviket i studien. Dette bekreftes også med en høy VaR på 1,51 %, bare marginalt lavere enn indeksens VaR.

Et viktig resultat fra denne sektoren er at beta for porteføljen ligger rundt null som indikerer at avkastningen er uavhengig i forhold til markedet selv ved et lite antall par i porteføljen.

6.1.3 Consumer goods



Figur 15 Graf, Consumer goods

	Consumer goods 5 %	Consumer goods 10 %	
Antall par	6	7	
Antall posisjoner	15	18	
Antall handler	60	72	
Gj.snitt holdingperiode	28,6	26,2	
Største gevinst	29,37 %	9,69 %	
Største tap	-26,09 %	-32,44 %	
Gevinst %	53,33 %	27,78 %	
Tap %	46,67 %	72,22 %	
W/L ratio	1,143	0,385	
	Consumer goods 5 %	Consumer goods 10 %	IYK
Periodeavkastning	0,0155	-0,1267	0,0622
Årlig avkastning	0,0313	-0,2553	0,1264
Sharpe-ratio (Årlig)	0,2648	-1,7319	1,1831
Information Ratio	-0,6565	-1,9566	-
VaR	-0,0081	-0,0142	-0,0109
Conditional VaR (ES)	-0,0110	-0,0221	-0,0145
Standardavvik	0,1114	0,1485	0,1053
Max Drawdown	0,0649	0,1427	0,0503
Beta	-0,0394	-0,0583	1

Tabell 5 Consumer goods

Sektoren Consumer goods inneholder selskaper innen bransjene klær og sko, tobakk, leker samt motoriserte kjøretøy. Som referanseindeks til sektoren har vi benyttet oss av iShares ETF ticker: IYK som repliserer indeksen Dow Jones US Consumer Goods.

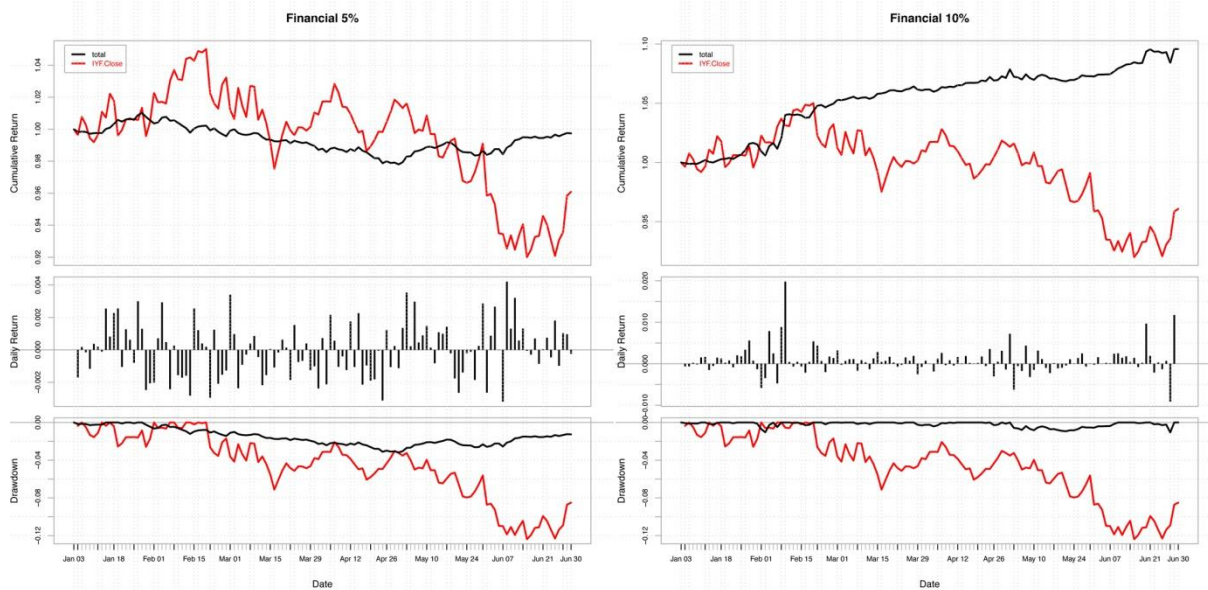
Samlet testet vi 25 aksjer innenfor 4 undersektorer. I 5 % gruppen fant vi 6 par som resulterte i 15 posisjoner i løpet av perioden. For 10 % gruppen fant vi 7 par som resulterte i 18 posisjoner. Gjennomsnittlig holdingperiode per par var henholdsvis 28,6 og 26,2 handelsdager for de to porteføljene. En stor forskjell mellom gruppene er antall gevinster der 5 % gruppen oppnår positiv avkastning i 53,33 % av posisjonene, mens 10 % gruppen kun oppnår positiv avkastning i 27,78 % av posisjonene.

Ser vi på grafene viser den at 5 % gruppen tenderer svakt i positiv retning mens 10 % gruppen tenderer i negativ retning. Periodeavkastningen ble 1,55 % for 5 % gruppen og – 12,67 % for 10 % gruppen hvor sistnevnte representerer den svakeste prestasjonen i studien. Begge gruppene presterer langt dårligere enn referanseindeksen som oppnår en periodeavkastning på 6,22 %. Naturlig nok blir dermed prestasjonsmålene til porteføljene lave. Sett i forhold til referanseindeksen viser det seg at indeksen både ga bedre avkastning samt lavere risiko uttrykt i standardavvik enn de to gruppene.

Standardavviket i 5 % gruppen er 0,1114, mens for 10 % gruppen ble det noe høyere med 0,1485 slik at førstnevnte presterte klart best av de to. Videre er VaR 0,0081 for gruppen 5 %, mens den er 0,0142 for gruppen 10 %. Max drawdown bekrefter også at det i denne sektoreren var stor volatilitet med henholdsvis 6.49 % og 14.27 % hvor sistnevnte er det høyeste i studien.

Selv om begge gruppene viser ulike prestasjoner er det viktig å påpeke at begge gruppene indikerer relativt større risiko sammenlignet med sektorer hvor vi har flere par i porteføljen. Det kan derfor tyde på at det er en diversifiseringsgevinst ved å holde flere par i en portefølje. Dette er i tråd med studien til Gatev et al. (2006) som også finner at standardavviket til en portefølje faller når man øker antall par i porteføljen.

6.1.4 Financials



Figur 16 Graf, Financials

	Financials 5 %	Financials 10 %	
Antall par	53	49	
Antall posisjoner	139	137	
Antall handler	556	548	
Gj.snitt holdingperiode	30,6	32,2	
Største gevinst	28,11 %	184,62 %	
Største tap	-21,60 %	-22,96 %	
Gevinst %	56,12 %	65,69 %	
Tap %	43,88 %	34,31 %	
W/L ratio	1,279	1,915	
	Financials 5 %	Financials 10 %	IYF
Periodeavkastning	-0,0026	0,0958	-0,0392
Årlig avkastning	-0,0053	0,1931	-0,0796
Sharpe-ratio (Årlig)	-0,2763	3,7643	-0,5266
Information Ratio	0,5210	1,8228	-
VaR	-0,0025	-0,0029	-0,0180
Conditional VaR (ES)	-0,0028	-0,0051	-0,0231
Standardavvik	0,0256	0,0508	0,1546
Max Drawdown	0,0318	0,0105	0,1238
Beta	-0,0023	-0,0150	1

Tabell 6 Financials

Financials er en sektor som består av selskaper innenfor bransjene bank, finans og forsikring.

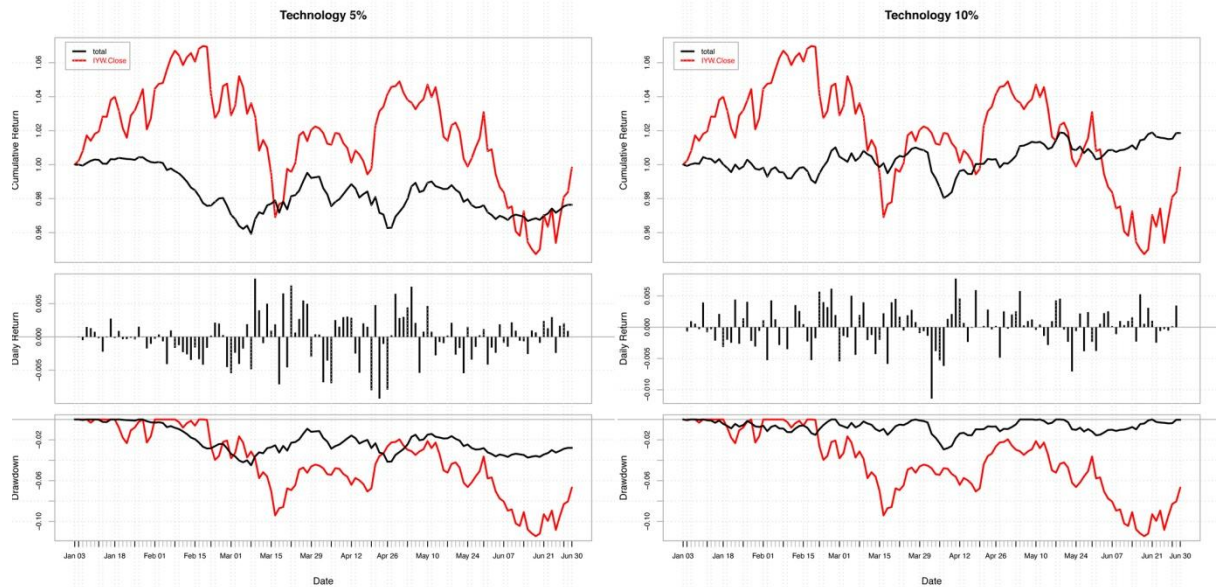
Samlet testet vi 384 aksjer fordelt på 30 undersektorer. I 5 % gruppen fant vi 53 par som resulterte i 139 posisjoner i løpet av perioden. For 10 % gruppen fant vi 49 par som resulterte i 137 posisjoner. Gjennomsnittlig holdingperiode per par var henholdsvis 30,6 og 32,2 dager. Som referanseindeks har vi brukt iShares ETF ticker: IYF som repliserer indeksen Dow Jones US Financial Sector.

Periodeavkastningen ble -0,26 % for 5 % gruppen og 9,58 % for 10 % gruppen hvor sistnevnte presterte best i studien. Standardavviket er meget lavt for begge gruppene på 0,0256 og 0,0508. Man ser av grafen at 5 % gruppen tenderer svakt negativt etter toppen i starten av perioden. 10 % gruppen viser derimot en sterk positiv trend igjennom hele perioden, dog med noen marginale drawdowns underveis. Denne kombinasjonen av en kontinuerlig positiv trend kombinert med svært lav volatilitet er et prakt eksemp lar på hvordan en god handelsstrategi skal fungere. Win/loss ratio er positiv for begge gruppene selv om 10 % gruppen skiller seg ut med gevinst i hele 65.69 % av posisjonene.

Ettersom referanseindeksen presterte dårligere enn begge gruppene ble også prestasjonsmålene til begge gruppene bedre enn indeksens. Det er mest interessant å se på 10 % gruppens prestasjonsmål ettersom denne presterte best i studien. Vi oppnådde her en Sharpe ratio på 3,76 og en Information Ratio på 1,82.

Et funn er at 10 % gruppen som har prestert klart best i studien, også er den gruppen med høyest gjennomsnittlig holdingperiode per posisjon. Forklaringen på dette kan være så enkel som at kursbevegelsene i denne gruppen passer handelsstrategien godt, som igjen medfører at færre posisjonene blir løst inn ved 3 standardavvik i forhold til de andre gruppene, samtidig som at vinnerposisjoner holdes lenger. I utgangspunktet er det naturlig å tro at gruppen med 5 % er mer kointegrerte som igjen betyr en bedre reversaleffekt gjenspeilet ved en lavere holdingperiode. Problemet som oppstår ved hurtig reversal er man mister noe av avkastningen, fordi posisjonen også lukkes raskere på bakgrunn av at indikatoren og handlere gelen består av standardavvik og glidende gjennomsnitt. Man kan derfor argumentere for at 5 % gruppen kutter gevinstposisjonene for tidlig, mens 10 % gruppen holder dem lenger. Dette er en refleksjon over hvorfor 10 % har lenger holding periode og bedre avkastning enn 5 % gruppen, men om dette er korrekt er vanskelig å bevise.

6.1.5 Technology



Figur 17 Graf, Technology

	Technology 5 %	Technology 10 %
Antall par	49	44
Antall posisjoner	137	124
Antall handler	548	496
Gj.snitt holdingperiode	24,2	26,5
Største gevinst	32,89 %	26,33 %
Største tap	-47,79 %	-47,01 %
Gevinst %	52,55 %	55,65 %
Tap %	47,45 %	44,35 %
W/L ratio	1,108	1,255

	Technology 5 %	Technology 10 %	IYW
Periodeavkastning	-0,0236	0,0186	-0,0017
Årlig avkastning	-0,0476	0,0376	-0,0034
Sharpe-ratio (Årlig)	-0,9469	0,7237	-0,0330
Information Ratio	-0,1917	0,3276	-
VaR	-0,0054	-0,0053	-0,0186
Conditional VaR (ES)	-0,0074	-0,0066	-0,0228
Standardavvik	0,0522	0,0495	0,1589
Max Drawdown	0,0449	0,0294	0,1145
Beta	-0,0086	0,0263	1,0000

Tabell 7 Technology

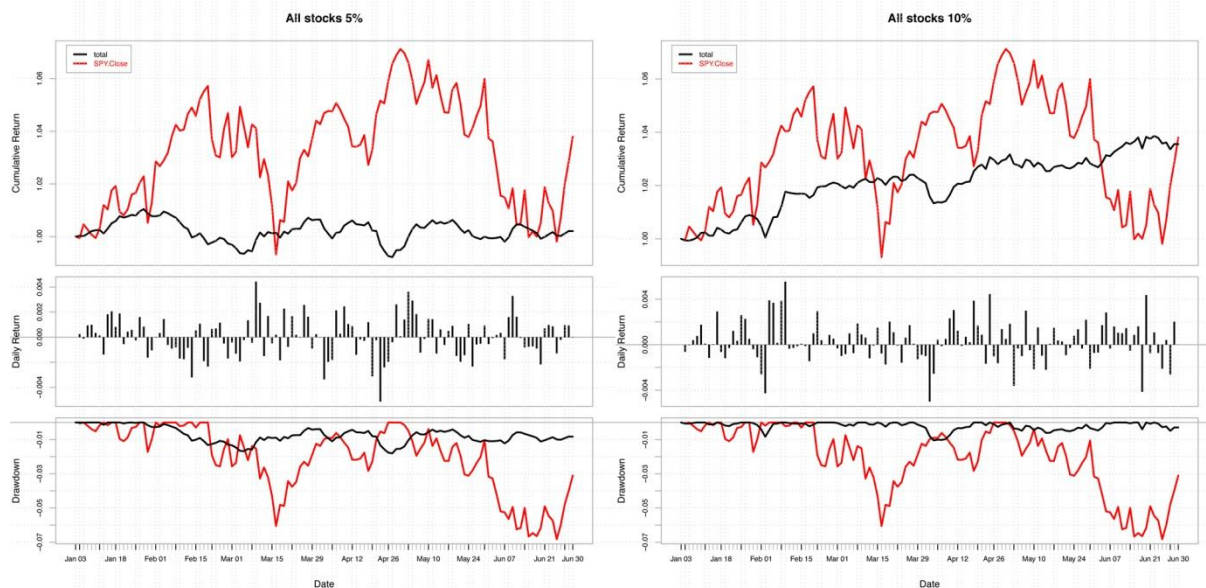
Technology består av teknologibedrifter som blant annet arbeider med programvare, IT og kommunikasjon. Som referanseindeks har vi brukt iShares ETF ticker: IYW som repliserer indeksen Dow Jones US Technology.

Totalt sett testet vi 464 aksjer som er fordelt på 30 undersektorer. I 5 % og 10 % gruppen fant vi henholdsvis 49 og 44 aksjepar som medførte totalt 137 og 124 posisjoner i perioden. 10 % gruppen presterte klart best med positiv avkastning på 1,86 %, mens 5 % gruppen måtte ta et tap på -2,36 %. Prestasjonsmålene Sharpe- og Information ratio gjenspeiler dette med svakt positiv for 10 % gruppen og svakt negativ for 5 % gruppen. Man kan se av tabellen at 5 % gruppen inneholdt det paret som ga høyest avkastning på 32,89 % mens 10 % gruppen viser seg mer konsistent med en høyere win/loss ratio. Verdt å merke seg inneholdt begge gruppene parene med høyest tap i hele studien med henholdsvis -47,79 % og -47,01 %.

Strategien hadde som forventet lavere risiko enn referanseindeksen. Vi ser at standardavviket til 5 % gruppen er 0,0522, mens det bare er marginalt lavere for 10 % gruppen med 0,0495. Dette er bare en drøy tredjedel av standardavviket til referanseindeksen. VaR var på henholdsvis 0,0054 og 0,0053 for 5 % og 10 % gruppen. Maksimal drawdown for 5 % gruppen ble på hele 4,49 %, mens 10 % gruppen er lavere med 2,94 %. Alt i alt ser det ut til at begge gruppene har relativt lav risiko sammenliknet med referanseindeksen. Et godt eksempel på dette er å se på grafen med indeksen og 10 % gruppen fra midten av perioden. Her ser vi at 10 % gruppene viser en stabil positiv trend med liten volatilitet, mens indeksen viser klart høyere volatilitet.

Videre ser man at begge gruppene viser en beta rundt null, som antyder at avkastningen i stor grad er uavhengig av markedsutviklingen.

6.1.6 Alle parene



Figur 18 Graf, alle parene

	Alle 5 %	Alle 10 %	Alle 10 % - Financials 10 %	
Antall par	161	161	112	
Antall posisjoner	443	454	317	
Antall handler	1772	1816	1268	
Gj.snitt holdingperiode	19,6	23,4	23,3	
Største gevinst	42,86 %	184,62 %	56,04 %	
Største tap	-47,79 %	-47,59 %	-47,59 %	
Gevinst %	53,27 %	57,71 %	54,26 %	
Tap %	46,73 %	42,29 %	45,74 %	
W/L ratio	1,140	1,365	1,186	
	Alle 5 %	Alle 10 %	Alle 10 % - Financials 10 %	S&P 500
Periodeavkastning	0,0021	0,0356	0,0092	0,0380
Årlig avkastning	0,0042	0,0717	0,0186	0,0772
Sharpe-ratio (Årlig)	0,0966	2,4849	0,5253	0,5926
Information Ratio	-0,5035	0,0221	-0,3971	-
VaR	-0,0023	-0,0025	-0,0027	-0,0164
Conditional VaR (ES)	-0,0031	-0,0035	-0,0042	-0,0190
Standardavvik	0,0245	0,0281	0,0320	0,1273
Max Drawdown	0,0180	0,0105	0,0167	0,0683
Beta	-0,0393	-0,0245	-0,0215	1,0000

Tabell 8 Alle parene

Deskriptiv statistikk

Det viser seg at pairs trading er en veldig handelsintensiv investeringsstrategi. Samlet sett fant vi i gruppen 5 % 161 par som ifølge våre tester egnet seg for handel. I perioden gikk vi inn i til sammen 443 posisjoner. Ettersom hver posisjon består av fire handler, gjorde vi dermed 1772 handler. Den gjennomsnittlige holdingperioden for hver posisjon var 19,6 handelsdager. Fordelingen av antall gevinster og tap var rimelig jevnt fordelt med en win/loss ratio på 1,14. Samtidig var største gevinst 42,86 %, mens største tap ble -47,79 %.

Det var ikke store forskjeller i antall posisjoner inngått mellom gruppen 5 % og gruppen 10 %. I den sistnevnte gruppen åpnet vi 454 posisjoner, som tilsvarer 1816 handler. Det kan derimot virke som at den 10 % gruppen har dårligere mean-reverting egenskaper, ettersom den gjennomsnittlige holdingperioden er 23,4 handelsdager. Dette representerer en økning på nesten fire handelsdager i denne gruppen i forhold til gruppen 5 %. Videre synes det som at forholdet mellom antall gevinster og tap har økt i porteføljen til 10 % gruppen som følge av at win/loss ratio er økt til 1,365. Samtidig er det største tapet på en posisjon 47,59 %, omtrent det samme som i 5 % gruppen. Den største gevinsten er på hele 184,6 %.

Den opplagte forskjellen denne studien sammenliknet med Gatev et al. (2006) er hvor lenge posisjoner er åpne og gjennomsnittlig antall åpning av posisjoner per par. I deres studie fant de i gjennomsnitt at den gjennomsnittlige holdingperioden på et par var 3,97 måneder, mer i tråd med en "medium term" investeringsstrategi. I vår studie er den gjennomsnittlige holdingperioden på rundt en måned i gjennomsnitt, som må kunne anses som en kort tidshorisont. Videre har vi i studien samlet 322 par i begge gruppene, med totalt 897 posisjoner åpnet i perioden. Vi kan derfor si at et par gjennomsnittlig gjorde 2,79 "roundtrips", noe som er markant høyere enn 1,62 "roundtrips" som Gatev et al. (2006) finner for alle parene i hans portefølje.

Bakgrunnen for forskjellene kan være basert i handelsindikatoren. Gatev et al. (2006) bruker i sin studie en handelsindikator basert på historisk standardavvik på spreaden som blir estimert i oppbyggingsfasen av hvert enkelt par. Ved at vi bruker glidende gjennomsnitt på 50 dager resulterer nok dette i at vår indikator gir flere handelssignaler.

Avkastning og prestasjonsmål

I porteføljen med alle aksjene i gruppen 5 % oppnådde vi en periodeavkastning på 0,21 %, noe som tilsvarer en årlig avkastning på 0,42 %. Ettersom vi oppnådde så lav avkastning, gjenspeiles dette også i prestasjonsmålene Sharpe ratio og Information ratio, som begge er betydelig lavere enn de samme forholdstallene til referanseindeksen.

Vi oppnådde overraskende resultater i porteføljen med alle parene i gruppen 10 %.

Periodeavkastningen ble 3,56 %, godt hjulpet av bra resultater fra de finansielle aksjene.

Dette tilsvarer en årlig avkastning på 7,17 %. Med tanke på at referanseindeksen oppnådde en årlig avkastning på 7,72 % var dette et veldig bra resultat.

Ettersom vi oppnådde en betydelig avkastning med lav risiko i forhold til referanseindeksen ble naturlig nok prestasjonsmålene bra. Sharpe ratio på 2,48 på porteføljen i motsetning til 0,59 på referanseindeksen. Information ratio var også positiv, noe som indikerer at vi oppnådde meravkastning.

For å sammenlikne med tidligere studier, ser vi at Gatev et al. (2006) finner en gjennomsnittlig månedlig avkastning på 0,614 % for alle parene i sin studie, dersom man legger den mest konservative måten å regne ut avkastningen på til grunn. Dette tilsvarer en gjennomsnittlig årlig avkastning på 7,37 % i perioden 1963 til 2002. Do & Faff (2009) finner på sin side en gjennomsnittlig månedlig avkastning på 0,07 % i perioden 2003 til 2008, som i følge dem ikke er signifikant annerledes enn null. Dette tilsvarer en årlig avkastning på 0,84 %.

Vi ser fra våre resultater at avkastningen i gruppen 5 % er mest i tråd med Do & Faff (2009), mens resultatene i gruppen 10 % er sammenliknbar med funnene til Gatev et al. (2006).

Forskjellen skyldes i hovedsak avkastningen til porteføljen med parene i sektoren Financials i gruppen 10 % som oppnådde en årlig avkastning på 19,31 %. I vår studie var dette den porteføljen som gjorde det aller best. Ettersom vi bare har en periode i vår studie, er det vanskelig å trekke noen konklusjoner om hvorvidt en slik avkastning er representativ, eller om det dreier seg om tilfeldigheter. Porteføljen skiller seg så klart fra de øvrige funnene våre slik at det kan være nyttig å se hvilke resultater man oppnår uten denne.

Dersom vi ser bort fra denne porteføljen vil de øvrige parene i gruppen 10 % gi en årlig avkastning på 1,86 % og en Sharpe ratio på 0,5253. Dette er altså langt lavere enn hva Gatev et al. (2006) sine estimater, men fremdeles drøye ett prosentpoeng høyere enn Do & Faff

(2009). Samtidig er det viktig å huske på at avkastningen vil variere fra periode til periode. Selv i studien til Do & Faff (2009) varierer den månedlige avkastningen fra et minimum på -2,72 %, til et maksimum på 3,19 %. Dette kan tyde på at vi ikke har noe grunnlag for å konkludere med at funnene i vår studie skiller seg fra deres.

Risikoparametre

Det kan virke som at vi oppnådde en markedsnøytral strategi i porteføljen 5 % ettersom beta ikke er langt fra null. Når vi ser på porteføljens risikoparametre finner vi et veldig lavt årlig standardavvik på 0,0245, mens VaR og Conditional VaR er hhv. -0,0023 og -0,0031. VaR impliserer at gjennomsnittlig vil vi en gang hver 20 dag tape minimum 0,23 % av porteføljen. Det største daglige tapet i perioden for denne porteføljen ble 0,51 %.

Maksimal drawdown er 0,018, noe som også er lavere enn referanseindeksen. Strategien innebærer dermed mindre volatilitet i forhold til å holde markedsporteføljen, noe vi også forventet.

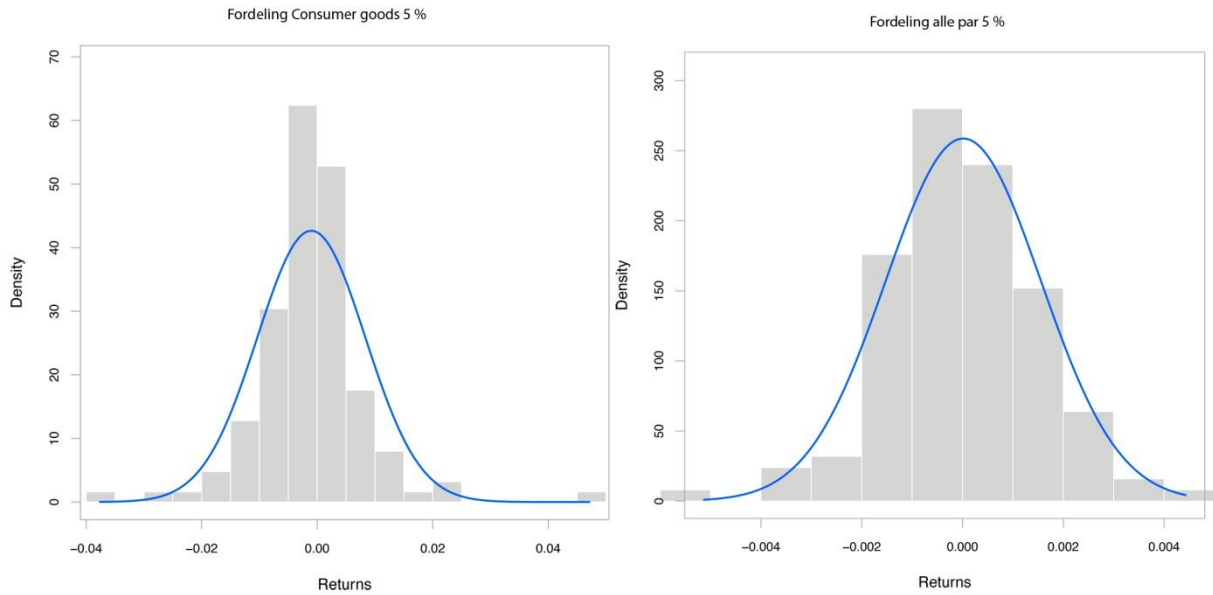
Også i porteføljen 10 % virker det som vi var markedsnøytrale med en beta nær null. Samtidig var det årlige standardavviket på 0,0281, mens VaR og Conditional var hhv. 0,0025 og 0,0035. Maksimal drawdown ble i denne porteføljen 0,0105.

Gatev et al. (2006) finner i sin studie at den daglige Value at Risk (5 %) i porteføljen for alle parene er -0,00202. Dette stemmer godt overens med vår studie.

Fordelingen til avkastningen

Et interessant funn fra vår studie er fordelingen til avkastningen. I Gatev et al.'s (2006) studie finner de en skewnesskoeffisient på 3,42 og kurtosiskoeffisient på 25,25 på fordelingen til alle parene ved handel samme dag som handelssignalet utløses. I Do & Faff (2009) sin studie finner de tilsvarende en skewnesskoeffisient på 0,41 og kurtosiskoeffisient på 6,93. Fordelingen til avkastningen i de to studiene har altså en høy kurtosiskoeffisient.

Vi finner like tendenser i vår studie der vi har få aksjer i porteføljen. For eksempel har fordelingen til avkastningen til porteføljen Consumer goods i gruppen 5 % en positiv skewness på 0,61 og kurtosis på 6,67.



Figur 19 Fordeling

Ser vi derimot på fordelingen til porteføljene som inneholder alle parene i en gruppe virker fordelingen til avkastningen å nærme seg normalfordelingen. Avkastningen til porteføljen med alle parene innenfor 5 % gruppen har en skewnesskoeffisient på $-0,073$ og en kurtosiskoeffisient på $0,42$, noe som er nært normalfordelt. Dette gjør at man ikke trenger å gjøre vurderinger hvorvidt utregninger vi har gjort som forutsetter normalfordeling er representative. Som et eksempel nevner Gatev et al. (2006) at Sharpe ratio kan gi et feilaktig bilde dersom man har en distribusjon med et negativt tredjemoment.

Han finner videre at selv om man øker antall par i en portefølje, har dette egenskapen at minimumsavkastningen øker mens maksimumsavkastningen holder seg relativt stabil. Dette representert med en økende skewnesskoeffisient. Basert på data fra vår studie finner vi at det motsatte er tilfelle, nemlig en avtagende skewnesskoeffisient, slik at vi kan ikke konkludere på samme måte som Gatev et al. (2006).

6.2 Transaksjonskostnader

Når man tester investeringsstrategier vil naturligvis transaksjonskostnader ha stor innvirkning på resultatet. Dette gjelder særlig handelsintensive strategier slik som den vi har testet.

Det er ikke tatt høyde for transaksjonskostnader ved utregning av avkastning i vår studie. Vi ønsker derfor å diskutere hvorvidt strategien er robust for transaksjonskostnader. Gatev et al. (2006) hadde i sin studie to handleregler. For den første gjorde de handler samme dag som handlesignalet indikerte at de burde inngå en posisjon, mens for den andre ventet de med handelen til dagen etter når handlesignalet indikerte åpning.

De tar en konservativ tilnærming for å estimere transaksjonskostnadene. Her ser de for seg en ekstrem tilnærming ved at når man åpner posisjonene er den relative vinneren ask pris og taperen er bid pris. Ved å vente en dag er det lik sannsynlighet for at prisene skal være henholdvis bid eller ask, og dette vil da redusere avkastningen med halvparten av spreaden til kjøps- og salgsposisjon. Når parene konvergerer forutsetter de at vinneren er bid, mens taperen er ask. Ved igjen å vente en dag vil man redusere avkastningen med halvparten av bid-ask spreaden til aksjene. Dette impliserer at den reduserte avkastningen med å vente en dag transaksjonskostnadene til et par.

De fant i studien at periodeavkastningen falt med 324 basispoeng dersom man ventet en dag. Ut fra dette mente de at forskjellene i avkastningen gjenspeilet gjennomsnittlig bid-ask spread som igjen kan forklares som transaksjonskostnader.

Siden de i gjennomsnitt inngikk to posisjoner per par i en seks måneders periode, impliserer dette ifølge dem en gjennomsnittlig transaksjonskostnad per posisjon på 162 basispoeng, eller en estimert bid-ask spread på 81 basispoeng. Dette var høyere enn estimer funnet i tidligere studier av Peterson and Fialkowski (1994) i Gatev et al. (2006) som fant en spread på 37 basispoeng. Av denne grunn mente Gatev et al. (2006) at deres konservative estimering i verste fall er for høy. Ved å se på studiene vi tidligere refererte til kan det virke som at transaksjonskostnadene for en institusjonell aktør ligger et sted mellom 74 og 162 basispoeng per par.

Metoden Gatev et al. (2006) brukte for å vurdere hvorvidt strategien var robust for transaksjonskostnader var dermed å ta den gjennomsnittlige seks måneders fully invested avkastningen og trekke transaksjonskostnadene fra denne. I studiet fant de en avkastning

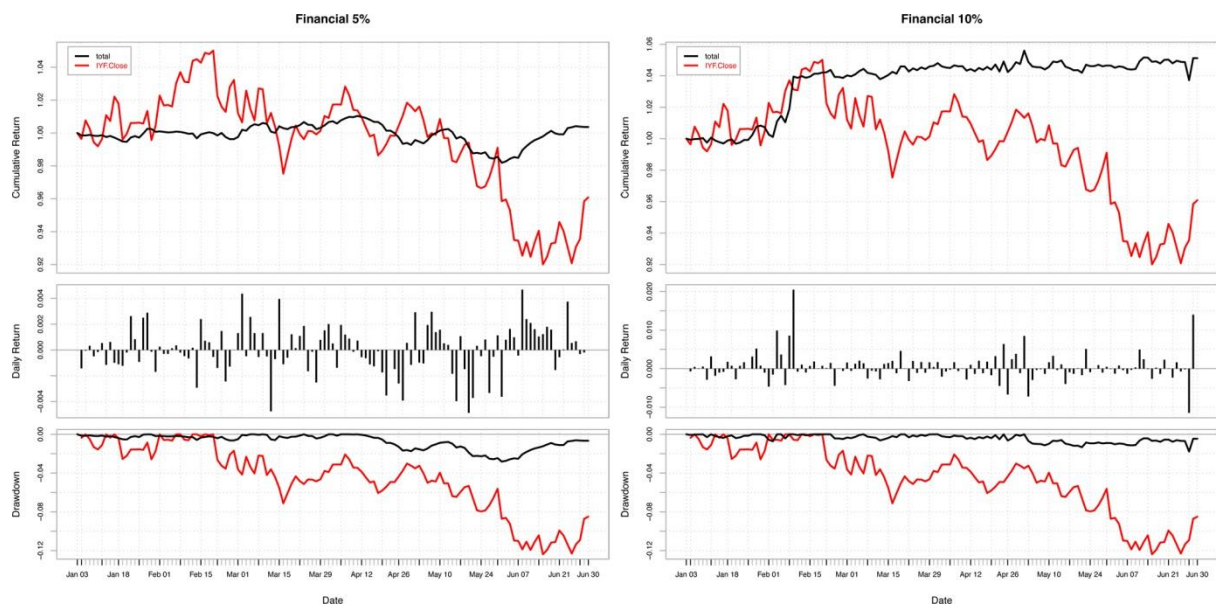
som varierte fra 437 til 549 basispoeng som de trakk fra de 324 basispoengene de fant at avkastningen falt med. På bakgrunn av at de fremdeles fikk et overskudd på 113 til 225 basispoeng konkluderte de med at deres strategi var robust for transaksjonskostnader.

I vår studie inngikk vi gjennomsnittlig 2,8 posisjoner i de seks månedene vi handlet. Dersom vi legger Gatev et al. (2006) til grunn med 162 basispoeng i transaksjonskostnad per posisjon, vil den gjennomsnittlige transaksjonskostnaden vår være 453,6 basispoeng eller 4,536 %. Etter å ha gått gjennom alle parene viser det seg at det bare finnes ett par som ikke handlet i perioden. Av denne grunn vil dermed avkastningen vi kom fram til, være svært lik begge målene på avkastningen til Gatev et al. (2006). Ettersom alle parene i porteføljen er vektet likt, kan vi forstå periodeavkastningen som gjennomsnittlig avkastning per par i perioden. For at avkastningen vi fant i studien vår skal være robust for transaksjonskostnader, må dermed periodeavkastningen være minimum 4,536 %. Vi finner i vår studie bare en portefølje som presterer bedre enn dette, nemlig porteføljen til gruppen 10 % i sektoren Financials. Det kan derfor virke som at vår strategi totalt sett ikke er robust for transaksjonskostnader.

Dermed blir spørsmålet hvordan man kan gjøre strategien robust for transaksjonskostnader, noe vi vil diskutere i neste kapittel.

6.2.1 Sensitivitetsanalyse transaksjonskostnader

Intuitivt kan man tenke seg at dersom en gjør færre handler vil transaksjonskostnadene også bli lavere. Vi har i kapittel 4.6.1 diskutert valg av lengde på glidende gjennomsnitt i indikatoren. Ved å velge et lengre tidsrom for beregning av glidende gjennomsnitt, ser vi at holdingperioden per par blir lengre, som igjen reduserer hyppigheten av handler. En annen måte vil være å vente med å inngå posisjoner til aksjene divergerer til et større antall standardavvik enn det vi bruker. Man vil da teoretisk sett få en større gevinst per par som igjen vil dekke inn transaksjonskostnadene.



Figur 20 Graf, Financials SMA150

	Financials 5% SMA150	Financials 10% SMA150
Periodeavkastning	0,0037	0,0512
Antall posisjoner	82	71
Antall par	53	49
Roundtrips	1,55	1,45
Holdingsperiode	40,0	42,8

Tabell 9 Financials SMA150

I studien fant vi at gjennomsnittlige transaksjonskostnader per par var 453,6 basispoeng, dette med utgangspunkt i estimeringen av transaksjonskostnader i studien til Gatev et al. (2006). Vi har forsøkt å øke lengden på glidende gjennomsnitt for å se hvilken effekt dette vil få for transaksjonskostnadene. I eksempelet vist i tabell 9, har vi brukt 150 dagers glidende gjennomsnitt, dette tilsvarer i overkant av 7 måneder med handelsdager. Som forventet øker holdingsperioden per par markant. Vi får her en gjennomsnittlig holdingsperiode påhenholdsvis 40 og 42,8 handelsdager for gruppene 5 % og 10 %, mot tidligere hhv. 30,6 og 32,2 dager. Dette tilsvarer en økning i holdingsperioden på rundt 33 % for begge gruppene.

Som en konsekvens av lengre holdingsperiode, faller også gjennomsnittlig antall roundtrips per par i perioden. Ved å bruke 50 dagers glidende gjennomsnitt i indikatoren hadde vi for denne sektoren i gjennomsnitt 2,6 og 2,9 roundtrips per par i perioden for gruppen 5 % og 10 %, ikke ulikt gjennomsnittet vi fant for alle parene. Som vi ser i tabell 9 har antallet

roundtrips falt markant til henholdsvis 1,55 og 1,45 i perioden ved å øke glidende gjennomsnitt til 150 dager. Det impliserer at transaksjonskostnadene faller fra 453,6 basispoeng til 243 basispoeng, altså en reduksjon på 44 %. Dette er stemmer bra overens med antall handler vi gjør, ettersom vi gjør 41 % færre handler i 5 % gruppen mot 48 % færre handler i 10 % gruppen.

Dersom holdingperioden øker vil ceteris paribus, avkastningen øke ettersom transaksjonskostnadene reduseres. Dessverre er det ikke så enkelt, ettersom vi får en annen avkastning når vi endrer indikatoren. For gruppen 10 % får vi en periodeavkastning på 5,12 %. Trekker vi fra transaksjonskostnadene får vi en nettoavkastning på 2,69 %. Hvis vi ser på eksempelet med 50 dagers glidende gjennomsnitt har vi en periodeavkastning på 9,58 %. Med transaksjonskostnader estimert til 453,6 basispoeng, tilsvarer dette en nettoavkastning på 5,04 % som er bedre enn i tilfellet med 150 dagers glidende gjennomsnitt.

Allikevel må vi si at forskjellen mellom de to transaksjonskostnadene på 210,6 basispoeng er stor. Hvis man tenker seg at man har en langsiktig investeringshorisont vil en årlig besparelse i transaksjonskostnader på 421,2 basispoeng kunne utgjøre forskjellen mellom tap og gevinst.

6.3 Oppsummering

Vi vil i dette kapitlet gjøre en oppsummering av de viktigste funnene i vår studie.

Antall par i porteføljene har variert ettersom hvilken sektor de tilhører. Porteføljen til sektoren Conglomerates inneholdt bare to par, men vi oppnådde allikevel en beta nær null. Dette kan tyde på at pairs trading er en markedsnøytral strategi uavhengig av antall par i porteføljen. Videre har vi funnet indikasjoner på at det er en diversifiseringsgevinst ved å øke antall par i en portefølje, ettersom både porteføljens standardavvik og VaR reduseres når man introduserer flere par i porteføljen. Investeringsstrategien har for øvrig lav risiko sammenliknet med å holde referanseindeksen målt i både standardavvik og VaR.

Ved å bruke den handelsindikatoren vi presenterte i kapittel 4.4 resulterer dette i en veldig handelsintensiv strategi. Vi ser at holdingperioden ligger rundt 20 handelsdager. For gruppen 10 % fant vi at gjennomsnittlig holdingperiode var lengre enn for 5 % gruppen. Den korte holdingperioden fører naturlig nok til flere handler, noe som øker transaksjonskostnadene. Ved å øke lengden på glidende gjennomsnitt i indikatoren fra 50 dager til 150 dager ble transaksjonskostnadene redusert med 44 %.

For porteføljen med alle parene i gruppen 5 % oppnådde vi i studiet nær nullavkastning, noe som stemmer overens med resultater fra studien til Do & Faff (2009), mens i gruppen 10 % ble periodeavkastningen 3,56 %.

7. Avslutning

7.1 Oppsummering og konklusjon

Teori om markedseffisiens er essensielt for å forstå hvorfor det er vanskelig å oppnå unormal avkastning i aksjemarkedet. Vår tilnærming ble derfor ikke å stille spørsmålstegn med hypotesen, men å se på om avkastningen til handelsstrategien er sammenfallende med tidligere forskning.

Gatev et al. (2006) hevder i sin studie at mulighetene ved pairs trading skyldes en kompensasjon til arbitrasjører for å håndheve "The Law of One Price". Arbitrasje er derfor ikke ment til å slå markedet, men å oppnå positiv avkastning uavhengig av markedsrisiko ved å gjenopprette likevekter.

Vårt utgangspunkt var å finne aksjer som er kointegrerte. Vi sorterte aksjeparne i to grupper på følgende måte: Alle aksjeparne som var signifikante på 1 % og 5 % nivå i Johansen testen og i tillegg var signifikante i både PP og ADF på 5 % nivå ble delt inn i en gruppe som vi ga navnet 5 % gruppen. De parene som ble signifikante på minimum 10 % nivå i alle testene ble gruppert i den andre gruppen som fikk navnet 10 % gruppen. I begge gruppene fikk vi 161 aksjepar.

Vi spekulerte derfor i at aksjepar som divergerer fra likevekten igjen vil konvergere, ved å selge den overprisede og å kjøpe den underprisede med like store posisjoner på begge sider. Handleregelen vår er basert på mispricing uttrykt i standardavvik på spreaden. Glidende gjennomsnitt var 50 dager og posisjonene åpnet på 2 standardavvik. Vi lukket posisjonene ved konvergering til likevekten og på slutten av handelsperioden, eller hvis aksjene divergerte utover 3 standardavvik.

Problemstillingen vår er som følger:

"Gir handelsstrategien pairs trading meravkastning kombinert med lav risiko eller er avkastningen avtagende jamfør tidligere forskning?"

I 5 % gruppen fikk vi en årlig avkastning på 0,42 %. Dette resultatet samsvarer med Do & Faff (2009) som oppnådde nær nullavkastning i sin studie. 10 % gruppen presterte langt bedre med en årlig avkastning tilsvarende 7,17 %. Dette resultatet er mer i samsvar med Gatev et al. (2006). Siden vi bare har en periode i studien er det vanskelig å gi en forklaring på hvorfor 10 % gruppen presterer såpass mye bedre enn 5 % gruppen. Dette kan skyldes tilfeldigheter, ettersom Do & Faff (2009) fant stor variasjon i den månedlige avkastningen fra -2,72 %, til 3,19 %. Dette kan tyde på at vi ikke har noe grunnlag for å konkludere med at funnene i vår studie skiller seg fra deres. Et fellestrekk ved hele studien er at strategien er markedsnøytral, med langt lavere risiko uttrykt i standardavvik og VaR enn referanseindeksen. Dette er også i samsvar med studiene til Gatev et al. (2006) og Do & Faff (2009).

Resultatene våre er totalt sett ikke lønnsomme justert for transaksjonskostnader. Ved å øke glidende gjennomsnitt til 150 dager så reduserer vi transaksjonskostnadene med 44 %. Har man derfor en lang investeringshorisont kan dette utgjøre forskjellen mellom tap og gevinst. Ser vi derimot på 10 % gruppen innenfor finanssektoren så er strategien vært attraktiv på bakgrunn av positiv avkastning samt svært lav risiko. Det bør derfor gjøres videre forskning på om dette er tilfeldigheter eller ikke.

I forhold til andre studier så har vi stort sett valgt å se på Do & Faff (2009), samt Gatev et al. (2006). Begge disse studiene er omfattende på bakgrunn av antall aksjer og aksjepar. I et slikt perspektiv kan man si noe om pairs trading ved store datasett, men kan man på bakgrunn av dette generalisere pairs trading som helhet? Dette kan man trolig ikke, da en trader som oftest vil spesialisere seg innenfor en sektor eller et antall aksjepar som vedkommende føler seg komfortabel med.

Som konklusjon kan det tyde på at strategien totalt sett gir nullavkastning som er sammenfallende med Do & Faff (2009). Studien til Gatev et al. (2006) er noe eldre, men også denne viser en avtagende lønnsomhet. Vi henviser derfor til videre forskning i kapittel 7.3.

7.2 Kritikk av studiet

I en oppgave som denne vil det naturlige spørsmålet være i hvilken grad strategien kan overføres til virkelighetens verden. I vårt tilfelle er dette mulig. Det finnes tilleggspakker til R som gjør det mulig å koble programmet opp mot en internettbasert megler, og handlere ville da automatisk blitt utført etter hvert som handelssignalene oppstår. Naturligvis kan man også benytte seg av andre løsninger som gir samme effekt.

Når det er sagt er det viktig å klargjøre at vi har brukt sluttkurser, som vil si den siste handelen som blir gjennomført i sluttauksjonen hver dag. I den sammenheng kan man selvsagt argumentere for at handelssignal basert på intradagpriser er en mer realistisk tilnærming. Dette er derfor noe som må tas i betraktning når man skal vurdere resultatene i studien. Vårt utgangspunkt er allikevel at utslagene ved bruk av intradagpriser vil være minimale med tanke på den gjennomsnittlige holdingperioden til parene i studiet. Hvis man tenker seg handelsstrategier som er beregnet kun på daytrading eller opp til noen dager, så vil bruk av intradagpriser naturlig nok være mer essensielt.

Som nevnt i kapittel 4.3.3 viser Whistler (2004) til metoder hvor man gjør undersøkelser vedrørende fundamentale forhold i selskaper, og at man på bakgrunn av funnene i undersøkelsene kan luke ut par som ikke egner seg til handel. På bakgrunn av det store utvalget av par i vår studie, var det ikke hensiktsmessig for oss å undersøke fundamentale forhold mellom selskapene i parene vi fant. Det er grunn til å tro at de som trader i den virkelige verden gjør slike undersøkelser. I følge Grøtte (2002) følger han i underkant av 10 par som han handler med. Bredt utvalg av par gjør dette at vår studie kanskje ikke er sammenlignbar med hvordan tradere i den virkelige verden opptrer, men på den andre siden gjør det at vi kan sammenligne vår studie med tidligere studier.

7.3 Forslag til videre forskning

I studien har vi testet strategien over de seks første måneder i 2011. Det ville vært meget interessant å se hvordan resultatene ville blitt om vi hadde testet strategien over en lengre tidsperiode. I den sammenheng måtte man da brukt den samme metoden for hver nye handelsperiode. Dette betyr altså testing av par for kointegrasjon i 12 måneder forut for hver seks måneders handelsperiode. For og ytterligere få flere perioder kan man suksessivt flytte seg en måned framover med testing og handel.

Et sentralt poeng som leseren bør være klar over er at metoden vår ikke inneholder stop-loss ved prosentvis tap, men ved divergering utover +/- 3 standardavvik. Selv om dette stopper en del tapsposisjoner åpner det allikevel mulighet for betydelige tap. Det ville derfor være meget interessant å teste ut hvordan resultatene ville blitt ved å stoppe tapsposisjoner ved ulike nivå for stop-loss, som ved 10 %, 15 % og 20 %.

En annen metode er å kjøre en backtest av strategien i de samme 12 måneder som man finner parene, for så kun å bruke de parene som presterer hensiktsmessig i handelsperioden. Målet ville da være å se om resultatene bedrer seg eller ikke.

I sakens natur er det kun fantasien som begrenser mulighetene for videre forskning. Alle parametere kan endres og optimaliseres for å finne den beste metoden og tilhørende personlige preferanser.

8. Litteraturliste

- Avellaneda, M., & Lee, J. H. (2010). Statistical arbitrage in the US equities market. *Quantitative Finance*, 10(7), 761-782.
- Björk, T. (1998). *Arbitrage theory in continuous time*. Oxford: Oxford University Press.
- Bodie, Z., Kane, A., & Marcus, A. J. (2009). *Investments* (8th ed.). Boston, Mass.: McGraw-Hill.
- Bondarenko, O. (2003). Statistical Arbitrage and Securities Prices. *Review of Financial Studies*, 16(3), 875-919.
- Brown, S. J. (2011). The efficient markets hypothesis: The demise of the demon of chance? *Accounting & Finance*, 51(1), 79-95. doi: 10.1111/j.1467-629X.2010.00366.x
- Bruce, N. L. (1990). Fads, Martingales, and Market Efficiency: National Bureau of Economic Research, Inc.
- Chan, E. P. (2009). *Quantitative trading: how to build your own algorithmic trading business*. Hoboken, N.J.: Wiley.
- Chordia, T., Roll, R., & Subrahmanyam, A. (2008). Liquidity and market efficiency. *Journal of Financial Economics*, 87(2), 249-268.
- Damodaran, A. (2003). *Investment philosophies: successful strategies and the investors who made them work*. Hoboken, N.J.: Wiley.
- Do, B., & Faff, R. (2009). Does simple pairs trading still work? *Financial Analysts Journal*, 66(4), 1-13.
- Do, B. H., & Faff, R. W. (2011). Are Pairs Trading Profits Robust to Trading Costs? *SSRN eLibrary*. doi: 10.2139/ssrn.1707125
- Dubil, R. (2004). *An arbitrage guide to financial markets*. Chichester: Wiley Finance.
- Engle, R. F., & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, 47(2), 427-465.
- Gatev, E., Rouwenhorst, K. G., & Goetzmann, W. N. (2006). *Pairs trading: performance of a relative value arbitrage rule* (Vol. 7032). Cambridge, Mass.: NBER.
- Grøtthe, O. (2002). *Aksjekjøp og daytrading: metode, psykologi, risiko og strategier*. [Lysaker]: Hegnar media.
- Keim, D. B. (1983). Size-related anomalies and stock return seasonality. Further empirical evidence. *Journal of Financial Economics*, 12(1), 13-32.
- Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, 49(5), 1541-1578.
- Lütkepohl, H., & Krätzig, M. (2004). *Applied time series econometrics*. Cambridge: Cambridge University Press.
- Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *The Journal of Economic Perspectives*, 17(1), 59-82.
- Peter Carl, B. G. P., Kris Boudt, Eric Zivot. (2012). PerformanceAnalytics: Econometric tools for performance and risk analysis.
- Pole, A. (2007). *Statistical arbitrage: algorithmic trading insights and techniques*. Hoboken, N.J.: Wiley.
- Reverre, S. (2001). *The complete arbitrage deskbook*. New York: McGraw-Hill.
- Ross, S. A., Westerfield, R. W., Jordan, B., Roberts, G. (2007). *Fundamentals of Corporate Finance* (Sixth Canadian ed.): McGraw-Hill Ryerson.
- Setterberg, H. (2007). *Swedish post-earnings announcement drift and momentum return*. Center for Financial Analysis and Managerial Economics in Accounting. Stockholm School of Economics.

- Shleifer, A. (2000). *Inefficient markets: an introduction to behavioral finance*. Oxford: Oxford University Press.
- Shleifer, A., Vishny, Robert W. (1997). The Limits of Arbitrage. *The Journal of Finance*, 52(1), 35-55.
- U.S Department of the Treasury (n.d). Resource Center. Daily Treasury Yield Curve Rates. Retrieved 22.04.2012 from <http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2011>
- Varian, H. R. (1987). The Arbitrage Principle in Financial Economics. *The Journal of Economic Perspectives*, 1(2), 55-72.
- Vidyamurthy, G. (2004). *Pairs trading: quantitative methods and analysis*. Hoboken, N.J.: Wiley.
- Whistler, M. (2004). *Trading pairs: capturing profits and hedging risk with statistical arbitrage strategies*. Hoboken, N.J.: Wiley.
- Zivot, E., & Wang, J. (2006). *Modeling financial time series with S-PLUS*. New York: Springer.

Vedlegg 1 Script for å få data inn i R. Eksempel sektoren Agricultural i Basic Materials.

```
mon <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/MON.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
mon$V1 <- mon$V3 <- mon$V4 <- mon$V5 <- mon$V7 <- NULL
colnames(mon) <-c("Dato", "MON")
pot <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/POT.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
pot$V1 <- pot$V3 <- pot$V4 <- pot$V5 <- pot$V7 <- NULL
colnames(pot) <-c("Dato", "POT")
syt <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/SYT.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
syt$V1 <- syt$V3 <- syt$V4 <- syt$V5 <- syt$V7 <- NULL
colnames(syt) <-c("Dato", "SYT")
mos <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/MOS.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
mos$V1 <- mos$V3 <- mos$V4 <- mos$V5 <- mos$V7 <- NULL
colnames(mos) <-c("Dato", "MOS")
agu <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/AGU.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
agu$V1 <- agu$V3 <- agu$V4 <- agu$V5 <- agu$V7 <- NULL
colnames(agu) <-c("Dato", "AGU")
cf <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/CF.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
cf$V1 <- cf$V3 <- cf$V4 <- cf$V5 <- cf$V7 <- NULL
colnames(cf) <-c("Dato", "CF")
smg <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/SMG.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
smg$V1 <- smg$V3 <- smg$V4 <- smg$V5 <- smg$V7 <- NULL
colnames(smg) <-c("Dato", "SMG")
cmp <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/CMP.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
cmp$V1 <- cmp$V3 <- cmp$V4 <- cmp$V5 <- cmp$V7 <- NULL
colnames(cmp) <-c("Dato", "CMP")
tnh <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/TNH.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
tnh$V1 <- tnh$V3 <- tnh$V4 <- tnh$V5 <- tnh$V7 <- NULL
colnames(tnh) <-c("Dato", "TNH")
avd <- read.table("/Users/andreas/Documents/Stockdata/Basicmaterial/Agricultural/AVD.csv",
  header=FALSE, sep=";", na.strings="NA", dec=".", strip.white=TRUE)
avd$V1 <- avd$V3 <- avd$V4 <- avd$V5 <- avd$V7 <- NULL
colnames(avd) <-c("Dato", "AVD")
#Har hentet script fra http://www.stockhistoricaldata.com/ og det til å hente data fra Yahoo
Finance.
agricultural <- data.frame
agricultural <- merge(mon,pot,by="Dato")
rm(mon,pot)
agricultural <- merge(agricultural,syt,by="Dato")
rm(syt)
agricultural <- merge(agricultural,mos,by="Dato")
rm(mos)
agricultural <- merge(agricultural,agu,by="Dato")
rm(agu)
agricultural <- merge(agricultural,cf,by="Dato")
rm(cf)
agricultural <- merge(agricultural,smg,by="Dato")
rm(smg)
agricultural <- merge(agricultural,cmp,by="Dato")
rm(cmp)
agricultural <- merge(agricultural,tnh,by="Dato")
rm(tnh)
agricultural <- merge(agricultural,avd,by="Dato")
rm(avd)
agricultural <-na.omit(agricultural)
write.csv(agricultural, file= "agricultural.csv")
```

Vedlegg 2 Johansen test av en hel sektor. Eksempel Property and Casualty Insurance.

```
library(urca)
result <- list()
j <- 0
startindex <- which(pcinsurance$Dato>"20100100")[1]
sluttindex <- which(pcinsurance$Dato>"20110100")[1]-1
for(i1 in 2:46){
  for (i2 in (i1+1):47){
    j <- j+1
    a1 <- pcinsurance[startindex:sluttindex,i1]
    a2 <- pcinsurance[startindex:sluttindex,i2]
    a3 <- cbind(a1,a2)
    colnames(a3) <- c(colnames(pcinsurance)[i1], colnames(pcinsurance)[i2])
    result[[j]] <- ca.jo(a3, type = c("eigen"), ecdet = c("none"), K = 10,
spec=c("transitory"), season = NULL, dumvar = NULL)
  }
}
for (j in 1:1035){
  print(summary(result[[j]]))
}
```

Vedlegg 3 Eksempel output Johansen test.

```
#####
# Johansen-Procedure #
#####

Test type: maximal eigenvalue statistic (lambda max) , with linear trend

Eigenvalues (lambda):
[1] 0.081569696 0.005161012

Values of teststatistic and critical values of test:

      test 10pct  5pct  1pct
r <= 1 |  1.25  6.50  8.18 11.65
r = 0  | 20.59 12.91 14.90 19.19

Eigenvectors, normalised to first column:
(These are the cointegration relations)

      HTH.11  AMSF.11
HTH.11 1.000000 1.00000000
AMSF.11 1.303256 -0.09938318

Weights W:
(This is the loading matrix)

      HTH.11  AMSF.11
HTH.d -0.002493368 -0.009520242
AMSF.d -0.123693611 -0.008397578
```

Vedlegg 4 ADF og PP-test av spreaden mellom to aksjer. Eksempel Aksjeparet HTH.AMSF.

```
Date <- strptime(pcinsurance$Dato,format="%Y%m%d")
hth.amsf <- data.frame
hth.amsf <- pcinsurance[c("Dato","HTH","AMSF")]
hth.amsf <-cbind(hth.amsf,Date)
hth.amsf$Dato <- NULL
hth.amsf <-hth.amsf[ ,c(3,1:2)]
hth.amsf.xts <- xts(hth.amsf[,-1], order.by= hth.amsf[,1])
rm(hth.amsf)
library(PairTrading)
price.pair <- hth.amsf.xts[,1:2]["2010-01-01::2010-12-31"]
reg <- EstimateParameters(price.pair, method = lm)
str(reg)
plot(reg$spread)
IsStationary(reg$spread, 0.05)
```

Vedlegg 5 Eksempel output ADF og PP-test

```
List of 3
 $ spread      :An 'xts' object from 2010-01-04 to 2010-12-31 containing:
  Data: num [1:252, 1] 0.0599 0.0287 0.0179 0.0126 0.0242 ...
 - attr(*, "dimnames")=List of 2
  ..$ : NULL
  ..$ : chr "AMSF"
 Indexed by objects of class: [POSIXct,POSIXt] TZ:
 xts Attributes:
List of 2
 ..$ tclass: chr [1:2] "POSIXct" "POSIXt"
 ..$ tzone : chr ""
 $ hedge.ratio: num -0.416
 $ premium   : num 3.86
 > plot(reg$spread)
 > IsStationary(reg$spread, 0.05)
 PP.test adf.test
 TRUE FALSE
```


Vedlegg 6 Script for testing av handelsstrategi. Eksempel fra alle 5 %.

```
#by Garrett See
#code borrowed heavily from existing quantstrat demos

#This is a simple pairs trading example intended to illustrate how you can extend
#existing quantstrat functionality. Also, it uses addPosLimits to specify
#levels and position limits, and shows how to pass a custom order sizing function to osFUN

#Note that it would be easier to build a spread first and treat it as a single instrument
#instead of dealing with a portfolio of stocks.

## given 2 stocks, calculate their ratio. If the ratio falls below it's
# 2 stdev band, then when it crosses back above it, buy stock 1 and sell stock 2.
# If the ratio rises above it's 2 stdev band, then when it crosses back below
# it, sell stock 1 and buy stock 2. If the ratio crosses it's moving average,
# then flatten any open positions.

# The Qty of Stock A that it buys (sells) = MaxPos / lvlsl
# The Qty of Stock B that it sells (buys) = MaxPos * Ratio / lvlsl
pdf(file="PairTrade-alle50.pdf", onefile=TRUE)
try(rm(list=.blotter[["portfolios"]], pos=.blotter), silent=TRUE)
try(rm(list=.blotter[["accounts"]], pos=.blotter), silent=TRUE)
try(rm(list=.strategy[["order_books"]], pos=.strategy), silent=TRUE)
try(rm("portfolios","accounts", pos=.blotter), silent=TRUE)
try(rm("order_books", pos=.strategy), silent=TRUE)
ls(envir=.blotter)
ls(envir=.strategy)
try(rm("initDate", "endDate", "startDate", "initEq", "SD", "N", "symb1", "symb2", "M",
      "portfoliol.st", "account.st", "pairStrat", "out1", "ret1"), silent=TRUE)

require(quantstrat)
initDate = '2011-01-03'
endDate = '2011-06-30'
#StartTrade = '2011-01-03'
startDate = '2010-10-21'
#xstartDate <- index(CEO[which(index(CEO)==StartTrade)-100])
initEq = 1000000
#SD = 2
#N = 100

#MaxPos = 100000 #max position in stockA;
#max position in stock B will be max * ratio, i.e. no hard position limit in Stock B
lvlsl = 1 #how many times to fade; Each order's qty will = MaxPos/lvlsl

#symb1<-'FII'
#symb2<-'CLMS'
#portfoliol.st <- 'pair1'
#account.st <- 'pairs'

symb1 <- c('GFI','BVN','NUE','SCHN','CEO','CEO','SNP','OXY','ECA','ECA',
  'ECA','SU','APC','APC','SWN','SWN','SWN','CHK','STR','NFX','NFX','PXP','CRK',
  'FST','SM','PVA','HES','VLO','TSO','IOC','IOC','EROC','CLMT','HAL','BHI','PDS','SIAL',
  'SXT','WPZ','WPZ','ENB','ENB','EP','KMR','DO','BBG','BBG','BBG','MWE','HERO','GIFI',
  'CBE','GE','ALV','KO','SNE','UA','HBI','CPO',
  'IVZ','FII','MS','BBD','BBD','UBS','CS','C','TD','BPO','CXW','HXM','RYN','DEI','NNN',
  'FSP','BDN','FPO','PSA','ARE','USB','NTRS','FITB','FMER','ONB','MBFI','MBFI',
  'VLY','VLY','NPBC','TRST','SIVB','UMPQ','ESS','CPT','BRE','HME','SIM','SPG','VNO',
  'KIM','REG','HPT','BRK-A','BRK-A','ACE','AIG','RNR','RNR','THG','VR','VR','CB',
  'CPWR','BMC','LFT','ADP','VSAT','EFII','ZBRA','NICE','LOGI','BMI','MTSC','CPHD',
  'IIVI','RSTI','MTD','TRMB','WAT','BHE','USMO','LEAP','CEL','SBAC','PTNR','NIHD','NIHD',
  'USM','VIV','PCS','PT','VIP','TU','MBT','TKC','TKC','CHU','AMX','DCM',
  'DCM','VOD','CHL','CHL','NTT','SHEN','DLB','SNDA','VRSN','CRM','SYMC','YHOO')
symb2 <- c('HMY','HMY','HSC','HAYN','ECA','FST','ECA','CNQ','EPD','STR',
  'XTEX','FST','KWK','ATPG','STR','CRK','PVA','GDP','GMXR','GPOR','BEXP','KWK','PVA',
  'GPOR','PQ','GMXR','CLMT','DK','CVI','CLMT','XTXI','DK','XTXI','WFT','NGS','TESO',
  'NEU','NEU','PPO','KWR','MMP','WES','CPNO','NGLS','VQ','MWE','BBEP','PSE','EVEF',
  'VQ','LNG','DHR','GY','TRW','HANS','HAR','OXM','ZQK','FLO',
  'AMP','CLMS','HGT','BMA','BFR','LYG','BCS','KEY','KEY','CXW','CIM','CIM','FUR','FUR',
  'FUR','CUZ','FUR','FUR','ARE','BMR','PVTB','PVTB','CHFC','PVTB','PVTB','EMBI','WTFC',
  'SUSQ','WBS','WFD','WBS','EWBC','HAFc','BRE','AEC','SUI','MAA','PHH','UBA','FRT','REG',
  'EPR','IRC','THG','PTP','THG','EIG','THG','ERIE','ERIE','MRH','TWGP','ACE',
  'TRAK','TRAK','NTCT','BLKB','ELMG','SSYS','SSYS','SSYS','SSYS','ALOG','BMI','BMI','BMI',
```

```

'CPHD','BMI','BMI','BMI','PLXS','UTSI','UTSI','UTSI','UTSI','UTSI','UTSI','USMO','UTSI','UTSI
',
'UTSI','UTSI','NIHD','NIHD','UTSI','UTSI','VIP','UTSI','UTSI','UTSI','VIP','UTSI','UTSI','MBT
',
'NZT','CNSL','CTS','ARBA','SNDA','SNDA','SNDA','UNTD')
M <-length(symb1)
if (M != length(symb2)) stop('symb1 and symb2 must have equal length\n')
portfoliol.st<-paste(symb1,symb2,sep=".")
account.st<-paste(symb1,symb2,sep=" ")
.blotter[["portfolios"]]<-paste("portfolio",portfoliol.st,sep=".")
.blotter[["accounts"]]<-paste("account",account.st,sep=".")
.strategy[["order_books"]]<-paste("order_book",portfoliol.st,sep=".")
pairStrat<-list()
ret1<-list()
out1<-list()

calcRatio <- function(x) { #returns the ratio of close prices for 2 symbols
x1 <- get(x[1])
x2 <- get(x[2])
rat <- Cl(x1) / Cl(x2)
colnames(rat) <- 'Ratio'
rat
}
getRatio <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Ratio <- last(portf$Ratio[toDate])
print(timestamp)
as.numeric(Ratio)
}
calcPrice <- function(x) { #returns the price of stock A
x1 <- get(x[1])
price <- Cl(x1)
colnames(price) <- 'Price'
price
}
getxPrice <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Price <- last(portf$Price[toDate])
as.numeric(Price)
}
calcDelta <- function(x) { #returns the ratio of close prices for 2 symbols
x1 <- get(x[1])
x2 <- get(x[2])
spread <- Cl(x1) - Cl(x2)
mean50 <-SMA(spread,n=50)
sd50 <- runSD(spread,n=50,sample=TRUE,cumulative=FALSE)
delta <- (spread-mean50)/sd50
colnames(delta) <- 'Delta'
delta <- as.xts(delta)
delta
}
getDelta <- function(portfolio, timestamp) {
portf <- getPortfolio(portfolio)
toDate <- paste(":", timestamp, sep="")
Delta <- last(portf$Delta[toDate])
as.numeric(Delta)
}
calcFinish <- function(x) {
finish <- index(mktdata)=="2011-06-29"
finish <- as.numeric(finish)
names(finish) <- index(mktdata)
finish <- as.xts(finish)
colnames(finish) <- c("Finish")
finish
}
osSpreadMaxPos <- function (data, timestamp, orderqty, ordertype, orderside, portfolio,
symbol, ruletype, ..., orderprice)
{
portf <- getPortfolio(portfolio)
legside <- portf$pair[symbol] #"long" if symbol=symb1[i], "short" if symbol=symb2[i]
cat("legside=",legside,"\n")
if (legside != "long" && legside != "short") stop('pair must contain "long" and "short"')
ratio <- getRatio(portfolio, timestamp)
}

```

```

price <- getxPrice(portfolio, timestamp)
MaxPos <- initEq/price
if (orderqty=='all') qty <- MaxPos
if (mktdata[timestamp,"cross.p2"]==1) qty = -qty
# addPosLimit(portfolio=portfolio, timestamp=timestamp, symbol=symbol,
# maxpos=MaxPos, longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)
pos <- getPosQty(portfolio, symbol, timestamp)
cat("pos=",pos,"qty=",qty,"\n")
if((pos> 0.5) | (pos< -0.5)) qty<-0
# PosLimit <- getPosLimit(portfolio, symbol, timestamp)
# qty <- orderqty
if (legside == "short") {#symbol is 2nd leg
  ## Comment out next line to use equal ordersizes for each stock.
  # addPosLimit(portfolio=portfolio, timestamp=timestamp,
  # symbol=symbol, maxpos=MaxPos*ratio, longlevels=lvls,
  # minpos=-MaxPos*ratio, shortlevels=lvls)
  #TODO: is it okay that MaxPos and lvls come from .GlobalEnv ?
qty <- -qty*ratio #switch orderqty for Stock B
}
if (qty > 0) orderside = "long"
if (qty < 0) orderside = "short"
# orderqty <- osMaxPos(data=data,timestamp=timestamp,orderqty=qty,
# ordertype=ordertype,orderside=orderside,portfolio=portfolio,
# symbol=symbol,ruletype=ruletype, ...)
orderqty <- round(qty,0)
cat("orderqty=",orderqty,"\n")

#Add the order here instead of in the ruleSignal function
if (!is.null(orderqty) & !orderqty == 0 & !is.null(orderprice)) {
  addOrder(portfolio = portfolio, symbol = symbol,
    timestamp = timestamp, qty = orderqty, price = as.numeric(orderprice),
    ordertype = ordertype, side = orderside, status = "open", ... = ...)
}
return(0) #so that ruleSignal function doesn't also try to place an order
}

for (i in 1:M) {
  getSymbols(c(symb1[i], symb2[i]), from=startDate, to=endDate, adjust=TRUE)

  currency("USD")
  stock(symb1[i], currency="USD", multiplier=1)
  stock(symb2[i], currency="USD", multiplier=1)

#Initialize Portfolio, Account, and Orders
  initPortf(name=portfoliol.st[i], c(symb1[i],symb2[i]), initDate=initDate)
  initAcct(account.st[i], portfolios=portfoliol.st[i], initDate=initDate, initEq=initEq)
  initOrders(portfolio=portfoliol.st[i],initDate=initDate)
  mktdata<-get(symb1[i])
#create a slot in portfolio for symb1[i] and symb2[i] to make them available to osFUN
  pair <- c('long','short')
  names(pair) <- c(symb1[i],symb2[i])
  .blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$pair <- pair

  # Create initial position limits and levels by symbol
  # allow 3 entries for long and short.
  #addPosLimit(portfolio=portfoliol.st[i], timestamp=initDate, symbol=symb1[i],
  #maxpos=MaxPos, #longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)
  #addPosLimit(portfolio=portfoliol.st[i], timestamp=initDate, symbol=symb2[i],
  #maxpos=MaxPos, #longlevels=lvls, minpos=-MaxPos, shortlevels=lvls)

  # Create a strategy object
  pairStrat[[i]] <- strategy(paste('pairStrat',symb1[i],symb2[i],sep="."))

  Ratio <- calcRatio(c(symb1[i],symb2[i]))
  #let's go ahead and put this in a slot in portfolio
  .blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Ratio <- Ratio
  #and make a function to get the most recent Ratio

#Get the price from one stock
  Price <- calcPrice(c(symb1[i]))
  #let's go ahead and put this in a slot in portfolio
  .blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Price <- Price

```

```

Delta <- calcDelta(c(symb1[i],symb2[i]))
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Delta <- Delta
#and make the most recent Ratio

Finish <- calcFinish(c(symb1[i],symb2[i]))
.blotter[[paste('portfolio',portfoliol.st[i],sep='.')]]$Finish <- Finish
#and make a function to get the most recent Ratio

# Create an indicator - BBands on the Ratio
pairStrat[[i]] <- add.indicator(strategy = pairStrat[[i]], name = "calcDelta",
arguments = list(x=c(symb1[i],symb2[i])))
pairStrat[[i]] <- add.indicator(strategy = pairStrat[[i]], name = "calcFinish",
arguments = list(x=c(symb1[i],symb2[i])))

applyIndicators(strategy=pairStrat[[i]],mktdata=get(symb1[i])) #for debugging

# Create signals - buy when crossing lower band from below, sell when crossing upper band from
above, flatten when crossing mavg from above or from below
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=-2, relationship="lte",
cross=TRUE), label="cross.m2")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=-3, relationship="lte", cross=TRUE),
label="cross.m3")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=2, relationship="gte",
cross=TRUE), label="cross.p2")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=3, relationship="gte",
cross=TRUE), label="cross.p3")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=0,relationship="gte",
cross=TRUE), label="cross.mid.fa")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Delta", threshold=0, relationship="lte",
cross=TRUE), label="cross.mid.fb")
pairStrat[[i]] <- add.signal(strategy = pairStrat[[i]], name = "sigThreshold",
arguments= list(column="Finish", threshold=0, relationship="gt",
cross=FALSE), label="sigfinish")

# Create entry and exit rules for longs and for shorts. Both symbols will get the same
buy/sell signals, but osMaxPos will reverse those for the second symbol.
# orderqty's are bigger than PosLimits allow. osMaxPos will adjust the orderqty down to 1/3
the max allowed. (1/3 is because we are using 3 levels in PosLimit)
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.m2", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL, osFUN='osSpreadMaxPos'), type='enter' )
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.m3", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.p2", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL, osFUN='osSpreadMaxPos'), type='enter')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.p3", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.mid.fb", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="cross.mid.fa", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')
pairStrat[[i]] <- add.rule(strategy = pairStrat[[i]], name='ruleSignal',
arguments = list(sigcol="sigfinish", sigval=TRUE, orderqty='all',
ordertype='market', orderside=NULL), type='exit')

#applySignals(strategy=pairStrat[[i]],
mktdata=applyIndicators(strategy=pairStrat[[i]],mktdata=get(symb1[i]))) #for debugging

out1[[i]]<-try(applyStrategy(strategy=pairStrat[[i]], portfolios=portfoliol.st[i]))

updatePortf(Portfolio=portfoliol.st[i],
Dates=paste("::",as.Date(Sys.time()),sep=''))

```

```

updateAcct(account.st[i],Dates=paste(startDate,endDate,sep="::"))
updateEndEq(account.st[i],Dates=paste(startDate,endDate,sep="::"))
getEndEq(account.st[i],Sys.time())

chart.Posn(Portfolio=portfoliol.st[i],Symbol=symb1[i])
chart.Posn(Portfolio=portfoliol.st[i],Symbol=symb2[i])
chartSeries(Delta,name=paste("Delta",portfoliol.st[i],sep=" "))

ret1[[i]] <- PortfReturns(account.st[i])
ret1[[i]]$total <- rowSums(ret1[[i]])
#ret1

if("package:PerformanceAnalytics" %in% search() ||
    require("PerformanceAnalytics",quietly=TRUE)) {
  getSymbols("SPY", from='2011-01-03')
  SPY.ret <- Return.calculate(SPY$SPY.Close)
  tmp <- merge(SPY.ret,ret1[[i]]$total,all=FALSE)
  charts.PerformanceSummary(cbind(tmp[,2],tmp[,1]),main=portfoliol.st[i],geometric=FALSE,
    wealth.index=TRUE)
}
}
ExcessReturn <- ret1[[1]]$total
for (i in 2:M) {
  ExcessReturn <- ExcessReturn+ret1[[i]]$total
  ExcessReturn <- ExcessReturn["2011-01-01::2011-06-30"]
}
ExcessReturn <- ExcessReturn/M

#zzPort<-getPortfolio(portfoliol.st[i])
#str(zzPort)
#zzAcct<-getAccount(account.st[i])
#str(zzAcct)
#str(pairStrat[[i]])
require(sandwich)
require(lmtest)
ExcessReturn.lm<-lm(as.vector(ExcessReturn) ~1)
ExcessReturn.lm.coefstest<-coefstest(ExcessReturn.lm,df=3,
  vcov=NeweyWest(ExcessReturn.lm,lag=4,prewhite=FALSE))
ExcessReturn.lm.coefstest

portefolje <-
  blotter:::.getByPortf(.blotter[.blotter[["accounts"]][1]],Attribute="Realized.PL",
  Dates=index(mktdata))
for (i in 2:M) {
  portefolje <-
  cbind(portefolje,blotter:::.getByPortf(.blotter[.blotter[["accounts"]][i]],Attribute="Realized.PL",
  Dates=index(mktdata)))
}

portefoljel <- (initEq+portefolje - initEq)/initEq
portefoljel <- portefoljel["2011-01-01::2011-06-30"]
x <- table(portefoljel,exclude=0)
x <- as.data.frame(x)
x <- na.omit(x)
posisjoner <-sum(x$Freq) #antall lukninger av posisjoner, ergo antall posisjoner åpnet.
y <- x$portefoljel
y <- as.numeric(levels(y)[y])
stortap <- min(y) #største tap
storgev <- max(y) #største gevinst
snitt <- mean(y) #gjennomsnittlig avkastning
zmin<- table(y[y<0])
zmin <-as.data.frame(zmin)
zpos<- table(y[y>0])
zpos <- as.data.frame(zpos)
zb <- sum(zmin$Freq)
zc <- sum(zpos$Freq)

win <- (zc/posisjoner)*100 #prosent gevinst
tap <- (zb/posisjoner)*100 #prosent tap
wlratio <- win/tap #win/loss ratio
handler <- posisjoner*4

portefolje3 <-

```

```

blotter:::.getByPortf(.blotter[.blotter[["accounts"]][1]],Attribute="Short.Value",
Dates=index(mktdata))
for (i in 2:M) {
  portefolje3 <-
  cbind(portefolje3,blotter:::.getByPortf(.blotter[.blotter[["accounts"]][2]],Attribute="Short.Value",
Dates=index(mktdata)))
}

portefolje3 <- portefolje3["2011-01-01::2011-06-30"]
portefolje3 <- table(portefolje3,exclude=0)
portefolje3 <- as.data.frame(portefolje3)
holding <- sum(portefolje3$Freq)/posisjoner

res <- cbind(M,posisjoner, handler, holding, stortap, storgev, win, tap, wlratio)
colnames(res) <- c("Antall Par","Antall posisjoner", "Antall handler", "Gj.snitt
holdingperiode", "Største tap", "Største gevinst","Gevinst %", "Tap %", "W/L ratio")

getSymbols('SPY', from='2011-01-03', to='2011-06-30')
getSymbols('^FVX', from='2011-01-03', to='2011-06-30')
Rf <- FVX$FVX.Close
Rb <- SPY$SPY.Close
Rb <-Return.calculate(Rb, method="compound")
return.cumulative <- Return.cumulative(ExcessReturn, geometric=FALSE)
return.ann <- Return.annualized(ExcessReturn, geometric=FALSE, scale=252)
sharpe <- SharpeRatio.annualized(ExcessReturn, geometric=FALSE, Rf=0.0018/252)
var <- VaR(ExcessReturn, method=c("historical"),portfolio_method=c("single"))
es <- ES(ExcessReturn, method=c("historical"),portfolio_method=c("single"))#conditionalVAR
sdev <- sd.annualized(ExcessReturn, scale=252)
maxdraw <- maxDrawdown(ExcessReturn, weights=NULL, geometric = FALSE, invert=TRUE)
table.CAPM(ExcessReturn, Rb, scale = 252, Rf = 0.0018/252, digits = 4)
capm <- CAPM.beta(ExcessReturn, Rb, Rf=0.0018/252)
charts.PerformanceSummary(cbind(ExcessReturn,Rb), geometric=FALSE, wealth.index=TRUE,
main="All stocks 5%")
table.Stats(ExcessReturn)
info <- InformationRatio(ExcessReturn, Rb, scale=252)

res2 <- cbind(return.cumulative, return.ann, sharpe, var, es, sdev,maxdraw, capm, info)
colnames(res2) <- c("Periode avkastning", "Årlig avkastning", "Sharpe-ratio (Årlig)", "VaR",
"Conditional VaR", "Standardavvik (Årlig)", "Max Drawdown", "Beta", "Information Ratio")
rownames(res2) <- NULL

return.spy <- Return.cumulative(Rb, geometric=FALSE)
ann.spy <- Return.annualized(Rb, geometric=FALSE, scale=252)
sharpe.spy <- SharpeRatio.annualized(Rb, geometric=FALSE, Rf=0.0018/252)
var.spy <- VaR(Rb, method=c("historical"),portfolio_method=c("single"))
es.spy <- ES(Rb, method=c("historical"),portfolio_method=c("single"))#conditionalVAR
StdDev(Rb)
sdev.spy <- sd.annualized(Rb, scale=252)
maxdraw.spy <- maxDrawdown(Rb, weights=NULL, geometric = FALSE, invert=TRUE)
capm.spy <- CAPM.beta(Rb, Rb, Rf=0.0018/252)

res3 <- ExcessReturn.lm.coefest

res4 <- cbind(return.spy, ann.spy, sharpe.spy, var.spy, es.spy, sdev.spy,maxdraw.spy,
capm.spy)
colnames(res4) <- c("Periode avkastning", "Årlig avkastning", "Sharpe-ratio (Årlig)", "VaR",
"Conditional VaR", "Standardavvik (Årlig)", "Max Drawdown", "Beta")
rownames(res3) <- NULL

con <- file("tradingstats.csv","w")
write.table(res, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res2, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res3, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
write.table(res4, file=con,sep="," ,dec=".",row.names=FALSE,col.names=TRUE,quote=FALSE)
close(con)

dev.off(which = dev.cur())

```

Vedlegg 7 Finne antall par som faktisk handler.

```
port <- portefoljel
port <- na.omit(port)
colSums(port)
colSums(port) != 0
port <- port[, colSums(port) != 0]
antallpar <- ncol(port)

port1 <- port[,1]
for (i in 2:antallpar) {
  port1 <- port1+port[,i]
}
port <- port/antallpar
fullyinvested <- Return.cumulative(port, geometric=FALSE)
fullyinvested
```