



LUNDS
UNIVERSITET

Hur påverkar marknadsvärde och book-to-market
aktiers idiosynkratiska risk? - En empirisk
undersökning av Stockholmsbörsen 1999 - 2017

Erik Enfors & Peter Sjöström

Oktober 2017

Kandidatuppsats
Nationalekonomiska Institutionen

Handledare: Dag Rydorff

Sammanfattning

Denna uppsats undersöker om variablerna i Fama och French (1992) kan användas för att relativt Capital Asset Pricing Model bättre förstå individuella aktiers idiosynkratiska risk framgent. De ytterligare faktorerna, utöver marknadsbeta, är företagens marknadsvärde och book-to-market. Vår analys visar att det finns ett signifikant samband mellan faktorerna i Fama och French (1992) och den idiosynkratiska risken hos enskilda aktier. Företag med högt marknadsvärde eller book-to-market har en markant högre idiosynkratisk risk relativt mittenportföljen. Företag med lågt marknadsvärde eller book-to-market har också ett positivt, men begränsat, samband med idiosynkratisk risk. Det senare i motstridighet med författarnas förväntningar. Vidare uppstår problematik med Maximum likelihood-skattningen eftersom de skattade parametervärdena visar sig bero på optimeringsalgoritmens startvärden, vilket indikerar att globalt maximum inte hittas. Förslag på hur det kan hanteras samt hur framtida studier kan utveckla vår modell diskuteras.

Nyckelord: Idiosynkratisk risk, Marknadsvärde, Book-to-market, Relativt värde, CAPM, Maximum likelihood-skattning

Tabeller

1	Data som inhämtas från Datastream	19
2	Aktier vars utdelningsdag sammanföll med utdelning av aktie i annat bolag	23
3	Alla portföljer som kombinationer av våra förklarande variabler bildar.	26
4	Resultatet av Jarque-Beratest för de olika portföljernas ej marknadsberoende avkastning	29
5	Sammanställning av resultat från Matlabs optimeringsalgoritm <i>fminunc</i> . Resultaten är sorterade efter fallande funktionsvärde.	30
6	Optimering av parametrar i den begränsade modell som motsvarar CAPM	31
7	Optimering av parametrar i begränsad modell med endast en extra parameter i förhållande till CAPM	31
8	Sammanställning av resultat för ML-skattning av modell begränsad genom att en parameter sätts till noll	31
9	Sammanställning av resultat för ML-skattning av modell begränsad genom att en parameter sätts till noll	32
10	De teoretiska kvantilerna för de olika portföljerna givet de uppskattade parametrarna i modellen. Alla värden anges i procent.	35
11	Procentuell andel av observationerna i respektive portfölj som understiger motsvarande teoretiskt kvantilvärde givet modellens uppskattade parametrar	35
12	Historisk varians i respektive portfölj	37
13	Beräknat värde för respektive beta. Baseras på historisk volatilitet i respektive portfölj	38
A1	Företag som är med i urvalet	46
A2	Aktier som ej är med i denna studie men som är med i Johnsson (2017). Dessa aktier exkluderas från urvalet för att de är banker eller för att det redan finns en mer omsatt aktie från samma företag i urvalet.	52
A3	Aktier som är med i denna studie som ej är med i Johnsson (2017)	52
A4	Företag som varit listade på Stockholmsbörsen sedan 1997 men som inte är med i urvalet	53
B1	Avkastning som är $\leq -60\%$	55
B2	Avkastning som är $\geq 50\%$	55

Figurer

1	PTBV för bolag inom 4 olika branscher	20
2	Standardavvikelsen i PTBV för bolag inom 4 olika branscher . .	21
3	Marknadsavkastningar och glidande årsvarians i vårt urval	40

Innehåll

1	Inledning	6
1.1	Bakgrund	6
1.2	Mål och syfte	8
1.3	Data och metod	8
1.4	Avgränsningar	9
1.5	Uppsatsens disposition	9
2	Teori och litteraturgenomgång	9
2.1	Finansiell risk	9
2.2	Riskminskning genom diversifiering	10
2.3	Systematisk och idiosynkratisk risk	10
2.4	Capital Asset Pricing Model	11
2.5	Fama-Frenchs trefaktormodell	12
2.5.1	Storleksfaktorn: Small minus big	13
2.5.2	Värdefaktorn: High minus low	13
2.6	Value at Risk	14
2.7	ML-skattning av modellparametrar	14
2.8	Jarque-Bera test av empirisk fördelning	15
2.9	Likelihood ratio-test av nästlade modeller	16
2.10	Tidigare studier	17
3	Data	17
3.1	Bearbetning av data	19
3.2	Antaganden	21
3.3	Beräkning av överavkastning	22
3.4	Beräkning av marknadens överavkastning	23
3.5	Definition av förklarande variabler	24
4	Metod	24
4.1	Tidsvarierande α och β	25
4.2	Skattning av parametrar	26
4.2.1	Parametrisering av sannolikhetsfördelning	27
4.2.2	Maximering av likelihoodfunktionen	28
4.3	Utvärdering av modellen	28

5	Resultat och analys	29
5.1	Resultat av Jarque-Beratest	29
5.2	Resultat av ML-skattning	29
5.3	Resultat av LR-test	32
5.4	Ekonomisk tolkning av resultat	32
5.5	Jämförelse av kvantiler	34
5.6	Beräknade varianser och parametervärden	36
5.7	Heteroskedasticitet	39
6	Slutsatser	42
	Bilaga A Företagslistor	46
	Bilaga B Osannolika observationer	55
	Bilaga C Matlabkod	57

1 Inledning

1.1 Bakgrund

Aktiemarknaden är idag glödhet. Sverige är en av de ekonomier i världen som återhämtat sig bäst efter finanskrisen 2008 vilket i sig har varit med och drivit den positiva utvecklingen på den svenska aktiemarknaden. Även det penningpolitiska läget har varit en förklarande faktor till utvecklingen. Den 12 februari 2015 fick vi historiskt låga och spektakulära -0,1 procent i reporänta, vilken sedan stabiliserats på -0,5 procent. Den expansiva penningpolitiken innehåller också kvantitativa lättnader som i slutet av 2017 kommer att uppgå till hela 290 miljarder kronor, exklusive återinvesteringar, sedan starten i februari 2015 (Riksbanken, 2017b, s. 9)

I denna starka ekonomi med höga företagsvärderingar, som delvis är underblåst av penningpolitiken, är intresset för aktiemarknaden stort. I Sverige äger 40,8 procent av hushållen aktier direkt och 66,2 procent av hushållen om även indirekt ägande inkluderas. Indirekt ägande inkluderar fonder och pensionssparande. Dessa siffror placerar Sverige bland de länder som har högst deltagande av hushållen i aktiemarknaden (Guiso och Sodini, 2013, s. 1455). Hushållen är per definition inte professionella investerare och därför kan överanvändning och marknadsföring av överdrivet förenklade verktyg komma att drabba dessa investerare hårdare än mer sofistikerade investerare. Således är det relevant att undersöka de enkla metoderna som många icke-professionella investerare använder och förstå hur de relaterar till de finansiella grundpelarna förväntad avkastning och risk (Piketty, 2014) (Gabaix och Maggiori, 2015).

Den vanligaste modellen för att beräkna en akties förväntade avkastning är idag Capital Asset Pricing Model (CAPM), som fått stort genomslag både på marknaden och internt på företag. En annan vanlig modell är multifaktor-modellen i Fama och French (1992) som förutom marknadsbeta (samvariation med marknadsportföljen) även använder företagets marknadsvärde (storleksfaktor) och book-to-market relation (värdefaktor) för att beräkna förväntad avkastning. Båda modellerna förenklar marknadens prissättningsmekanismer, men med skillnaden att Fama och French (1992) bättre förklarar den historiska avkastningen för aktier.

Risk i sin tur delas in i systematisk risk och idiosynkratisk risk. Systematisk risk är risk som är icke-diversifierbar och den påverkar därför den avkastning som en investerare kräver som kompensation för att bära risken. Idiosynkratisk risk

beror däremot på företagsspecifika faktorer. Några exempel på sådana faktorer är FoU, patent och juridiska processer. Detta är risker som genom diversifiering kan elimineras i en portfölj av tillgångar. Eftersom det är en "frivillig" typ av risk så påverkar den inte avkastningskraven på en aktie enligt både CAPM och Fama och French (1992). Därför är det intressant att undersöka vad som driver idiosynkratisk risk för att kunna ta bättre beslut i termer av relationen mellan förväntad avkastning och risk.

Diversifiering är både historiskt och idag en självklarhet för många investerare. Därför kan det te sig onödigt att närmare utforska komponenterna i idiosynkratisk risk. Men perfekt diversifiering, alltså allokering i den s.k. marknadsportföljen, är ett teoretiskt optimum som för många inte är möjligt. Många investerare innehar en betydande mängd idiosynkratisk risk, till exempel genom aktier eller optioner i arbetsgivarens företag som en del av sitt pensionspaket eller för att sammanföra de anställdas och ägarnas intressen (Perold, 2004, s. 18).

Idiosynkratisk risk härstammar också ofta från ett fenomen kallat "home bias", vilket innebär att investerare generellt är överexponerade mot aktiemarknaderna i sina respektive hemländer. Perold (2004) använder exemplet med USA, i december 2003, där endast tio procent av de publikt handlade aktierna ägdes av utländska investerare, samtidigt som marknadsportföljen dikterar ett ägandeskap av 47,8 procent i amerikanska aktier. Allokeringen till svenska aktier ska enligt marknadsportföljen i Perold (2004) uppgå till 1,1 procent av hela portföljen.

Perfekt diversifiering är också ofta praktiskt utmanande eftersom det, beroende på både direkta och indirekta kostnader, kan vara kostsamt. Exempel på sådana kostnader är courtage och skatter på köp av utländska aktier. En annan anledning till den ofullständiga diversifieringen kan vara okunskap eller vissa mänskliga beteenden (s.k. bias). CAPM gör en rad antaganden, och möjligen är det när dessa antaganden inte möter den enskilda investerarens verklighet och beteenden som diskrepansen mellan teori och praktik uppstår (Perold, 2004). Exempel på sådana diskrepanser kan till exempel härstamma från de orealistiska antagandena att alla investerare är rationella och att det inte existerar några skatter eller transaktionskostnader.

Campbell et al. (2001a) beskriver slutligen att den idiosynkratiska risken utgör majoriteten av den totala risken hos en tillgång. Vid låg grad av diversifiering är det således en stor källa till finansiell risk.

Sammanfattningsvis är idiosynkratisk risk ett högst intressant område inom

den finansiella ekonomin och trots den teoretiska möjligheten till diversifiering så brister det ofta i praktiken. Sambandet mellan förväntad avkastning och den finansiella risken är idag välundersökt. Risk är centralt vid finansiellt beslutsfattande och målet bör alltid vara att optimera sina positioner för att få högsta möjliga förväntad avkastning för en given nivå av risk. Därför måste även investerarnas förståelse för risk anses angelägen.

Vår hypotes är att storleksfaktorn och värdefaktorn kan förklara en del av den idiosynkratiska risken som CAPM beräknar. Det baseras på att Fama och French (1992), jämfört med CAPM, har högre förklaringsgrad för historiska avkastningar. I enlighet med teorin om effektiva marknader så förväntar vi oss att se en större mängd total risk i de fall Fama och French (1992) har observerat en relativt hög avkastning och motsatsen vid en relativt låg avkastning. Bartram et al. (2016) visar på en positiv korrelation mellan systematisk risk och idiosynkratisk risk, vilket medför att högre total risk bör medföra högre idiosynkratisk risk. Mer specifikt är således vår hypotes att mängden idiosynkratisk risk är negativt korrelerad med marknadsvärde och positivt korrelerad med book-to-market. Mot denna bakgrund så hoppas denna uppsats kunna validera hela eller delar av vår hypotes och därmed komplettera kunskapen om idiosynkratisk risk hos investerare som idag endast använder CAPM.

1.2 Mål och syfte

Målet med denna uppsats är att komplettera de verktyg som CAPM erbjuder investerare för att beräkna idiosynkratisk risk. Denna uppsats syftar därför specifikt till att bestämma sambandet mellan marknadsvärde (storleksfaktorn) och book-to-market (värdefaktor) med en enskild akties idiosynkratiska risk då den idiosynkratiska risken beräknas baserat på Capital Asset Pricing Model (CAPM).

1.3 Data och metod

Metoden i denna uppsats bygger på en kvantitativ studie. Dataunderlaget består framför allt av dagliga avkastningar, marknadsvärde och price-to-book relationen för aktier på Stockholmsbörsen under perioden 1997-01-01 - 2017-06-30. Diverse ekonometriska verktyg används, men i huvudsak används en Maximum likelihood-skattning av respektive parameters påverkan på den idiosynkratiska risken.

1.4 Avgränsningar

Denna uppsats begränsar sin analys till att bedöma sambandet mellan den information som finns tillgänglig vid ingången av varje tidsperiod. Vi undersöker således inte sambandet mellan våra förklarande variabler och den idiosynkratiska risken a posteriori.

1.5 Uppsatsens disposition

Kapitel 2 innehåller en redogörelse för den teoretiska grund och tidigare studier som akademien idag producerat inom ämnet. Kapitel 3 beskriver vilken och hur data samlas in, samt hur databehandlingen görs och varför. Kapitel 4 redogör för den metod som används och i kapitel 5 presenteras resultat och analys. Slutligen så sammanfattar vi våra slutsatser i kapitel 6 och ger förslag på framtida studier på området.

2 Teori och litteraturgenomgång

I detta avsnitt redogörs för den teori som är relevant för förstå både undersökningens metod och analysen av dess resultat. Avsnitten 2.1 till 2.5 behandlar den ekonomiska teorin kring risk, CAPM och Fama och French (1992). Sedan fokuserar avsnitten 2.6 till 2.8 på den teori som ligger till grund för vår ekonometriska metod. Avslutningsvis så redogör avsnitt 2.9 för tidigare studier på området.

2.1 Finansiell risk

Begreppet finansiell risk innefattar investerarens osäkerhet kring eventuella avkastningar och den potentiella förlust som investerare riskerar. Den finansiella risken kan härstamma från en eller flera olika källor av osäkerhet. Eftersom källorna till osäkerhet kan vara svåra att urskilja och särskilja så fokuserar man ofta istället på den uppmätta osäkerheten. Det görs under antagandet att avkastningen, ofta betecknat \tilde{r} , är en stokastisk variabel med ett väntevärde, $E[\tilde{r}]$, och en varians, $Var(\tilde{r})$. Variansen blir då ett mått på sannolikheten för avvikelser från avkastningens väntevärde. Även om sådana avvikelser kan vara positiva likväl som negativa så menar von Neumann och Morgenstern (1944) att rationella investerare som söker att maximera väntevärdet av sin upplevda nytta bör eftersträva att minimera avkastningens varians. I ljuset av von Neumann och

Morgenstern (1944) är det naturligt att likställa avkastningens varians, numera oftast dess standardavvikelse eller s.k. volatilitet, och dess finansiella risk.

2.2 Riskminskning genom diversifiering

Redan långt innan von Neumann och Morgenstern (1944) formulerade sina teorier om riskaversa investerare så hade riskbärande institutioner och investerare anammat diversifiering som ett sätt att minska risken. Den berömda förmaningen att inte lägga alla ägg i samma korg myntades av författaren de Cervantes redan 1605 (de Cervantes, 1605). Vare sig det gällde att lägga ägg i olika korgar eller att fördela last på flera olika fartyg vid handel i Fjärran Öster så var syftet med dessa strategier för dåtidens investerare att minska den maximala förlusten. Idén om att diversifiera bort risk höll dock i sig och fick stöd i den statistiska teori som senare utvecklades. Under förutsättningen att en mängd olika tillgångars avkastningar är okorrelerade med varandra, d.v.s. att $Cov(\tilde{r}_i, \tilde{r}_j) = 0$ för $i \neq j$, kan risken avsevärt minskas med bibehållen förväntad avkastning genom att investeringen delades upp mellan tillgångarna. Matematiskt är argumentet följande

$$Var\left(\sum_{i=1}^N \frac{1}{N} \tilde{r}_i\right) = \frac{1}{N^2} \sum_{i=1}^N Var(\tilde{r}_i) + 2 \sum_{1 \leq i < j \leq N} Cov(\tilde{r}_i, \tilde{r}_j) = \frac{\bar{\sigma}^2}{N} \xrightarrow{N \rightarrow \infty} 0 \quad (1)$$

där $\bar{\sigma}^2$ är den genomsnittliga variansen för de olika tillgångarnas avkastning. Alltså minskar risken ju fler tillgångar som investeringen delas upp mellan.

2.3 Systematisk och idiosynkratisk risk

Antagandet att olika tillgångars avkastning är okorrelerade ifrågasätts av Markowitz (1959) som menar att tillgångar genom ett gemensamt beroende av breda ekonomiska trender är delvis korrelerade. Det innebär att risken inte helt och hållet kan diversifieras bort genom att investera i oändligt många tillgångar. Markowitz (1959) menar att varje tillgångs slumpmässiga avkastning i sin tur kan delas upp i två nya slumpmässiga variabler, varav den ena är beroende av ekonomiska trender och korrelerad med andra tillgångars avkastning och den andra är företagsspecifik och helt okorrelerad med andra tillgångars avkastning. Risken som är kopplad till den första delen är då den risk som i någon mening är gemensam för alla tillgångar och benämns som *systematisk risk*. Risken kopplat

till den andra variabeln benämns som företagsspecifik risk, eller *idiosynkratisk risk*. Hur stor andel av en enskild tillgångs totala risk som består av systematisk risk respektive idiosynkratisk risk beror på avkastningens känslighet för variationer i det bredare ekonomiska klimatet.

2.4 Capital Asset Pricing Model

Baserat på arbetet i Markowitz (1959) så formulerar Sharpe (1964) Capital Asset Pricing Model, eller CAPM.¹ Sharpe (1964) menar att en rationell investerare bör eftersträva den högsta möjliga förväntade avkastningen för varje risknivå eller, omvänt, den lägsta risken för varje given förväntad avkastning. Detta kan formuleras som att varje rationell investerare bör investera i den tillgång eller portfölj som har högst Sharpekvot, definierad enligt följande

$$S = \frac{E[\tilde{r}] - r_f}{\sigma} \quad (2)$$

där $E[\tilde{r}]$ är tillgångens förväntade avkastning, r_f är den riskfria räntan och σ är avkastningens standardavvikelse.

En konsekvens av slutsatserna i Sharpe (1964) är att i en effektiv marknad blir investerare inte kompenserade för att bära idiosynkratisk risk eftersom denna är möjlig att diversifiera bort utan extra kostnad och därmed en risk som i någon mån är ”frivillig” att bära. Däremot bör investerare bli kompenserade för den del av tillgångens risk som kan härledas till den systematiska risken, d.v.s. den risk som är gemensam för hela marknaden. Därför måste, enligt Sharpe (1964), följande samband råda på en effektiv marknad

$$E[\tilde{r}] = r_f + \beta(E[\tilde{r}_M] - r_f) \quad (3)$$

där $E[\tilde{r}]$ är tillgångens förväntade avkastning, $E[\tilde{r}_M]$ är marknadens förväntade avkastning, r_f är den riskfria räntan och β definieras enligt ekvation (4) och är ett mått på tillgångens känslighet gentemot marknadsrisken.

$$\beta = \frac{Cov(\tilde{r}, \tilde{r}_M)}{Var(\tilde{r}_M)} \quad (4)$$

Läsaren bör dock notera att definitionen av β i Sharpe (1964) i praktiken inte skiljer sig från det β -värde som erhålls med en enkel OLS-regression av

¹W. Sharpe, J. Treynor, J. Lintner och J. Mossin utvecklade oberoende av varandra det ramverk som kom att kallas CAPM.

följande samband

$$r_t - r_f = \alpha + \beta(r_{M,t} - r_f) + \epsilon_t \quad (5)$$

Här är r_t och $r_{M,t}$ tillgångens respektive marknadens observerade avkastning vid tidpunkt t . Detta eftersom den äkta kovariansen mellan tillgången och marknaden, samt marknadsavkastningens variation, är okänd. Därmed måste β i Sharpe (1964) uppskattas med stickprovsvariansen enligt ekvation (6).

$$\begin{aligned} \beta_{Sharpe} &= \frac{Cov(\tilde{r}, \tilde{r}_M)}{Var(\tilde{r}_M)} = \frac{\frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})(r_{M,t} - \bar{r}_M)}{\frac{1}{N-1} \sum_{t=1}^N (r_{M,t} - \bar{r}_M)^2} = \\ &= \frac{\sum_{t=1}^N (r_t - \bar{r})(r_{M,t} - \bar{r}_M)}{\sum_{t=1}^N (r_{M,t} - \bar{r}_M)^2} = \beta^{OLS} \end{aligned} \quad (6)$$

I ekvation (5) inkluderas interceptet α . Enligt Sharpe (1964) måste denna koefficient, dock under den effektiva marknadshypotesen, vara lika med 0. Residualen, ϵ_t , är skillnaden mellan den avkastning som modellen i Sharpe (1964) förutspår och den verkliga avkastningen. I detta ramverk kan $\tilde{\epsilon}$ tolkas som den slumpmässiga del av avkastningen som inte förklaras av marknaden och $Var(\tilde{\epsilon})$ är då ett mått på tillgångens idiosynkratiska risk.

2.5 Fama-Frenchs trefaktormodell

CAPM blev snabbt accepterad inom akademien och är än idag en viktig grundsten inom portföljvalsteori. Modellen förutsätter att priser på en komplett marknad kommer att justeras så att varje tillgångs pris är satt så att den förväntade avkastningen bara kan förklaras av tillgångens β -värde. Detta är något som ifrågasätts av forskare som menar att marknadens risk inte är den enda källan till icke-diversifierbar risk. Om så är fallet så bör investerare kompenseras även för denna risk genom en högre förväntad avkastning. Efter att ha observerat att mindre bolag ofta har en högre avkastning än stora bolag använder Banz (1981) marknadsvärde som förklarande variabel för historisk avkastning. Denna två-faktormodell visar sig ha en högre förklaringsgrad än CAPM och väckte därför akademins intresse av att finna ytterligare källor till icke-diversifierbar risk. Flera andra forskare har kompletterat Banz (1981) arbete med studier som visar på en rad olika faktorer som hjälper till att förklara historiska avkastningar. Bland dessa faktorer återfinns bland annat skuldsättningsgrad, lönsamhet, PE-tal och book-to-market ratio.

Fama och French (1992) genomför en omfattande studie där många av dessa faktorer, inklusive marknadsrisk, testas i en multivariabelmodell. Eftersom många av dessa variabler är korrelerade med varandra visar det sig att endast två nya variabler förblir relevanta efter att regressionsanalyserna är genomförda. De kvarstående variablerna är marknadsvärde och ett mått på företagets relativa värde i form av kvoten mellan bokfört värde² och marknadsvärde (book-to-market eller i denna uppsats benämnt som värdefaktor). Tre-faktorsmodellen som Fama och French (1992) presenterar använder marknadsavkastningen som en förklarande variabel men även avkastningen från två självbalanserade portföljer som proxy för avkastning för de faktorer som de ämnar att undersöka.

2.5.1 Storleksfaktorn: Small minus big

Precis som Banz (1981) noterar Fama och French (1992) att små bolag historiskt tenderar att ha en högre genomsnittlig avkastning än stora bolag. Denna skillnad kvarstår efter att avkastningen justerats efter känslighet mot marknadsrisken. Det måste alltså finnas ytterligare risk som investerare blir kompenserade för. Eftersom risken i sig själv inte går att observera så måste avkastningen som är kopplat till storleksrisken uppskattas genom något som är möjligt att observera. Fama och French (1992) väljer att år för år dela in företagen med avseende på storlek. Därefter skapar de en självbalanserad portfölj som består av långa positioner av små bolag och korta positioner av stora bolag. Denna portfölj, eller faktor, benämns som Small-minus-big (SMB) och dess avkastning kan sedan användas som förklarande variabel för att försöka fånga den variation i avkastning som kan hänföras till storleksrisken.

2.5.2 Värdefaktorn: High minus low

I jakten på ytterligare riskfaktorer finner Fama och French (1992) att bolag med högt bokfört värde i förhållande till marknadsvärde (även benämnt book-to-market eller relativt värde) tenderade att ge en högre genomsnittlig avkastning än de bolag som hade liten mängd bokfört värde i förhållande till marknadsvärde. Enligt Fama och French (1992), kan denna variabls förklarande effekt bero på att företag med osäkra framtidsutsikter bestraffas med en högre kapitalkostnad. De menar samtidigt att variabeln kan täcka en mängd olika riskfaktorer som var och en för sig är svåra att hitta proxys för. Fama och French

²På svenska ofta benämnt eget kapital.

(1992) visar också att denna faktor har större förklarande effekt på avkastningarna än storleksfaktorn. Den förklarande variabeln skapas även i det här fallet genom avkastningarna för en självbalanserande portfölj med långa positioner i bolag som har högt book-to-market och korta i bolag med lågt book-to-market. Namnet på faktorn blir således High-minus-low (HML).

2.6 Value at Risk

Konsekvensen av att äga en tillgång med hög risk är att både sannolikheten för en avkastning som skiljer sig från det förväntade värdet och storleken på avvikelsen är högre än för en tillgång med låg risk. För många investerare är det viktigt att uppskatta vilken potentiell nedsida man är exponerad mot via sin investering. Mer formellt kan denna uppskattning formuleras som vilket avkastning en investering inte bör understiga, given en viss sannolikhet. Detta värde benämns som *Value at Risk* eller VaR och anges tillsammans med en önskad säkerhetsnivå, exempelvis 95% eller 99%. Matematiskt kan $\text{VaR}_{95\%}$ beskrivas som det värde som uppfyller

$$0.05 = P(\tilde{r} \leq \text{VaR}_{95\%}) \quad (7)$$

Detta värde beror både på den slumpmässiga variabelns varians och fördelning, varför det är av stor vikt att antaganden om dessa är korrekta för att sannolikheten att underskrida VaR inte ska skilja sig från den önskade säkerhetsnivån.

2.7 ML-skattning av modellparametrar

Om vi antar att den slumpmässiga variabeln, \tilde{X} , tillhör en kontinuerlig sannolikhetsfördelning som beror på ett ändligt antal parametrar, $\boldsymbol{\theta} = [\theta_1, \dots, \theta_n]^T$, så kan dess täthetsfunktion betecknas $f(x|\boldsymbol{\theta})$ (Rawlings et al., 1998).

Antar vi vidare att $\boldsymbol{x} = [x_1, \dots, x_n]$ är en sekvens av observationer från samma fördelning så kan den simultana fördelningens täthetsfunktion skrivas som

$$f(\boldsymbol{x}|\boldsymbol{\theta}) = f(x_n|x_{n-1}, x_{n-2}, \dots, x_1, \boldsymbol{\theta}) \cdot f(x_{n-1}|x_{n-2}, x_{n-1}, \dots, x_1, \boldsymbol{\theta}) \cdot \dots \cdot f(x_1|\boldsymbol{\theta}) \quad (8)$$

Om observationerna dessutom är oberoende kan uttrycket förenklas till

$$f(\boldsymbol{x}|\boldsymbol{\theta}) = f(x_n|\boldsymbol{\theta}) \cdot f(x_{n-1}|\boldsymbol{\theta}) \cdot \dots \cdot f(x_1|\boldsymbol{\theta}) = \prod_{i=1}^n f(x_i|\boldsymbol{\theta}) \quad (9)$$

Detta uttryck är alltså den simultana täthetsfunktionen för en sekvens av observationer \mathbf{x} givet ett antal parametrar $\boldsymbol{\theta}$. Genom att istället se observationerna, \mathbf{x} , som givna och låta parametrarna, $\boldsymbol{\theta}$, variera kan vi definiera likelihoodfunktionen, $\mathcal{L}(\boldsymbol{\theta}|\mathbf{x})$, enligt

$$\mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) = \prod_{i=1}^n f(x_i|\boldsymbol{\theta}) \quad (10)$$

Sannolikhetsfördelningens parametrar kan då skattas genom s.k. Maximum Likelihood-skattning (ML-skattning), d.v.s. genom att finna de parametrar som maximerar sannolikheten för den givna sekvensen av observationer (Rawlings et al., 1998). De skattade parametrarna kan alltså uttryckas som

$$\hat{\boldsymbol{\theta}}_{MLE} = \max_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}|\mathbf{x}) \quad (11)$$

2.8 Jarque-Bera test av empirisk fördelning

Teorin bakom ML-skattning grundas på att sannolikhetsfördelningen för den slumpmässiga variabeln är känd så att dess fördelningsfunktion på ett lämpligt sätt kan parametreras. När den verkliga fördelningen inte är känd krävs ett antagande rörande denna. Vanliga sådana antagande i finansiella modeller är att avkastningen över en tidsperiod är normalfördelad, med eller utan tidsberoende varians, eller t-fördelad. Målet är att fördelningen så bra som möjligt ska stämma överens med den verklighet som den empiriska datan är hämtad ifrån. En metod för att testa om den empiriska datan kan antas vara observationer från en normalfördelad variabel är att utföra ett s.k. Jarque-Bera test. Testet mäter om den empiriska fördelningen har en skevhet och toppighet som stämmer överens med en normalfördelad variabel och testvärdet definieras som

$$JB = \frac{n}{6}(S^2 + \frac{1}{4}(C - 3)^2) \quad (12)$$

där S och C ges av

$$S = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^{3/2}} \quad \text{och} \quad C = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} \quad (13)$$

Enligt Jarque och Bera (1980) är JB asymptotiskt fördelad enligt $\chi^2(2)$ under förutsättning att datan verkligen är normalfördelad. Detta gör att hy-

potesen att datan är normalfördelad kan förkastas med signifikansnivå α ifall $JB > \chi_\alpha^2(2)$.

2.9 Likelihood ratio-test av nästlade modeller

Om två modeller är konstruerade så att den ena modellen kan ses som ett specialfall av den andra genom att dess parametrar begränsas benämns dessa som nästlade modeller. Sådana modeller kan jämföras mot varandra genom ett s.k. Likelihood ratio-test (LR-test). Formellt testas ifall hypotesen H_0 kan förkastas till fördel för H_1 där de båda hypoteserna formuleras enligt

$$H_0 : \boldsymbol{\theta} \in \boldsymbol{\Omega}_0$$

$$H_1 : \boldsymbol{\theta} \in \boldsymbol{\Omega}_0^C$$

där $\boldsymbol{\theta}$ betecknar modellens verkliga parametrar och $\boldsymbol{\Omega}_0$ betecknar den begränsade mängden av möjliga parametrar som undersöks medan $\boldsymbol{\Omega}_0^C$ betecknar alla möjliga parametrar som inte ingår i den begränsade mängden. Unionen av dessa båda mängder, $\boldsymbol{\Omega} = \boldsymbol{\Omega}_0 \cup \boldsymbol{\Omega}_0^C$, täcker därmed in den totala mängd av möjliga parametrar i vår modell.

Genom att definiera sannolikhetskvoten (eng. Likelihood ratio) som

$$\Lambda(\boldsymbol{x}) = \frac{\sup\{\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{x}) : \boldsymbol{\theta} \in \boldsymbol{\Omega}_0\}}{\sup\{\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{x}) : \boldsymbol{\theta} \in \boldsymbol{\Omega}\}} \quad (14)$$

fås ett mått på hur mycket mer sannolika observationerna, \boldsymbol{x} , är i den obegränsade modellen (Kendall et al., 1991). I ekvation (14) står sup för supremumfunktionen som, enkelt uttryckt, är en något vidgad version av maximumfunktionen och innebär att kvoten tas mellan likelihoodfunktionens högsta möjliga värde för de två olika mängderna av tillåtna parametrar.

För nästlade modeller gäller att $-2 \ln \Lambda(\boldsymbol{x})$ under H_0 asymptotiskt är χ^2 -fördelad med samma antal frihetsgrader som skillnaden i antal obegränsade parametrar i de två modellerna (Wilks, 1938). Teststatistiken för ett LR-test är alltså

$$D = -2 \ln \Lambda(\boldsymbol{x}) = -2 \ln \frac{\mathcal{L}(\boldsymbol{\theta}_0|\boldsymbol{x})}{\mathcal{L}(\boldsymbol{\theta}|\boldsymbol{x})} = 2(\ln \mathcal{L}(\boldsymbol{\theta}|\boldsymbol{x}) - \ln \mathcal{L}(\boldsymbol{\theta}_0|\boldsymbol{x})) \quad (15)$$

och H_0 kan förkastas med signifikansnivå α ifall $D > \chi_\alpha^2(f)$.

2.10 Tidigare studier

Det finns ett antal studier gjorda på idiosynkratisk risk med olika fokus. Bartram et al. (2009) och Bartram et al. (2012) försöker förklara varför amerikanska företag generellt har högre idiosynkratisk risk än företag i övriga världen. De gör detta genom att testa olika faktorer förklaringsförmåga och jämföra dem mellan olika länder. Bartram et al. (2016) undersöker sambandet mellan systematisk risk och idiosynkratisk risk. Författarna visar att det finns en positiv relation mellan risktyperna, även då de kontrollerar för storleksfaktorn och värdefaktorn. En potentiell förklaring enligt författarna är att det beror på att olika företag har olika förmåga att kapitalisera på en högre nivå av osäkerhet.

Gempesaw (2014) undersöker bland annat storleksfaktorn och värdefaktor och den volatilitet som trefaktormodellen i Fama och French (1992) inte förklarar. De undersöker om denna oförklarade volatiliteten kan dölja en riskfaktor som påverkar den avkastning som investerare kräver av en tillgång, vilket Ang et al. (2006) menar i sin artikel. Gempesaw (2014) testar detta genom att bilda diversifierade portföljer och finner då att den icke-förklarade volatiliteten inte förklarar följande avkastningar på portföljnivå och således inte döljer systematisk risk.

3 Data

Denna uppsats bygger vidare på data från Johnsson (2017). Data hämtas primärt från databasen Datastream. Johnsson (2017) uppskattar att det under perioden januari 1997 till december 2015 varit 659 stycken aktier listade på Stockholmsbörsen, exklusive preferensaktier. Av dessa används 220 aktier i vår undersökning, varav 198 också ingår i Johnsson (2017) och 22 tillkommer. De bolag som tillkommer är bolag som i Johnsson (2017) exkluderats på grund av att data inte fanns tillgänglig för 24 månader eller mer vid tidpunkten för undersökningen³. Detta 24 månaders lagg är nödvändigt eftersom både värdefaktorn och marknadsbeta beräknas baserat på förgående års data, medan den undersökta volatiliteten är från det aktuella året. Denna studie har tillgång till data för ytterligare ett år och därför tillkommer 22 aktier och även ytterligare finansiell data för året december 2015 till december 2016 för de bolag som inte avnoterats. De noteringar som gjorts under 2016 och 2017 på Stockholmsbörsen kommer ej att undersökas närmare eftersom de inte möjligen kan uppfylla kravet på att ha

³Aktierna BULTEN B, CREADES A och SANITEC OY har exkluderats i Johnsson (2017) trots att de, vid tidpunkten för undersökningen, varit noterade längre än 24 månader

varit noterade minst 24 månader vid tidpunkten för denna uppsats.

Utvidgningen av de befintliga aktielistorna från Johnsson (2017) görs genom att vi kontrollerar de aktier som uteslutits på grund av att de vid december 2015 inte varit noterade i minst 24 månader. Av dessa aktier så återfanns de 22 aktier som vi i denna undersökning inkluderar. Kontroll görs även för avnoteringar, både av aktier som inkluderats av Johnsson (2017) och de som adderats till vår undersökning. Denna kontroll görs genom en genomgång av pressmeddelanden från Nasdaq Stockholm för perioden 2015-06-01 till 2017-06-30. Pressmeddelandena offentliggörs en tid innan avnoteringen sker, varför även pressmeddelandena innan årsskiftet 2015/2016 inkluderas. Inget av de bolag som inkluderas i denna undersökning hade avnoterats under perioden (Nasdaq, 2015) (Nasdaq, 2016).

Tidsserierna kan inte tillförlitligt sträckas längre bak i tiden än januari 1997 eftersom varken Nasdaq OMX eller annan tillförlitlig källa har information om vilka aktier som noterats, avnoterats eller bytt namn innan dess (Johnsson, 2017). Problematiken kring att inkludera en längre tidsserie uppstår således i form av dålig tillförlitlighet beträffande att aktien verkligen varit noterad på Stockholmsbörsen och att relevant data visas för rätt bolag vid rätt tidpunkt. Med bakgrund mot detta resonemang har vi likt Johnsson (2017) valt att begränsa vår undersökning till tidigast januari 1997. Det bör också noteras att vår undersökning i praktiken börjar först i juni 1998 på grund av lagget.

Räntan på Svenska statsskuldväxlar hämtas från Riksbankens hemsida. För denna undersökning används den kortaste tillgängliga löptiden, vilket är en månad (Riksbanken, 2017a).

För att beräkna värdefaktorn så behöver vi veta marknadspris och bokfört värde. Datatypen "Price-to-book-value" (PTBV) finns i Datastream endast tillgänglig för delar av urvalet. För resterande aktier beräknas bokfört värde med hjälp av datatyperna "Bokfört värde per aktie" och "Utställda aktier" enligt följande

$$BV = \text{Bokfört värde per aktie} * \text{utställda aktier} \quad (16)$$

där BV är totalt bokfört värde. För en komplett förteckning av vilken data som inhämtas, se tabell 1.

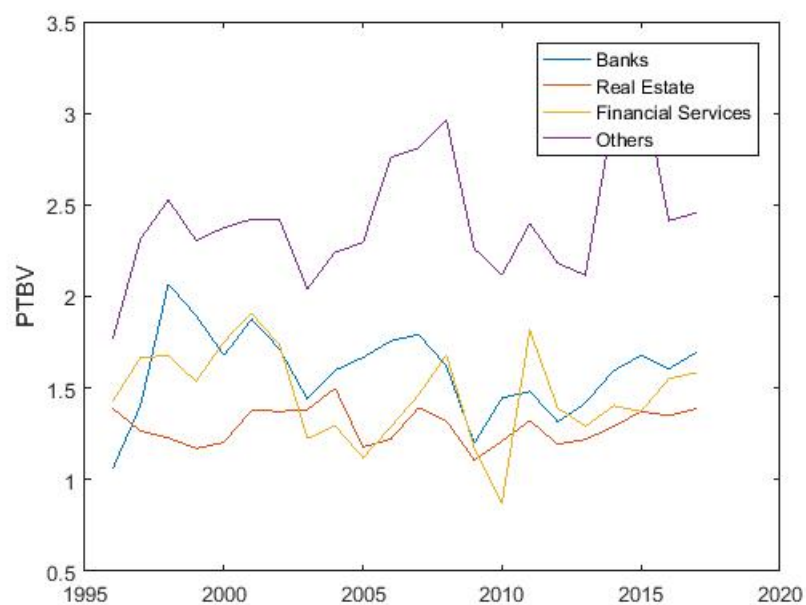
Dataserie	Förkortning	Förklaring	
OMX Large Cap	LC	Listar noterade bolag med ett börsvärde över 1 miljard EUR	
OMX Mid Cap	MC	Listar noterade bolag med ett börsvärde mellan 150 miljoner EUR till 1 miljard EUR	
OMX Small Cap	SC	Listar noterade bolag med ett börsvärde under 150 miljoner EUR	
Ränta Statsskuldväxlar	Rf	Ränta på Svenska stats-skuldväxlar med löptiden 1 månad	
Datatyper	Förkortning	Förklaring	Frekvens
Pris (Justerat)	P	Officiell stängningskurs, justerad för kapitalåtgärder som exempelvis utdelningar och återköp. Anges i SEK.	Dagligen
Pris (Ojusterat)	UP	Officiell stängningskurs. Anges i SEK.	Dagligen
Utdelning	UDDE	Utdelning per aktie, sista handelsdagen innan utdelning. Anges i SEK.	Dagligen
Marknadsvärde	MV	Totalt marknadsvärde, d.v.s. officiell stängningskurs multiplicerat med antalet utställda aktier. Anges i SEK.	Dagligen
Price-to-book value	PTBV	Marknadsvärde genom bokfört värde	Årsslut
Bokfört värde per aktie	WC05491	Bokfört värde per aktie	Årsslut
Antal aktier	WC05301	Antal utställda ordinarie aktier	Årsslut

Tabell 1: Data som inhämtas från Datastream

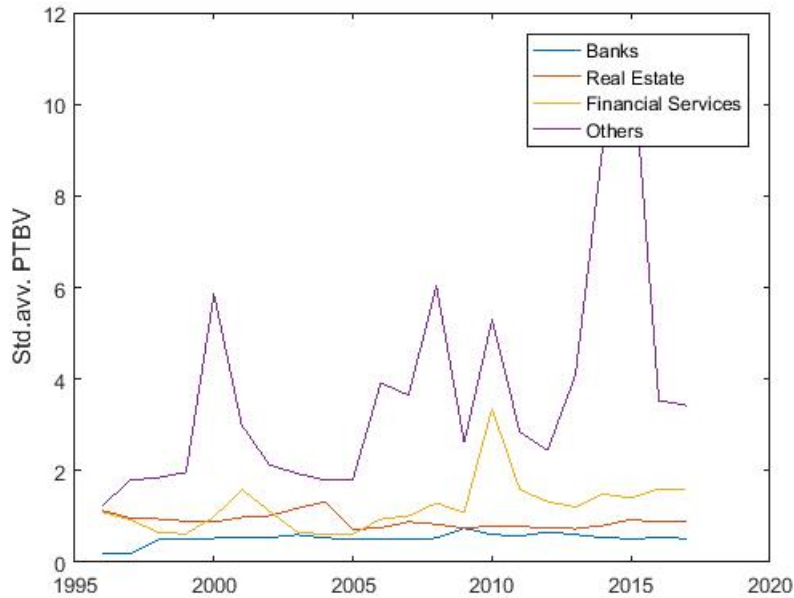
3.1 Bearbetning av data

I likhet med Fama och French (1992) så väljer vi att exkludera banker i vårt urval. Detta eftersom de relativt andra branscher har en ovanligt låg värdefaktor i kombination med liten standardavvikelse i värdefaktor mellan bolag inom branschen under ett år. Även företag som Datastream klassificerar tillhöra branscherna "Financial Services" och "Real Estate Investments and Services" undersöks. De bedöms dock, trots en ovanligt låg värdefaktor, rimliga att inkludera i urvalet eftersom standardavvikelsen i värdefaktorn mellan bolag inom en bransch är cirka dubbelt så stor som i banksektorn. Det indikerar att det finns flera

olika affärsmodeller och operativa verksamheter som skiljer sig mellan bolagen inom de två branscherna. För banker är vår analys istället att de är kraftigt begränsade till en affärsmodell som bygger på att de lånar ut sina tillgångar, vilket medför den låga standardavvikelsen i värdefaktorn. För jämförelse mellan branscherna, se figur 1 och 2.



Figur 1: PTBV för bolag inom 4 olika branscher



Figur 2: Standardavvikelsen i PTBV för bolag inom 4 olika branscher

Vidare har flertalet av bolagen flera aktier där likviditeten ofta skiljer sig väsentligt. För att våra resultat inte ska påverkas av handel i dessa illikvida, ofta röststarka aktier, så används endast den mest likvida aktien för varje företag. Vi anser detta särskilt motiverat eftersom det ofta är större företag som har flertalet noterade aktier, vilket skulle kunna resultera i ett systematiskt fel i resultatet.

3.2 Antaganden

Schwartz och Whitcomb (1977) visar att aktiers avkastning ofta uppvisar autokorrelation. För att minimera dess påverkan på resultatet är ett alternativ att arbeta med en veckofrekvens på aktiernas avkastningar i syfte att begränsa problemet. Dock gör vi avvägningen att fördelen som en daglig frekvens medför, i termer av fler datapunkter som resulterar i mer robusta resultat av vår analys, väger upp för problematiken med autokorrelation. Vidare görs alltså antagandet att avkastningens fördelning inte uppvisar autokorrelation.

Även heteroskedasticitet hos avkastningarnas residualer har dokumenterats vara ett problem av bland annat Morgan (1976). Initialt kommer vi, i likhet med

Campbell et al. (2001a), att göra antagandet att avkastningarnas residualer är homoskedastiska, även om antagandets potentiella påverkan diskuteras djupare i kapitel 5.

Ytterligare en utmaning i datainsamlingen är förekomsten av så kallat Survivorship Bias. Elton et al. (1996) undersöker och beskriver detta fenomen hos fonder. Problemet uppstår när tillgångar försvinner ur datamaterialet, ofta eftersom de inte presterar tillräckligt bra eller, som vid fallet med aktier, för att företag blir uppköpta och avnoterade. Det medför att de kvarvarande tillgångarna potentiellt har, jämfört med de tillgångar som försvunnit, annan karaktäristik och alltså blir urvalet med tiden snedvridet. I urvalet i Elton et al. (1996) ser författarna att knappt 65 procent av aktierna var kvar i urvalet efter 18 år. Det stämmer väl överens med Norberg (2013) som menar att den genomsnittliga livslängden för företag i S&P-500 har minskat från 61 år på 1950-talet till 18 år 2013. Problemet med survivorship bias är större i undersökningar som fokuserar på tillgångars avkastning eftersom det ofta är just otillräcklig avkastning som gör att en tillgång försvinner ur datamaterialet. Vi bedömer dock ändå att det kan ha en viss effekt på vår data. Till exempel kan aktier handlas och sedan avnoteras, vilket döljer delar av den idiosynkratiska volatilitet som en investerare skulle ha haft om inte börsen agerat (SVD Näringsliv, 2013). Framgent kommer vi inte att korrigera för eventuellt survivorship bias eftersom detta vore för resurskrävande för att rymmas inom ramarna för denna undersökning och eftersom problemet är begränsat vid studier av idiosynkratisk volatilitet.

3.3 Beräkning av överavkastning

För att få en rättvisande bild av avkastningen för alla aktier i vår undersökning så tas hänsyn till både prisutvecklingen och direktavkastningen i form av utdelningar. De inhämtade prisuppgifterna är justerade för kapitalåtgärder så som aktiesplit, omvänd split eller återköp som påverkar den faktiska avkastningen. Prisuppgifterna är dock inte justerade för utdelningar, varför de justerade priserna används för att beräkna den dagliga avkastningen för alla dagar utom de handelsdagar då aktierna handlas utan rätt till utdelning. På utdelningsdagar adderas utdelningen till det ojusterade stängningspriset som sedan jämförs med det ojusterade stängningspriset föregående dag. Vid tre tillfällen sammanfaller den ordinarie utdelningsdagen med en utdelning av aktier i andra bolag. Eftersom Datastream inte tar upp värdet av dessa utdelningarna av aktier så

är vi tvungna att korrigera för detta. Vid dessa tillfällen så sätts avkastningen manuellt till 0,0 procent. För information om aktie, datum, händelse och ursprunglig avkastning, se tabell 2. För alla handelsdagar subtraheras sedan den riskfria dagsräntan. Då räntan för statsskuldväxlar med löptid mindre än ett år anges i termer av enkel årsränta så beräknades dagsräntan $r_{SSVX_{dag,t}}$ enligt följande

$$r_{SSVX_{dag,t}} = \left(1 + \frac{r_{SSVX_{år,t}}}{12}\right)^{1/30} - 1 \quad (17)$$

där $r_{SSVX_{dag,t}}$ är dagsräntan dag t och $r_{SSVX_{år,t}}$ är den enkla årsräntan i dag t.

Aktie	Datum	Händelse	Dagsavkastning
Bure Equity	2001-04-27	Delar ut aktier i Observer AB till befintliga aktieägare i Bure Equity	-25,5 procent
Haldex	2011-06-09	Exklusive rätt till aktier i avknoppningen Concentric som noterades på Stockholmsbörsen den 16 juni 2011	-31,1 procent
XANO Industri B	2014-05-09	Delar ut aktier i AGES till befintliga aktieägare i XANO	-44,7 procent

Tabell 2: Aktier vars utdelningsdag sammanföll med utdelning av aktie i annat bolag

3.4 Beräkning av marknadens överavkastning

Eftersom det för perioden inte finns något tillgängligt brett index för Stockholmsbörsen som inkluderade utdelningar⁴, till exempel OMX Stockholm GI, så väljer vi att skapa ett marknadsindex bestående av alla de bolag som ingår i vår studie. Syftet med detta är att i bästa möjliga mån täcka in hela marknaden

⁴OMX Stockholm GI finns tillgängligt för som tidigast 2002-12-30. Även smalare index som OMX Stockholm 60 och OMX Stockholm 30 finns inte tillgängliga från 1997 (Nasdaq, 2017).

i ett index med dagliga datapunkter. Marknadens dagliga avkastning beräknas som den dagliga procentuella förändringen av det totala marknadsvärdet

$$r_{M,t} = \frac{\sum_{i=1}^N MV_{i,t} - \sum_{i=1}^N MV_{i,t-1}}{\sum_{i=1}^N MV_{i,t-1}} \quad (18)$$

Eftersom antalet företag som ingår i urvalet varierar över tid så ingår endast de företag som har ett registrerat marknadsvärde för både dag t och dag $t - 1$ i beräkningen av marknadens avkastning. Marknadens överavkastning beräknas sedan genom att subtrahera den genomsnittliga riskfria dagsräntan.

3.5 Definition av förklarande variabler

På samma sätt som i Fama och French (1992) delar vi upp hela tidsperioden i vår undersökning i årslånga perioder som sträcker sig från 1 juli år y till 30 juni år $y + 1$. Detta för att säkerställa att data hämtat från företagets årsredovisning varit kända för alla investerare för hela den period de används som förklarande variabler. Vi delar sedan upp företagen i 9 grupper baserat på företagets storlek och relativt värde. Storlek definieras som marknadsvärde på dagen före varje årsperiods början. Relativt värde, RV , definieras som kvoten mellan det bokförda värdet i den senaste årsredovisningen och marknadsvärdet

$$RV = \frac{BV}{MV} \quad (19)$$

Då den tillgängliga datatypen anges som Price-to-Book value, d.v.s.

$$PTBV = \frac{P}{BV/(\text{utställda aktier})} \quad (20)$$

kan vi använda sambandet $RV = 1/PTBV$.

4 Metod

Merparten av tidigare forskning har fokuserat på att hitta en avkastningsmodell som har en så hög förklaringsgrad, d.v.s. R^2 -värde, som möjligt. Det innebär att den modell vars residualer har lägst varians anses vara bäst. Exempelvis visar Fama och French (1992) att de kan öka förklaringsgraden för de observerade avkastningarna genom att införa två ytterligare förklarande variabler i CAPM. Det innebär att en del av det som av CAPM anses vara idiosynkratisk risk

istället förklaras som icke-diversifierbar risk kopplat till de två faktorerna i Fama och French (1992). Vår undersökning ämnar inte att öka förklaringsgraden i termer av R^2 -värde. Syftet är istället att undersöka om företagens storlek och värde har ett direkt samband med den idiosynkratiska risken. Vi använder här, och framöver, idiosynkratisk risk som den är definierad i CAPM. Sambandet undersöks genom en modell som ser ut enligt följande

$$\begin{aligned}\tilde{R} &= \alpha + \beta \tilde{R}_M + \tilde{\epsilon} \\ \text{Var}(\tilde{\epsilon}) &= (1 + \beta^S \tilde{X}_S + \beta^B \tilde{X}_B + \beta^H \tilde{X}_H + \beta^L \tilde{X}_L) \sigma^2\end{aligned}\tag{21}$$

I modellen är \tilde{R} överavkastning, \tilde{R}_M är marknadens överavkastning och $\tilde{\epsilon}$ är avkastningarnas residualer. Variablerna \tilde{X}^S , \tilde{X}^B , \tilde{X}^H och \tilde{X}^L är binära variabler som indikerar huruvida ett företag klassas som litet eller stort respektive med högt eller lågt relativt värde. Parametrarna β^S , β^B , β^H , β^L och σ skattas med hjälp av MLE, se avsnitt 2.7. Motiveringen av användandet av denna modell baseras på hypotesen att de faktorer som Fama och French (1992) noterade var förknippade med högre avkastning även borde vara förknippade med högre risk enligt den effektiva marknadshypotesen. Vidare så kan sannolikheten för en viss avvikelse från marknadens avkastning för osofistikerade investerare vara av större intresse än tillgångens känslighet för faktorerna i Fama och French (1992). Vi argumenterar även för att denna modell är lättare att tolka för dessa investerare. I detta avsnitt redogörs närmare för den metod som används i undersökningen.

4.1 Tidsvarierande α och β

Enligt tidigare resonemang är vi i denna undersökning intresserade av variansen på residualen, $\text{Var}(\tilde{\epsilon})$. Då denna beror på både α och β i vår modell så är skattningen av dessa av stor vikt. Dessa värden tillåts variera år för år då det är rimligt att anta att ett företags känslighet för marknadsrisken förändras i takt med att företaget utvecklas. Viktigt att notera är att undersökningen ämnar att modellera den osäkerhet en investerare står inför vid ingången av varje tidsperiod. Därför används α och β skattade med avkastningsdata från föregående år i analysen av avkastningen under följande år. Denna begränsning innebär att vi frångår OLS-estimatet av marknads- β vilket i sin tur innebär att residualernas varians är högre i vår analys än i regressioner som görs a posteriori. Vi anser dock att detta är önskvärt eftersom förändringen av marknads- β från

	Högt relativt värde	Mitten	Lågt relativt värde
Högt marknadsvärde	Högt/Högt	Högt/Mitt	Högt/Lågt
Mitten	Mitt/Högt	Mitt/Mitt	Mitt/Lågt
Lågt marknadsvärde	Lågt/Högt	Lågt/Mitt	Lågt/Lågt

Tabell 3: Alla portföljer som kombinationer av våra förklarande variabler bildar.

år till år är en källa till osäkerhet som i allra högsta grad påverkar risken kopplad till en investering.

Brytpunkterna för de olika portföljerna bestäms till 30 respektive 70 procent. Det innebär att till exempel bolag i småbolagsportföljer tillhör de bolag med 30 procent lägst marknadsvärde, medan storbolagen tillhör de bolag med 30 procent högst. Det bör även noteras att mittenportföljen således är 10 procentenheter större än de båda andra. Resultaten av denna segmentering blir ett antal kombinationer av stora, medel och små företag med högt, medel eller lågt relativt värde. Varje företag tilldelas sedan ett värde av antingen 0 eller 1 på var och en av variablerna X_S , X_B , X_H och X_L beroende på i vilken grupp, eller portfölj, de hamnar i sorteringen. Ett värde på 1 för respektive variabel betyder att företaget tillhör de små (S) eller stora (B) bolagen respektive de med högt (H) eller lågt (L) relativt värde. För samtliga av dessa portföljer, se tabell 3.

4.2 Skattning av parametrar

Vi använder oss av ML-skattning av våra modellparametrar (se avsnitt 2.7 för mer detaljerad teorigenomgång). Tillvägagångssättet för ML-skattningen i vår modell beskrivs nedan mer ingående.

För varje företag i vårt urval definieras, ϵ_i enligt vår modell

$$\epsilon_{i,t} = R_i - \alpha_{i,t} - \beta_{i,t} * R_M \quad (22)$$

Här betecknar subscriptet i det i:e företaget och t betecknar den t:e dagen i vår undersökning. Vi antar att $\epsilon_{i,t}$ är en observation från en underliggande sannolikhetsfördelning. Vi antar vidare att denna slumpmässiga variabel är oberoende både i tiden och mellan de olika företagen, d.v.s. att $Cov(\tilde{\epsilon}_{i,t}, \tilde{\epsilon}_{j,s}) = 0$ för alla $i \neq j$ och $t \neq s$. Det är ett starkt antagande, men som bedöms vara en nödvändig kompromiss i denna undersökning.

Genom detta antagandet kan vi definiera likelihoodfunktionen som

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^N \prod_{t=1}^T f(\epsilon_{i,t}|\boldsymbol{\theta}) \quad (23)$$

De parametrar vi söker är de som maximerar likelihoodfunktionen.

4.2.1 Parametrering av sannolikhetsfördelning

Med hjälp av Jarque-Beras hypotestest av en normalfördelad variabel kan vi förkasta hypotesen att $\tilde{\epsilon}_p$ är normalfördelad för alla portföljer p (se tabell 4 i avsnitt 5.1 för utförligare resultat av detta test). Istället antar vi att den slumpmässiga variabeln tillhör en skalad t-fördelning. En skalad t-fördelad variabel, \tilde{T} , har en varians som beror på både skalningsfaktor och fördelningens parameter, ν , genom förhållandet $Var(\tilde{T}) = s^2 \frac{\nu}{\nu-2}$, där s är skalningsparametern (Jackman, 2009). Vi väljer vi att definiera skalningsparametern genom sambandet

$$s^2 = \frac{\nu-2}{\nu} (1 + \beta^S \tilde{X}_S + \beta^B \tilde{X}_B + \beta^H \tilde{X}_H + \beta^L \tilde{X}_L) \sigma^2 \quad (24)$$

och bibehåller därmed uttrycket för modellens varians från (21), enligt nedan.

$$\begin{aligned} Var(\tilde{\epsilon}) &= s^2 \frac{\nu}{\nu-2} = \frac{\nu}{\nu-2} \frac{\nu-2}{\nu} (1 + \beta^S \tilde{X}_S + \beta^B \tilde{X}_B + \beta^H \tilde{X}_H + \beta^L \tilde{X}_L) \sigma^2 = \\ &= (1 + \beta^S \tilde{X}_S + \beta^B \tilde{X}_B + \beta^H \tilde{X}_H + \beta^L \tilde{X}_L) \sigma^2 \end{aligned} \quad (25)$$

Variabeln $\tilde{\epsilon}$ har således en sannolikhetsfördelning vars täthetsfunktion parametreras av s och ν enligt

$$f(\epsilon|s, \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{\pi \nu} s} \left(1 + \frac{1}{\nu} \left(\frac{\epsilon}{s} \right)^2 \right)^{-\frac{\nu+1}{2}} \quad (26)$$

där $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$ betecknar Gamma-funktionen. Täthetsfunktionen beror alltså på parametern ν men även på β^H , β^L , β^S , β^B och σ^2 genom den beroende parametern s .

4.2.2 Maximering av likelihoodfunktionen

Genom att sätta in $f(\epsilon|s, \nu)$ i ekvation (23) och logaritmera båda sidor så får vi ett uttryck för den funktion vi vill maximera, nämligen

$$\ln \mathcal{L}(\boldsymbol{\theta}) = \sum_{i=1}^N T \left(\ln \left(\Gamma \left(\frac{\nu+1}{2} \right) \right) - \ln \left(\Gamma \left(\frac{\nu}{2} \right) \sqrt{\pi \nu s} \right) \right) - \frac{\nu+1}{2} \sum_{t=1}^T \ln \left(1 + \frac{1}{\nu} \left(\frac{\epsilon_{i,t}}{s} \right)^2 \right) \quad (27)$$

Vi använder Matlabs inbyggda funktion *fminunc* för att finna de parametrar $\boldsymbol{\theta} = [\beta^H, \beta^L, \beta^S, \beta^B, \nu, \sigma^2]^T$ som maximerar $\ln \mathcal{L}(\boldsymbol{\theta})$, eller snarare minimerar funktionen $-\ln \mathcal{L}(\boldsymbol{\theta})$. I appendix C redovisar vi vår databehandling, log-likelihoodfunktionen och den skattning med högst funktionsvärde (se resultatavsnittet 5.2) för mer information om hur metoden exekverades.

Numeriska optimeringsalgoritmer kan vara känsliga för den valda startpunkten för algoritmen, särskilt när många parametrar uppskattas. Därför körs optimeringsalgoritmen ett antal gånger med olika initiala värden på parametrarna.

4.3 Utvärdering av modellen

Utvärdering av modellen sker i två steg. Först jämförs modellen med CAPM genom ett LR-test. Testet visar ifall de adderade parametrarna ökar sannolikheten för våra observationer tillräckligt mycket för att styrka att de är signifikant skilda från 0.

I det andra steget jämförs likelihooden av vår modell med en begränsad modell där vi i tur och ordning sätter en av de adderade parametrarna till 0. Ifall LR-testet visar att någon av parametrarna är överflödiga så stryks denna och samma procedur upprepas tills alla kvarvarande parametrar signifikant ökar sannolikheten för våra observationer.

För den modell som klarat testerna så, och alltså endast har signifikanta parametrar, genomförs även en kvalitativ jämförelse av den av modellen implikerade fördelningen och den empiriska fördelningen av våra observationer. Jämförelsen går ut på att undersöka om andelen observationer som underskrider ett givet kvantilvärde för respektive portfölj, baserat på den varians som beräknas med hjälp av modellen, överensstämmer med motsvarande kvantil.

Slutligen så beräknas både den historiska stickprovsvariansen för respektive portfölj samt beta baserat på denna varians. Beräknade beta kan sedan relateras till våra skattade parametrar och jämföras för validering av resultatet.

5 Resultat och analys

Detta kapitel inleds med att presentera teststatistiken för vårt Jarque-Beratest. Sedan redogör vi för resultatet av ML-skattningarna av parametrarna i den modell som specificeras i ekvation (21) i avsnitt 5.2. Vidare så tolkas och analyseras resultaten i avsnitt 5.4 och 5.5. Slutligen så jämförs parameterskattningen med den beräknade historiska stickprovsvariansen i avsnitt 5.6 och antagandet om homoskedasticitet och dess påverkan analyseras i avsnitt 5.7.

5.1 Resultat av Jarque-Beratest

Hypotesen att de olika portföljernas avkastningar är normalfördelade kan förkastas med en 95% signifikansnivå. Resultatet av testet redovisas i tabell 4.

Portfölj	Testvärde	P-värde
Mid/Mid	$1.6 \cdot 10^7$	≤ 0.05
High/Mid	$1.7 \cdot 10^6$	≤ 0.05
Low/Mid	$6.1 \cdot 10^6$	≤ 0.05
Mid/Small	$1.4 \cdot 10^8$	≤ 0.05
Mid/Big	$1.1 \cdot 10^6$	≤ 0.05
High/Small	$6.9 \cdot 10^7$	≤ 0.05
High/Big	$1.1 \cdot 10^6$	≤ 0.05
Low/Small	$6.5 \cdot 10^7$	≤ 0.05
Low/Big	$6.6 \cdot 10^6$	≤ 0.05

Tabell 4: Resultatet av Jarque-Beratest för de olika portföljernas ej marknadsberoende avkastning

5.2 Resultat av ML-skattning

Vår undersökning visar att den empiriska datan ger stöd för en modell av den typ som vi specificerat i ekvation (21). Det kan dock konstateras att de optimala parametervärdena varierar kraftigt beroende på valda startvärden för optimeringsalgoritmen. För att ge en tydlig bild på detta beroende visas i tabell 5 resultatet av optimeringsalgoritmen för ett antal olika startvärden. Dessa värden är ett urval av de startvärden som användes i hopp om att hitta den bästa skattningen av parametrarna. I tabell 5 redovisas även värdet för funktionen $\ln \mathcal{L}(\hat{\theta})$ för de skattade parametrarna $\hat{\theta}$. Storleken på detta värde indikerar

bättre matchning av parametrar för vår empiriska data.

Vi konstaterar att variationen i skattningarna är ett stort problem i vårt resultat som bör undersökas vidare. Förslagsvis kan en teknik kallad "Simulated annealing" användas. Det är en probabilistisk teknik för att bestämma globala optimum. Tekniken bedöms dock för komplex för att inkluderas i denna uppsats och lämnas till framtida undersökningar. Fortsättningsvis kommer vi att använda skattningen med högst funktionsvärde (högst upp i tabell 5). Om inget annat anges kommer denna skattning att benämnas som vår "skattning" och "parametrarna" som samma skattnings resultat.

Startvärden [$\beta^H, \beta^L, \beta^S, \beta^B, \nu, \sigma^2$]	Optimerade parametervärden [$\beta^H, \beta^L, \beta^S, \beta^B, \nu, \sigma^2$]	Funktionsvärde $\ln \mathcal{L}(\hat{\theta})$
[-0.45, 0.45, -0.45, 0.45, 3.0, $10.0 \cdot 10^{-3}$]	[8.54, 0.90, 9.30, 0.40, 2.0, $33.0 \cdot 10^{-3}$]	$2.94 \cdot 10^6$
[-0.04, 0.54, 1.28, -0.53, 3.0, $0.7 \cdot 10^{-3}$]	[1.73, 1.08, 1.51, 2.51, 2.0, $12.0 \cdot 10^{-3}$]	$2.93 \cdot 10^6$
[-0.05, 0.20, 1.00, -0.40, 3.0, $100 \cdot 10^{-3}$]	[-0.05, 0.20, 1.00, -0.40, 3.0, $0.3 \cdot 10^{-3}$]	$2.82 \cdot 10^6$
[-0.45, 0.45, -0.45, 0.45, 6.0, $1000 \cdot 10^{-3}$]	[-0.35, 2.67, -0.35, 1.35, 2.0, $4.4 \cdot 10^{-3}$]	$2.81 \cdot 10^6$
[0.00, 0.00, 0.00, 0.00, 5.0, $100 \cdot 10^{-3}$]	[-0.00, -0.00, 1.00, 1.00, 5.0, $0.2 \cdot 10^{-3}$]	$2.77 \cdot 10^6$
[-0.20, 0.20, -0.20, 0.20, 5.0, $100 \cdot 10^{-3}$]	[-0.20, 0.20, -0.20, 0.20, 5.0, $0.3 \cdot 10^{-3}$]	$2.71 \cdot 10^6$

Tabell 5: Sammanställning av resultat från Matlabs optimeringsalgoritm *fminunc*. Resultaten är sorterade efter fallande funktionsvärde.

För att kunna testa de olika parametrarnas signifikans genom ett LR-test genomfördes ett antal ML-skattningar av de nästlade modellerna, d.v.s. modeller där en eller flera av parametrarna begränsas. I tabell 6 redovisas resultatet av ML-skattningen av den begränsade modell som motsvarar CAPM. I denna skattning sätts $\beta^H = \beta^L = \beta^S = \beta^B = 0$ och således optimeras likelihoodfunktionen enbart med avseende på ν och σ^2 . Även i denna optimering användes ett antal olika startvärden i ett försök att hitta den globalt optimala skattningen. I tabell 7 redovisas resultatet av ML-skattningen för de fyra modeller där tre av parametrarna $\beta^H, \beta^L, \beta^S, \beta^B$ begränsas till att vara noll. I tabell 8 redovisas resultatet av optimeringen av parametrarna för fyra modeller som begränsas genom att en av parametrarna $\beta^H, \beta^L, \beta^S, \beta^B$ sätts till att vara noll. I de två senare modellerna används samma startvärden för de respektive parametrarna som används i det försök som ger bäst resultat för den obegränsade modellen.

Startvärden	Optimerade parametervärden	Funktionsvärde
$[\nu, \sigma^2]$	$[\nu, \sigma^2]$	$\ln \mathcal{L}(\hat{\theta})$
$[3.0, 10 \cdot 10^{-3}]$	$[2.0, 9.8 \cdot 10^{-3}]$	$2.84 \cdot 10^6$
$[3.0, 1000 \cdot 10^{-3}]$	$[2.4, 0.5 \cdot 10^{-3}]$	$2.82 \cdot 10^6$
$[3.0, 100 \cdot 10^{-3}]$	$[3.0, 3.0 \cdot 10^{-3}]$	$2.80 \cdot 10^6$
$[6.0, 1000 \cdot 10^{-3}]$	$[5.5, 0.2 \cdot 10^{-3}]$	$2.7 \cdot 10^6$

Tabell 6: Optimering av parametrar i den begränsade modell som motsvarar CAPM

Inkluderade parametrar	Startvärden	Optimerade parametervärden	Funktionsvärde
$[\beta^H, \nu, \sigma^2]$	$[-0.45, 3.0, 10 \cdot 10^{-3}]$	$[1.97, 2.0, 3.2 \cdot 10^{-3}]$	$2.85 \cdot 10^6$
$[\beta^L, \nu, \sigma^2]$	$[0.45, 3.0, 10 \cdot 10^{-3}]$	$[0.69, 2.0, 16.4 \cdot 10^{-3}]$	$2.85 \cdot 10^6$
$[\beta^S, \nu, \sigma^2]$	$[-0.45, 3.0, 10 \cdot 10^{-3}]$	$[9.10, 10.33, 0.12, 2.2 \cdot 10^{-3}]$	$2.89 \cdot 10^6$
$[\beta^B, \nu, \sigma^2]$	$[0.45, 3.0, 10 \cdot 10^{-3}]$	$[0.69, 2.0, 16.4 \cdot 10^{-3}]$	$2.85 \cdot 10^6$

Tabell 7: Optimering av parametrar i begränsad modell med endast en extra parameter i förhållande till CAPM

Inkluderade parametrar	Startvärden	Optimerade parametervärden	Funktionsvärde
$[\beta^H, \beta^L, \beta^S, \nu, \sigma^2]$	$[-0.45, 0.45, -0.45, 3.0, 10 \cdot 10^{-3}]$	$[10.34, 0.72, 10.93, 2.0, 7.3 \cdot 10^{-3}]$	$2.92 \cdot 10^6$
$[\beta^H, \beta^L, \beta^B \nu, \sigma^2]$	$[-0.45, 0.45, 0.45, 3.0, 10 \cdot 10^{-3}]$	$[1.95, 0.88, 0.29, 2.0, 18.6 \cdot 10^{-3}]$	$2.89 \cdot 10^6$
$[\beta^H, \beta^S, \beta^B \nu, \sigma^2]$	$[-0.45, -0.45, 0.45, 3.0, 10 \cdot 10^{-3}]$	$[1.97, 2.0, 3.2 \cdot 10^{-3}]$	$2.85 \cdot 10^6$
$[\beta^L, \beta^S, \beta^B \nu, \sigma^2]$	$[0.45, -0.45, 0.45, 3.0, 0.01]$	$[1.08, 2.84, 0.29, 2.0, 23.0 \cdot 10^{-3}]$	$2.92 \cdot 10^6$

Tabell 8: Sammanställning av resultat för ML-skattning av modell begränsad genom att en parameter sätts till noll

5.3 Resultat av LR-test

De olika parametrarnas signifikans testas med hjälp av nio olika hypotestest enligt den metod som beskrivs i avsnitt 2.9. I det första testet jämförs den obegränsade modellen med den begränsade modellen som motsvarar CAPM. Sedan genomförs fyra test där den obegränsade modellen jämförs med de modeller där en enda parameter satts till noll. Till sist testas de fyra modeller där tre av parametrarna satts till noll mot den modell som motsvarar CAPM. Resultatet av testen redovisas i tabell 9.

Ordning	Test	D-värde	P-värde
1	Modellen mot CAPM	$2.018 \cdot 10^5$	≤ 0.05
2	Modellen minus β^S mot CAPM	$4.223 \cdot 10^4$	≤ 0.05
3	Modellen minus β^B mot CAPM	$1.044 \cdot 10^5$	≤ 0.05
4	Modellen minus β^H mot CAPM	$1.762 \cdot 10^5$	≤ 0.05
5	Modellen minus β^L mot CAPM	$4.276 \cdot 10^4$	≤ 0.05
6	β^S mot CAPM	$2.564 \cdot 10^4$	≤ 0.05
7	β^B mot CAPM	$2.343 \cdot 10^4$	≤ 0.05
8	β^H mot CAPM	$2.564 \cdot 10^4$	≤ 0.05
9	β^L mot CAPM	$2.343 \cdot 10^4$	≤ 0.05

Tabell 9: Sammanställning av resultat för ML-skattning av modell begränsad genom att en parameter sätts till noll

Resultatet visar tydligt att de obegränsade modellerna signifikant ökar sannolikheten för våra observationer. Det innebär att parametrarna, tillsammans och var och en för sig, är signifikanta.

5.4 Ekonomisk tolkning av resultat

Enligt vår hypotes så förväntar vi oss att, relativt mittenportföljen, se större idiosynkratisk volatilitet för småbolag och lägre för storbolag. Resultatet från våra ML-skattningar visade sig istället vara motstridigt. Två av de sex skattningar visar på resultat som var i linje med det förväntade, tre av sex visade på

ett motsatt samband och en visade på en förhöjd volatilitet för både stor- och småbolagsportföljerna relativt mittenportföljen.

Vidare har fyra av sex skattningar av småbolagens beta positivt tecken, vilket betyder att de flesta skattningar visar på en större volatilitet i småbolag relativt mittenportföljen. Mest förvånande är dock att fem av sex skattningar av storbolagens beta också har positivt tecken, vilket innebär att storbolagen också är förknippade med större risk relativt mittenportföljen. För en matematisk utvidgning av detta resultat, se kapitel 5.5.

De skattade parametervärdena uppvisar även en hög varians i absolutbeloppen. Det är dock inte motstridigt med den ekonomiska tolkningen eftersom skattningarna även inkluderar olika värden för $\tilde{\sigma}^2$ som kan kompensera för skillnaden i skalningseffekten i de olika parametervärdena. Det är enkelt uttryckt endast relationen mellan varje par av parametrar som är intressanta för tolkningen av resultatet.

Gällande värdefaktorn är vår hypotes att, relativt mittenportföljen, se en högre volatilitet för de bolag med ett högt relativt värde och lägre volatilitet för de bolag med lågt relativt värde. Även för denna faktor så visade sig resultaten vara motstridiga. För två av skattningarna är resultaten i linje med det förväntade och för fyra skattningar är relationen mellan parameterparet för värdefaktor motstridigt med det förväntade. För hälften av skattningarna kan vi dessutom notera att parameterparet har samma tecken, vilket indikerar en u-form på parameterns påverkan på den idiosynkratiska volatiliteten. Även detta fenomen utreds mer matematiskt i kapitel 5.5.

Den skattade variansen, $\tilde{\sigma}^2$, för mittenportföljen antar värden mellan 0,0002 och 0,0330 med fyra av sex skattningar under 0,0044. Motsvarande dagsvolatilitet blir då att skattningarna antar värden mellan 1,41 procent och 18,17 procent, med fyra av fem skattningar under 6,63 procent. Det kan jämföras med endast HENNES MAURITZ B (Storbolag på Large Cap Stockholm) och KNOWIT (Småbolag på Small Cap Stockholm) som i vår data har en standardavvikelse på 1,9 respektive 3,5 procent.

Vårt resultat visar en misstänksamt hög volatilitet i mittenportföljen, varför vi vill förstå vad som kan driva den höga volatiliteten. Exempel på vad som kan driva upp den skattade volatiliteten är s.k. "Outliers", alltså avkastningar som under det rådande antagandet om t-fördelade avkastningar är väldigt osannolika. Denna vidare analys inleds därför med att manuellt gå igenom avkastningsdatan och leta efter osannolika avkastningar. Vi bestämmer denna gräns till avkastningar som överskred 50 procent eller underskred negativa 60 procent.

För en fullständig förteckning över dessa observationer se appendix B. Eftersom vår data innehåller många outliers så väljer vi att fördjupat testa antagandet att enskilda aktiers avkastningar är t-fördelade genom att jämföra de av vår modell implicerade teoretiska kvantilerna med de empiriska kvantilerna för varje portfölj enligt beskrivning i 4.3. Denna analys genomförs i avsnitt 5.5.

Genom LR-testet så fastställdes att våra resultat är robusta och att de undersökta faktorerna har en statistiskt signifikant påverkan, med p-värde under 5 procent för samtliga test, på enskilda aktiers idiosynkratiska risk. Det innebär att vi i huvudsak fick det resultat som vi förväntat oss och att användningen av storleksfaktorn och värdefaktorn för att bättre bedöma en akties idiosynkratiska risk framgent är möjlig. För en anställd som får en del av sin tjänstepension i form av aktier i arbetsgivarens företag, och således innehar en betydande mängd idiosynkratisk risk, kan resultatet i vår uppsats användas för att bättre förstå vilken risk som denne är exponerad mot, endast med hjälp av två enkla och lättillgängliga nyckeltal. Således kan vi rekommendera att komplettera CAPM med våra två faktorer.

Baserat på vår antagna t-fördelning och den bästa skattningens parametervärden så kan vi också bilda oss en uppfattning om avkastningars teoretiska fördelning, se avsnitt 5.5 för ytterligare information. Denna fördelning återfinns i tabell 10 och är att jämföra med s.k. value-at-risk beräkningar. I termer av avkastning så kan vi se att för portföljen småbolag med höga relativa värden så är 95 procent value-at-risk -3,5 procent. Motsvarande value-at-risk för portföljen storbolag med lågt relativt värde är -2,8 procent. Tabell 10 erbjuder således en överskådlig sammanfattning av resultatets praktiska implikationer för en investerare.

5.5 Jämförelse av kvantiler

Givet vår skattning i avsnitt 5.2 kan skalningsfaktorn, s , beräknas (se ekvation 24). Då t-fördelningen definieras entydigt av ν och s så kan den idiosynkratiska risken, $Var(\tilde{\epsilon})$, beräknas för de olika portföljerna genom sambandet $Var(\tilde{\epsilon}) = s^2(\frac{\nu}{\nu-2})$. Dessutom kan vi med hjälp av inversen av t-fördelningens fördelningsfunktion beräkna kvantiler för $\tilde{\epsilon}$ för respektive portfölj. Värdet på kvantilerna tolkas som det värde som variabeln understiger med given sannolikhet. I tabell 10 anges tio olika kvantiler och deras respektive värde för de nio olika portföljerna.

Portfölj	0.1%	1%	5%	10%	25%	75%	90%	95%	99%	99.9%
High/Mid	-18.5	-5.8	-2.4	-1.6	-0.7	0.7	1.6	2.4	5.8	18.5
Low/Mid	-17.4	-5.4	-2.3	-1.5	-0.6	0.6	1.5	2.3	5.4	17.4
Mid/Small	-20.2	-6.3	-2.6	-1.7	-0.7	0.7	1.7	2.6	6.3	20.2
Mid/Big	-13.7	-4.3	-1.8	-1.2	-0.5	0.5	1.2	1.8	4.3	13.7
High/Small	-26.8	-8.4	-3.5	-2.3	-1.0	1.0	2.3	3.5	8.4	26.8
High/Big	-22.3	-7.0	-2.9	-1.9	-0.8	0.8	1.9	2.9	7.0	22.3
Low/Small	-26.0	-8.1	-3.4	-2.2	-1.0	1.0	2.2	3.4	8.1	26.0
Low/Big	-21.4	-6.7	-2.8	-1.8	-0.8	0.8	1.8	2.8	6.7	21.4
Mid/Mid	-5.6	-1.7	-0.7	-0.5	-0.2	0.2	0.5	0.7	1.7	5.6

Tabell 10: De teoretiska kvantilerna för de olika portföljerna givet de uppskattade parametrarna i modellen. Alla värden anges i procent.

Under förutsättning att den t-fördelning som vår modell implicerar för tillgångarnas avkastningarna i de olika portföljerna, borde den empiriska datan stämma överens med de beräknade kvantilerna i tabell 10. Det innebär att vi förväntar oss att den andel av de totala observationerna för en given portfölj stämmer överrens med motsvarande kvantilvärde. Som en del av utvärderingen av vår modell jämför vi våra observationer med modellens kvantiler. I tabell 11 listas andelen observationer som understiger motsvarande kvantilvärde från tabell 10. Andelen anges i procent.

Portfölj	0.1%	1%	5%	10%	25%	75%	90%	95%	99%	99.9%
High/Mid	0.008	0.75	7.60	14.76	28.57	65.77	78.00	84.45	91.56	92.76
Low/Mid	0.075	2.29	12.73	21.32	35.15	66.09	77.67	84.55	93.79	96.53
Mid/Small	0.130	2.26	10.50	17.33	29.37	63.84	74.74	80.73	88.56	91.36
Mid/Big	0.021	1.15	8.89	16.42	31.45	65.81	79.27	86.35	94.09	95.46
High/Small	0.025	0.84	6.66	13.45	27.83	72.77	85.36	91.40	97.82	99.26
High/Big	0.001	0.19	3.17	8.07	23.69	76.63	90.63	95.41	98.60	98.91
Low/Small	0.07	2.01	10.11	17.39	31.15	69.11	80.49	86.29	93.77	96.55
Low/Big	0.023	0.55	4.72	10.42	26.29	74.94	88.95	94.17	98.33	98.98
Mid/Mid	0.745	7.26	14.53	16.53	18.76	23.02	25.14	26.83	31.24	35.59

Tabell 11: Procentuell andel av observationerna i respektive portfölj som understiger motsvarande teoretiskt kvantilvärde givet modellens uppskattade parametrar

Resultatet visar att den empiriska datan avviker från de teoretiska kvantilerna. Avvikelseernas storlek varierar mellan de olika portföljerna och de olika kvantilerna, men vi kan konstatera att sju av nio portföljer har en betydligt högre andel utfall under den teoretiska fördelningens 5%-kvantil än de 5% som vi förväntat oss. Vi ser även att andelen utfall som överstiger både 99%- och 99.9%-kvantilen är större än vad kvantilen antyder. Det finns alltså skäl att anta att t-fördelningen inte till fullo beskriver den äkta sannolikheten för de mest extrema utfallen. Detta kan bero på att det i vårt utfall finns företag vars aktievärdering är ytterst känslig för enskilda företagsnyheter, exempelvis bioteknikföretag vars aktievärdering påverkas starkt av forskningsresultat. Dessa extrema utfall är dock inte knutet till en särskild bransch eller tidsperiod. En uttömmande lista av de företag och tillfällen då den dagliga avkastningen varit extrem finns i appendix B.

5.6 Beräknade varianser och parametervärden

Eftersom våra skattade parametervärden är inkoncisa och delvis inte överensstämde med forskningen eller vår hypotes så valde vi att beräkna den historiska dagliga idiosynkratiska stickprovsvariansen för respektive portfölj direkt ur vår data. Vi hoppas med hjälp av denna analys kunna kontrollera våra skattade parametervärden genom att jämföra med respektive portföljs beräknade värde.

Inledningsvis så beräknas variansen för respektive portfölj. Resultaten finns listade i tabell 12. Värt att lyfta fram är variansen för mittenportföljen (i tabell 12 portfölj [0,0,0,0]) som antar värdet $0.7 \cdot 10^{-3}$. Genom kvadratroten ur variansen får vi den dagliga volatiliteten i denna portfölj, vilket är 2,6 procent. Än vanligare är det att ange volatilitet på årsbasis, vilket för vår mittenportfölj är 41,6 procent ⁵.

⁵Volatilitet på årsbasis beräknas genom den dagliga volatiliteten gånger roten ur 252, under antagandet att det är 252 handelsdagar under ett år.

Portfölj [$\beta^H, \beta^L, \beta^S, \beta^B$].	Varians
[0, 0, 0, 0,]	$0.7 \cdot 10^{-3}$
[1, 0, 0, 0,]	$0.5 \cdot 10^{-3}$
[0, 1, 0, 0,]	$0.9 \cdot 10^{-3}$
[0, 0, 1, 0,]	$1.5 \cdot 10^{-3}$
[0, 0, 0, 1,]	$0.3 \cdot 10^{-3}$
[1, 0, 1, 0,]	$1.2 \cdot 10^{-3}$
[1, 0, 0, 1,]	$0.3 \cdot 10^{-3}$
[0, 1, 1, 0,]	$2.1 \cdot 10^{-3}$
[0, 1, 0, 1,]	$0.4 \cdot 10^{-3}$

Tabell 12: Historisk varians i respektive portfölj

Den beräknade variansen i mittenportföljen kan även jämföras med skattningen av variansen ($\tilde{\sigma}^2$). Baserat på hur variansen beräknas i ekvation 21 så ser vi att variansen ingår i alla olika portföljer och den får således stor inverkan på den idiosynkratiska risk som kan beräknas för respektive portfölj baserat på våra skattade parametrar. Den skattade variansen varierar mellan $0.2 \cdot 10^{-3}$ till $33.0 \cdot 10^{-3}$. Vi kan således konstatera att den beräknade variansen hamnar inom det intervall som våra skattningar bildar, vilket indikerar ett någorlunda rimligt resultat av skattningarna.

För storleksfaktorn så får vi de beräknade varianserna $1.5 \cdot 10^{-3}$ och $0.3 \cdot 10^{-3}$ för små respektive stora företag. Den ekonomiska tolkningen att små bolag alltså förväntas ha större idiosynkratisk risk stämmer väl med forskningen på området och vår logiska grund för valet av att undersöka just storleksfaktorn. För värdefaktorn så kunde variansen beräknas till $0.5 \cdot 10^{-3}$ och $0.9 \cdot 10^{-3}$ för hög respektive låg. Motsvarande ekonomiska tolkning, att bolag med högt relativt värde ska vara mindre riskfyllda, stämmer inte överens med forskningen och vår hypotes. Det är dock i linje med de skattade parametervärdena som tolkas i avsnitt 5.4, och vi ser således i båda dessa analyser att resultatet inte är vad som förväntas. För de blandade portföljerna är en kontroll av resultaten svårare eftersom man då måste bedöma de olika parametrarnas relativa signifikans för idiosynkratisk risk.

För att lättare kunna kontrollera våra resultat, som uttrycks i parametervärden, kan vi använda respektive portföljs varians för att beräkna respektive parametervärde. Från vår modell (ekvation 21) ser vi att

$$Var(\tilde{\epsilon}) = (1 + \beta^X)\sigma^2 \quad (28)$$

där β^X är β för parameter X, $Var(\tilde{\epsilon})$ är residualens varians, σ^2 är variansen för mittenportföljen och alla parametrar förutom X är satta till 0. Om vi för våra beräknade varianser ändrar noteringen så att $Var(P_{mid/mid})$ är vår beräknade varians i mittenportföljen och istället för $Var(\tilde{\epsilon})$, där epsilon tillhör en enskild tillgång, så låter vi variansen för portfölj X benämnas $Var(P_X)$. Då kan β^X beräknas enligt

$$\beta^X = Var(P_X)/Var(P_{mid/mid}) - 1 \quad (29)$$

Resultaten av beräkningarna i ekvation 29 hittas i tabell 13.

För storleksfaktorn ser vi att stora bolag mer än halverar variansen i mittenportföljen. Små bolag däremot avviker ännu kraftigare från mittenportföljen, med ett beta på 1,2 och således en ökad mängd idiosynkratisk risk. Detta stämmer väl överens med den ekonomiska teorin och vad vi förväntade oss att se. Intressant är dock att notera skillnaderna i absolutbeloppen av de båda betavärdena, vilket indikerar att effekten på idiosynkratisk risk avtar i och med att företag växer. Det känns också ur ett ekonomiskt perspektiv rimligt eftersom även de största företagen alltid kommer att möta idiosynkratiska utmaningar. Denna slutsats förstärks ytterligare av att Campbell et al. (2001b) finner att idiosynkratisk risk utgör majoriteten av den totala finansiella risken.

Beta [$\beta^H, \beta^L, \beta^S, \beta^B$]	Beräknat värde
[1, 0, 0, 0,]	-0,34
[0, 1, 0, 0,]	0,29
[0, 0, 1, 0,]	1,21
[0, 0, 0, 1,]	-0,58
[1, 0, 1, 0,]	0,73
[1, 0, 0, 1,]	-0,61
[0, 1, 1, 0,]	2,07
[0, 1, 0, 1,]	-0,42

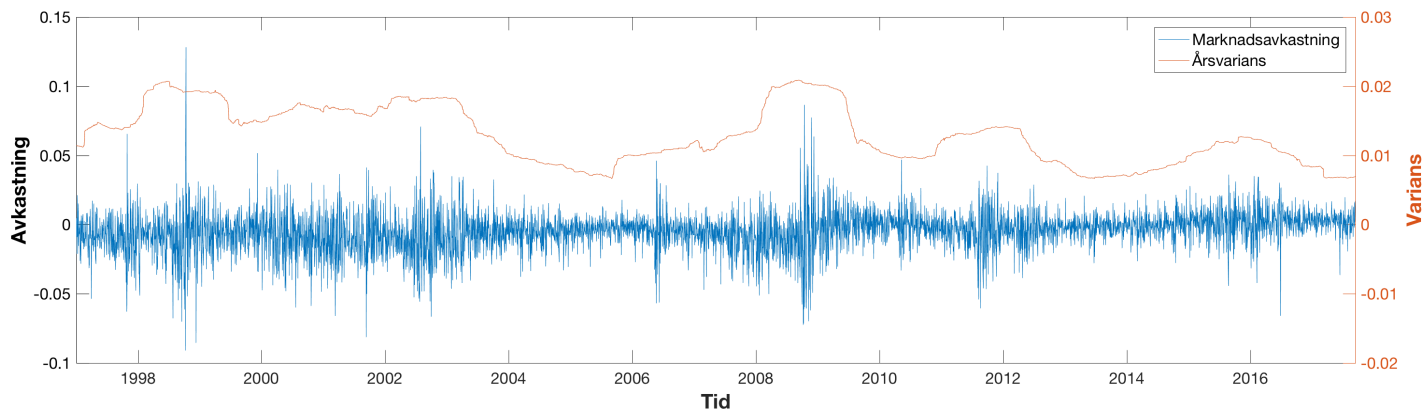
Tabell 13: Beräknat värde för respektive beta. Baseras på historisk volatilitet i respektive portfölj

Värdefaktorn uppvisar ett liknande samband med idiosynkratisk volatilitet som storleksfaktorn. Aktier med lågt relativt värde uppvisade en cirka 30 procent högre idiosynkratisk risk, medan aktier med ett högt relativt värde visade ungefär lika stor motsatt effekt jämfört med mittenportföljen. Att företag som värderas nära sina bokförda tillgångar, och därmed har ett högt relativt värde, generellt anses mer riskfyllda är i enlighet med motiveringen av resultaten i Fama och French (1992) men stämmer inte överens med det vi ser i datan. Vi är dock delvis skeptiska till motiveringen i Fama och French (1992), som menare att företag med högt relativt värde generellt bör anses vara i trångmål och straffas med en högre kapitalkostnad, *ceteris paribus*. Vi såg till exempel i figur 1 att fastighetsbolag uppvisade särskilt låga Price-to-book value och således högt relativt värde. Det betyder inte nödvändigtvis att de företagen är i trångmål, utan är i det fallet endast ett resultat av deras tillgångsintensiva affärsmodell. Vi ställer oss således skeptiska till den logiska grunden bakom värdefaktorns påverkan, även om detta onekligen är en anomali som vi inte förväntas oss.

Sammanfattningsvis kan vi i datan se indikationer på att det samband som vi förväntade se gäller för storleksfaktorn men inte för värdefaktorn i vårt urval. Denna delanalys är värdefull för att hjälpa oss tolka resultaten i vår huvudsakliga analys samt i syfte att fungera som beslutsunderlag åt eventuellt fortsatta studier på området. Däremot ska resultaten i denna delanalys ses med en viss skepsis eftersom metodiken är bristfällig ur ett ekonometriskt perspektiv. Exempel på en sådan brist är att variablernas signifikans inte testas.

5.7 Heteroskedasticitet

Under antagandet att avkastningarnas felterm är homoskedastiska, alltså att $\text{Var}(\epsilon)$ är konstant, så låter vi vår Maximum Likelihood-skattning leta efter de optimala parametervärdena för hela dataunderlaget samtidigt. Resultatet som presenteras blir då i någon mån ett snitt av hur verkligheten har sett ut under perioden från 1997. Eftersom att målet med vår undersökning är att erbjuda investerare verktyg för att bedöma den idiosynkratiska risken framgent för en aktie, och eftersom vi misstänker att vårt antagande kan vara för starkt, så plottade vi glidande årsvolatilitet och dagliga marknadsavkastningar för vårt urval under perioden. Se figur 3.



Figur 3: Marknadsavkastningar och glidande årsvariens i vårt urval

I figur 3 så ser vi tydligt att den observerade volatiliteten hos avkastningarna varierar kraftigt över tid. Det indikerar att standardavvikelsen hos $\text{Var}(\tilde{\epsilon})$ inte är konstant utan varierar över tid, vilket i ekonometrin benämns heteroskedasticitet. Särskilt hög volatilitet kan vi observera runt sekelskiftet 2000, då IT-bubblan uppstod och sprack, samt i samband med finanskrisen 2008-2009. I framtida studier som rymmer en stokastisk volatilitetsmodell för att modellera avkastningarna så visar vår initiala analys att det finns anledning att testa för heteroskedasticitet i det aktuella urvalet.

Matematiskt så innebär denna misstanke att σ^2 i ekvation (21) skulle behöva tillåtas att variera över tid. Om relationen mellan parameterparen också visar sig variera över tid så bör även berörda beta i (21) tillåtas variera över tid, även om det kraftigt skulle komma att påverka användarbarheten och möjligheten till intuitiv förståelse för respektive parameters påverkan på idiosynkratisk risk. Till exempel kan en Generalized Autoregressive Conditional Heteroskedasticity (GARCH) modell användas för att låta σ^2 variera över tid. Simpelt uttryckt så låter man då förväntningarna på σ^2 imorgon baseras på var man trodde att σ^2 skulle vara igår tillsammans med vad σ^2 de facto var igår.

Implementering av en GARCH-modell i vår undersökning kommer tyvärr att påverka hur användarvänliga resultaten i vår undersökning blir. För en mindre sofistikerad investerare kan införandet av en differentialekvation verka skrämmande och således skada vårt syfte. Samtidigt är våra misstankar om heteroskedasticitet starkt baserade på forskning inom området, bland annat Morgan (1976), och figur 3. Därför råder vi framtida studier att testa för heteroskedas-

ticitet och vidare utveckla vår undersökning genom att tillåta för en stokastisk volatilitetsmodell, då gärna en GARCH som erbjuder en bra avvägning mellan funktion och komplexitet.

6 Slutsatser

Resultatet av vår analys är att både storleksfaktorn och värdefaktorn förklarar en del av den idiosynkratiska risken som CAPM beräknar för enskilda aktier. Vi bedömer således vårt syfte vara uppfyllt och att uppsatsens resultat kan användas för att bättre förstå vilken idiosynkratisk risk en enskild aktie kommer att ha framgent. Validering av resultatets signifikans visar på ett robust resultat för samtliga parametrar. Dock varierar skattningarna baserat på våra startvärden för optimeringsalgoritmen, vilket är problematiskt. Framtida studier föreslås därför använda tekniken ”Simulated annealing” för att hitta ett globalt optimum.

För storleksfaktorn kan det observeras att storbolagen har lägre idiosynkratisk volatilitet relativt småbolagen, vilket är i linje med vår hypotes baserat på resultaten i Fama och French (1992). Dock finner vi att beta för storbolagen optimeras till ett positivt värde, vilket implikerar en högre volatilitet relativt mittenportföljen. Detta resultat är inte i linje med det förväntade.

För värdefaktorn visar vår analys ett förvånande resultat. Dels så visar skattningen att bolag med höga relativa värden är betingade med lägre idiosynkratisk risk relativt bolag med lågt relativt värde. Baserat på Fama och French (1992) borde motsatt samband gälla eftersom det höga relativa värdet är en proxy för att företag befinner sig i trångmål och bestraffas med en högre kapitalkostnad. Uppsatsen diskuterar dock kritiskt resonemangen kring värdefaktorn i Fama och French (1992).

Resultaten får oss också att ifrågasätta huruvida enskilda aktiers dagliga avkastning på ett tillfredställande sätt kan modelleras med t-fördelningen. Anledningen är att kursdrivande nyheter vid publicering kan öka sannolikheten för extrema avkastningar, både positiva och negativa.

Avslutningsvis dras slutsatsen att uppsatsens antagande om homoskedasticitet är värt att undersöka vidare. Vid plottning av marknadsavkastningen och glidande årsvarians så framgår det tydligt att den observerade variansen skiljer sig över tid, vilket indikerar att avkastningarna istället uppvisar heteroskedasticitet. Simplicitet i metodval och resultat är av yttersta vikt för denna uppsats syfte, varför en mer komplex modellering av avkastningarna inte används. Framtida studier inom området uppmanas dock att undersöka urvalet för heteroskedasticitet och eventuellt använda sig av en stokastisk volatilitetsmodell, förslagsvis en GARCH-modell som använder sig av en autokorrelerad volatilitet.

Referenser

- Ang, A. , Hodrick, R. J. , Xing, Y. och Zhang, X. (2006), The cross-section of volatility and expected returns. *Journal of Finance*, 61(1):259–299.
- Banz, R. W. (1981), The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9(1):3 – 18.
- Bartram, S. M. , Brown, G. och Stulz, R. M. (2009), Why do foreign firms have less idiosyncratic risk than U.S. firms? *National Bureau of Economic Research*.
- Bartram, S. M. , Brown, G. och Stulz, R. M. (2012), Why are U.S. stocks more volatile? *Journal of Finance*, 67(4):1329–1370.
- Bartram, S. M. , Brown, G. och Stulz, R. M. (2016), Why does idiosyncratic risk increase with market risk? *CESifo Working Paper Series*, (6560).
- Campbell, J. Y. , Lettau, M. och Malkiel, B. G. (2001a), Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. *The Journal of Finance*, 56(1):1 – 43.
- Campbell, R. , Huisman, R. och Koedijk, K. (2001b), Optimal portfolio selection in a value-at-risk framework. *Journal of Banking and Finance*, 25(9):1789 – 1804.
- de Cervantes, M. (1605). *Don Quixote*. Francisco de Robles.
- Elton, E. J. , Gruber, M. J. och Blake, C. R. (1996), Survivorship bias and mutual fund performance. *The Review of Financial Studies*, (4):1097.
- Fama, E. F. och French, K. R. (1992), The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465.
- Gabaix, X. och Maggiori, M. (2015), International liquidity and exchange rate dynamics. *Quarterly Journal of Economics*, 130(3):1369 – 1420.
- Gempesaw, D. (2014). Does idiosyncratic volatility proxy for a missing risk factor? Evidence from using portfolios as test assets. Masteruppsats vid Department of Economics, Miami University.
- Guiso, L. och Sodini, P. (2013). *Household Finance: An Emerging Field*, volume 2. Elsevier.

- Jackman, S. (2009). *Bayesian Analysis for the Social Sciences*. John Wiley & Sons Incorporated.
- Jarque, C. och Bera, A. (1980), Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6:255 – 259.
- Johnsson, O. (2017). Fama-Frenchs femfaktormodell på den svenska aktiemarknaden. Kandidatuppsats vid Nationalekonomiska institutionen, Lunds universitet.
- Kendall, M. G. , Sir, Stuart, A. och Ord, J. K. (1991). *Kendall's advanced theory of statistics*. London : Edward Arnold.
- Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press.
- Morgan, I. G. (1976), Stock prices and heteroscedasticity. *The Journal of Business*, 49(4):496– 508.
- Nasdaq. Pressmeddelanden, Sverige (2015). URL <http://www.nasdaqomxnordic.com/news/marketnotices/stockholm>.
- Nasdaq. Pressmeddelanden, Sverige (2016). URL <http://www.nasdaqomxnordic.com/news/marketnotices/stockholm>.
- Nasdaq. Index, Sverige (2017). URL <http://www.nasdaqomxnordic.com/index>.
- Norberg, J. . En frihetlig revolution (2013). URL <http://www.di.se/artiklar/2014/12/13/debatt-en-frihetlig-revolution/>.
- Perold, A. F. (2004), The capital asset pricing model. *The Journal of Economic Perspectives*, 18(3):3–24.
- Piketty, T. (2014). *Capital in the Twenty-First Century*. Belknap Press: An Imprint of Harvard University Press.
- Rawlings, J. O. , Pantula, S. G. och Dickey, D. A. (1998). *Applied regression analysis: a research tool*. Springer texts in statistics. Springer.

- Riksbanken. Ränta, statsskuldväxlar (2017a). URL <http://www.riksbank.se/sv/Rantor-och-valutakurser/Sok-rantor-och-valutakurser/?g2-SECBREPOEFF=onfrom=1997-01-01to=2017-09-27f=DaycAverage=Averages=Commasearch>.
- Riksbanken (2017b), Penningpolitisk rapport september 2017. (1):31.
- Schwartz, R. A. och Whitcomb, D. K. (1977), The time-variance relationship: Evidence on autocorrelation in common stock returns. *The Journal of Finance*, 32(1):41–55.
- Sharpe, W. F. (1964), Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.
- SVD Näringsliv. Därför kan den billiga aktien stå dig dyrt (2013). URL <https://www.svd.se/darfor-kan-den-billiga-aktien-sta-dig-dyrt>.
- von Neumann, J. och Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.
- Wilks, S. S. (1938), The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, (1):60.

A Företagslistor

Tabell A1: Företag som är med i urvalet

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
AARHUSKARLSHAMN	06-09-12	17-06-30
ABB LTD N	99-06-22	17-06-30
ACANDO B	97-01-03	17-06-30
ACTIVE BIOTECH	97-01-03	17-06-30
ADDNODE B	99-06-11	17-06-30
ADDTECH B	01-09-04	17-06-30
AFRICA OIL CORP.	14-07-02	17-06-30
ALFA LAVAL	02-05-21	17-06-30
ALLTELE AB	09-06-16	17-06-30
ANOTO GROUP	00-06-19	17-06-30
ARCAM	12-06-19	17-06-30
ARCTIC PAPER	12-12-21	17-06-30
ARISE	10-03-25	17-06-30
ASSA ABLOY B	97-01-03	17-06-30
ASTRAZENECA	99-04-07	17-06-30
ATLAS COPCO A	97-01-03	17-06-30
ATRIUM LJUNGBERG B	97-01-03	17-06-30
AVEGA GROUP B	10-12-17	17-06-30
AXFOOD	97-06-30	17-06-30
BACTIGUARD HOLD	14-06-23	17-06-30
BALDER B	99-10-13	17-06-30
BE GROUP	06-11-27	17-06-30
BEIJER ALMA B	97-01-03	17-06-30
BEIJER ELECTRONICS	00-06-14	17-06-30
BEIJER REF. B	97-01-03	17-06-30
BERGS TIMBER B	97-01-03	17-06-30
BESQAB	14-06-13	17-06-30
BETSSON B	97-01-07	17-06-30
BILIA A	97-01-03	17-06-30
BILLERUD KORSNÄS	01-11-21	17-06-30
BIOGAIA A	01-11-21	17-06-30
BIOINVENT INTERN	01-06-13	17-06-30
BIOTAGE	00-07-03	17-06-30
BJÖRN BORG	07-05-08	17-06-30
BLACK EARTH FARM.	09-06-23	17-06-30
BOLIDEN	99-05-04	17-06-30
BONG	97-01-07	17-06-30
BOULE DIAGNOSTICS	11-06-27	17-06-30

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
BTS GROUP	01-06-07	17-06-30
BULTEN	11-05-23	17-06-30
BURE EQUITY	97-01-03	17-06-30
BYGGMAX GROUP	10-06-03	17-06-30
CASTELLUM	97-05-26	17-06-30
CATENA	06-04-27	17-06-30
CAVOTEC	11-10-21	17-06-30
CELLAVISION	07-05-29	17-06-30
CLAS OHLSON B	99-10-06	17-06-30
CLOETTA B	09-02-17	17-06-30
COM HEM HOLDING	14-06-18	17-06-30
CONCENTRIC	11-06-17	17-06-30
CONCORDIA MAR. B	97-01-03	17-06-30
CONSILIUM B	97-01-03	17-06-30
COREM PROPERTY GR.	09-06-25	17-06-30
CRAD B	14-12-17	17-06-30
CREADES A	13-12-09	17-06-30
CTT SYSTEMS	97-11-12	17-06-30
DEDICARE	11-05-05	17-06-30
DGC ONE	08-06-17	17-06-30
DIÖS FASTIGHETER	97-01-03	17-06-30
DORO	07-11-15	17-06-30
DUNI	97-01-03	17-06-30
DUROC B	97-01-03	17-06-30
ELANDERS B	09-06-02	17-06-30
ELECTRA GRUPPEN	97-01-03	17-06-30
ELECTROLUX B	97-01-03	17-06-30
ELEKTA B	97-01-03	17-06-30
ELOS B	12-11-08	17-06-30
ENDOMINES	97-01-03	17-06-30
ENEA	00-10-11	17-06-30
ENIRO	14-06-12	17-06-30
EPISURF MEDICAL	97-01-03	17-06-30
ERICSSON B	10-02-22	17-06-30
EWORK SCANDINAVIA	97-01-03	17-06-30
FABEGE	97-05-14	17-06-30
FAGERHULT	97-01-03	17-06-30
FAST PARTNER	97-05-13	17-06-30
FEELGOOD SVENSKA	14-06-27	17-06-30
FENIX OUTDOOR INTL	00-04-20	17-06-30
FINGERPRINT CARDS B	10-01-20	17-06-30

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
FORMPIPE SOFTWARE	14-06-11	17-06-30
G5 ENTERTAINMENT	97-01-03	17-06-30
GETINGE	08-10-04	17-06-30
GHP SPECIALTY CARE	97-01-03	17-06-30
GUNNEBO	97-01-03	17-06-30
HALDEX	97-01-03	17-06-30
HEBA B	14-03-24	17-06-30
HEMFOSA FASTIGHETER	97-01-03	17-06-30
HENNES & MAURITZ B	97-01-03	17-06-30
HEXAGON B	08-06-10	17-06-30
HEXPOL B	99-04-13	17-06-30
HIQ INTERNATIONAL	07-10-22	17-06-30
HMS NETWORKS	97-01-03	17-06-30
HOLMEN B	97-01-03	17-06-30
HUFVUDSTADEN A	06-06-14	17-06-30
HUSQVARNA B	99-01-05	17-06-30
I A R SYSTEMS GROUP	05-12-09	17-06-30
ICA GRUPPEN	97-01-03	17-06-30
INDUSTRIVÅRDEN C	05-10-06	17-06-30
INDUTRADE	97-01-03	17-06-30
INTELLECTA B	02-06-10	17-06-30
INTRUM JUSTITIA	97-01-03	17-06-30
INVESTOR B	14-09-29	17-06-30
INWIDO	08-07-10	17-06-30
ITAB SHOP CONCEPT B	97-01-03	17-06-30
JM	97-01-03	17-06-30
KABE HUSVAGNAR B	06-02-24	17-06-30
KAPPAHL	98-04-03	17-06-30
KARO PHARMA	11-04-18	17-06-30
KAROLINSKA DEVELOP.	04-06-08	17-06-30
KINDRED GROUP SDR	97-01-03	17-06-30
KINNEVIK B	97-01-03	17-06-30
KLÖVERN B	97-11-11	17-06-30
KNOW IT	99-04-15	17-06-30
KUNGSLEDEN	01-09-06	17-06-30
LAGERCRANTZ GR. B	97-06-26	17-06-30
LAMMHULTS D. GRP.	97-01-03	17-06-30
LATOUB B	97-01-03	17-06-30
LIFCO B	06-12-04	17-06-30
LINDAB INT.	08-12-10	17-06-30
LOOMIS B	97-01-03	17-06-30

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
LUNDBERG B	03-10-03	17-06-30
LUNDIN PETROLEUM	99-03-15	17-06-30
MALMBERGS ELEKTR. B	97-01-03	17-06-30
MEDIVIR B	00-05-30	17-06-30
MEKONOMEN	06-12-07	17-06-30
MELKER SCHÖRLING	11-12-28	17-06-30
MICRO SYSTEMATION B	99-06-16	17-06-30
MIDSONA B	97-01-03	17-06-30
MIDWAY HOLDINGS B	04-03-31	17-06-30
MILLICOM INTL.CELU.S.	10-06-21	17-06-30
MQ HOLDING	98-05-22	17-06-30
MSC KONSULT B	97-09-19	17-06-30
MTG B	98-02-16	17-06-30
MULTIQ INT.	00-03-10	17-06-30
MYCRONIC	97-01-03	17-06-30
NCC B	07-05-18	17-06-30
NEDERMAN HOLDING	99-06-08	17-06-30
NET INSIGHT B	09-01-15	17-06-30
NETENT B	13-04-11	17-06-30
NEUROVIVE PHARMA.	97-12-12	17-06-30
NEW WAVE B	97-06-18	17-06-30
NIBE INDUSTRIER B	02-06-20	17-06-30
NOBIA	97-01-03	17-06-30
NOLATO B	08-07-21	17-06-30
NORDIC MINES	04-06-24	17-06-30
NOTE	99-07-01	17-06-30
NOVOTEK B	14-12-05	17-06-30
NP3 FASTIGHETER	10-06-28	17-06-30
OASMIA PHARMA.	10-07-22	17-06-30
ODD MOLLY INTL.	97-01-08	17-06-30
OEM INTERNATIONAL B	13-07-03	17-06-30
OPUS GROUP	05-11-10	17-06-30
OREXO	04-03-25	17-06-30
ORIFLAME HOLDING	97-01-03	17-06-30
ORTIVUS B	97-01-03	17-06-30
PEAB B	13-12-02	17-06-30
PLATZER FASTIGHETER	99-06-24	17-06-30
POOLIA B	99-12-14	17-06-30
PRECISE BIOMETRICS	98-06-02	17-06-30
PREVAS B	97-01-03	17-06-30
PRICER B	97-10-17	17-06-30

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
PROACT IT GROUP	98-12-17	17-06-30
PROBI	97-06-23	17-06-30
PROFILGRUPPEN B	10-12-16	17-06-30
QLIRO GROUP	97-01-03	17-06-30
RATOS B	97-01-03	17-06-30
RAYSEARCH LABS.B	14-04-04	17-06-30
RECIPHARM	06-12-19	17-06-30
REJLERS B	06-11-29	17-06-30
REZIDOR HOTEL GROUP	97-01-03	17-06-30
ROTTNEROS	98-06-22	17-06-30
SAAB B	13-04-09	17-06-30
SAGAX B	97-01-03	17-06-30
SANDVIK	13-12-10	17-06-30
SANITEC OY	97-01-03	15-02-27
SAS	97-01-03	17-06-30
SCA B	14-06-30	17-06-30
SCANDI STANDARD	12-06-14	17-06-30
SEAMLESS DIST.	99-03-04	17-06-30
SECTRA B	97-01-03	17-06-30
SECURITAS B	97-05-27	17-06-30
SEMCON	01-02-01	17-06-30
SENSYS GATSO GROUP	97-01-03	17-06-30
SINTERCAST	97-01-03	17-06-30
SKANSKA B	97-01-03	17-06-30
SKF B	97-01-03	17-06-30
SKISTAR B	98-12-04	17-06-30
SOFTRONIC B	97-01-03	17-06-30
SSAB A	97-01-03	17-06-30
STOCKWIK FÖRV	00-04-26	14-03-26
STRAX	01-05-07	17-06-30
STUDSVIK	06-09-18	17-06-30
SVEDBERGS B	98-09-23	17-06-30
SW. ORPHAN BIOVIT.	97-10-06	17-06-30
SWECO B	97-01-03	17-06-30
SWEDISH MATCH	08-06-13	17-06-30
SWEDOL B	07-10-15	17-06-30
SYSTEMAIR	97-01-03	17-06-30
TELE2 B	00-06-13	17-06-30
TELIA COMPANY	13-05-03	17-06-30
TETHYS OIL	14-11-27	17-06-30
THULE GROUP	97-07-04	17-06-30

Företagsnamn	Inkl. f.r.o.m.	Exkl. per den
TRACTION B	05-11-09	17-06-30
TRADEDOUBLER	14-12-01	17-06-30
TRANSCOM WW	97-01-03	17-06-30
TRELLEBORG B	04-11-22	17-06-30
UNIFLEX B	97-01-03	17-06-30
VBG GROUP	97-01-03	17-06-30
VICTORIA PARK B	14-05-20	17-06-30
VIKING SUPPLY SHIPS	05-05-24	17-06-30
VITEC SOFTWARE GR. B	97-01-03	17-06-30
VITROLIFE	11-07-05	17-06-30
VOLVO B	01-06-27	17-06-30
VOSTOK NEW VENT.	97-01-03	17-06-30
WALLENSTAM B	07-07-05	17-06-30
WIHLBORGS FASTIGH.	97-01-03	17-06-30
XANO INDUSTRI B	97-01-03	17-06-30
ÅF B	97-01-03	17-06-30
ÖRESUND	97-01-03	17-06-30

Tabell A2: Aktier som ej är med i denna studie men som är med i Johnsson (2017). Dessa aktier exkluderas från urvalet för att de är banker eller för att det redan finns en mer omsatt aktie från samma företag i urvalet.

Företagsnamn

ATLAS COPCO B
AVANZA BANK HOLD.
B&B TOOLS B
EAST CAPITAL EXPL.
ELECTROLUX A
ENQUEST
ERICSSON A
ETRION
HANDELSBANKEN A
HANDELSBANKEN B
HOLMEN A
HUFVUDSTADEN C
HUSQVARNA A
INDUSTRIVÄRDEN A
INVESTOR A
KINNEVIK A
KLÖVERN A
MIDSONA A
MIDWAY HOLDINGS A
MOBERG PHARMA
MTG A
NCC A
NORDEA BANK
NORDNET B
ORTIVUS A
RATOS A
SAGAX A
SCA A
SEB A
SEB C
SEMAFO
SKF A
SSAB B
SWECO A
SWEDBANK A

Företagsnamn

TELE2 A
TIETO
TRIGON AGRI
VICTORIA PARK
VOLVO A

Tabell A3: Aktier som är med i denna studie som ej är med i Johnsson (2017)

Företagsnamn

BACTIGUARD HOLD
BESQAB
BIOGAIA A
BULTEN
COM HEM HOLDING
CRAD B
CREADES A
EPISURF MEDICAL
FENIX OUTDOOR INTL
G5 ENTERTAINMENT
HEMFOSA FASTIGHETER
INWIDO
KLÖVERN B
LIFCO B
NP3 FASTIGHETER
RECIPHARM
SANITEC OY
SCANDI STANDARD
THULE GROUP
TRANSCOM WW
VICTORIA PARK B

Tabell A4: Företag som varit listade på Stockholmsbörsen sedan 1997 men som inte är med i urvalet

ABB A	ABB B	ACADEMEDIA
ACAP INVEST A	ACAP INVEST B	ACRIMO B
ACSC	AEROCRINE B	AINAX
AKZO NOBEL SDB	ALCATEL	ALIMAK GROUP
ALLENEX	ALLIANCE OIL SDB	ALTIMA
ARETE	ARTEMA MEDICAL B	ASPIRO
ASTICUS	ASTRA A	ASTRA B
ATLANTICA FRB.	ATTENDO	AU SYSTEM
AUDIODEV B	AUTOFILL	AUTOLIV SDB
AUTOLIV	AVAILO	AVESTAPOLARIT
AXIS	BALLINGSLOV INTL.	BAYER SDB
BENIMA FERATOR E.	BIOLIN SCIENTIFIC	BIOPHAUSIA A
BLACKPEARL RESC.	BOSTADS AB DROTT	BRAVIDA HOLDING
BRINOVA FASTIGHETER	BUFAB	CAMURUS
CAPINORDIC	CAPIO	CARL LAMM HOLDING
CARL LAMM	CARLI GRY	CASHGUARD B
CELLPOINT SDB	CISION	CLX COMMUNICATIONS
COASTAL CONTACTS	COLLECTOR	COLUMNA
COOR SERVICE MAN.	CYBERCOM GROUP	CYNCRONA B
D CARNEGIE & CO B	D CARNEGIE & CO	DAGON
DIAL NXT GROUP	DIMENSION	DIN BOSTAD SVERIGE
DOMETIC GROUP	DUSTIN GROUP	ELTEL
EMPIRE B	EOLUS VIND B	EPSILON B
FORCENERGY B	FORENINGS BKN. A	FORENINGS BKN.B
FRANGO B	FRILUFTSBOLAGET E	FRONTLINE
GANT COMPANY	GIBECK B	GLOBAL HEALTH PARTN.
GLOCALNET	GOTIC A	GOTIC B
GRANGES	GUIDE KONSULT B	GULLSPANGS KRAFT B
GUNNEBO INDUSTRIER	HANSA MEDICAL	HEMSTADEN BOSTADS
HEMTEX	HEXATRONIC GROUP	HIFAB GROUP
HOIST FINANCE	HOIST INTL. B	HOME PROPERTIES
HQ	HQ FONDER	HUFVUDSTADEN INTL.
HUMAN CARE H C	IMAGE SYSTEMS	IMG INDE.MEDIA GP. B
IMMUNE PHARMA	INVIK&COB	INVISIO COM.
IPC	JAAKKO POYRY GP.	JC
JEEVES INFO.SYSTEMS	JOBLINE INT.	KANTHAL B
KARLSHAMNS	KAUPTHING BANK	KINNEVIK IND. A
KINNEVIK IND. B	KIPLING HOLDING	KLOVERN A
KLOVERN B	KVAERNER SDB A	LAWSON SOFTWARE
LBI INTERNATIONAL	LILJEHOLMEN A	LILJEHOLMEN B
LINDAB B	LODET FASTIGHETS B	LOGICA
LUCARA DIAMOND	LUNDGRENS B	LUNDIN GOLD
LUNDIN MINING	LUXONEN SDB	M2 FASTIGHETER

M2S SVERIGE B	MALDATA B	MARIEBERG TID.A
MAXIM PHARMS	MEDICOVER HOLDING	MERTIVA
S &	METRO INTL.SDB A	METRO INTL.SDB B
MIND	MOGUL	MUNKSJ...
NACKEBRO	NAN RESOURCES	NARKES ELECTRISKA
NATURKOMPANIET	NAXS NORDIC ACS. BUY.	NEONET
NETONNET	NETWISE B	NGEX RESOURCES
NILORNGRUPPEN B	NISCAYAH GROUP B	NK CITY FASTIGHETS
NOBEL BIOCARE	NOBEL BIOCARE	NOBINA
NOKIA SDB	NOKIA	NORDAX GROUP
NORDBANKEN	NORDIC SER.PTNS.H. B	NORDSM.& THULIN B
NORSK HYDRO SDB	NOVACAST TECHS.B	NOVESTRA
OMI CORPORATION SDB	OPCON	OPTIMA BATTERIES B
OPTIMAIL A	ORC GROUP	OSCAR PROPERTIES
OSTGOTA ENSK. BANK.	OXIGENE	PA RESOURCES B
PANDOX	PEAB INDUSTRI B	PEAK PERFORMANCE B
PERBIO SCIENCE	PERGO	PHARMACIA SDB
PLATZER FTGH. B	POWERWAVE TECH.	PRONYX
PROSOLVIA B	PROTECT DATA	PROVOBIS B
PSI GROUP	QMED	READSOFT B
RKS B	RNB RETAIL AND BR.	RÖRVIKS GRUPPEN B
SAINT GOBAIN SDB	SAK I	SANDBLOM & ST. B
SAPA	SCANDIC HOTELS	SCANDIC HOTELS GR.
SCANDINAVIA ONLINE	SECURITAS DIRECT	SHELTON PETROLEUM
SIAB A	SIAB B	SIFAB
SIGMA B	SKOOGS B	SOCIETE EURO A SDB
SOCIETE EURO B SDB	SOLITAIR KAPITAL	SORB INDUSTRI
SPCS SCANDINVN. PC S.	SPCSGRUPPEN	SPIRA
SPORTAMORE	STADSHYPOTEK A	STORA A
STORA B	STORA ENSO A	STORA ENSO R
STORHEDEN B	SVITHOID TANKERS B	SYNGENTA
TANGANYIKA OIL SDB	TECHNOLOGY NEXUS	TELELOGIC
TERRA MINING	TILGIN	TOBII
TRANSMODE	TRENTION	TRIBONA
TRIO INFO.SYSTEMS	TROAX GROUP	TRYCKINVEST I NORDEN
TRYGGHANSA B	UNIBET GROUP SDB	UNITED TANKERS
UTFORS	VBB A	VBB B
VENCAP INDUSTRIER	VENUE RETAIL GROUP B	VERIMATION
VICTORIA PARK A	WISE GROUP	VISION PARK
VOSTOK GAS SDB		

B Osannolika observationer

Tabell B1: Avkastning som är $\leq -60\%$

Företagsnamn	Datum
ABB	22 Oktober 2002
Active Biotech	16 april 2015, 5 maj 2017
Biovent Intern	5 jun 2012, 11 jul 2012
Corem Property	30 oktober 2007
Fingerprint Cards	6 juni 2005
Karo Pharma	14 feb 2012
Nordic Mines	6 juli 2013
Sensys Gatso	10 oktober 2001

Tabell B2: Avkastning som är $\geq 50\%$

Företagsnamn	Datum
Addnode	15 Jun 2002, 26 Oct 2002
Alltele	02 Mar 2007, 27 Sep 2007
Anoto Group	22 Sep 2001
Arcam B	28 May 2004, 05 Sep 2016
Betsson	07 Sep 2001, 25 Sep 2001
Bjorn Borg	19 May 2017
Corem Property group	3 Oct 2002, 24 Jan 2004, 27 Sep 2007
CTT Systems	20 Aug 2003, 22 Aug 2005, 15 Dec 2006
DGC ONE	7 jun 2017
Duroc	04 Feb 2000, 13 Jan 2017
Endomines	14 Nov 2015
ENEA	19 Oct 2002, 16 May 2017
Eniro	25 april 2009
Episurf Medical	29 juni 2011
Fast Partner	25 Feb 2000
Feelgood Svenska	8 juli 2016
Fingerprint Cards	14 Oktober 1998, 22 jan 2000
Formpipe software	16 feb 2005
IAR systems	10 okt 2002
Micro Systemation	21 Mar 2001, 30 Jun 2001, 3 Jul 2002, 28 Jun 2003
Micro Systemation	12 Aug 2003, 22 Dec 2003, 16 Jan 2004
MSC group	30 nov 2002, 24 aug 2016
Multiq International	3 jun 2003

Företagsnamn	Datum
Neurovive Pharma	8 dec 2009, 29 jun 2017
Nordic Mines	9 tillfällen under 2014 , 2015
Ortivus	8 juni 2013, 12 juni 2013
Precise Biometrics	10 nov 2007, 4 april 2009, 26 maj 2012, 3 sept 2014
Pricer	3 ggr under 1998 1999 , en gång 2004
Probi	10 maj 2003
Raysearch labs	29 feb 2000
Sanitec	14 oktober 2014
Sintercast	1997
Softtronic	2 maj 2003
Stockwick förvaltnin	9 ggr under vår tidsperiod
Studsvik	20 april 2016
Tradedoubler	20 may 2015
Vitec Software	5 ggr under perioden 2000 03

C Matlabkod

Databehandling och ML-skattning

```
1 %% Initializing import setup from Excel-file.
2 filename = 'Uppsatsdata_klar.xlsx';
3 dataRange = 'B2:HM5400';
4 headerRange = 'B1:HMI';
5 dateRange = 'A1:A5400';
6 t = datetime(xlsread(filename,1,dateRange),'ConvertFrom',
7             'Excel');
8 [~,companyNames] = xlsread(filename,1,headerRange);
9 clear dateRange;
10 clear headerRange;
11
12 %% Calculating returns
13 sheet = 'Adjusted price';
14 P = xlsread(filename,sheet,dataRange);
15 R = (P(2:end,:) - P(1:end-1,:)) ./ P(1:end-1,:);
16 % Adjusting for dividends
17 sheet = 'Dividend return';
18 D = xlsread(filename,sheet,dataRange);
19 D(find(D < -0.2)) = 0; % Ignoring strange dividends
20 R(D(2:end,:) ~ = 0) = D(D(2:end,:) ~ = 0);
21
22 %% Calculating index returns
23 sheet = 'Market value';
24 MV = xlsread(filename,sheet,dataRange);
25 Rm = zeros(length(MV) - 1, 1);
26 Pm = zeros(length(MV), 1) + 100;
27 for i = 1:length(MV) - 1
28     validCols = intersect(find(~isnan(MV(i,:))), find(~
29                             isnan(MV(i+1,:))));
30     Rm(i) = sum(MV(i+1,validCols)) / sum(MV(i,validCols))
31           - 1;
32     Pm(i+1) = Pm(i) * (1 + Rm(i));
33 end
```

```

31 clear validCols;
32 clear i;
33
34 %% Adjusting for risk free rate
35 sheet = 'Risk free rate';
36 Rf = xlsread(filename, sheet, dataRange);
37 nan_indices = find(isnan(Rf));
38 for i = 1:length(nan_indices)
39     Rf(nan_indices(i)) = Rf(nan_indices(i)-1);
40 end
41 clear nan_indices
42 Rf = (1+ Rf/12).^ (1/30) -1;
43 Rm = Rm-Rf(2:end);
44 R = R-Rf(2:end)*ones(1, size(R,2));
45
46 %% Setting up explanatory variables
47 PTBV = xlsread(filename, 'PTBV', 'B2:HM23');
48 PTBV_years = xlsread(filename, 'PTBV', 'A2:A23');
49 %Limits for ordering
50 low = 30;
51 high = 70;
52 %Set start and end year
53 startYear = 1999;
54 endYear = 2017;
55
56 B = zeros(size(R));
57 alpha = zeros(size(R));
58 X_H = zeros(size(R));
59 X_L = zeros(size(R));
60 X_S = zeros(size(R));
61 X_B = zeros(size(R));
62
63 for i = find(PTBV_years == startYear):find(PTBV_years ==
        endYear)-1
64     current_year_rows = find(t >= datetime(PTBV_years(i)
        ,7,1) & t < datetime(PTBV_years(i)+1,7,1));

```

```

65
66 %Sort first on value then size
67 high_value_companies = find(PTBV(i,:) <= prctile(PTBV
    (i,:),low));
68 low_value_companies = find(PTBV(i,:) >= prctile(PTBV(
    i,:),high));
69 mid_value_companies = setdiff(setdiff(linspace(1,size
    (R,2),size(R,2)),high_value_companies),
    low_value_companies);
70 X_H(current_year_rows, high_value_companies) = 1;
71 X_L(current_year_rows, low_value_companies) = 1;
72 for j = 1:size(R,2)
73     if ismember(i, high_value_companies)
74         if MV(current_year_rows(1)-1,j) <= prctile(MV
            (current_year_rows(1)-1,
            high_value_companies),low)
75             X_S(current_year_rows, j) = 1;
76         elseif MV(current_year_rows(1)-1,j) > prctile
            (MV(current_year_rows(1)-1,
            high_value_companies),high)
77             X_B(current_year_rows, j) = 1;
78         end
79     elseif ismember(i, low_value_companies)
80         if MV(current_year_rows(1)-1,j) <= prctile(MV
            (current_year_rows(1)-1,
            low_value_companies),low)
81             X_S(current_year_rows, j) = 1;
82         elseif MV(current_year_rows(1)-1,j) > prctile
            (MV(current_year_rows(1)-1,
            low_value_companies),high)
83             X_B(current_year_rows, j) = 1;
84         end
85     else
86         if MV(current_year_rows(1)-1,j) <= prctile(MV
            (current_year_rows(1)-1,
            mid_value_companies),low)

```

```

87         X_S(current_year_rows , j) = 1;
88         elseif MV(current_year_rows(1)-1,j) > prctile
            (MV(current_year_rows(1)-1,
            mid_value_companies), high)
89             X_B(current_year_rows , j) = 1;
90         end
91     end
92 end
93
94 for k = 1:size(R,2)
95     prev_year_rows = find(t>= datetime(PTBV_years(i)
            -1,7,1) & t< datetime(PTBV_years(i),7,1));
96     cov_matrix = nancov(R(prev_year_rows , k), Rm(
            prev_year_rows));
97     B(current_year_rows , k) = cov_matrix(1,2)/
            cov_matrix(2,2);
98     alpha(current_year_rows , k) = mean(R(
            prev_year_rows , k))-B(current_year_rows , k)*
            mean(Rm(prev_year_rows));
99     end
100
101 end
102
103
104 %% Running MLE estimation
105 startRow = find( t == datetime(1999,7,1) );
106 endRow = find( t == datetime(2017,6,30) );
107
108
109 %% Run MLE to find parameters
110 startVals = [0.45, -0.45, 0.45, -0.45, 3, 0.1];
111 [theta , minval] = fminunc(@(x) -loglikelihood(R(startRow:
            endRow,:) , Rm(startRow:endRow) , alpha(startRow:endRow
            ,:) , B(startRow:endRow,:) , X_H(startRow:endRow,:) , X_L
            (startRow:endRow,:) , X_S(startRow:endRow,:) , X_B(
            startRow:endRow,:) , x) , startVals);

```

Log-likelihoodfunktionen

```
1 function [ll] = loglikelihood(R, Rm, alpha, B, XH, XL,  
    XS, XB, theta)  
2 %Evaluates likelihood function with t(v) noise  
3 % The input theta is a vector [Beta_H, Beta_L, Beta_S ,  
    Beta_B , v, sigma2]  
4 v= theta(length(theta)-1);  
5 sigma2 = theta(length(theta));  
6 ll = 0;  
7 for i = 1:size(R,2)  
8     X=[XH(:,i), XL(:,i), XS(:,i), XB(:,i)];  
9     s2 = ((v-2)/v)*(1+X*theta(1:4)')*sigma2;  
10    T = (R(:,i)-alpha(:,i)-Rm.*B(:,i));  
11    G=gamma((v+1)/2)./(gamma(v/2)*sqrt(pi*v*s2));  
12  
13    n=sum(~isnan(T));  
14    ll = ll + nansum(log(G))-((v+1)/2)*nansum(log(1+T  
        .^2./(v*s2)));  
15 end  
16 %% Ensuring positive variance  
17 if sigma2 <= 0  
18     ll = -1.0e+10;  
19 end  
20  
21 end
```