Stockholms universitet

Risk-return relationships in Nordic Stock Exchanges

Felix Martinsson

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2016:27
Matematisk statistik
Augusti 2016

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2016:27**
http://www.math.su.se

Risk-return relationships in Nordic Stock Exchanges

Felix Martinsson*

August 2016

Abstract

The report examines the relationship between return and the risk measures beta, volatility, Value at Risk, skewness and kurtosis with robust linear regression. The analysis includes the long as well as the short term relationsship for all non-delisted stocks listed on the Stockholm, Helsinki and Copenhagen Stock Exchanges over the period 1989-2015. Beta seems to have no long-term relationship with return while higher risk in terms of volatility and Value at Risk seems to be associated with lower return. The short term relationship is ambigous but seems at least not positive. The report thus rejects the assertion of a positive risk-return relationsship by standard theories such as CAPM. This is consistent with earlier research which more often than not renounces a positive relationsship, particulary in later ones.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: felix.martinsson@gmail.com. Supervisor: Mathias Lindholm, Filip Lindskog, Joanna Tyrcha.

Contents

1 Introduction

I am grateful for all help and feedback provided by my supervisors Mathias Lindholm, Filip Lindskog and Joanna Tyrcha and for my education at the Department of Mathematics at University of Stockholm. I would also like to express my gratitude to NASDAQ OMX Nordic's easily accessible historical stock prices and to friends, girlfriend and family who have read and given me response.

The purpose of this report is to investigate the relationship between return and risk of the stocks listed at NASDAQ OMX Nordics. The conventional risk measures beta, (β) and volatility (σ) will be used but also the more exotic risk measures skewness (γ), kurtosis (κ) and Value at Risk (VaR) will be used. Three approaches will be taken: a retrospective model, a prospective model and monthly (time series) model which are stated as different regression models. The monthly model uses panel data, i.e. a time series dimension and a cross-sectional dimension, and two other use cross-sectional data. The regression models will be carried out for all risk measures[1] on two periods, three stock exchanges and their union, three caps and their union, see Figure 1. This result in a total of 448 regressions.

$$\left\{\begin{array}{c}\text{Retrospective}\\ \text{Prospective}\\ \text{Monthly}\end{array}\right\}\times\left\{\begin{array}{c}\gamma\\ \kappa\\ \text{VaR}\\ \sigma\\ \beta\end{array}\right\}\times\left\{\begin{array}{c}\text{1987-2015}\\ \text{2008-2015}\end{array}\right\}\times\left\{\begin{array}{c}\text{Stockholm}\\ \text{Helsinki}\\ \text{Copenhagen}\\ \text{All}\end{array}\right\}\times\left\{\begin{array}{c}\text{Large}\\ \text{Mid}\\ \text{Small}\\ \text{All}\end{array}\right\}$$

Figure 1: *The combinations of regressions that will be carried out*

Volatility indicates how much the stock is moving. Low volatility suggests low variability in daily return while high volatility suggests the opposite. Volatility may be divided up into market volatility, the volatility due to market movement, and idiosyncratic volatility, the additional volatility due to the movement of the specific stock. Idiosyncratic volatility as well as total volatility has been used extensively in earlier research. Only total volatility is to be considered in this report.

Beta quantifies in what degree the stock moves with the market. Negative beta stocks are negatively correlated with the market while positive beta stocks are positively correlation. A low beta enables an investor to reduce the market risk through diversification. Beta is together with volatility the most common risk measure and has been used widely in research. Definitions and interpretations for the other risk measures are provided in the methodology section.

[1] Except beta for the monthly model

2 Earlier research

M. Baker, B. Bradley, J. Wurgler wrote: *"While there are many candidates for the greatest anomaly in finance, perhaps the most worthy is the long-term success of low volatility and low beta stock portfolios."*[2] Their study of US stocks from 1968 to 2009 implied a negative relationship between return and volatility as well as beta.

A paper from 2011 by David Blitz, Pim van Vliet and Bart van der Grient[3] reports that different studies have given positive as well as a neutral and negative relationships between return and volatility, and claim that the primary cause is the use of different methods. They also investigated the effect of survivorship bias by comparing the return spread between the 5th and 1st quantile of the stocks ordered by volatility. The spread was -1.3% when all available stocks were included and 5.5% when excluding non-survivors, a quite dramatic difference.

I. Mathur, G. Pettengill and S. Sundaram found a strong positive relationship between return and beta in 65 years of US stock market data.[4]. L. Martellini also found a positive volatility-return relationship in his study with data from the US stocks, however according to Blitz et al this was due to survivorship bias since only surviving stocks were included.[5][6] F. Fu made use of EGARCH models to estimate the unsystematic volatility and found a positive relationship.[7]

M.J. Brennan and F. Li found no relationship between beta and return 1931-1969 but a negative relationship since, driven solely by Large Cap.[8] Their assessment of cause is the growing share of institutional holdings and shrinking share of households' holdings. The result is in line with the 2013 claim by S. Fernando and P.D. Nimal that earlier studies more often found positive risk-return relationship than later. They themselves investigated the relationship at the Tokyo and Colombo (Sri Lanka) exchanges and found a positive relationship between beta and return.[9]

[2] Baker et al., 'Benchmarks as Limits to Arbitrage: Understanding the Low Volatility Anomaly'

[3] Blitz et al., 'Is the Relation Between Volatility and Expected Stock Returns Positive, Flat or Negative?'

[4] Mathur et al., 'The Conditional Relation Between Beta and Return'

[5] Martellini, 'Towards the Design of Better Equity Benchmarks'.

[6] Blitz et al., 'Is the Relation Between Volatility and Expected Stock Returns Positive, Flat or Negative?'

[7] Fu, 'Idiosyncratic Risk and the Cross-Section of Expected Returns'.

[8] Brennan and Li, 'Agency and Asset Pricing'.

[9] Fernando and Nimal, 'The Conditional Relation Between Beta and Return: Evidence from Japan and Sri Lanka'.

3 Data

The stock values were retrieved from NASDAQ OMX Nordic's website[10] 2016-03-15. The stock exchanges included in NASDAQ OMQ Nordic are those of Stockholm, Helsinki, Copenhagen and Iceland. Since Iceland had very few companies listed and a remarkably low trading volume, it was excluded. The data was separated into caps, roughly a partition of stocks by market capitalization. The data used ranged from 1987-01-01 to 2015-12-31. Some stocks were listed later and for these all available data were used. All stock values were already adjusted for dividends. Some stocks were in addition to A-stocks also traded in B, C, R or preference stocks where the main difference is that only A-stocks give voting rights. In those cases, the A-stocks were excluded. A few companies were traded on two or more stock markets and in these cases, the stock with most data was included, and the others excluded.

Most stocks have only been listed a small part of the period, however, only stocks with at least two years of data were included. The first observation of the Helsinki Exchange was 1997-01-03 and the first of the Copenhagen Exchange was 2000-11-07. For the Stockholm exchange, just 10 companies listed 2015-12-31 were also listed whole 1987. It is obvious for the Copenhagen and Helsinki exchange, and very plausible for the Stockholm exchange, that there are missing old observations.

Delisted companies were not include, which may create a bias in favour of surviving companies, i.e survivorship bias. This may make the relationship more positive since high risk companies have higher probability of a bankruptcies and thus being delisted. The return for high-risk companies could then be skewed by excluding the ones that have failed. E.g. 2011-2015 in the Stockholm Exchange, an average of ten companies were delisted yearly, which is a considerable amount considering the number of companies used.[11]

Two data periods will be used, the whole period of data containing values from 1987 to 2015, and the current recession containing the values from 2008 to 2015. [12] The reason for dividing the data is primarily for investigating whether the relation between return and risk changes differ according to the business cycle but also due to incomplete data and survivorship bias. By limiting one data period to include only the most recent years, these issues should be reduced.

A descriptive data analysis for giving an overview of the data follows.

[10] http://www.nasdaqomxnordic.com/shares/historicalprices

[11] nasdaqomx.com/transactions/markets/nordic/corporate-actions/stockholm/changes-to-the-list (160506)

[12] ekonomifakta.se/Fakta/Ekonomi/Tillvaxt/hogkonjunktur-eller-lagkonjunktur/ (2016-06-08)

The data consisted of stock values of 502 different stocks with a total of 1,709,132 observations. See Table 1 for the composition of the stocks into exchanges and caps. The mean number of observation per stock were 3,417. See Table 2 for how the mean number of observations differ due to country and cap. Notice that the differences are small, even though the time frames for the data differ substantially by the exchanges.

Table 1: *Number of small, medium and large caps in the different stock exchanges*

	STO	HEL	CPH	Total
Large	72	25	28	125
Medium	82	37	17	136
Small	100	56	85	241
Total	254	118	130	502

Table 2: *Mean number of observations of small, medium and large caps in the different stock exchanges*

STO	HEL	CPH	Total
4,522	3,423	4,147	4,201
2,846	3,133	3,702	3,115
3,269	2,856	3,514	3,180
3,488	3,707	3,014	3,417

See Table 3 for the different partitions' contribution to the total number of observations. Partitions with a higher percentage of the total number of observations automatically weigh higher in the coming regression analysis.

Table 3: *Percentage of total number of observations of small, medium and large caps in the different stock exchanges*

	STO	HEL	CPH	Σ
Large	19%	6%	6%	31%
Medium	14%	8%	3%	25%
Small	19%	11%	14%	44%
Σ	52%	25%	23%	100%

See Table 4 and Table 5 for the mean return of different partitions. Remember that less data were available for Helsinki and Copenhagen, which makes direct comparison impossible. Stockholm has done fairly well over the whole period while the Helsinki and Copenhagen exchanges have had returns around zero. This is primarily due to the sharper rise for Stockholm since the great recession crash.

Note that Small Cap has done far worse than Large and Mid Cap for all exchanges. This may be due to change of cap. If Small Cap companies that fared well moved to Mid Cap, the remaining Small Cap would have lesser return. This likely bias would seriously impact the data and the soundness of the results.

See Table 6 and Table 7 for the volatility of different caps and exchanges. As expected, Small Cap has much higher volatility than Medium and Large Cap. The exchanges are quite similar. Copenhagen has the highest volatility, especially during the recession. Stockholm and Helsinki have smaller volatility during the recession compared to the whole period while Copenhagen has about the same. The reason may be the absence of Copenhagen data for the tremendous growth

Table 4: *Total mean return*

	STO	HEL	CPH	All
Large	12.7%	5.9%	10.8%	10.9%
Medium	12.2%	3.6%	6.6%	9.2%
Small	-2.4%	-5.1%	-5.7%	-4.2%
Total	6.6%	0.0%	-0.5%	3.0%

Table 5: *Reccesion mean return*

STO	HEL	CPH	Total
11.3%	2.1%	8.2%	8.8%
12.4%	-0.6%	-1.5%	7.1%
-2.3%	-7.8%	-13.5%	-7.5%
6.3%	-3.4%	-7.2%	0.5%

and crash during the Dot-com bubble.

Table 6: *Total mean volatility*

	STO	HEL	CPH	Total
Large	38.7%	37.1%	36.4%	37.9%
Medium	42.3%	36.8%	34.4%	39.8%
Small	58.7%	58.2%	60.8%	59.3%
Total	47.7%	46.9%	52.1%	48.6%

Table 7: *Recession mean volatility*

STO	HEL	CPH	Total
35.4%	38.1%	34.6%	35.8%
39.2%	35.6%	35.7%	37.8%
53.4%	52.4%	61.0%	55.8%
43.7%	44.0%	52.0%	45.9%

See Figure 2, Figure 4 and Figure 6 for the percentages of the companies that were listed in 2015-12-31 and also listed earlier. These show that most of the data comes from the current recession, especially for Copenhagen.

See Figure 3, Figure 5 and Figure 7 for the development of the three exchanges are shown throughout the time period. As expected, there is much volatility and clear trends.

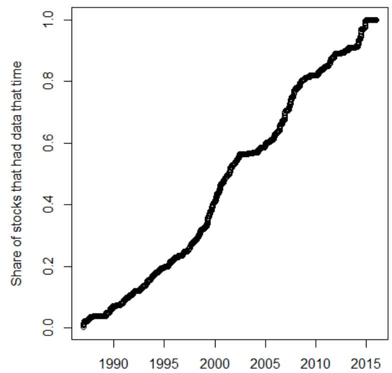

__Figure 2:__ Share of stocks with data from that year for Stockholm

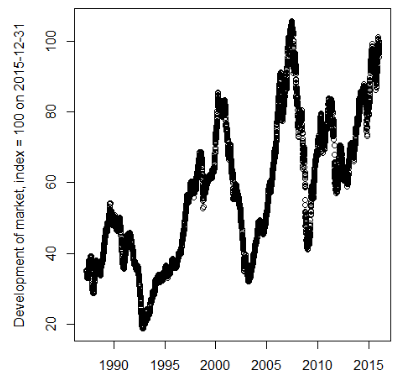

__Figure 3:__ Development of Stockholm

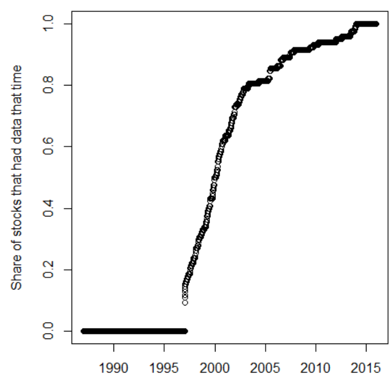

__Figure 4:__ Share of stocks with data from that year for Helsinki

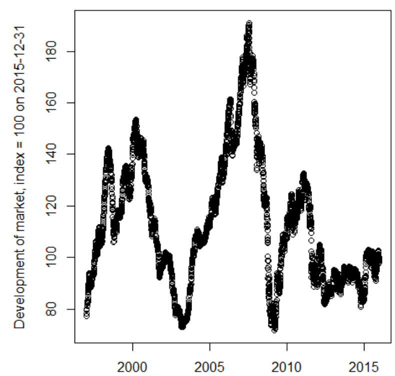

__Figure 5:__ Development of Helsinki

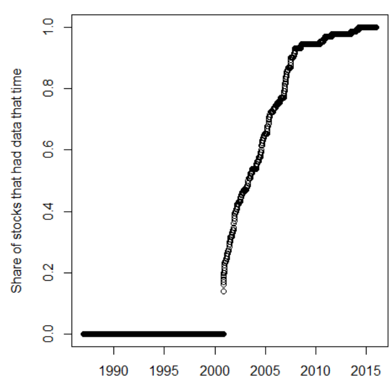

__Figure 6:__ Share of stocks with data from that year for Copenhagen

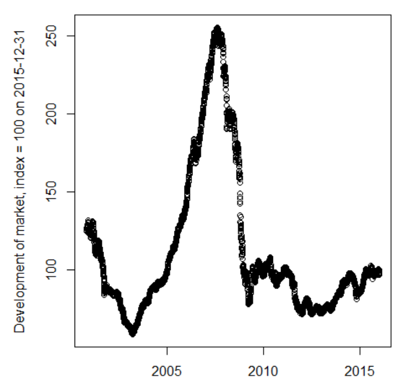

__Figure 7:__ Development of Copenhagen

4 Theoretical framework

Regression analysis is a statistical procedure for quantifying the relationship amongst variables, in this paper risk and return. In simple linear regression, there is one explanatory variable and one response variable. The models in this report will all be simple linear regressions, thus most theory will be limited to only cover the case with one explanatory variable.

The return will be denoted y, the risk (any measure) as x and ϵ as error. The subscript i will denote a specific stock and the subscript t a specific month. The intercept parameter will be denoted as α, the slope parameter as θ and the vector (α, θ) as $\boldsymbol{\theta}$.

4.1 The Classical Linear Regression Model

The classical linear regression model is described by the following equation.[13]

$$y_i = \alpha + \theta x_i + \epsilon_i \qquad i = 1, 2, 3...$$

The goal is to find the estimates $\hat{\alpha}$ and $\hat{\theta}$ that form the regression line that fits the observations the best.

$$y = \alpha + \theta x \tag{\textit{Regression line}}$$

The estimates are those which minimizes the sum of the squared residuals. This is called least square estimation. The estimates are found by differentiating the sum in respect to the parameters, setting the partial derivates to zero and solving the obtained equations.

$$(\hat{\alpha}, \hat{\theta}) = \arg \min_{\alpha,\theta} \sum_i \epsilon^2 = \arg \min_{\alpha,\theta} \sum_i (y_i - \alpha - \theta x_i)^2 \tag{\textit{Least square estimates}}$$

Behind the model are several assumptions important to its validity and theoretical justification. These are (1) constant variance or homoscedasticity as opposed to heteroscedasticity, (2) zero expectation of the residuals, (3) uncorrelated errors and (4) normal distributed errors, which implies (1)-(3). When (4) is fulfilled (and ergo all the others), least square estimation is equivalent to maximum likelihood estimation. Regardless whether (4) is fulfilled, if (1)-(3) are, the least squares estimates are the unbiased estimates with lowest variance according to the Gauss-Markov theorem.

[13] Fahrmeir et al., *Regression Models, Methods and Applications*

ASSUMPTIONS BEHIND
THE CLASSICAL LINEAR REGRESSION MODEL

$$V(\epsilon_i) = \sigma^2 < \infty \qquad i = 1, 2, 3... \qquad (Assumption\ 1)$$

$$E(\epsilon_i) = 0 \qquad i = 1, 2, 3..., n \qquad (Assumption\ 2)$$

$$cor(\epsilon_i, \epsilon_j) = 0 \qquad i \neq j \quad i, j = 1, 2, 3... \qquad (Assumption\ 3)$$

$$\epsilon_i \sim \mathcal{N}(0, \sigma^2) \qquad i.i.d. \quad i = 1, 2, 3... \qquad (Assumption\ 4)$$

4.2 Errors-in-variables

In the ordinary linear regression model, the assumption is that there are measurement errors in the explanatory variable, hence the error term. However, the response variable is assumed to be measured correctly. When this is not the case the so called errors-in-variables is present.[14] A linear regression model with errors-in-variables taken into account follows, with ϵ_i^x and ϵ_i^y being independent error terms.

$$y_i = \alpha + \theta(x_i + \epsilon_i^x) + \epsilon_i^y \qquad i = 1, 2, 3, \ldots$$

It is possible to work with this model, estimate the parameters, carry out statistical tests and create confidence interval. However, it is hard and rather complicated, e.g. consistent parameter estimation requires the value of $\mathrm{Var}(\epsilon_i^x)/\mathrm{Var}(\epsilon_i^y)$. Given the complications, it is very common to use ordinary regression models even though errors-in-variables is present.[15]

R. Davidson and J.G. MacKinnon have written that there is nothing wrong with an ordinary model for establishing the relationship between variables, but that the estimates will be inconsistent and biased downwards.[16] J. Hausman calls this downward bias the "*Iron law of Econometrics*" due to how common measurement errors are.[17] K. Donghceol points out that the problem is more severe

[14] Sundberg, *Lineära Statistiska Modeller*, p 98.
[15] Hausman, 'Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left'.
[16] Davidson and MacKinnon, *Econometric Theory and Methods*, p 310-311.
[17] Hausman, 'Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left'.

in multiple regression when some but not all of the explanatory variables are measured with errors.[18]. S. Durvasula, S. Sharma, S and K. Carter have shown that the t-statistic has a downward bias if errors-in-variables is present.[19] This implies lower p-values than what would have been obtained without measurement errors and so forth any errors-in-variables causes the results to be more conservative. Using ordinary regression thus increases type II errors, but not type I.

4.3 Robust estimation

A problem with least square estimation is the large sensibility to non-constant variance.[20] One way to overcome this is to use a robust estimator, which yields different estimates more robust to non-constant variance. According to F. Mosteller and J.W. Tukey, a robust estimator satisfies two conditions: it is resistant to outliers and effective under a wide range of circumstances.[21] When the assumption of homoscedasticity fails, a robust estimator can be an alternative to the least square estimator. There are many robust regression methods, this report will only consider the so called *M-estimators*, a generalization of the least square estimator.

Other M-estimators than the least square estimator may be more effective if the assumptions for least squares fail. Each estimator in the class have a corresponding objective function, $\rho(\epsilon)$, assumed to be convex and differentiable in this report. Define the influence function $\psi(\epsilon) = \delta\rho(\epsilon)/\delta\theta$ and the weight function as $w(\epsilon_i) = \psi(\epsilon_i)/\epsilon_i$. The estimates are obtained by minimizing the sum of the residuals in the objective function.

$$(\hat{\alpha}, \hat{\theta}) = \arg\min_{\alpha,\theta} \sum_{i=1}^{n} \rho(\epsilon_i) = \arg\min_{\alpha,\theta} \sum_{i=1}^{n} \rho(y_i - \alpha - \theta x_j) \qquad (\textit{M-estimates})$$

Unlike least squares, when $\rho(\epsilon) = \epsilon^2$, there is no closed form solution for minimizing the sum in the general case, however the technique iteratively reweighted least squares may be used. Since ρ is convex, the sum is also convex.[22] Since every local minimum for a convex function is a global minimum,[23] the estimates are obtained by differentiate the sum and setting the partial derivates to zero, see Equation 1.

$$\frac{\delta}{\delta\theta}\Big(\sum_{i=1}^{n} \rho(\epsilon_i)\Big) = \sum_{i=1}^{n} \psi(\epsilon_i) x_i = \sum_{i=1}^{n} w_i(\epsilon_i)\epsilon_i x_i = 0 \qquad (1)$$

[18] Dongcheol, 'The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns'.

[19] Durvasula et al., 'Correcting the t statistic for measurement error'

[20] Fahrmeir et al., *Regression Models, Methods and Applications*, p 160-163.

[21] Mosteller and Tukey, *Data Analysis and Regression: A Second Course in Statistics*, p 203–209.

[22] Sydsaeter et al., *Further Mathematics for Economic Analysis*, 59.

[23] Ibid., 104.

Solving Equation 1 is equivalent to minimizing Equation 2.

$$f(\theta) = \sum_{i=1}^{n} w_i(\epsilon_i)\epsilon^2 \tag{2}$$

Proof: *Assume that the parameter vector when Equation 2 is minimized is $\hat{\boldsymbol{\theta}}$. For all i, define (the constant) C_i to be equal to $w\epsilon_i$. Define $g(\boldsymbol{\theta}) = \sum_{i=1}^{n} C_i\epsilon_i^2$. Differentiate $g(\boldsymbol{\theta})$ with respect to θ at point $g(\hat{\boldsymbol{\theta}})$ and set the partial derivates to zero, $\sum_{i=1}^{n} C_i\epsilon_i x_i = 0$. This is equivalent to Equation 1 if $C_i = w(\epsilon_i)$, thus a solution is found.*

Thus, the estimates are the ones that minimizes Equation 2. This is done by the following algorithm, which is globally convergent[24] and written in the general matrix form, and thus not limited to two variables. No further proof or explanation is given.

ITERATIVELY REWEIGHTED LEAST SQUARES ALGORITHM

Define $\boldsymbol{\theta}^{(t)}$ as the parameter vector for iteration t, $\boldsymbol{X}$ as the design matrix, $\boldsymbol{W}^{(t)}$ as square matrix with the weights on its diagonal for the iteration t, and $\boldsymbol{Y}$ the vector of y-values.

1. Start with the initial parameters
2. Calculates the weights and residuals from the parameters
3. Solve for new parameters by the following equation

$$\boldsymbol{\theta}^{(t)} = \left(\boldsymbol{X'W}^{(t-1)}X\right)^{-1}\boldsymbol{X'W}^{(t-1)}Y$$

4. Repeat 2-3 until convergence, thus $\boldsymbol{\theta}^{(t)} \approx \boldsymbol{\theta}^{(t-1)}$

[24] Chaudhury, ‘On the convergence of the IRLS algorithm in Non-Local Patch Regression’.

4.4 The Huber Estimator

The most common M-estimator is the *Huber estimator*, proposed by Huber in 1964.[25][26] Without any further comment, this will be the estimator used in this report. The Huber estimator functions follow.

$$\rho(\epsilon) = \begin{cases} \epsilon^2/2 & |\epsilon| \leq k \\ k \cdot (|\epsilon| - k) & |\epsilon| > k \end{cases} \qquad (\textit{Huber objective function})$$

$$\psi(\epsilon) = \begin{cases} \epsilon & |\epsilon| \leq k \\ k \cdot \mathrm{sign}(\epsilon) & |e| > k \end{cases} \qquad (\textit{Huber influence function})$$

$$w(\epsilon) = \begin{cases} 1 & |e| \leq k \\ k/|\epsilon| & |e| > k \end{cases} \qquad (\textit{Huber weight function})$$

where k is a constant called the *tuning constant*. For $|\epsilon| \leq k$ the estimation is equivalent to the least square objective function but for larger residuals the residuals are weighted smaller with the Huber estimator. In choice of the tuning constant, there is a trade off between efficiency in the case of homoskedasticity and robustness to outliers. A usual value of the tuning constant is $1.345\hat{\sigma_\epsilon}$.[27] This gives an effectiveness of 95% in the case of homoskedasticity.[28] Since a function with a continuous increasing derivate is convex, and the influence function is increasing and continuous, the estimator sum is convex. Hence, every stationary point of $\sum_{i=1}^{n} \rho(\epsilon_i)$ is a global minimum and iterated reweighted least squares may be used for finding the minimum.

---

[25] Yu et al., 'Robust Linear Regression: A Review and Comparison'

[26] Huber, 'Robust estimation of a location parameter'.

[27] $\hat{\sigma_\epsilon}$ is the robust standard deviation estimate equal to $\frac{\text{Median Absolute Residual}}{0.6745}$

[28] Defined as $E(\theta - \hat{\theta}_{leastsquare})^2/E(\theta - \hat{\theta}_{Hubor})^2$.

4.5 Statistical tests

Statistical tests will be used for testing model assumptions. Short descriptions follow, please refer to the biography for proof and further details.[29]

The Breusch-Pagan test is is used for detecting heteroskedasticity in linear regression models. Let y be the response variable and x the sole explanatory variable. Let ϵ_i be the residuals of a simple linear regression with parameters estimated by least squares. ϵ_i^2 are fitted into a new regression model as response variable with x as explanatory variable, see Equation 3.

$$\epsilon_i^2 = \alpha + \theta x_i + \eta \tag{3}$$

The parameters in Equation 3 are estimated by least squares. The zero hypothesis of homoscedasticity is tested against the alternative hypothesis of heterscedasticity. The coefficient of determination times the sample size is under the zero hypothesis approximately chi-square distributed with one degree of freedom, see Equation 4, which gives a p-value for the test.

$$nR^2 \sim \chi_1^2 \tag{4}$$

Woolridge test is used for detecting auto-correlation in panel data. It is a test robust for heteroscedasticity and non-parametric, hence suitable for stock data. Let i denote stock and t a day with y and x as variables like before. Define, $\Delta y_{i,t} = y_{i,t+1} - y_{i,t}$ and $\Delta x_{i,t} = x_{i,t+1} - x_{i,t}$. Firstly $\Delta y_{i,t}$ is regressed on $\Delta x_{i,t}$, see Equation 5. The residuals from the regression are denoted as $\Delta\epsilon_{i,t}$.

$$\Delta y_{i,t} = \alpha + \theta\Delta x_{i,t} + \Delta\epsilon_{i,t} \tag{5}$$

The slope of the regression is estimated using least squares. In the case of no serial correlation, $cor(\Delta\epsilon_i, t, \Delta\epsilon_{i,t+1}) = -0.5$. By setting up a new regression, Equation 6, where $\Delta\epsilon_{i,t+1}$ is regressed against $\Delta\epsilon_{i,t}$ and with γ being the slope parameter, this can be tested by the zero hypothesis $H_0 : \gamma = -0.5$ against the alternative hypothesis $H_1 : \gamma \neq -0.5$. This is done be an ordinary t-test.

$$\Delta\epsilon_{i,t+1} = \Omega + \gamma\Delta x_{i,t} + \eta_{i,t} \tag{6}$$

[29] Breusch and Pagan, 'A simple test for Heteroscedasticity and Random Coefficient Variation'; Drukker, 'Testing for Serial Correlation in Linear Panel-Data Models'.

5 Methodology

5.1 Definition of model variables

The risk measures used as explanatory variables in the regression are all calculated by historical data of daily stock values. The regression analysis used will cover different time periods and the variables will be calculated from the stock data. The periods that the variables will be calculated for is the whole period of data, the recession period, the whole period except 2015, the recession except 2015, and single months. Thus the calculation of the return and risk will be for any of these given of time periods. Let $\bar{r_i}$ be the arithmetic mean for the stock return i for a given time period and let n be the number of days for that time period.

Simple and not compound return will be used. The literature review by Blitz et al mentioned earlier showed that of 16 similar studies, two used compound return, three used both and eleven of them used simple return[30]. As shown by Hudson and Gregouriou, it is not possible to easily compare simple and compound returns, and for the sake of generality it is best to use the same as most previous papers, i.e. simple return.[31] Simple return is also, by definition, the increase in wealth which is of interest. If $V_{i,t}$ is the price of stock i day t, the simple return for the stock day t is $V_{i,t+1}/V_{i,t} = r_{i,t}$. Return as response variable used in the regressions will be the annual return. The observed annual return over a given time period for the stock i is calculated as,

$$r_i = \prod_{t=1}^{n} r_{i,t}^{251/n} - 1 \qquad (\textit{Annual return})$$

Volatility as variable will be the one year normalised volatility. The observed (sample) volatility over a given time period for the stock i is calculated as,

$$\sigma_i = \sqrt{\frac{251}{n-1}\sum_{i=1}^{n}(r_{i,t} - \bar{r_i})^2} \qquad (\textit{Volatility})$$

Skewness is the third central moment and measures asymmetry in the return distribution. A distribution with a high median return, but high risk for great losses, would have a negative skew; while a distribution with lower median return, but high probability for great returns rather than losses, would have a positive skew. Skewness will not be normalised since there is not an intuitive interpretation apart from the sign, unlike return and volatility. The observed

[30] Blitz et al., 'Is the Relation Between Volatility and Expected Stock Returns Positive, Flat or Negative?'

[31] Hudson and Gregoriou, 'Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns'.

skewness over a given time period for the stock i is calculated as,

$$\gamma_i = \frac{\frac{1}{n}\sum_{t=1}^{n}(r_{i,t} - \bar{r_i})^3}{\sigma_i^3} \qquad (\textit{Skewness})$$

Kurtosis is the fourth central moment. While the variance measures how spread out data are kurtosis measures the origination of the variance, i.e. fatness or length of the tails. A dataset with no outliers but plenty of observation within one to two standard deviations from the mean may have equal variance as a dataset with all but a few data points centered around the mean, but the remains several standard deviations away, but the kurtosis would differ a lot. For the same reason as skewness, kurtosis will not be normalised to represent a one year equivalent. The observed kurtosis over a given time period for the stock i is calculated as,

$$\kappa_i = \frac{\frac{1}{n}\sum_{t=1}^{n}(r_{i,t} - \bar{r_i})^4}{\sigma_i^4} \qquad (\textit{Kurtosis})$$

Let $\rho_{i,M}$ be the correlation between the stock i and the market (denoted as M) over the given time period. Market in this context means the index calculated in the following way. For a given day, the return for all stocks with available values were calculated and the market return is calculated by taking the arithmetic mean. This was done for every day to create a market time serie. The same market was used for calculating beta for all stocks. The beta over a given time period for the stock i will be calculated as,

$$\beta_i = \rho_{i,M} \cdot \frac{\sigma_i}{\sigma_\mathrm{M}} \qquad (\textit{Beta})$$

Value at Risk at a 95% confidence level for daily returns is defined as the threshold that the returns exceed in 95% of the cases.[32] If the Value at Risk is 0.97% for a stock, it means that the daily return of the stock will be more than -3% for 95% of the days.

Value at Risk in this report will be the observed Value at Risk with some simplification, the calculation method follows. Let the floor (ceiling) of x be the largest (smallest) integer equal to or less (greater) than x. Value at Risk for periods in excess of one month will be calculated as the average of the floor and ceiling of the $0.025n$ worst daily outcome. For the monthly time frame, the estimate will be the worst day that month. Define $R_{(i)}$ as *ith* lowest return for that time frame. The Value at Risk over a given period longer than one month for the stock i is calculated as,

$$VaR_i = \frac{R_{\mathrm{floor}(0.025n)} + R_{\mathrm{ceiling}(0.025n)}}{2} \qquad (\textit{VaR, periods excess of a month})$$

[32]Tsay, *Analysis of Financial Time Series*, 327.

The monthly VaR for the stock i is calculated as,

$$VaR_i = R_{(1)} \qquad \textit{(VaR, monthly)}$$

5.2 Models set up

The relationship between return and the different risk measures will be investigated using linear regression for every risk measures. Return will be the response variable for all regressions. The risk measures will be the explanatory variables for the regressions, but only one at a time, i.e. simple and not multiple regression will be used.

Since risks as variables used is observed values and not exact quantifications of how risky the stocks really are, an errors-in-variables model would seem reasonable. However, grounding in the theoretical framework, this problem may not be very severe. A model without errors-in-variables will still show the relationship, even though the estimates will be biased downwards. This will increase false negatives which is tolerable. Thus, for simplicity, the regression models will not include any error term for the explanatory variable.

Three different models will be used for investigating the relationship between return and risk. These will be called the Retrospective Model, the Prospective Model and the Monthly model. All models will be used for all risk measures, except beta which will not be used in the Monthly model due to instability in observing the correlation between a stock and the market in just a month.

All models will be used for all datasets, namely the unmerged datasets Stockholm Small Cap, Stockholm Mid Cap, Stockholm Large Cap, Helsinki Small Cap, Helsinki Mid Cap, Helsinki Large Cap, Copenhagen Small Cap, Copenhagen Mid Cap and Copenhagen Large Cap, and the merged dataset, Stockholm (all caps), Helsinki (all caps), Copenhagen (all caps), Small Cap (all exchanges), Mid Cap (all exchanges), Large Cap (all exchanges), and the dataset containing all stocks.

This retrospective model is the main model. In this model, the return and risk will be observed over the time frames 1987-2015 and 2008-2015. The return will be regressed on the risk. In the prospective model, the risk will be observed over the time frames 1987-2014 and 2008-2014 while the return will be observed 2015. This will answer if the level of risk in the earlier period is correlated with later return. Let y_i be the observed return for stock i, θ and α the parameters and x_i the observed risk for stock i. The models are stated as,

$$y_i = \alpha + \theta x_i + \epsilon_i \qquad i = 1, 2, 3, ..., n \dots \qquad \textit{(Retro- \& Prospective Model)}$$

The Monthly Model will have a time series dimension as well as a cross-sectional dimension and thus consist of panel data. The data points will be the observed return and risk for all stocks included and for all month with complete data for that stock for the given time period. The arithmetic mean return for the total period of data for every stock will be used as offsets. Let $y_{i,t}$ be the observed return for stock i month t, $\bar{y}_i$ the arithmetic mean return for the stock i, θ and α the parameters, $x_{i,t}$ the observed risk for stock i month t and m_i the number of months with complete data for the stock i. The model is stated as,

$$y_{i,t} = \alpha + \bar{y}_i + \theta x_{i,t} + \epsilon_{i,t} \qquad i = 1, 2, 3..., n \quad t = 1, 2, 3, .., m_i \ \ (\textit{Monthly Model})$$

The Monthly Model differs from the others a great deal. If, for example, there would be a positive relationship between risk and return on long and medium term investments, but a negative on short term, the results from the methods could be opposites. The Monthly Model answers the questions whether months with higher (lower) return have higher (lower) risk, while the other models answers question the whether stocks with higher (lower) return have higher (lower) risk.

For all models the null hypothesis, $H_0 : \theta = 0$ will be tested against the two-sided alternative hypothesis $H_0 : \theta \neq 0$ for all markets for both time periods for all risk measures of interest. This will be done by either least squares or the Huber estimator, dependent on whether heterscedasticity is present, which will be determined in the model evaluation. For both methods an ordinary t-test will be used, with sample standard deviation and robust standard errors used respectively. Due to the central limit theorem, the t-statistic assumption of normal distributed mean is assumed to be fulfilled.

To sum up, three regression models will be used. Each model will be used with four-five risk measures as explanatory variable. This will be done for 16 datasets of stocks, two times, one for the period 1987-2015 and one for 2008-2015. The main result will essentially be 448 regressions with corresponding estimates and p-value. In the next page, step by step instructions of how the regressions is done is provided.

UTILISE THE DESCRIBED MODELS: STEP BY STEP

1. Choose the stocks to be included in the regression. For a regression on e.g. the dataset Stockholm Small, choose the subset of all stocks which belongs to both Stockholm and Small Cap.
2. Adjust the stock data according to the time period. For a regression on the whole period, all data is used. For a regression on the recession period, the data prior to 2008 is disregarded.
3. Determine the model, prospective, retrospective or monthly.
4. Calculate the returns to be used in the regression. The returns are calculated according to the definition provided in subsection 5.1.
   (a) For the retrospective model, calculate one return value y_i for every stock included over the chosen time period.
   (b) For the prospective model, calculate one return value y_i for every stock included over the year 2015.
   (c) For the monthly model, divide all stock data into month subsets for all months in the given time period. For every stock, calculate the monthly return if data for the whole month is available, y_{it}.
5. Choose the risk measure to be used in the regression, volatility, skewness, kurtosis, Value at Risk or beta (not for the monthly model).
6. Calculate the risk measures values to be used in the regressions. The measures are calculated according to the definition provided in subsection 5.1.
   (a) For the retrospective model, calculate one risk value x_i for every stock included over the chosen time period.
   (b) For the prospective model, calculate one risk value x_i for every stock included over the starting year of the period until 2014.
   (c) For the monthly model, divide all stock data into month subsets for all month in the given time period. For every stock, calculate the monthly risk value if data for the whole month is available, x_{it}.
7. Regress y_i on x_i for the retro- and prospective model. For the monthly model, regress $y_{i,t}$ on $x_{i,t}$ with mean stock returns $\bar{y}_i$ as offsets.
8. Estimate the parameters according to least squares or Huber estimator according to the risk measure (which is to be determined in section 6) and save the sloop estimate and corresponding p-value when testing if the slope parameter differs significantly from zero.

6 Model diagnostic

The aim of this study is to investigate any historical relationship between risk and return. Therefore no interest lays in the coefficient of determination, which answers the question whether the risk actually explains much of the difference in return or not. Nor is any backtesting applied, since the ambition is not to create a model for predict future stock increases.

Firstly, there will be inspection of the data for outliers. Secondly, the Breusch-Pagan test for detecting homoscedasticity will be employed. Due to the many regressions, this will only be applied on the Retrospective method for the whole period but for all datasets and risk measures. Thirdly, Woolridge test for serial correlation will be applied. This only regards the Monthly model, which has a time series dimension, and will only be employed on the dataset with all observations but for all risk measures. Fourthly, there will be a discussion of diagnostic plots, namely scatterplots, Normal Q-Q plots and Residuals vs Fitted plots. The plots will only be included and discussed for the whole period and full dataset, however, possible differences with other datasets and time periods will be mentioned later in the model evaluation which will sum up the diagnostic.

6.1 Outlier detection

Three extreme outliers were detected and excluded from the diagnostic and regressions. These were the Finish company Pohjois-Karjalan Kirjapaino (PKK1V), the Danish company Novo Nordisk (NOVO) and the Swedish company Trigon Agri (TAGR). PKK1V had a beta of more than 15 and a volatility more than 15 times larger than the 2nd largest. NOVO had a yearly return of 115%, more than three standard deviations larger than the return with 2nd highest return. TAGR had a volatility nearly three times larger than the 2nd largest remaining stock, which seriously impacted the testing for heteroscedasticity.

Exclusion of observations may be problematic since information is discarded. Even though the observations are extreme, they are still valid data points and it is a fact to acknowledge that some stocks have extreme volatility with discontinuous jumps.[33] However, the purpose of this report is to investigate if there is a general relationship between return and risk. Removing the extreme observations is thus justified.

6.2 The Breusch-Pagan test

The confidence level for the Breusch-Pagan test will be 0.01. Due to the large number of regressions, too many false positives would arise with a higher confidence level. It is important to focus on these that show strong evidence for heteroscedasticity. The test was employed for detecting heteroscedasticity for all

[33] Lahaye et al., 'Jumps, cojumps and macro announcements'

risk measures and all 16 datasets. In Table 8 the combinations of risk measures and datasets with heteroscedasticity according to the test are shown.

Table 8: *Risk measures and datasets with heteroscedasticity according to the Breusch-Pagan test on the Retrospective model on the whole time period*

	STO	HEL	CPH	All
Large				
Medium	γ			γ
Small	VaR, σ		σ	VaR, σ
All	γ, κ, σ	VaR, σ	γ, VaR, σ	γ, VaR, σ

Heteroscedasticity is absent in the regressions on beta for all datasets. Skewness, Value at Risk and volatility show significant heteroscedasticity on several datasets each. Furthermore, for all three risk measures, this includes the especially important dataset with all the observations. The complete dataset for Stockholm is the only one where there is significant heteroscedasticity present for kurtosis.

6.3 Woolridge test for serial correlation

For the same reason as the Breusch-Pagan test, a confidence level of 0.01 was used. Woolridge test was employed for the Monthly model be done for the time frame 1987-2016, for all risk measures and for the merged dataset with all stocks included. No autocorrelation was found for any of the risk measures and datasets.

6.4 Diagnostic plots

Three types of diagnostic plots is provided. Scatterplots with the regression line plotted shows the observations and the fitted regression line. A Normal Q-Q plot compare the actual residuals to the theoretical normal distribution. The y-axis shows deviation of residuals from the fitted values and the x-axis shows the expected deviation if the set of residuals were to be normal distributed. Residuals vs Fitted plot shows the fitted value of an observation on the x-axis and the residual for that observation on the y-axis. A trend line shows any trend by dependence of fitted value and residuals.

Due to simplicity, either least squares or robust regression is used for all diagnostic plots for that risk measure. With the result from the Breusch-Pagan test in mind, least square will be used when regressing on beta and kurtosis and the Huber estimator for volatility, skewness and VaR. Only diagnostic plots for the dataset with all observations for the whole period are included and discussed. Any deviance of the plots of the regressions not provided will be discussed in the model evaluation. A discussion of the diagnostic plots of the regressions on all risk measures for all three period follows, with plots coming after.

Retrospective model

The Normal Q-Q plots show very strong evidence of fatter tails than the normal distribution for the regressions on skewness, volatility and Value at Risk, see Figure 16, Figure 22 and Figure 34. The Normal Q-Q plots show that the regressions on kurtosis and beta look more normal, see Figure 10 and Figure 28 even though the tails are substantially fatter than if the were to be normal distributed, the differences is too big. It is no coincidence that The Breusch-Pagan test did not detect heteroscedasticity for kurtosis and volatility.

The scatterplots for kurtosis and skewness, see Figure 26 and Figure 32, show that most observations are clustered together on the x-axis but with about 10 outliers for kurtosis and about thirty for skewnesss. When zooming in on the cluster (not included as plots), it is clear that the fits could be better. The Fitted vs Residuals plots, see Figure 30 and Figure 36 confirm this: the fit is not good for most of the observations. The plots cast doubt over the validity of these regressions.

The scatterplot for beta, see Figure 9, looks overall quite adequate for linear regression but with some problematic tendencies. There is one very troubling outlier and the residuals seem to be smaller for low-beta stocks. The Fitted vs Residuals plot for beta, see Figure 12, confirms this. The fit seems to be good for beta values with middle range values, a bit worse for others, but in general a decent fit.

The scatterplot for volatility, see Figure 14, show that the positive residuals tend to have greater absolute values than negative ones. Also, several outliers with very high volatility or return cause a bad fit. The Fitted vs Residuals plot confirm the heteroscedasticity suspicion and the overall fit is not good.

The scatterplot for Value at Risk, see Figure 20 shows a decent fit, but with some problem with positive residuals having larger absolute value than negative ones. The Residuals vs Fitted plot shows good fit for most observations with some sign of non-linearity with larger residuals for observation with smaller Value at Risk and smaller residuals for observations with larger Value at Risk.

Prospective model

It is striking how similar all the plots for the Prospective model is to the plots for the Retrospective model, scatterplots, Normal Q-Q and Fitted vs Residuals and for all risk measures. It is expected that the observed risk would be similar for both periods, since the only difference is the inclusion of one more year of data in the Retrospective model. However, the measure periods of the return were mutually exclusive, which could have affected the result from the different models considerable, but this does not seem to be the case. Due to the similarities, the statistical analysis of the Prospective model will rely on that for the Retrospective.

Monthly model

Regarding the diagnostic plots for the Monthly model, there is a huge amount of observations, namely 81,495. This makes the scatterplots hard to interpret in themselves. For all scatterplots, the observations fully cover the sample space where the observations are as most dense, thus the relative distribution of observations in the regions with highest density is not shown.

The Normal Q-Q plots for all risk types for the Monthly model show very strong evidence against the residuals being normal distributed, see Figure 40, Figure 41, Figure 46 and Figure 47.

In the scatterplot for volatility, see Figure 38, some tendency to heteroscedasticity and non-linear residuals is shown with larger residuals for greater volatility. In the scatterplot for kurtosis, see Figure 44, the same conclusion can be drawn. In the scatterplot for skewness, see Figure 45, there are outliers, which nearly all have higher skewness than expected, and judging by the plot, the model does not seem to work well. In the scatterplot for Value at Risk, see Figure 39, two abnormalities are shown. Firstly, the residuals are in general more negative when Value at Risk is small. Secondly, for larger Value at Risk, the residuals tend to become larger.

The Fitted vs Residuals plots for kurtosis for the Monthly model show no tendency to heteroscedasticity, see Figure 48. For the regression on volatility, there is a small tendency to increased residuals with higher fitted volatility values, see Figure 42. The opposite is true for the regression on Value at Risk but more evident, the residuals seem to be smaller for values with larger fitted Value at Risk values, see Figure 43. The regression on skewness seems to have stable residuals, except for two outliers, see Figure 49.

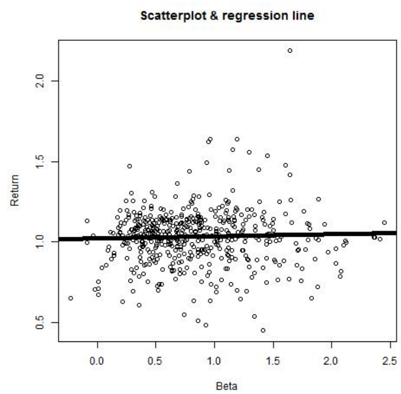

Figure 8: *Retrospective, beta*

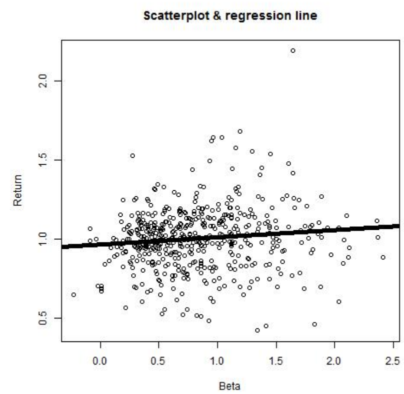

Figure 9: *Prospective, beta*

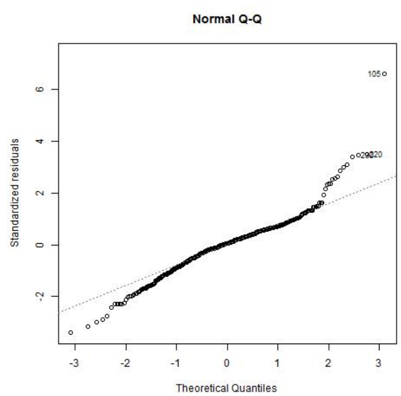

Figure 10: *Retrospective, beta*

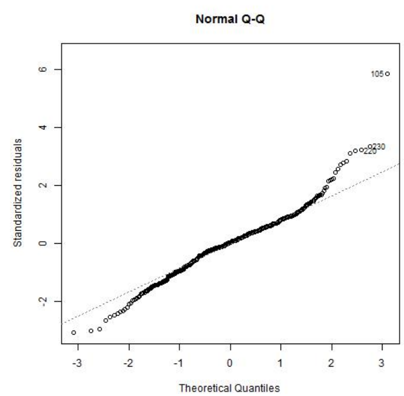

Figure 11: *Prospective, beta*

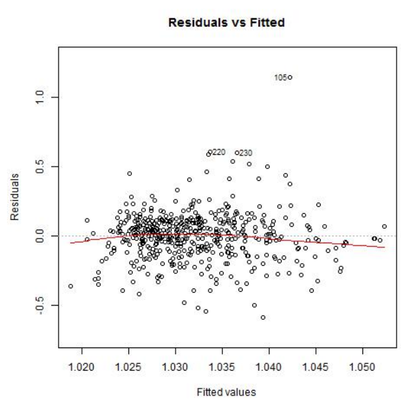

Figure 12: *Retrospective, beta*

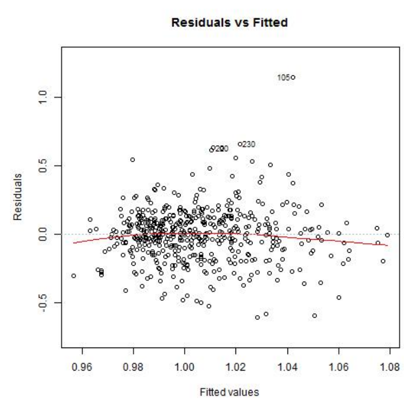

Figure 13: *Prospective, beta*

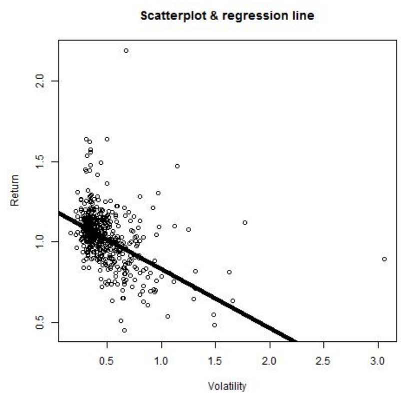

Figure 14: *Retrospective, volatility*

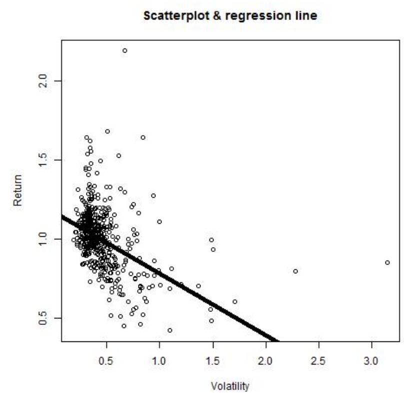

Figure 15: *Prospective, volatility*

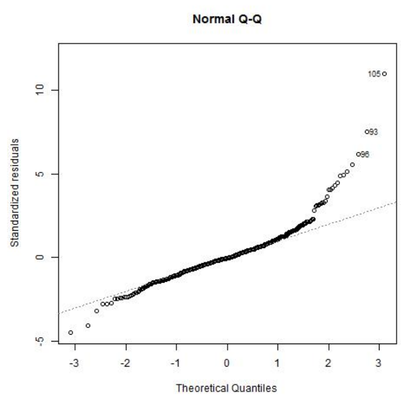

Figure 16: *Retrospective, volatility*

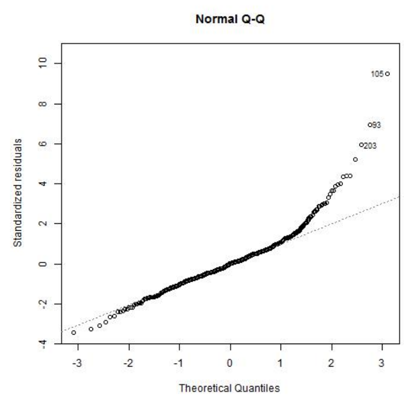

Figure 17: *Prospective, volatility*

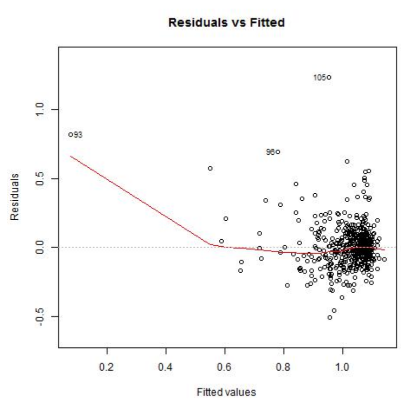

Figure 18: *Retrospective, volatility*

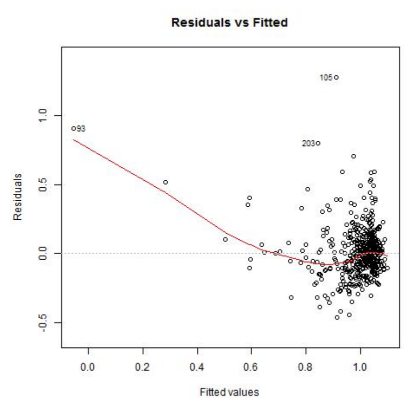

Figure 19: *Prospective, volatility*

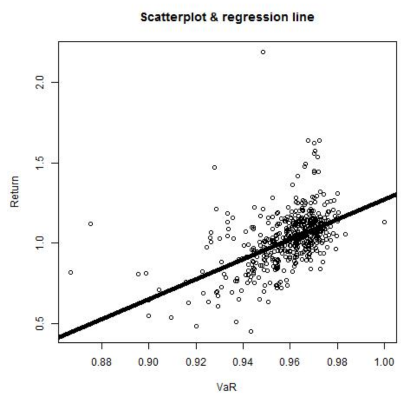

Figure 20: *Retrospective, VaR*

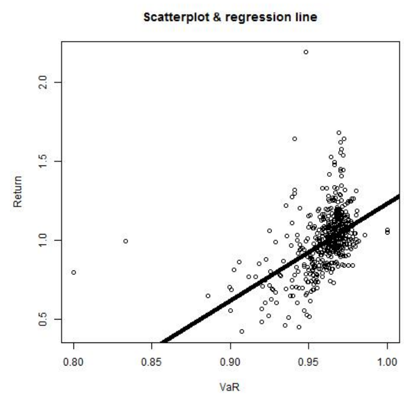

Figure 21: *Prospective, VaR*

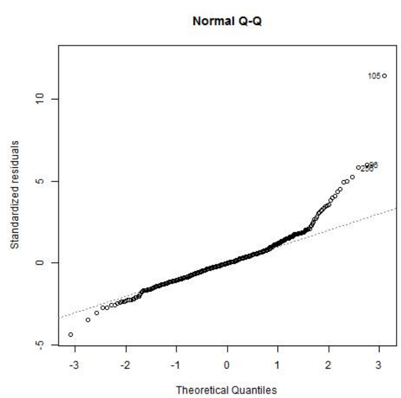

Figure 22: *Retrospective, VaR*

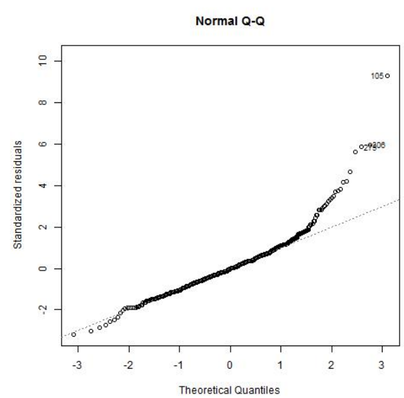

Figure 23: *Prospective, VaR*

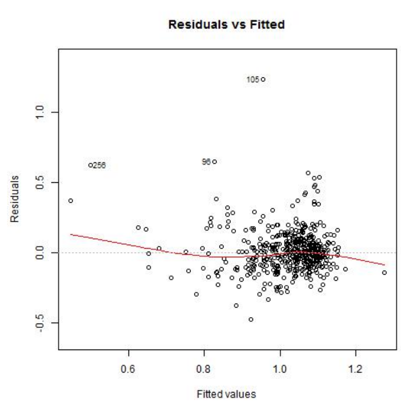

Figure 24: *Retrospective, VaR*

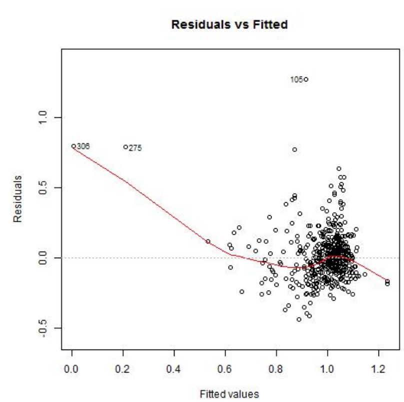

Figure 25: *Prospective, VaR*

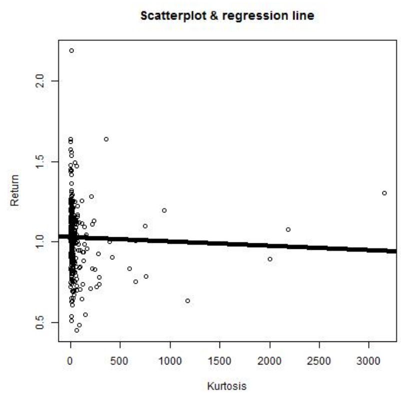

Figure 26: *Retrospective, kurtosis*

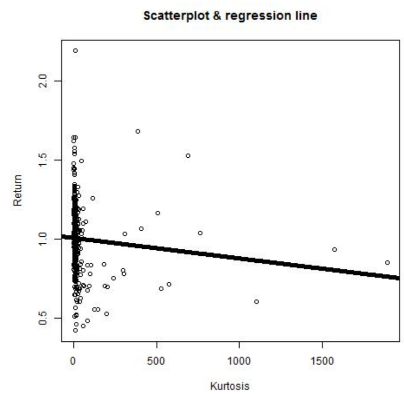

Figure 27: *Prospective, kurtosis*

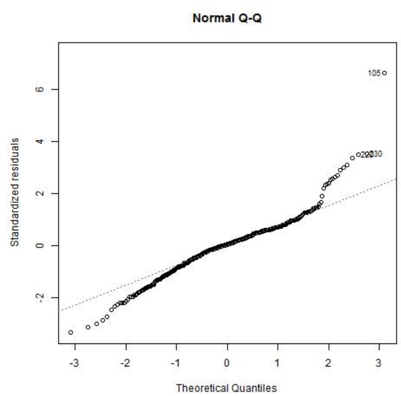

Figure 28: *Retrospective, kurtosis*

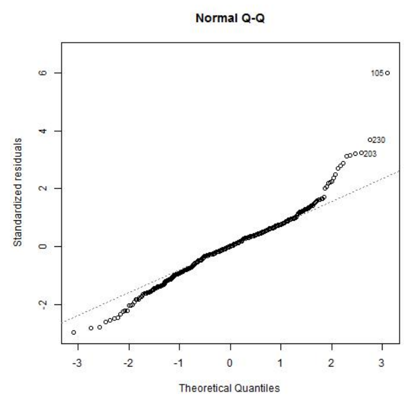

Figure 29: *Prospective, kurtosis*

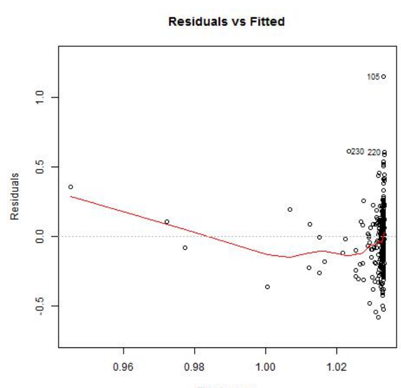

Figure 30: *Retrospective, kurtosis*

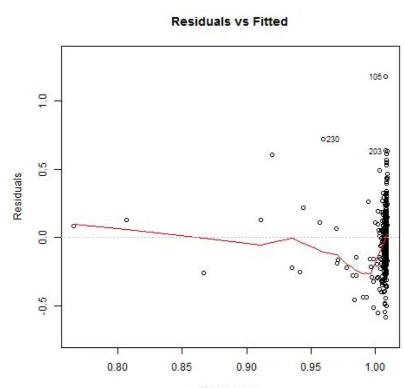

Figure 31: *Prospective, kurtosis*

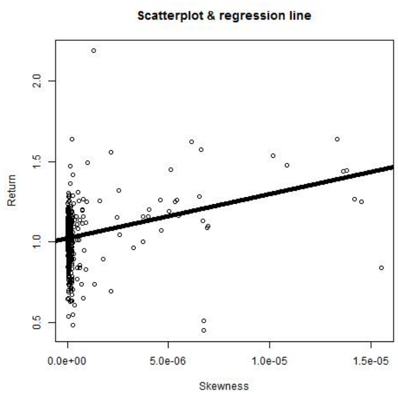

Figure 32: Retrospective, skewness

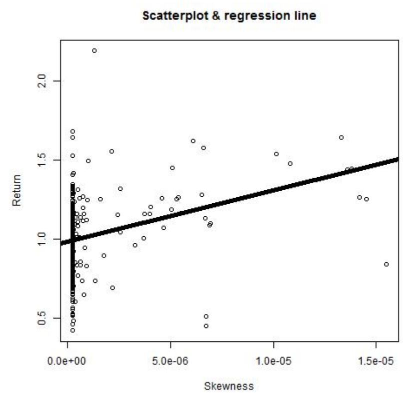

Figure 33: Prospective, skewness

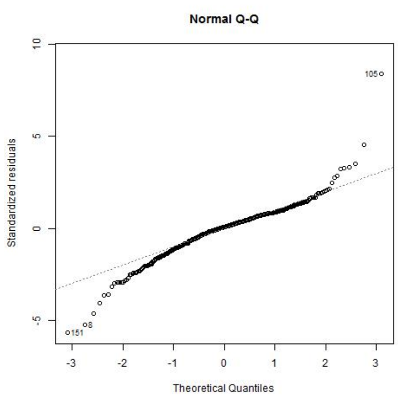

Figure 34: Retrospective, skewness

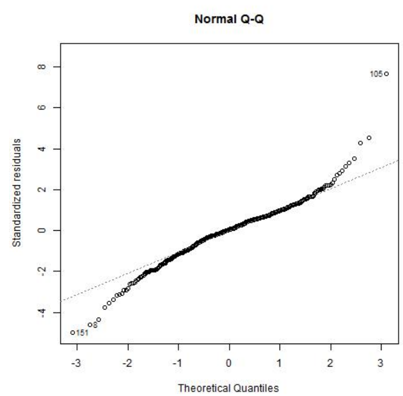

Figure 35: Prospective, skewness

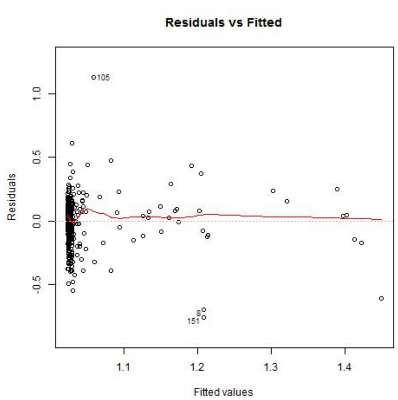

Figure 36: Retrospective, skewness

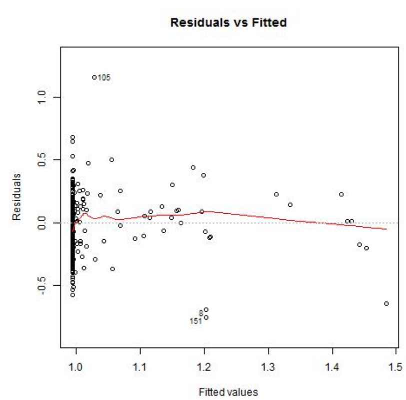

Figure 37: Prospective, skewness

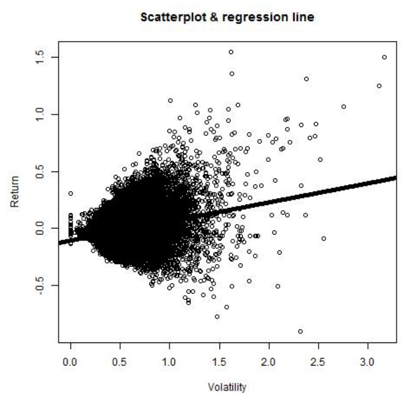

Figure 38: *Monthly, volatility*

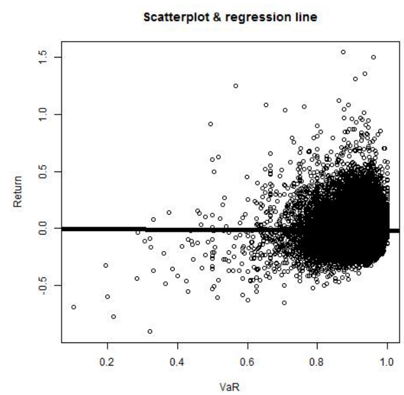

Figure 39: *Monthly, VaR*

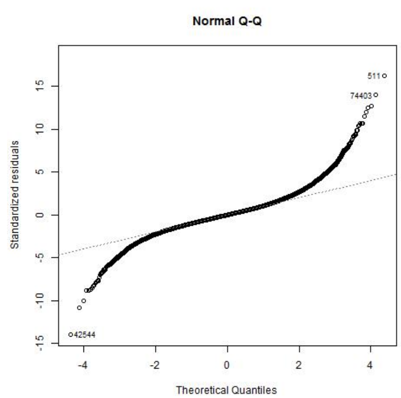

Figure 40: *Monthly, volatility*

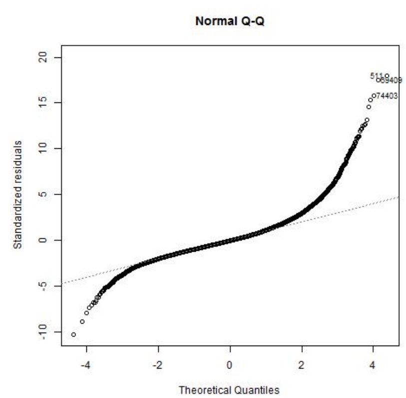

Figure 41: *Monthly, VaR*

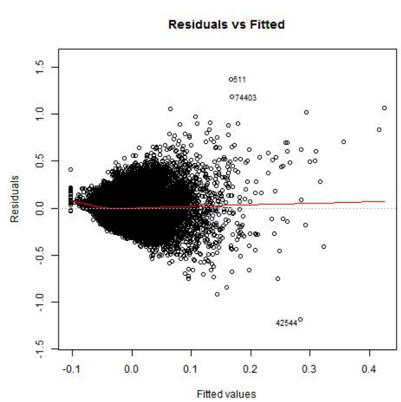

Figure 42: *Monthly, volatility*

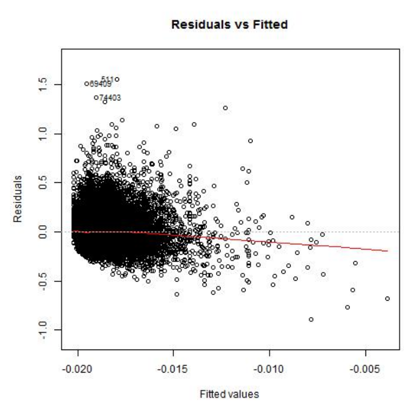

Figure 43: *Monthly, VaR*

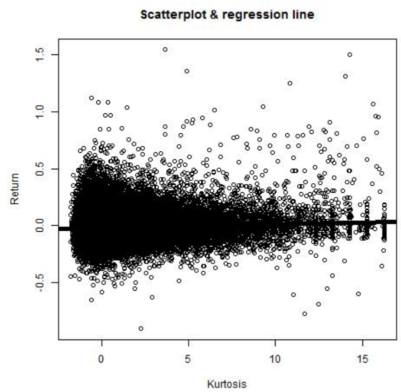

Figure 44: Monthly, kurtosis

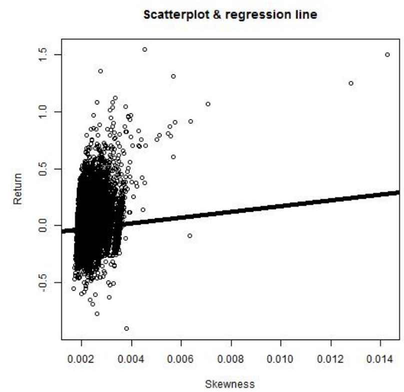

Figure 45: Monthly, skewness

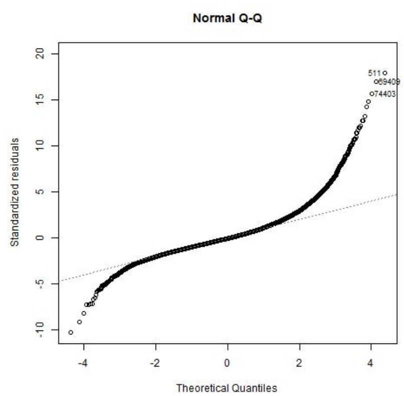

Figure 46: Monthly, kurtosis

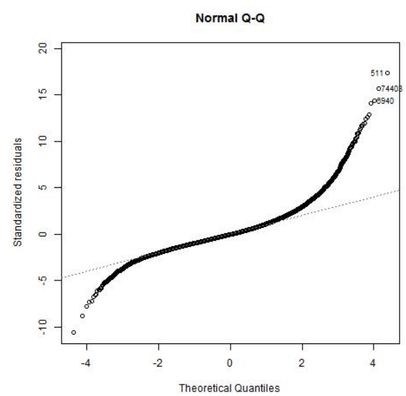

Figure 47: Monthly, skewness

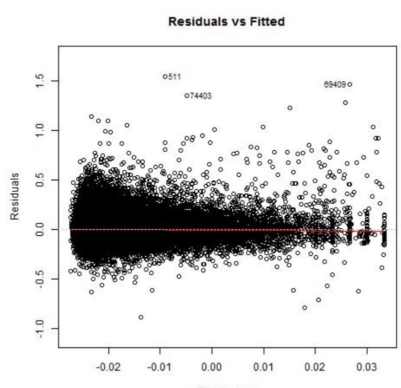

Figure 48: Monthly, kurtosis

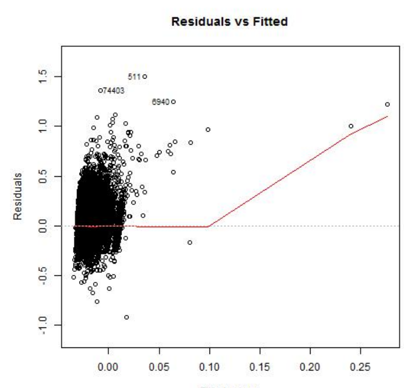

Figure 49: Monthly, skewness

7 Evaluations of the models

The diagnostic plots were based on least square estimation for beta and kurtosis and the Huber estimator for the other risk measures. This was due to the result of the Breusch-Pagan test. The Normal Q-Q plots support those choices. Beta and kurtosis for the Retro- and Prospective model seems to be not very far from normal distributed. However, kurtosis with the monthly model have an empirical distribution far from normal, which make the least squares estimates highly doubtful. Despite this, least squares will be used for the definitive result for these risk measures, however, the regression on kurtosis by the Monthly model will be discarded. The Huber estimator is to be used for the result for the Value at Risk, volatility and skewness.

When inspecting plots of regressions not provided, in general the diagnostic plots seem simililar. The differences between the periods are small and the similarities between the Retro- and Prospective model persist. In general, large cap seems to have the best fit and small cap the worst. For the Monthly model, the differences between caps and countries were even smaller.

For the Pro- and Retrospective model, all regressions had trend in residuals. The fit was best for beta. It was decent for Value at Risk but not very good for kurtosis, skewness and volatility. For the Monthly model the fit seems to be good for volatility and skewness. Kurtosis as said before is disregarded. Value at Risk has a clear trend in the residuals which damages the overall fit.

The overall conclusion of the models is that there are problems, but that the models generally work satisfactorily. A better fit would probably have been obtained with more careful investigations of the individual regressions. Also, comparing and selecting estimation method for specific regressions, including even more robust methods than the Huber estimator, could perhaps have improved the fit of individual regressions. Testing multiple regression with several risk measures or indicators for cap and country would also have been valuable. Despite the problems, the fit was in general good for the observations near the mass centra which makes it possible to draw conclusions from the result.

8 Results

In the coming pages the results for the three models will be presented. Positive estimates for kurtosis, volatility and beta are to be interpreted as a sign of positive relationship between return and risk, i.e. that risky stocks have seen larger returns. The opposite is true for VaR and skewness, i.e. with a positive estimate low-risk stocks have seen larger returns. The estimates will be presented under their symbol to the right the corresponding p-value for the zero hypothesis that the estimate differs from zero. Estimates significant on a 1% level which indicate a positive risk-return relationship will be coloured in blue and those which indicate a negative risk-return relationship will be coloured in red.

	γ	p	κ	p	VaR	p	σ	p	β	p
STOLARGE	1.4e+04	5.3e-03	7.1e-04	8.7e-03	7.7e-01	5.5e-01	1.3e-02	8.9e-01	1.3e-02	6.4e-01
STOMID	2.3e+04	4e-07	-3.8e-03	1.6e-02	6.9e+00	9.7e-05	-4.7e-01	2.1e-04	-5.5e-03	9.1e-01
STOSMALL	3.1e+04	5.1e-02	-8.1e-05	3.4e-01	7.3e+00	5.4e-10	-3.2e-01	2.7e-06	-1.3e-01	1.6e-03
HELLARGE	2.7e+04	2.6e-02	-1.6e-03	7.9e-02	6e+00	2.3e-02	-5.9e-01	1.3e-02	-5.1e-02	1.6e-01
HELMID	8.4e+04	6.2e-04	-1.3e-03	3.9e-01	-1.2e+00	4.9e-01	1.1e-01	4.7e-01	1.5e-02	6.5e-01
HELSMALL	6.4e+03	8e-01	5.2e-05	7.6e-01	2.7e+00	1.3e-04	-1.6e-01	1.8e-03	-4.9e-02	3.3e-01
CPHLARGE	2.3e+04	2.5e-01	6.4e-05	3.4e-02	6.1e+00	8.6e-03	1.3e-01	3.1e-01	-1.2e-01	1.5e-01
CPHMID	-1.5e+04	8.1e-01	-4.2e-03	2e-01	9.8e+00	2.4e-02	-7e-01	4.2e-02	-2.3e-01	8.6e-02
CPHSMALL	5e+04	2.3e-01	-2e-05	7.2e-01	5.3e+00	7.7e-09	-3.4e-01	5.3e-09	-4.2e-02	5.6e-01
STO	2.5e+04	5.2e-11	-1.3e-04	1.4e-01	7.5e+00	1.7e-23	-4.4e-01	1.2e-19	-2.7e-02	2.8e-01
HEL	2.3e+04	9.9e-02	-1.4e-04	3.2e-01	3.5e+00	1.5e-08	-2.2e-01	1.9e-07	2e-02	4.5e-01
CPH	3.6e+04	1.2e-01	2.4e-05	5.4e-01	6.2e+00	1.6e-17	-3.8e-01	4.6e-16	-7.4e-03	9e-01
LARGE	1.7e+04	2.3e-04	7.5e-05	3e-02	3.3e+00	1.3e-03	-1.9e-02	7.9e-01	-1.4e-02	4.7e-01
MID	2.6e+04	2.9e-10	-3.1e-03	9.5e-03	4.5e+00	3.5e-04	-2.8e-01	4.1e-03	4.1e-03	9e-01
SMALL	3.1e+04	1.5e-02	-4.2e-05	3.5e-01	5.2e+00	1.4e-18	-2.9e-01	3.3e-15	-5e-02	3.9e-02
ALL	2.7e+04	1e-14	-2.8e-05	4.3e-01	6.2e+00	4.6e-46	-3.7e-01	3e-37	1.3e-02	4.4e-01

Table 9: *Retrospective method for whole period*

Table 10: *Retrospective method for recession*

	γ	p	κ	p	VaR	p	σ	p	β	p
STOLARGE	1.7e+04	3.8e-03	6.5e-04	2e-02	1.7e+00	3.4e-01	3.1e-02	8.7e-01	2.3e-02	5.3e-01
STOMID	2.4e+04	1.5e-05	-7.8e-03	7.8e-04	1.1e+01	5.3e-06	-7.2e-01	4.6e-05	9.6e-03	8.5e-01
STOSMALL	2.5e+04	1.9e-01	2.1e-06	9.8e-01	8.3e+00	5.3e-09	-3.4e-01	1.2e-05	-1.2e-01	1.3e-02
HELLARGE	3.5e+04	1.2e-01	4.5e-04	8.4e-01	8.1e+00	4.9e-03	-8.5e-01	2.5e-03	-3.9e-02	5e-01
HELMID	1.2e+05	4e-04	2.6e-03	4.7e-01	2.1e+00	4.8e-01	4.3e-02	8.8e-01	1.8e-02	7e-01
HELSMALL	1.8e+04	5.7e-01	1.3e-04	5.6e-01	1.5e+00	1.2e-02	-1.3e-01	1.6e-02	3.6e-02	5e-01
CPHLARGE	2.9e+04	2.8e-01	1.9e-03	5.8e-02	6.5e+00	2.6e-02	-2.3e-01	3.3e-01	-1.4e-01	1.3e-01
CPHMID	4e+04	5.9e-01	-1.7e-02	2.4e-03	1.2e+01	1.1e-03	-8.5e-01	5.5e-04	-2.8e-01	7.4e-02
CPHSMALL	7.9e+04	6.8e-02	-9.8e-05	2.1e-01	4.5e+00	7.5e-09	-2.9e-01	4.6e-08	-1.2e-01	1e-01
STO	2.6e+04	1.4e-09	-3.4e-05	7.4e-01	8.8e+00	4.7e-21	-4.7e-01	1.4e-16	-1.5e-02	5.8e-01
HEL	4.1e+04	3.4e-02	-1.6e-05	9.3e-01	2.1e+00	1.7e-04	-1.7e-01	3.4e-04	5e-02	9.4e-02
CPH	6.2e+04	2.2e-02	-1.7e-04	4e-02	5.9e+00	3.5e-16	-3.7e-01	4.7e-14	-5.6e-02	3.9e-01
LARGE	2e+04	2.5e-04	7.5e-04	2.5e-03	4.8e+00	2.7e-04	-2.6e-01	3.6e-02	-1e-02	6.9e-01
MID	2.9e+04	1.3e-08	-6.8e-03	4.3e-04	8e+00	1.2e-05	-5.1e-01	5.1e-04	3.1e-02	4e-01
SMALL	4e+04	7.8e-03	-6.4e-05	3e-01	4.6e+00	4.6e-17	-2.8e-01	1.6e-13	4.5e-03	8.7e-01
ALL	3.2e+04	1.2e-14	-1.3e-04	4.2e-02	6.1e+00	1.9e-38	-3.9e-01	5.5e-32	4.6e-02	1.3e-02

	γ	p	κ	p	VaR	p	σ	p	β	p
STOLARGE	3.3e+03	9.7e-01	-7.1e-05	9.4e-01	1.1e+00	3.2e-01	-5.9e-02	4.4e-01	1.3e-02	7.6e-01
STOMID	-2.7e+05	1.2e-01	-8.2e-03	1.3e-01	4.9e+00	1.2e-02	-3.8e-01	1.5e-03	-9.3e-03	9.3e-01
STOSMALL	-2.4e+05	2.5e-04	-7.4e-04	6e-01	3.5e+00	9.5e-07	-2.6e-01	2.8e-07	-1.5e-01	1.1e-03
HELLARGE	-1.7e+05	6e-01	-6e-03	4.5e-01	3.3e+00	2e-02	-2.9e-01	2.3e-02	-5.9e-02	3.6e-01
HELMID	-7.7e+03	9.8e-01	-2.8e-03	4.3e-01	5.8e-02	9.7e-01	-1e-01	5.5e-01	2.4e-02	5.2e-01
HELSMALL	-3.5e+04	7.3e-02	-9.5e-04	3.1e-01	8.3e-01	5.3e-03	-1e-01	1e-03	-5.4e-02	3.2e-01
CPHLARGE	1.5e+04	9.5e-01	-2.9e-04	9.4e-01	5.5e+00	1.9e-01	-3.9e-01	3e-01	-2.1e-01	6.8e-03
CPHMID	-9.1e+05	1.7e-02	-1.4e-02	1.9e-03	1e+01	1.1e-02	-6.4e-01	8.9e-03	-2.6e-01	7.7e-02
CPHSMALL	-2.4e+05	5.2e-02	-1.6e-03	2.1e-01	4.5e+00	6.1e-09	-3.2e-01	8.2e-09	-7.3e-02	2.8e-01
STO	-2.7e+05	1.3e-07	-2e-03	1.4e-01	4.7e+00	1.2e-14	-3.2e-01	6.2e-16	-3.1e-02	4.9e-01
HEL	-4.5e+04	2.4e-03	-1.7e-03	4.3e-02	1.4e+00	2.3e-06	-1.5e-01	1.7e-08	3e-02	3.2e-01
CPH	-3.2e+05	4.4e-03	-3.5e-03	5.2e-03	5.4e+00	4.4e-14	-4.2e-01	3.4e-15	-4.7e-02	4e-01
LARGE	-3.3e+03	9.7e-01	-8e-06	9.9e-01	1.7e+00	6.5e-02	-1e-01	1.6e-01	-7.1e-03	8.1e-01
MID	-2.8e+05	5e-02	-8.3e-03	3.6e-02	3.6e+00	5.4e-03	-3.6e-01	5.2e-05	1.7e-02	8.1e-01
SMALL	-8.1e+04	2.4e-03	-1e-03	1.5e-01	2.6e+00	5.4e-14	-2.2e-01	1e-16	-6.4e-02	1.4e-02
ALL	-1.8e+05	1e-08	-2.2e-03	8.6e-03	3.5e+00	2.7e-28	-3e-01	8.7e-35	2.1e-02	4.1e-01

Table 11: *Prospective method for whole period*

Table 12: *Prospective method for recession*

	γ	p	κ	p	VaR	p	σ	p	β	p
STOLARGE	2.9e+05	7.3e-03	-5.4e-04	6.1e-01	-6.9e-01	5.8e-01	6.1e-02	4.9e-01	1.9e-02	6.8e-01
STOMID	-5.9e+04	7.9e-01	-7.9e-03	1.5e-01	3.4e+00	1.3e-01	-3e-01	4.6e-02	2.7e-03	9.8e-01
STOSMALL	-2e+05	8.9e-03	-1.1e-03	4.7e-01	3.4e+00	2.4e-05	-2.4e-01	3.6e-05	-1.5e-01	3.8e-03
HELLARGE	3.7e+05	3.8e-01	-8.3e-03	3.9e-01	6.5e+00	3.9e-04	-6.3e-01	4e-04	-5.1e-02	5.2e-01
HELMID	-1.1e+05	8.1e-01	-1.4e-03	7.7e-01	2.3e+00	3.9e-01	-2.3e-01	4.1e-01	2.8e-02	6e-01
HELSMALL	-4.3e+04	4.8e-02	-1.3e-03	2.3e-01	1.5e+00	3.5e-04	-1.3e-01	6.9e-04	2.4e-02	6.9e-01
CPHLARGE	-7.2e+04	8.3e-01	1.3e-03	7.6e-01	1.8e+00	7e-01	-3.1e-01	4.8e-01	-2.5e-01	5.4e-03
CPHMID	-5.6e+05	2.7e-01	-1.1e-02	5.6e-02	1e+01	1.1e-02	-5.5e-01	2e-02	-3.3e-01	4.3e-02
CPHSMALL	-1.9e+05	1e-01	-2.1e-03	1.1e-01	4e+00	4.9e-07	-2.8e-01	1.7e-06	-1.8e-01	7.7e-03
STO	-2e+05	5.7e-04	-2.3e-03	9e-02	4e+00	2.1e-10	-2.9e-01	4.8e-11	-2.4e-02	6.1e-01
HEL	-5e+04	1.2e-02	-1.9e-03	4.7e-02	2.2e+00	3.5e-08	-1.9e-01	2e-07	5.9e-02	8.5e-02
CPH	-2.9e+05	2.2e-02	-3.9e-03	4.1e-03	5.4e+00	1.6e-12	-3.9e-01	5.7e-11	-1.2e-01	5.4e-02
LARGE	2.3e+05	1.4e-02	-4.1e-04	6.5e-01	7.3e-01	5.5e-01	-3.8e-02	6.7e-01	-5.8e-03	8.5e-01
MID	-8.7e+04	6.6e-01	-7.7e-03	6.1e-02	2.3e+00	1.7e-01	-2.5e-01	3.2e-02	4.4e-02	5.6e-01
SMALL	-6e+04	9.9e-03	-1.4e-03	7.8e-02	2.6e+00	3.9e-12	-2.1e-01	1e-12	-1.8e-02	5.4e-01
ALL	-1e+05	6.1e-04	-2.5e-03	3.8e-03	3.6e+00	1.4e-22	-2.9e-01	3.1e-25	5.1e-02	5.7e-02

	γ	p	κ	p	VaR	p	σ	p
STOLARGE	6.7e-02	2.8e-45	2.1e+01	9.3e-27	1.3e-03	1.5e-04	4.7e-01	1e-108
STOMID	1.2e-01	1.3e-90	3.5e+01	8.1e-37	1.5e-03	3.6e-03	2.6e-01	4.7e-30
STOSMALL	1.6e-01	3.5e-190	6e+01	6.1e-97	7.2e-04	1.2e-01	1.7e-01	4.6e-21
HELLARGE	1.5e-01	5.2e-76	2.5e+01	1.1e-14	3.1e-03	2.2e-09	3.3e-01	7.7e-18
HELMID	4.9e-02	2.7e-14	1.4e+01	3.2e-08	2.2e-03	6e-08	5.2e-01	5.4e-74
HELSMALL	1e-01	1.3e-106	3.6e+01	1.4e-32	-4.4e-04	2.3e-01	1.1e-01	3.9e-10
CPHLARGE	1.5e-01	1.9e-67	2.4e+01	6.7e-12	1.7e-03	2.5e-03	4.4e-01	2.4e-28
CPHMID	1.1e-01	1.6e-19	2.3e+01	1.4e-05	3.3e-03	3.4e-04	2.4e-01	1.8e-05
CPHSMALL	1.2e-01	6.5e-160	6e+01	9.1e-99	-1.9e-03	4.6e-19	1.2e-01	8.9e-12
STO	1.9e-01	0e+00	4.3e+01	9e-149	3.2e-03	1.8e-32	-3.2e-02	8.9e-03
HEL	1.3e-01	1.4e-297	2.9e+01	5e-53	2.7e-03	3.3e-26	6.8e-02	2.2e-06
CPH	1.7e-01	0e+00	4.9e+01	4.1e-106	1.4e-03	3.9e-07	-7.1e-02	6.3e-06
LARGE	1e-01	2e-153	2.3e+01	6.5e-48	2.1e-03	6.8e-16	4.2e-01	1.5e-136
MID	8.1e-02	2.1e-79	2.7e+01	1.4e-45	2.7e-03	2.4e-17	3.7e-01	1.1e-100
SMALL	1.2e-01	0e+00	5.5e+01	1.4e-220	-1.1e-03	2.7e-07	1.4e-01	3.8e-40
ALL	1.7e-01	0e+00	4.1e+01	8.3e-297	3.4e-03	1.3e-102	-1.8e-02	2.8e-02

Table 13: *Monthly method for whole period*

Table 14: *Monthly method for recession period*

	γ	p	κ	p	VaR	p	σ	p
STOLARGE	4.6e-02	7.4e-05	2e+01	1.3e-03	-8.3e-04	2.6e-01	7.5e-01	1.4e-40
STOMID	2.5e-01	2.2e-59	3e+01	5.5e-07	2e-03	7.9e-02	1.8e-01	3e-03
STOSMALL	3.2e-01	4.1e-120	6.3e+01	8.8e-19	2.4e-03	2.6e-02	-9.5e-02	5e-02
HELLARGE	2.9e-01	2.6e-50	2.2e+01	8.1e-03	5e-04	6.9e-01	-6e-02	4.8e-01
HELMID	4.6e-02	1.2e-02	9.2e+00	1.7e-01	1.2e-03	2.5e-01	7.9e-01	6.8e-22
HELSMALL	1.7e-01	2.9e-56	3.8e+01	1.1e-08	4e-04	6.4e-01	-2.1e-01	6.9e-07
CPHLARGE	2e-01	1.2e-25	1.8e+01	1.6e-02	7.6e-04	5.6e-01	2.2e-01	1.8e-02
CPHMID	4.6e-01	7.6e-55	3.3e+01	2.5e-02	9.5e-03	7.3e-05	-8.8e-01	8.8e-11
CPHSMALL	2.5e-01	6.4e-145	6.7e+01	2.9e-29	-8.8e-03	6.4e-45	-1.9e-01	3.5e-07
STO	2.9e-01	1.5e-295	4.4e+01	6.6e-27	3.5e-03	2e-09	-1.8e-01	3.8e-09
HEL	2e-01	6.8e-122	2.9e+01	4.2e-11	3.5e-03	2e-08	-2.2e-01	1.3e-10
CPH	3.6e-01	0e+00	6.1e+01	3.7e-31	4.9e-04	4.1e-01	-6.6e-01	1.8e-73
LARGE	1.5e-01	3.3e-60	2e+01	4.1e-06	-7e-05	9e-01	3.7e-01	1.3e-18
MID	1.7e-01	4.1e-47	2.4e+01	1.8e-07	4.3e-03	1.3e-07	3.5e-01	8.3e-13
SMALL	2.5e-01	9e-303	6.1e+01	2e-51	-6e-04	2.1e-01	-2.4e-01	2.1e-20
ALL	3.1e-01	0e+00	4.7e+01	2e-64	5.9e-03	2.3e-61	-4.6e-01	2.5e-115

9 Discussion of the result

Even without the model problems and data issues, it is important to remember that unlike hard science, the field of finance is ever-changing. Investors alter their behaviour with increased knowledge and changed environments. Driven by more sophisticated technology and algorithms, better and more sensitive models, new research, globalization and more comprehensive regulations like Basel III, IFRS9 and Solvency II, financial patterns change over time. Conclusions from historical data may become outdated and old data obsolete.

9.1 Retro & Prospective model

Beta had a positive relationship with risk only for Total in the recession period with the retrospective model, however the p-value is just below the threshold. The relationship between beta and return is negative for Stockholm small cap for both periods and both models. There were also negative relationships the combined small stock dataset and Copenhagen small cap. However, the issue with changea of cap make this unreliable. Since beta was insignificant in most cases and positive in none and the fit was quite good: a conclusion can be drawn that beta in general does not seem to have a positive relationship in the Nordic Stock Exchanges.

The relationship between skewness and return seems in general to be negative or non-existent for the retrospective method and positive or non-existent for the prospective method. No conclusions can be drawn about the true relationship. Kurtosis too had ambiguous results and it is not possible to draw any conclusions.

Volatility and Value at Risk are similar, both by problems with model fit, and the result. Most datasets, including the dataset with all observations, show significant negative relationship between Value at Risk and volatility. Stocks with high-risk in terms of volatility and Value at Risk do not seem to pay off.

9.2 Monthly model

The estimates for regressions with the monthly model were generally significant, often with remarkably small p-values. No conclusion is drawn for kurtosis, since the parameters were estimated using least square, which as mentioned earlier makes the result unreliable. For skewness, the relationship with return was negative for all regressions with very small p-values. This indicates that months with higher skewness have lower return. This is intuitive since it seems reasonable that the short term distribution for stocks has a negative skew, i.e. great losses occur more frequently than the corresponding gains.

Value at Risk and volatility both show inconclusive results. For volatility for the whole period, the relationship is positive for all but Stockholm & Copenhagen

Total (negative) and the full dataset (non-significant). It is not possible to draw any conclusions.

9.3 Difference between caps, exchanges and periods

The result shows no clear evidence for any difference between the exchanges worth mentioning. The same largely goes for caps, however, there is some indicators that Small cap stocks have a more negative relationship than the others. But, as mentioned earlier, more successful Small cap companies may have moved cap. Thus, no conclusions about different relationships for different caps can be drawn.

9.4 Explanation of a possible low-risk anomaly

The purpose of this paper was to investigate the relationship rather than try to explain it. However, it could be worth mentioning something of what could be the cause of a possible low risk-anomaly. Bradley et al states two possible reasons.[34] Firstly, many investors are irrational with a preference for risk taking and overconfidence in their ability of picking successful stocks, which creates a preference for riskier stocks over safe ones. They thus have a preference for a lottery-like chance of very high return. Secondly, regulatory requirements make institutional investors pick stocks in a way that discourages investment in low-volatility stocks.

J. Karceski claims that mutual fund investors care more about good performance in bull market than in the bear market and in bull market high-risk stocks generally outperform low-risk stocks.[35] K. Hou and R. Loh evaluate different explanations and find that 78-84% of the relationship in volatility-sorted portfolios can be explained, where lottery preference is the single most important factor, with frictions in the markets coming secondly.[36]

10 Further research

All regression models contained only one risk measure. Regressions with multiple risk measures as variables would shine light on the interaction between the risk measures. A relationship between return and risk for some measures could be due to strong covariation with another measure. There could also be a covariation with some other factor, e.g. E.F. Fama and K.R. in their famous result explained the expected return with market capitalisation and book-to-market ratio.[37]

There were also a lot of specific choices that in various degree may have influenced the result considerably. This includes the choice of market index for beta

[34] Baker et al., 'The Low Beta Anomaly: A Decomposition into Micro and Macro Effects'
[35] Karceski, 'Returns-Chasing Behavior, Mutual Funds, and Beta's Death.'.
[36] Hou and Loh, 'Have we solved the idiosyncratic volatility puzzle?'.
[37] Fama and French, 'Common risk factors in the returns on stocks and bonds'.

where many other approaches are possible, including turnover-weighted average or using a well known index. A. Damodaran calculated beta for Disney with two different indices.[38] Damodaran found it to be 0.99 for Dow 30 and 1.13 for S&P. Even though the difference could seem small for a single stock, it could have profound effects on a regression if the index was to be biased. There was also the choice to use simple return instead of compound, which may have influenced the result. In this report, volatility was observed from historic data but it is also possible to use EGARCH models to estimate expected volatility like F. Fu did in his similiar study.[39]

The research in the literature review show large methodological differences between the papers. Nor do this report fully comply with any of the other papers, partly due to that the exact method is seldom explicitly written out in the paper. For sake of generality, as similiar methods as possible should be used. With more knowledge of methodological differences, perhaps more of the differences in conclusions among the papers could be explained. Like Blitz et al[40], the different methodologies can be used in the same paper and the result can be compared.

It is possible to use more data for being able to draw more general conclusion. For example, data from other stock exchanges and other asset classes such as bonds, derivates and real estate could be included and the time frame could be longer. However, there is a conflict between using more data and conducting more analysis and being able to carefully examine the model. This report has included a lot of regressions since several methods and datasets were included. As a consequence, fitting appropiate models for all of these was not possible, and the quality of the result suffered. Further research should beforehand be clear about the goal and how methodological choices could affect the scope substantially.

[38] Damodaran, 'Estimating Risk Parameters'.

[39] Fu, 'Idiosyncratic Risk and the Cross-Section of Expected Returns'.

[40] Blitz et al., 'Is the Relation Between Volatility and Expected Stock Returns Positive, Flat or Negative?'

References

Baker, M., Bradley, B. and Taliaferro, R., 'The Low Beta Anomaly: A Decomposition into Micro and Macro Effects', *Financial Analysts Journal* 70:2 (2014), 43–58.

Baker, M., Brendan, B. and Wurgler, J., 'Benchmarks as Limits to Arbitrage: Understanding the Low Volatility Anomaly', *Financial Analysts Journal* 67:1 (2011), 40–54.

Blitz, D., van Vliet, P. and van der Grient, B., 'Is the Relation Between Volatility and Expected Stock Returns Positive, Flat or Negative?', *Available at SSRN 1881503* (2011).

Brennan, M.J. and Li, F., 'Agency and Asset Pricing' (2008).

Breusch, T.S. and Pagan, A.R., 'A simple test for Heteroscedasticity and Random Coefficient Variation' (1979).

Chaudhury, K. N., 'On the convergence of the IRLS algorithm in Non-Local Patch Regression', *IEEE Signal Processing Letters* 20:8 (2013), 815–818.

Damodaran, A., 'Estimating Risk Parameters' (1999).

Davidson, R. and MacKinnon, J.G., *Econometric Theory and Methods* (Oxford, New York: Oxford University Press, 1999).

Dongcheol, K., 'The Errors in the Variables Problem in the Cross-Section of Expected Stock Returns' (1995).

Drukker, D., 'Testing for Serial Correlation in Linear Panel-Data Models' (2003).

Durvasula, S., Sharma, S. and Carter, K., 'Correcting the t statistic for measurement error' (2012).

Fahrmeir, L. et al., *Regression Models, Methods and Applications* (Heidelberg, Berlin: Springer, 2013).

Fama, Eugene F and French, Kenneth R, 'Common risk factors in the returns on stocks and bonds', *Journal of financial economics* 33:1 (1993), 3–56.

Fernando, S. and Nimal, P.D., 'The Conditional Relation Between Beta and Return: Evidence from Japan and Sri Lanka' (2013).

Fu, F., 'Idiosyncratic Risk and the Cross-Section of Expected Returns', *Journal of Financial Economics* 91:1 (2009), 24–37.

Hausman, J., 'Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left' (2001).

Hou, K. and Loh, R., 'Have we solved the idiosyncratic volatility puzzle?', *Journal of Financial Economics* 121:1 (2016), 167–194.

Huber, P.J., 'Robust estimation of a location parameter', *The Annals of Mathematical Statistics* 35:1 (1964), 73–101.

Hudson, R. and Gregoriou, A., 'Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns', *Available at SSRN 1549328* (2010).

Karceski, J., 'Returns-Chasing Behavior, Mutual Funds, and Beta's Death.' (2002).

Lahaye, J., Laurent, S. and Neely, C, 'Jumps, cojumps and macro announcements', *Journal of Applied Economics* 26:6, 893–921.

Martellini, L., 'Towards the Design of Better Equity Benchmarks'.

Mathur, I, Pettengill, G. and Sundaram, S., 'The Conditional Relation Between Beta and Return' (1995).

Mosteller, F. and Tukey, J.W, *Data Analysis and Regression: A Second Course in Statistics* (Reading, Massachusetts: Pearson, 1977).

Sundberg, R., *Lineära Statistiska Modeller* (Stockholm: University of Stockholm, 2015).

Sydsaeter, K. et al., *Further Mathematics for Economic Analysis*, 2nd edition.

Tsay, R.S., *Analysis of Financial Time Series* (Chicago, Illinois: Wiley, 2010).

Yu, Chun, Yao, Weixin and Bai, Xue, 'Robust Linear Regression: A Review and Comparison', *arXiv preprint arXiv:1404.6274* (2014).