

Norwegian School of Economics

Bergen, Spring, 2014

NHH



Back Testing Multi Asset Value At Risk

Norwegian Data

Gudmund Råheim

Supervisor: Professor Jørgen Haug

Master Thesis within the main profile of Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Abstract

This paper attempts to estimate Value At Risk (VaR) for a multi asset Norwegian portfolio, using some of the most popular estimation methods, Variance Covariance Method, Historical Simulation and Monte Carlo Simulation. The Variance Covariance Method is applied with both time varying and constant volatility.

Each VaR estimation method's accuracy is tested, using Kupiec's univariate testing framework, for multiple single points in the left tail of the portfolio's return distribution, and Pérignon and Smith's multivariate framework for a larger subset of the left tail.

It compares each method's overall results for the Norwegian portfolio with those found by Wu et al. (2012) on a similar Taiwanese portfolio.

And finally, based on the empirical testing, it attempts to draw a conclusion on which method is best suited for Norwegian data.

Table of Contents

ABSTRACT	2
1. OBJECTIVE	5
2. THEORY	6
2.1 CONTINUOUSLY COMPOUNDED RETURNS.....	6
2.2 VALUE AT RISK (VAR).....	6
2.3 VARIANCE COVARIANCE METHOD	6
2.3.1 <i>Historical Variance</i>	7
2.3.2 <i>Portfolio Variance</i>	7
2.3.3 <i>Co-Movement of Asset Returns (Covariance)</i>	7
2.3.4 <i>GARCH</i>	8
2.3.5 <i>The RiskMetrics Approach (EWMA)</i>	9
2.4 HISTORICAL SIMULATION	9
2.5 MONTE CARLO SIMULATION.....	10
2.5.1 <i>Lognormal Asset Price Model</i>	10
2.6 BACK TESTING.....	11
2.6.1 <i>Univariate Unconditional Coverage Test</i>	11
2.6.2 <i>Multivariate Unconditional Coverage Test</i>	12
3. DATA	14
4. APPLIED METHODOLOGY	15
4.1 VARIANCE COVARIANCE METHOD	15
4.1.1 <i>Historical Variance</i>	15
4.1.2 <i>GARCH(1,1)</i>	15
4.1.3 <i>EWMA (RiskMetrics)</i>	15
4.2 HISTORICAL SIMULATION	16

4.3	MONTE CARLO SIMULATION	16
4.4	BACK TESTING	17
5.	RESULTS.....	18
5.1	PORTFOLIO RETURN DATA.....	18
5.2	VARIANCE COVARIANCE METHOD.....	19
5.2.1	<i>Historical Variance</i>	19
5.2.2	<i>GARCH(1,1)</i>	20
5.2.3	<i>Risk Metrics (EWMA)</i>	21
5.3	HISTORICAL SIMULATION	22
5.4	MONTE CARLO SIMULATION	23
6.	DISCUSSION.....	24
6.1	PORTFOLIO RETURNS.....	24
6.2	VARIANCE COVARIANCE METHODS	24
6.2.1	<i>Historical Variance</i>	24
6.2.2	<i>GARCH(1,1)</i>	25
6.2.3	<i>RiskMetrics (EWMA)</i>	26
6.3	HISTORICAL SIMULATION	26
6.4	MONTE CARLO SIMULATION	27
6.5	OVERALL.....	28
6.6	WEAKNESSES AND POSSIBLE IMPROVEMENTS	28
7.	CONCLUSION	30
8.	REFERENCES	31
9.	APPENDIX.....	32
9.1	ESTIMATED GARCH(1,1) FUNCTIONS	32
9.2	VBA MONTE CARLO ESTIMATION	33

1. Objective

The objective of this paper is to examine by way of unconditional coverage tests, which of the three popular Value At Risk (VaR) estimation methods – Variance Covariance Method, Historical Simulation and Monte Carlo Simulation, give the best results for a multi asset Norwegian portfolio.

And further to compare the overall performance of each model on the Norwegian portfolio with that found by Wu et al. (2012) on a similar Taiwanese portfolio.

2. Theory

2.1 Continuously Compounded Returns

$$r = \ln \left[\frac{S_t}{S_{t-1}} \right]$$

r denotes the continuously compounded return of the asset S between the time $t-1$ and t (McDonald, 2006).

2.2 Value At Risk (VaR)

Value at Risk (VaR) is the maximum loss that can be expected with a given confidence level over a given time period (Hull, 2012). For a portfolio this means the difference between its current value and the value of the hypothetical portfolio generated by the negative return (loss) in the left tail of its return distribution, at the desired confidence level.

2.3 Variance Covariance Method

Assuming asset returns are normally distributed, the Value at Risk (VaR) can be found as the product of the assets current value, a set number of standard deviations depending on the desired confidence level and the square root of the desired time horizon (Wu et al., 2012)

$$VaR = Z_\alpha \cdot \sigma \cdot V \cdot \sqrt{t}$$

Where Z_α is the number of standard deviations dependent on the desired VaR confidence level.

2.3.1 Historical Variance

$$Var(r_i) = \frac{1}{M-1} \sum_{t=1}^M [r_{i,t} - E(r_i)]^2$$

$Var(r_i)$ is an expression of the sample variance of the return series r_i , M denotes the sample period and $E(r_i)$ is the mean return over the sample period (Benninga, 2008)

$$Cov(r_i, r_j) = \frac{1}{M-1} \sum_{t=1}^M [r_{i,t} - E(r_i)] \cdot [r_{j,t} - E(r_j)]$$

$Cov(r_i, r_j)$ is an expression of the sample covariance between the return series r_i and r_j , $E(r_i)$ and $E(r_j)$ is the respective mean return over the sample period and M denotes the sample period (Benninga, 2008)

2.3.2 Portfolio Variance

The variance of a return series is relatively easy to find as shown above, however estimating the variance of a portfolio containing two or more assets requires weighting each assets variance and their covariance, with their respective weights.

$$Var(r_p) = \sigma_p^2 = \sum_{i=1}^N (x_i)^2 \cdot Var(r_i) + 2 \cdot \sum_{i=1}^N \sum_{j=i+1}^N x_i x_j Cov(r_i, r_j)$$

Where x denotes the respective asset's current weight in the portfolio.

2.3.3 Co-Movement of Asset Returns (Covariance)

Estimation of the covariance elements in the portfolio variance becomes difficult with multivariate models. Even just a three asset model would require the estimation of 78 elements for the variance covariance matrix (Jorion, 2007). Harris et al. (2007) develop a simplified approach to calculating the covariance between return series, by combining each

return series $r_{1,t}$ and $r_{2,t}$ into two new series $r_{+,t} = r_{1,t} + r_{2,t}$ and $r_{-,t} = r_{1,t} - r_{2,t}$, the volatility of these new series can then be estimated by normal methods, for example GARCH(1,1), which will give the following variance

$$\sigma_{+,t}^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2 + 2\sigma_{12,t}$$

$$\sigma_{-,t}^2 = \sigma_{1,t}^2 + \sigma_{2,t}^2 - 2\sigma_{12,t}$$

These can then be combined into an estimate for the covariance of the original series

$$\sigma_{12,t} = \left(\frac{1}{4}\right)(\sigma_{+,t}^2 - \sigma_{-,t}^2)$$

2.3.4 GARCH

The historical variance approach assumes volatility is homoscedastic, which means it's constant over time, for many return series this assumption will be incorrect. Because of this several models for calculating time varying volatility have been conceived, one of these is GARCH or the generalized autoregressive conditional hetroskedastic model.

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1}$$

h_t expresses the conditional variance of a return series, where r_{t-1} and h_{t-1} is return and conditional variance of the previous period respectively and α_0 is a constant (Jorion, 2007). If the scaled residuals $\varepsilon_t = r_t / \sqrt{h_t}$, are assumed to be normally distributed and independent, the α_0 , α_1 and β parameters can be estimated by maximizing the logarithm of the likelihood function ($\max F(\alpha_0, \alpha_1, \beta | r)$) (Jorion, 2007).

$$\max F(\alpha_0, \alpha_1, \beta | r) = \sum_{t=1}^T \ln f(r_t | h_t) = \sum_{t=1}^T \left[\ln \frac{1}{\sqrt{2\pi h_t}} - \frac{r_t^2}{2h_t} \right]$$

Where f denotes the normal density function, r_t and h_t denote the return and conditional variance at time t.

2.3.5 The RiskMetrics Approach (EWMA)

This is a special case of the GARCH(1,1) model, where α_0 equals zero and α_1 and β sum to unity, the conditional variance is still given by h_t , but the model is expressed as follows (Jorion, 2007)

$$h_t = \lambda h_{t-1} + (1 - \lambda) r_{t-1}^2$$

JP Morgan through their development and testing of this approach found that the best values for lambda is 0.94 for daily return data and 0.97 for monthly return data (Jorion, 2007).

One of the main benefits of the approach is that the covariance between return series can be easily estimated

$$h_{12,t} = \lambda h_{12,t-1} + (1 - \lambda) r_{1,t-1} \cdot r_{2,t-1}$$

Where $h_{12,t}$ is the conditional covariance between return series r_1 and r_2 (Jorion, 2007).

2.4 Historical Simulation

Historical simulation only relies on the previous price movements of the assets in a portfolio, it is in other words non-parametric and requires no assumptions about market variables or the portfolio's true return distribution (Jorion, 2007).

It is implemented by using the historical returns of the assets over the sample period and weighting them with their current weights, to construct a hypothetical historical return distribution for the current portfolio. Value at risk is then found as the difference between the current value and the value generated at the desired confidence level in the hypothetical return distribution (Jorion, 2007).

$$HS_t = [a_1 \quad \dots \quad a_n] \begin{bmatrix} r_{1,t-1} & \dots & r_{1,t-k} \\ \vdots & \ddots & \vdots \\ r_{n,t-1} & \dots & r_{n,t-k} \end{bmatrix}$$

HS_t denotes the set of “historical” portfolio value scenarios generated for time t , based on historical price movements, a_i is the current value of asset i and $r_{i,t-j}$ is the return of asset i at time $t-j$, n is the number of assets in the portfolio and k is the length of the sample window.

2.5 Monte Carlo Simulation

Using Monte Carlo simulation to estimate the value at risk (VaR) for a portfolio consists of generating hypothetical future price paths for each asset, or generating the hypothetical future movements for the market variables that the portfolio is dependent on. These hypothetical values are then combined into possible future portfolios, the difference between today’s portfolio value and that of each hypothetical value will generate the portfolio’s return distribution. VaR can then be calculated using the price movement in the left tail of the distribution that corresponds to the desired VaR confidence level, this is equivalent to the difference between today’s portfolio value and the appropriate p percentile worst hypothetical portfolio (Hull, 2012). For example calculating VaR(5%) with 1000 hypothetical portfolio values, would mean finding the difference between today’s portfolio and the 50th lowest hypothetical portfolio.

This method requires pricing models for all the market variables relevant to the portfolio’s assets; or, alternatively each asset requires its own pricing model.

2.5.1 Lognormal Asset Price Model

If an asset has a lognormal price distribution, its continuously compounded returns will be normally distributed, with a mean of $(\alpha - \delta - 0.5\sigma^2)t$ and a variance of $\sigma^2 t$, where α is the asset’s average yearly return, δ its yearly dividend rate and σ^2 its yearly return variance (McDonald, 2006).

$$\ln(S_t/S_0) \sim N\left[(\alpha - \delta - 0.5\sigma^2)t, \sigma^2 t\right]$$

The asset price at time t , can then be expressed as a function of the current asset value.

$$S_t = S_0 e^{(\alpha - \delta - 0.5\sigma^2)t + \sigma\sqrt{t} \cdot z}$$

Where z is a normally distributed random variable, with a mean of zero and a variance of one ($z \sim N(0,1)$) (McDonald, 2006)

2.6 Back Testing

Back testing of VaR models consist of procedures that check if the empirically observed violation rates, which means the number of times where the actual loss is higher than the VaR estimate, is not statistically different from the desired VaR confidence level

For example with a VaR confidence level of 95% and a time horizon of one day, the actually loss should exceed the estimated VaR 5% of the time, when the sample size is sufficiently large. (Jorion, 2007)

2.6.1 Univariate Unconditional Coverage Test

The unconditional coverage test developed by Kupiec (1995) uses a log-likelihood ratio to test whether the empirical violation rate is statistically different with a 95% confidence level, from the expected violation rate at the given VaR confidence level. This test has a chi-square distribution when T is large, with one degree of freedom (Jorion, 2007).

$$H_0: \hat{p} = p$$

$$H_A: \hat{p} \neq p$$

H_0 denotes the null hypotheses to be tested, where \hat{p} is the empirically found violation rate and p is the desired VaR confidence level.

$$LR_{UC} = -2 \cdot \ln \left[(1-p)^{T-N} \cdot p^N \right] + 2 \cdot \ln \left\{ \left[1 - \left(\frac{N}{T} \right) \right]^{T-N} \cdot \left(\frac{N}{T} \right)^N \right\}$$

LR_{UC} is the formal expression of the test, where N is the number of violations, T is the number of observations in the estimation period and p is the desired VaR confidence level (Jorion, 2007).

2.6.2 Multivariate Unconditional Coverage Test

Pérignon and Smith (2008) expand Kupiec's univariate test to a multivariate framework, where multiple empirical violation rates are tested against their expected counterparts. This test also has a chi-square distribution when T is large, but the degrees of freedom depend on the number of tested violation rates.

$$LR_{MUC} = 2 \left\{ \left[n_0 \ln(1 - 1'\theta^*) + \sum_{i=1}^K n_i \ln(\theta_i^*) \right] - \left[n_0 \ln(1 - 1'\theta) + \sum_{i=1}^K n_i \ln(\theta_i) \right] \right\}$$

$$LR_{MUC} = \left\{ \ln \left[\frac{1 - 1'\hat{\theta}}{1 - 1'\theta} \right]^{n_0} + \sum_{i=1}^K \ln \left(\frac{\hat{\theta}_i}{\theta_i} \right)^{n_i} \right\}$$

$$n_0 = T - \sum_{i=1}^K n_i$$

Where T is the total number of observations in the estimation period, n_0 is the number of observations where none of the VaR estimates are violated.

$$J_{it} = \begin{cases} 1, & \text{if } -VaR(p_{i+1}) < R_t \leq -VaR(p_i) \\ 0, & \text{otherwise} \end{cases}$$

$VaR(p_{i+1})$ denotes a more extreme VaR estimate than $VaR(p_i)$ and R_t is the actual loss at time t .

$$n_i = \sum_{t=1}^T J_{it}$$

$$\hat{\theta}_i = \frac{n_i}{T}$$

$$\theta_i = p_i - p_{i+1}$$

θ_i denotes a segment of the left tail in the distribution, n_i denotes the number of violations in this segment and $\hat{\theta}_i$ is the empirical found size of the segment.

$$\theta = [\theta_1 \quad \dots \quad \theta_K]$$

$$\hat{\theta} = [\hat{\theta}_1 \quad \dots \quad \hat{\theta}_K]$$

θ and $\hat{\theta}$ are the 1 by K matrix composed by the desired segments of the left tail and their empirical counterpart, respectively. The formal null and alternate hypotheses are then given as

$$H_0: \hat{\theta} = \theta$$

$$H_A: \hat{\theta} \neq \theta$$

3. Data

The portfolio is based on the Norwegian OBX index, 5 years Norwegian government bonds (ST5X) and a non interest bearing foreign exchange element, with a placement in US dollars. 2251 daily closing price observations, from January 3rd 2001 to December 18th 2009, were gathered from NHH's "Børsprosjektet" for the OBX index and the ST5X bonds. To stay consistent the same dates were used for the Norwegian NOK to US Dollars exchange rate, these were gathered from "forex-history.net" (<http://forex-history.net/main/>). The data used was updated last on the 25th of May 2014.

The 2251 price observations were turned into 2250 daily continuously compounded return observations, where the initial 250 were used as start-up values for the various simulations and estimations. The remaining 2000 observations, from January 7th 2002 to December 18th 2009, were used in the analysis.

The initial portfolio consisted of equal placement in each asset of 10 million NOK, the portfolio was not balanced in any way during the estimation period, so its weights changed only according to the market price fluctuations of the assets.

4. Applied Methodology

4.1 Variance Covariance Method

4.1.1 Historical Variance

A sample size of 250 previous return observations was used to calculate each asset's variance and their covariance at time t .

4.1.2 GARCH(1,1)

The covariance between the return series was estimated using the method developed by Harris et al. (2007). The initial variance for each return series and the initial values for the series constructed based on the Harris et al. (2007) method, was set to zero, then $\alpha_0, \alpha_1, \beta$ was estimated by maximizing the sum of the log likelihood functions $F(\alpha_0, \alpha_1, \beta | r)$, based on the last 2000 variance and corresponding return values. See appendix 9.1 for the estimated GARCH(1,1) models.

4.1.3 EWMA (RiskMetrics)

The initial variance for each return series was set to zero, the variance and covariance for the return series was calculated based on JP Morgans recommended lambda for daily data of 0.94 (Jorion, 2007). The first 250 variance and covariance estimates (including the initial values) were discarded and the remaining 2000 were used to calculate the portfolio VaRs.

4.2 Historical Simulation

A sample size of 250 daily return observations was used, meaning 250 scenarios were generated for the portfolio, on each day in the estimation period.

$$HS_t = \begin{bmatrix} a_{OBX} & a_{ST5X} & a_{NOK_USD} \end{bmatrix} \begin{bmatrix} r_{OBX,t-1} & \cdots & r_{OBX,t-250} \\ r_{ST5X,t-1} & \ddots & r_{ST5X,t-250} \\ r_{NOK_USD,t-1} & \cdots & r_{NOK_USD,t-250} \end{bmatrix}$$

HS_t generates the 250 hypothetical portfolio values at time t .

4.3 Monte Carlo Simulation

All three assets were assumed to have log normal price distributions, and the log normal pricing model was used to estimate hypothetical one period (here one day) ahead price paths. The pricing model was implemented using the historical daily variance, based on the previous 250 trading days and the mean return over the same sample period.

5000 hypothetical portfolio values were generated for each day in the estimation period, consisting of 2000 days, meaning a total of 2 million hypothetical portfolio values were generated. Each hypothetical portfolio was based on a single normally distributed random draw.

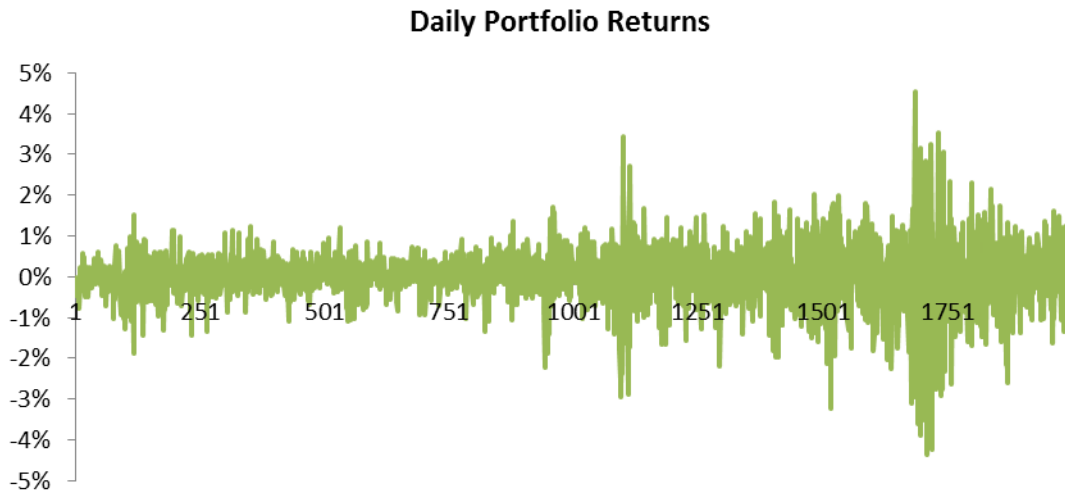
4.4 Back Testing

Both tests are evaluated on a 95% confidence level for the log likelihood ratio. For the univariate test this means the null hypothesis is rejected if the statistic (LR) is higher than 3.84. The multivariate test is applied with the 1%, 2.5% and 5% VaR significance levels; this gives three degrees of freedom and a rejection value of approximately 7.81. These rejection values are found using Excel's CHISQ.INV function, with the 95% confidence level and the respective degrees of freedom. The p-values listed in the results were calculated using Excel's CHISQ.DIST for the respective LR statistic and degrees of freedom.

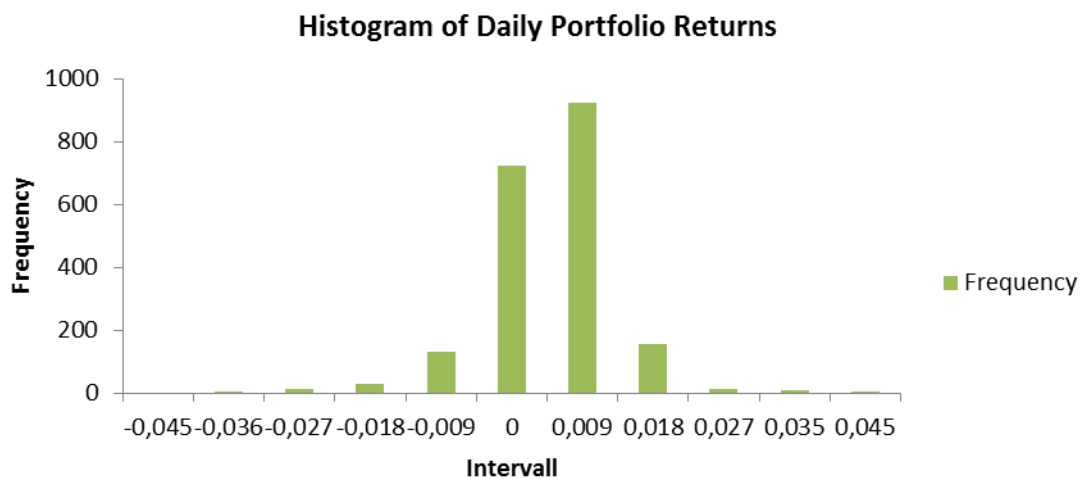
Both tests are applied on different sample sizes from 250 VaR estimates, increasing with 250 up to and including the full set of 2000 estimates.

5. Results

5.1 Portfolio Return Data



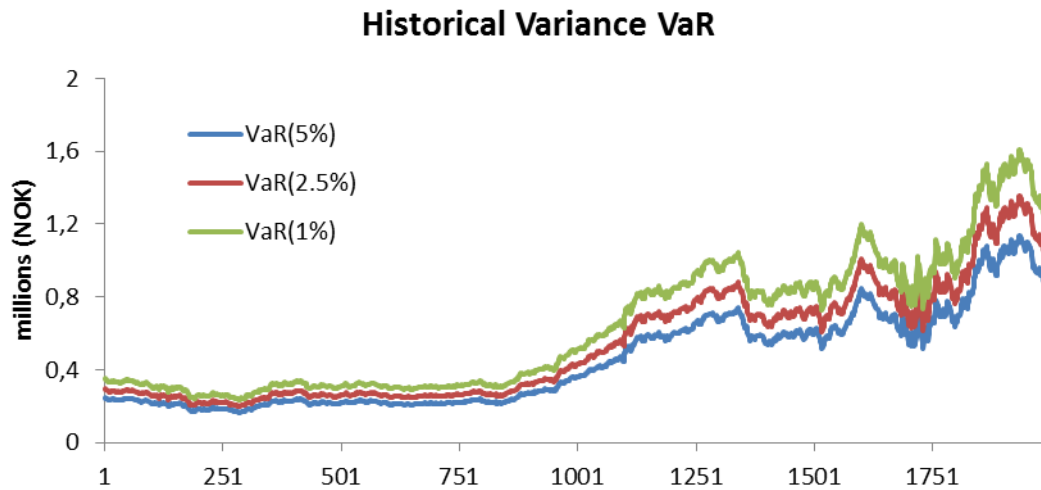
This figure show the daily returns for the Norwegian portfolio from the 7th of January 2002 to the 18th of December 2009. The returns look fairly stable for the first half of the sample; however they become much more unstable during the second half.



This figure shows the distribution of the portfolio returns, the distribution looks fairly symmetrical.

5.2 Variance Covariance Method

5.2.1 Historical Variance



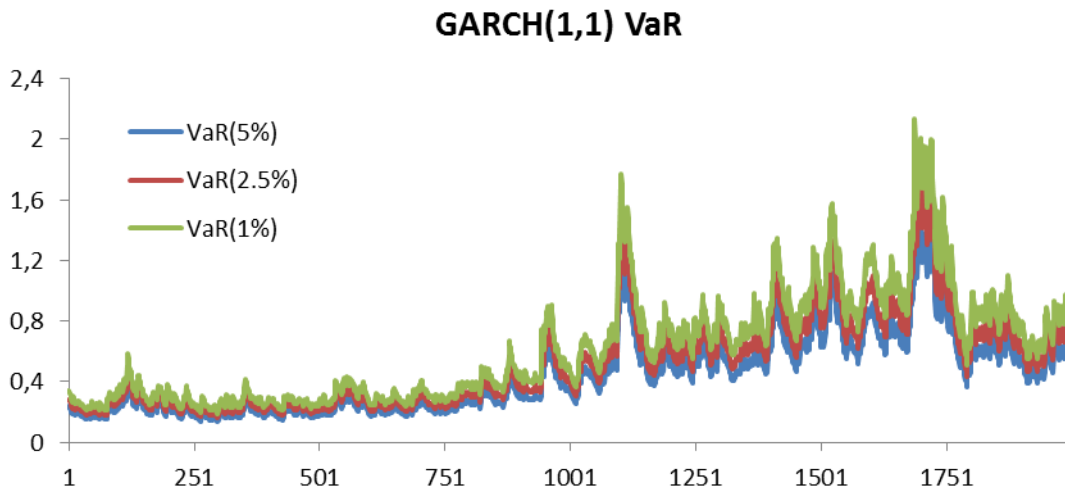
This figure shows the Historical Variance VaRs of the portfolio for 2000 days and 1%, 2.5% and 5% confidence levels.

Unconditional Coverage Tests (Hist. Variance)

Sample size (T)	250	500	750	1000	1250	1500	1750	2000	
Univariate	p=5%	7,60 %	4,60 %	5,07 %	5,00 %	5,28 %	7,80 %	6,69 %	5,95 %
		3,09	0,17	0,01	0,00	0,20	21,31	9,51	3,59
		(0,079)	(0,678)	(0,933)	(1)	(0,652)	(0)*	(0,002)*	(0,058)
	p=2,5%	5,20 %	3,20 %	3,07 %	3,20 %	3,60 %	5,53 %	4,74 %	4,20 %
		5,73	0,92	0,92	1,85	5,47	42,32	28,71	19,75
		(0,017)*	(0,336)	(0,337)	(0,174)	(0,019)*	(0)*	(0)*	(0)*
	p=1%	2,40 %	1,60 %	1,87 %	2,30 %	2,24 %	3,67 %	3,14 %	2,75 %
		3,56	1,54	4,53	12,49	14,36	64,01	51,78	41,90
		(0,059)	(0,215)	(0,033)*	(0)*	(0)*	(0)*	(0)*	(0)*
Multivariate	5,93	4,44	5,77	17,29	18,15	65,76	53,96	46,53	
	(0,115)	(0,217)	(0,124)	(0,001)*	(0)*	(0)*	(0)*	(0)*	

This table shows the results of univariate and multivariate convergence tests for the Historical Variance VaR method. The sample periods are from 250 to 2000 days. Violation rate is defined as the ratio of violations to observed samples and is displayed as a percentage figure. The LR statistics with their corresponding p-values in parentheses are displayed for each violation rate. * indicates p-values that allow rejection of the null hypotheses at the 5% confidence level.

5.2.2 GARCH(1,1)



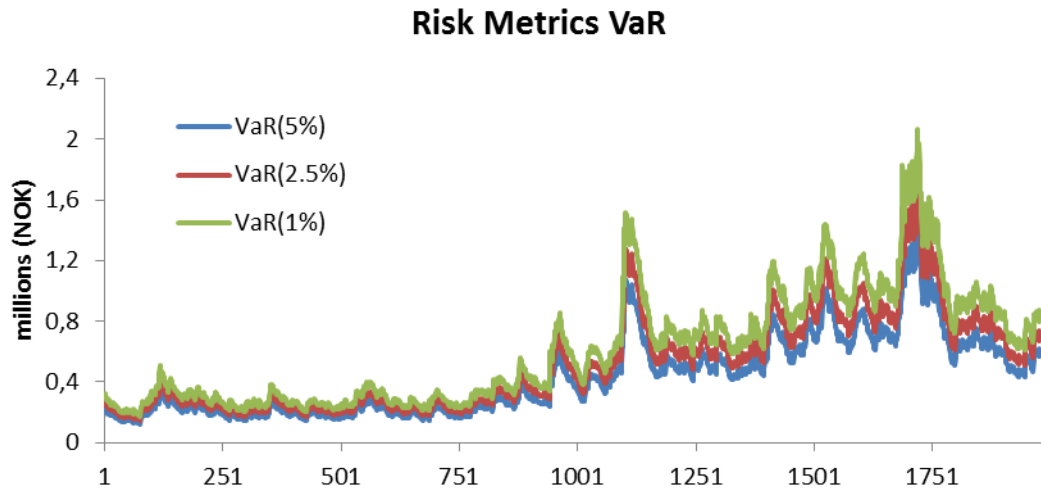
This figure shows the GARCH(1,1) VaRs of the portfolio for 2000 days and 1%, 2.5% and 5% confidence levels.

Unconditional Coverage Tests (GARCH)

Sample size (T)	250	500	750	1000	1250	1500	1750	2000	
Univariate	p=5%	6,80 %	4,60 %	5,20 %	4,70 %	4,80 %	6,60 %	5,66 %	5,55 %
		1,54	0,17	0,06	0,19	0,11	7,38	1,53	1,23
		(0,215)	(0,678)	(0,803)	(0,66)	(0,744)	(0,007)*	(0,216)	(0,267)
	p=2,5%	4,80 %	3,20 %	3,33 %	3,10 %	3,28 %	4,00 %	3,43 %	3,30 %
		4,29	0,92	1,94	1,37	2,85	11,75	5,56	4,78
		(0,038)*	(0,336)	(0,164)	(0,241)	(0,092)	(0,001)*	(0,018)*	(0,029)*
	p=1%	3,20 %	2,20 %	1,73 %	1,80 %	2,00 %	2,20 %	1,89 %	1,85 %
		7,73	5,42	3,34	5,23	9,78	16,26	11,00	11,67
		(0,005)*	(0,02)*	(0,068)	(0,022)*	(0,002)*	(0)*	(0,001)*	(0,001)*
Multivariate	7,98	9,18	4,67	9,17	15,68	17,35	11,50	12,14	
	(0,046)*	(0,027)*	(0,197)	(0,027)*	(0,001)*	(0,001)*	(0,009)*	(0,007)*	

This table shows the results of univariate and multivariate convergence tests for the GARCH(1,1) VaR method. The sample periods are from 250 to 2000 days. Violation rate is defined as the ratio of violations to observed samples and is displayed as a percentage figure. The LR statistics with their corresponding p-values in parentheses are displayed for each violation rate. * indicates p-values that allow rejection of the null hypotheses at the 5% confidence level.

5.2.3 Risk Metrics (EWMA)



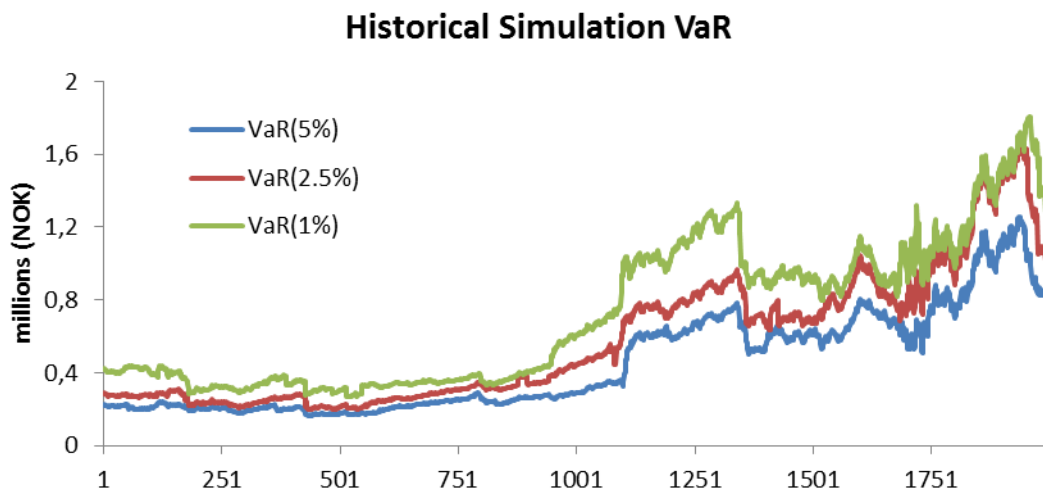
This figure shows the RiskMetrics VaRs of the portfolio for 2000 days and 1%, 2.5% and 5% confidence levels

Unconditional Coverage Tests (RiskMetrics)

Sample size (T)	250	500	750	1000	1250	1500	1750	2000	
Univariate	p=5%	7,20 %	5,00 %	5,60 %	5,30 %	5,28 %	7,07 %	6,06 %	5,75 %
		2,26	0,00	0,55	0,19	0,20	12,02	3,87	2,26
		(0,133)	(1)	(0,459)	(0,666)	(0,652)	(0,001)*	(0,049)*	(0,132)
	p=2,5%	5,60 %	3,60 %	4,00 %	3,80 %	4,00 %	4,80 %	4,11 %	3,75 %
		7,33	2,19	5,87	6,00	9,79	25,76	15,71	11,14
		(0,007)*	(0,139)	(0,015)*	(0,014)*	(0,002)*	(0)*	(0)*	(0,001)*
	p=1%	3,60 %	2,40 %	2,00 %	2,00 %	2,24 %	2,67 %	2,29 %	2,15 %
		10,23	7,11	5,87	7,83	14,36	28,89	21,43	20,10
		(0,001)*	(0,008)*	(0,015)*	(0,005)*	(0)*	(0)*	(0)*	(0)*
Multivariate	11,49	10,22	9,74	12,97	23,78	33,03	24,87	22,23	
	(0,009)*	(0,017)*	(0,021)*	(0,005)*	(0)*	(0)*	(0)*	(0)*	

This table shows the results of univariate and multivariate convergence tests for the RiskMetrics VaR method. The sample periods are from 250 to 2000 days. Violation rate is defined as the ratio of violations to observed samples and is displayed as a percentage figure. The LR statistics with their corresponding p-values in parentheses are displayed for each violation rate. * indicates p-values that allow rejection of the null hypotheses at the 5% confidence level.

5.3 Historical Simulation



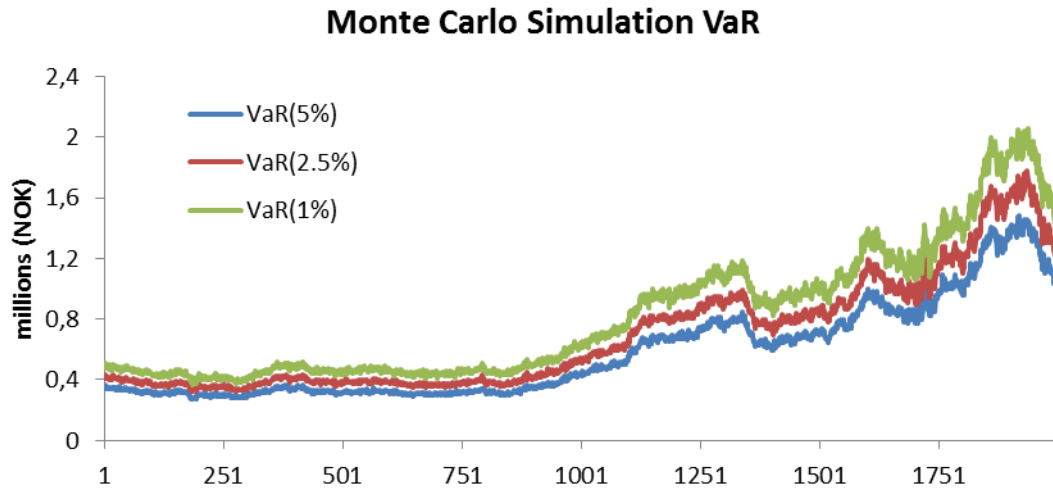
This figure shows the Historical Simulation VaRs of the portfolio for 2000 days and 1%, 2.5% and 5% confidence levels

Unconditional Coverage Tests (Historical Simulation)

Sample size (T)	250	500	750	1000	1250	1500	1750	2000	
Univariate	p=5%	7,20 %	4,80 %	5,33 %	5,20 %	5,60 %	7,93 %	6,80 %	6,00 %
		2,26	0,04	0,17	0,08	0,91	23,24	10,78	3,97
		(0,133)	(0,836)	(0,678)	(0,773)	(0,339)	(0)*	(0,001)*	(0,046)*
	p=2,5%	3,60 %	2,20 %	2,80 %	3,00 %	3,04 %	4,87 %	4,17 %	3,65 %
		1,09	0,19	0,27	0,96	1,40	27,12	16,75	9,52
		(0,295)	(0,661)	(0,606)	(0,326)	(0,237)	(0)*	(0)*	(0,002)*
	p=1%	0,80 %	0,80 %	0,80 %	1,40 %	1,52 %	2,53 %	2,17 %	1,90 %
		0,11	0,22	0,33	1,44	2,95	25,00	18,17	12,95
		(0,742)	(0,641)	(0,568)	(0,231)	(0,086)	(0)*	(0)*	(0)*
Multivariate	3,54	0,27	1,46	1,87	2,98	33,74	21,28	13,97	
	(0,316)	(0,965)	(0,69)	(0,601)	(0,395)	(0)*	(0)*	(0,003)*	

This table shows the results of univariate and multivariate convergence tests for the Historical Simulation VaR method. The sample periods are from 250 to 2000 days. Violation rate is defined as the ratio of violations to observed samples and is displayed as a percentage figure. The LR statistics with their corresponding p-values in parentheses are displayed for each violation rate. * indicates p-values that allow rejection of the null hypotheses at the 5% confidence level.

5.4 Monte Carlo Simulation



This figure shows the Monte Carlo Simulation VaRs of the portfolio for 2000 days and 1%, 2.5% and 5% confidence levels

Unconditional Coverage Tests (Monte Carlo Sim.)

Sample size (T)	250	500	750	1000	1250	1500	1750	2000
Univariate	p=5%							
	2,00 %	1,20 %	1,33 %	1,90 %	2,56 %	4,47 %	3,83 %	3,35 %
	6,07 (0,014)*	21,62 (0)*	29,61 (0)*	26,23 (0)*	18,93 (0)*	0,93 (0,335)	5,48 (0,019)*	12,91 (0)*
Univariate	p=2,5%							
	1,20 %	0,80 %	0,53 %	1,10 %	1,36 %	2,40 %	2,06 %	1,80 %
	2,14 (0,144)	8,03 (0,005)*	17,44 (0)*	10,14 (0,001)*	7,97 (0,005)*	0,06 (0,803)	1,50 (0,221)	4,45 (0,035)*
Univariate	p=1%							
	0,40 %	0,20 %	0,13 %	0,40 %	0,72 %	1,40 %	1,20 %	1,05 %
	1,18 (0,278)	4,81 (0,028)*	9,03 (0,003)*	4,71 (0,03)*	1,10 (0,295)	2,16 (0,142)	0,66 (0,415)	0,05 (0,824)
Multivariate	6,33 (0,097)	22,70 (0)*	30,41 (0)*	26,77 (0)*	20,22 (0)*	6,21 (0,102)	10,76 (0,013)*	18,19 (0)*

This table shows the results of univariate and multivariate convergence tests for the Monte Carlo Simulation VaR method. The sample periods are from 250 to 2000 days. Violation rate is defined as the ratio of violations to observed samples and is displayed as a percentage figure. The LR statistics with their corresponding p-values in parentheses are displayed for each violation rate. * indicates p-values that allow rejection of the null hypotheses at the 5% confidence level.

6. Discussion

6.1 Portfolio returns

The daily portfolio returns show clear signs of time varying volatility and volatility clustering, the clustering is especially clear between observations 1001 to 1251 and around observations 1501 and 1751. This roughly translates to the beginning and some of the most volatile points of the financial crisis.

The histogram of the continuously compounded portfolio returns seems fairly symmetrical, but doesn't really look normally distributed.

6.2 Variance Covariance Methods

Overall the VCV methods produce much better results for the Norwegian data than what they did for Taiwanese portfolio examined by Wu et al. (2012), the lack of none linear payout assets in the Norwegian data probably explains a lot of this difference.

The methods utilizing time varying volatility (GARCH and RiskMetrics), produce very close VaR estimates for all three significance levels, this in addition to the methods being rejected for the more extreme confidence levels seem to indicate that a normal distribution might be a bad assumption for the portfolio's return distribution, the portfolio's histogram seems to support this, a distribution with fatter tails might be more appropriate.

6.2.1 Historical Variance

The VaR estimates are slow to respond to changes in volatility; with the VaR estimates reaching their peaks well after the peaks in the variance have subsided.

The univariate test results seem to indicate that the method performs fairly well for 5% VaRs, with the null hypotheses only being rejected when the most volatile period is included in the sample. However for the 2.5% VaRs the null hypotheses is rejected for most sample sizes, except those containing the least volatile period. And for the most extreme VaR level (1%) the null hypotheses are rejected for most of the period. The decreasing effectiveness of the method is captured well by the multivariate test where the approach is rejected for most of the sample set.

The slow reaction to changes in volatility is one of the problems with historical variance, since each observation in the sample is given equal weight, risk will be underestimated when entering a period of higher volatility, and likewise risk will be overestimated when leaving a period of high volatility.

6.2.2 GARCH(1,1)

This method has the quickest and most extreme responses to changes in volatility, with sharp spikes in the VaR estimates that correspond to the large losses of the portfolio between 1001 to 1251 days, and around 1501 and 1751 days.

It does very well for the 5% VaRs and is only rejected for one sample size. It also does well with the 2.5% VaRs, but is rejected when the more volatile periods are included in the sample size. However the univariate test for the 1% VaRs and the multivariate test show that the approach isn't capturing the fatness of the return distribution's left tail.

Wu et al. (2012) had very little success with the GARCH(1,1) method, both testing frameworks reject the null hypotheses for all sample sizes and VaR significance levels. Their empirical violation rates are much higher than what should be observed, deviating with ten percentage points or more for many of the sample sizes. Compared to this, the method performs relatively well with the Norwegian data.

6.2.3 RiskMetrics (EWMA)

The method responds quickly to changes in portfolio volatility, we see corresponding spikes in the VaR estimates to the drastically increased variance between 1001 and 1251 days and right before and during day 1751. And generally higher VaR values during the financial crisis.

The univariate test shows that the method performs well at the 5% VaR significance level, it only underestimates risk during the most volatile period. However the approach is rejected for most sample sizes, when testing the more extreme VaR levels, it seems the method is unable to capture the actual thickness of the left tail in the return distribution. This is seen clearly from the multivariate test, where the approach is rejected for all sample sizes.

Wu et al. (2012) has more success with this approach than the GARCH(1,1), but the method is still rejected for most sample sizes and VaR significance levels, however the deviation between the expected violation rates and the observed is much smaller than for the GARCH(1,1) estimates. The Norwegian data show somewhat of the opposite effect, this approach produces slightly worse results than the GARCH(1,1) method, however the results are still better.

6.3 Historical Simulation

The VaR estimates seem to move similarly to those found under the Historical Variance, but with generally higher values especially during the periods with high return volatility. Further the more extreme VaR significance levels, usually result in more extreme VaR estimates; it would seem this model is better at capturing the true thickness of the left tail.

This is the only method with consistent univariate test results; it does very well for all three univariate tests and the multivariate test, however, the null hypotheses is rejected when the most volatile periods are included in the test sample. During the most volatile periods the approach seems to first underestimate risk when entering the turbulent period, and then overestimates it when volatility returns to normal. This is one of the problems with historical

simulation, since it can only simulate price movements within its sample window, it will be slower to react than the methods based on time varying volatility (Jorion, 2007).

Wu et al. (2012) achieved very good results with this approach, keeping the null hypotheses for all sample sizes and test specifications, all three sets of univariate tests and under the multivariate test. As such this is the only method that performs better with the Taiwanese data than with the Norwegian.

6.4 Monte Carlo Simulation

The VaR estimates move similarly to the Historical Simulation and Historical Variance estimates, but are generally higher; the different VaR significance levels are also more clustered than with Historical Simulation.

The method has by far the worst results with both univariate and multivariate tests and seems to systematically overestimate risk for all the sample sizes, only when the most volatile period is included is it accepted by the tests.

Several possible explanations for its bad performance, its accuracy relies heavily on the models used to estimate the portfolio's risk factors. In this paper it is assumed that all three assets have lognormal price distributions, from the results it might seem that the lognormal price formula is undervaluing the next period's asset price and thereby systematically generating undervalued hypothetical portfolio values. Using a more realistic volatility estimate than the historical variance might also improve its performance.

Monte Carlo simulation VaR estimates for the Taiwanese data, consistently overestimated the portfolio risk and the empirical violation rates were all zero percentage for all significance levels and all sample sizes. Because of this the null hypotheses were rejected for all but the shortest sample size and the more extreme VaR significance levels (Wu et al., 2012). Even though the empirical Norwegian results have a statistical significant deviation from the desired significance levels, the method still performs better with the Norwegian data set.

6.5 Overall

The univariate tests give inconsistent results for all estimation methods, except historical simulation, this clearly indicates the problem with just using Kupiec's test, and shows the value of the multivariate test. If only the 5% VaRs were tested, all the Variance Covariance methods would look fairly good, however including more extreme significance levels and using the multivariate test, shows their shortcomings for this data set.

None of the approaches are able to properly handle the massive increase in volatility during the financial crisis, even the Monte Carlo Simulation with its overestimation of risk, incurs more violations between 1250 and 1750 days than should be expected with the given VaR significance level.

6.6 Weaknesses And Possible Improvements

For the Historical Simulation the sample window was only 250 daily return observations, this is in the lower end of what most banks use, which typically is between 250 to 750 daily observations (Jorion, 2007). Increasing the window would have kept more of the extreme peaks in the sample; this in turn would probably have decreased the violation rates in the interval between the last 750 estimates, and possibly improved the performance of the method.

Both testing frameworks are based on unconditional testing, this means it is irrelevant for the test results if the violations are spread evenly over the sample period or clustered. However if a model shows clear signs of violation clustering it should be rejected (Jorion, 2007), so further testing with conditional coverage models would probably be beneficial.

The foreign exchange element is only exposed to exchange rate risk, and no interest rate risk, it seems unlikely that a real investor would hold a currencies position without getting some sort of interest on it. Including interest rate would have made the portfolio more realistic and made it more volatile, which in turn might have affected the results.

All three assets have linear payouts, a real portfolio is likely to include derivatives where the payouts are none linear, similar to the Taiwanese portfolio, further testing where Norwegian derivatives with none linear payouts were included, would help determine if the differences found between the papers, stem from difference between Norwegian and Taiwanese assets or if they come as a result of the portfolio's composition.

For the Variance Covariance estimates, absolute VaR was used; this means the expected mean return of the portfolio wasn't factored in. Whereas with the Historical Simulation and Monte Carlo Simulation the expected mean return was be factored into the generated hypothetical portfolios and in turn included in the VaR estimates. The effect of this discrepancy is not likely to be significant though because of the short horizon.

The risk factor of the bond element was based on its return series, an alternative and perhaps better way would be to measure the bond's risk using its duration. This is a possible area of future testing.

Basel II requires banks to keep three times the 10 day 1% VaR estimate as reserve capital (Jorion, 2007); it could be interesting to test if this requirement was violated during the estimation period.

7. Conclusion

The VCV methods and especially the once based on time varying volatility performed much better for the Norwegian assets, than for the Taiwanese, but they were still rejected under the multivariate framework for most of the estimation period. And they generally experienced very inconsistent results for the univariate test, using a return distribution with fatter tails might alleviate this somewhat.

Out of the five approaches used, the Monte Carlo Simulation was the least accurate, however it still performed better on the Norwegian data than on the Taiwanese, but it exhibited the same systematic overestimation of the portfolio's risk.

Overall I find that the best method is Historical Simulation, this is consistent with what Wu et al. (2012) found in their paper. The method is not perfect and is rejected by both the univariate and multivariate tests, when the estimates from the most volatile periods are included in the tested sample. But overall it is the model that performs best, and increasing the sampling window used to create the historical scenarios, might increase its accuracy.

8. References

Benninga, Simon. 2008. *Financial Modeling*. 3rd ed. Cambridge, Mass.: The MIT Press.

Harris, Richard D. F., Evarist Stoja, and Jon Tucker. 2007. *A Simplified Approach to Modeling the Co-Movement Of Asset Returns*. *The Journal of Futures Markets*, Volume 27, Issue 6, p. 575-598

Hull, John C. 2012. *Options, Futures, And Other Derivatives*. 8th ed. Pearson.

Jorion, Philippe. 2007. *Value At Risk: The New Benchmark for Managing Financial Risk*. 3rd ed. McGraw-Hill.

McDonald, Robert L. 2006. *Derivatives Markets*. 3rd ed. Pearson.

Pérignon, Christophe, and Daniel R. Smith. 2008. *A New Approach to Comparing VaR Estimation Methods*. SSRN Working papers series 20. April 2007.

<http://papers.ssrn.com/sol3/papers.cfm?abstract_id=981207> (6. feb 2014)

Wu, Po-Cheng, Cheng-Kun Kuo, and Chih-Wei Lee. 2012. *Evaluation Of Multi-Asset Value At Risk: Evidence From Taiwan*. *Global Journal Of Business Research* (6):23-34.

9. Appendix

9.1 Estimated GARCH(1,1) Functions

$$h_t = \frac{\alpha_0}{10^5} + \alpha_1 r_{t-1}^2 + \beta h_{t-1}$$

	OBX	ST5X	NOK/USD
α_0	0,34	0,01	0,03
α_1	0,10	0,07	0,03
β	0,89	0,92	0,96

This table shows the estimated GARCH(1,1) values for the variance of the three return series.

	OBX+ST5X	OBX-ST5X	OBX+N/U	OBX-N/U	ST5X+N/U	ST5X-N/U
α_0	0,35	0,36	0,68	0,36	0,04	0,03
α_1	0,10	0,10	0,09	0,07	0,04	0,03
β	0,89	0,89	0,88	0,92	0,96	0,96

This table shows the estimated GARCH(1,1) values for the return series constructed under the framework for modeling assets co-movement developed by Harris et al. (2007).

9.2 VBA Monte Carlo Estimation

```

Sub monte_carlo()
Dim avg, vol, norm, time, e, i, j
Dim obx_t, st5x_t, nok_usd_t
Dim obx0, st5x0, nok_usd0

time = 1
With Range("monte_carlo_portfolio")
For j = 1 To 2000
    obx0 = .Cells(j, 0).Value
    st5x0 = .Cells(j, 1).Value
    nok_usd0 = .Cells(j, 2).Value

    For i = 1 To 5000
        norm = Application.NormInv(Rnd, 0, 1)

        avg = Range("obx_250d_return").Cells(j, 1).Value / 250
        vol = Sqr(Range("obx_hist_var").Cells(j, 1).Value)
        e = Exp((avg - 0.5 * (vol * vol)) + vol * Sqr(time) * norm)
        obx_t = obx0 * e

        avg = Range("st5x_250d_return").Cells(j, 1).Value / 250
        vol = Sqr(Range("st5x_hist_var").Cells(j, 1).Value)
        e = Exp((avg - 0.5 * (vol * vol)) * time + vol * Sqr(time) * norm)
        st5x_t = st5x0 * e

        avg = Range("nok_usd_250d_return").Cells(j, 1).Value / 250
        vol = Sqr(Range("nok_usd_hist_var").Cells(j, 1).Value)
        e = Exp((avg - 0.5 * (vol * vol)) * time + vol * Sqr(time) * norm)
        nok_usd_t = nok_usd0 * e

        Range("output").Cells(j, i).Value = obx_t + st5x_t + nok_usd_t
    Next
Next
End With
End Sub

```

This is the VBA Subroutine used to calculate the hypothetical Monte Carlo Simulation portfolio values.