



Master Thesis, Spring 2012

Lund University
School of Economics and Management

The performance of time-varying volatility and regime switching models in estimating Value-at-Risk

Authors:

Alina Birtoiu 880221
Florin Dragu 861127

Supervisor:

Anders Vilhelmsson

Abstract

Markov Regime-Switching GARCH (MRS-GARCH) models have been gaining popularity due to their ability to account for shifts volatility regimes that tend to characterize returns series. Previous empirical studies have shown that this capacity to capture the volatility dynamics leads to a superior forecasting power of the MRS models. We investigate the performance of these models in quantifying and managing market risk for financial institutions. To this purpose, Klaasen's (2002) MRS-GARCH model both under the normal and t-distribution assumptions, was applied in order to calculate VaR for four real trading portfolios, obtained from American and European banks, as well as a mimicking portfolio built with constant weights. The results concluded that the MRS model has not had a consistent performance over the five data series, and, in many cases was outperformed by the more simplistic GARCH models. The model does perform better in terms of exceptions compared to the in-house VaR models employed by the banks analyzed. However, we must keep in mind that the MRS-GARCH specification can be sensitive to the length of the rolling window used, and a larger in-sample period could have added more precision to the variance forecasts.

Aknowledgements

We would like express our utmost gratitude to our thesis supervisor, Anders Vilhelmsson for his invaluable guidance, advice and enthusiasm throughout the development of this thesis.

Table of Contents

1.	Introduction and problem discussion.....	6
1.1.	Purpose.....	8
1.2.	Delimitations	8
1.3.	Structure of the paper	9
2.	Theoretical background.....	10
2.1.	Value-at-Risk.....	10
2.1.1.	Parametric VaR.....	11
2.1.2.	Non-parametric and semi-parametric methods	12
2.2.	Backtesting	14
2.2.1.	Christoffersen’s approach to backtesting	15
2.2.2.	GMM-duration backtesting.....	16
2.2.3.	Risk Map in backtesting.....	18
2.2.	Previous research	19
3.	Methodology	23
3.1.	Estimating volatility using standard GARCH models	23
4.	Data	29
5.	Empirical Results	34
5.1.	Separate portfolio analysis.....	36
5.1.1.	Bank of America	37
5.1.2.	Mimicking portfolio	40
5.1.3.	Danske Bank	42
5.1.4.	Deutsche Bank.....	44
5.1.5.	Swedbank	47
5.2.	Model analysis.....	49
5.2.1.	Parametric models	49
5.2.2.	Non-parametric and semi-parametric models.....	50
6.	Conclusions.....	54
7.	Future research	56
	References.....	57
	Appendix.....	61
	Appendix 1. Portfolio composition and statistic tests.....	61
	Appendix 2. Coefficient estimates.....	63

Appendix 3. VaR forecasts – graphical representation	68
Appendix 4. Risk Map graphical representation	76
Appendix 5. Average VaR	77
Appendix 6: GMM-duration backtesting results	81
Appendix 7. Selective Matlab Code.....	86

1. Introduction and problem discussion

Value-at-Risk is the most common quantitative measure for assessing market risk. A catalyst for its popularity was its introduction in the regulatory frameworks Basel I and II as a standardized measure for market risk and its employment in determining the capital requirements for banks' trading portfolios. The Basel Committee does not impose a standard VaR estimation model that banks have to apply, but allows each institution to calibrate its individual VaR model to its particular operations and portfolios. The recent financial crisis brought into the spotlight the way banks evaluate the risks in their trading portfolios and whether the models used by these institutions are the most appropriate ones. These are therefore the main reasons behind the current paper's focus on investigating the most accurate VaR models in the context of their employment in financial institutions.

The predictive accuracy of VaR models is an ongoing issue and numerous studies that have investigated this topic have been published over the past years. However, very few of these papers use real bank data to test the performance of Value-at-Risk models and to the best of our knowledge, no studies on VaR estimation based on regime switching time-varying volatility models fitted to real bank data have been published up to this point.

Therefore, the present study adds to the current research literature through the use of a regime switching GARCH-type model (MRS-GARCH) in estimating Value-at-Risk on real profit/loss data, extracted from four European and American banks as well as on a synthetic portfolio that mimics a prospective bank trading portfolio. A series of GARCH models without the regime-switching framework that have yielded good empirical results in terms of performance in forecasting the volatility of asset returns are also employed and used as a benchmark for the above volatility models. Moreover, the volatilities forecasted through the GARCH and MRS GARCH models are also used to compute semi-parametric Value-at-Risk measures by means of the volatility weighted historical simulation method along with the option implied volatility that is employed to compute parametric VaR through the HS-VIX method.

The efficiency of all these models is assessed through a series of backtesting methods. In addition to the traditional Christoffersen (1998) backtesting frameworks we also employ the Candelon et al. (2011) GMM-duration procedure and the Risk Map methodology proposed by Colletaz et al. (2011). Both the Risk Map and the GMM-duration backtest represent novel tools for validating VaR models, which provide the advantage of taking into account the

magnitude of the VaR exceptions as well as their frequency and the duration between violations.

There has been a large amount of controversy underlying the accuracy of VaR measures since its introduction, the model, receiving much criticism over the years. Its simplistic interpretation and implementation have been considered both an advantage and a drawback. Several empirical studies have questioned the use of VaR as a tool for understanding and managing market risk, thus shedding doubt on the entire Basel regulatory framework. Yet, despite the fact that alternative risk measures such as Expected Shortfall have been developed and are believed to be more efficient in determining market risk, VaR managed to stand tall in its position as the most widely used measure for market risk.

The initial RiskMetrics VaR model introduced by J.P. Morgan in 1994 was based on the assumption that asset returns follow a conditional normal distribution with zero mean and a variance that is an exponentially weighted moving average of historical squared returns. This assumption is often breached in an empirical context. Firstly, the distribution of asset returns commonly exhibits characteristics such as heavy tails (leptokurtosis) and skewness. Therefore, by using the normal distribution in determining VaR, and thus not taking into account these phenomena can lead to biased estimates. This motivates the use of the Student-t distribution which is able capture extreme events in VaR modeling. Secondly, the volatility dynamics of asset returns is usually characterized by volatility clusters and leverage effects. While the exponentially weighted moving average assumption manages to capture the time-varying variance to some extent, its performance is not always satisfactory. GARCH-type models have been used as a solution for capturing these particularities of time-varying volatility, since they employ an autoregressive structure in the conditional variance. However, an underlying assumption for GARCH models is that of diffusion processes in price dynamics. This leads to a persistence of individual shocks on volatility, which Lamoureux and Lastrapes (1990) attribute to the possible existence of structural breaks in the variance process. They prove that shifts in the unconditional variance can lead to a mis-estimation of the GARCH parameters in a way that entails high volatility persistence. An answer to this problem would be to incorporate GARCH-models in a regime-switching framework that allows accounting for the existence of two or more different regimes in the volatility. Such models were first developed by Cai (1994) and Hamilton and Susmel (1994) under an ARCH specification, following the introduction of Markov Regime Switching processes by

Hamilton(1989). Gray (1996) and Klaasen (2002) took these models a step further by combining a GARCH specification with a Markov Regime Switching process

Several empirical studies have been published that focus on testing regime switching GARCH models and GARCH models with a jump component on equity markets (Chan and Maheu (2002), Kim et al. (2003), Maheu and McCurdy (2004), Gau and Tang (2004), Marcucci (2005), Sajjad et al. (2008), and Nyberg and Wilhelmsson (2009), Liu (2011)), exchange rates (Klaasen(2002), Haas(2004)) or commodities (Chan and Young, 2009). The results derived in these papers show a better accuracy of MRS models compared to simple GARCH specifications which motivates our choice to assess the performance and possibility of employing such estimation models on trading portfolios of financial institutions.

1.1. Purpose

The purpose of this thesis is to evaluate the performance of regime-switching and time varying volatility models in the estimation of Value-at-Risk using real data from four commercial banks and a synthetic portfolio that mimics the trading portfolio of Bank of America.

1.2. Delimitations

Since the time limit set for this study and the availability to the data is rather restrictive, a number of delimitations are necessary.

Firstly, the access to the data was restrained due to the fact that financial institutions choose to selectively disclose their profit/loss data, VaR measures and portfolio composition on a daily basis to the wide public. Therefore the access to the high frequency data needed for the study is limited, as well as time consuming to collect. We were able to compile daily profit/loss data series from four commercial banks, using a graph data digitizing software applied on P/L graphs disclosed by four banks in their annual reports or their risk management reports. However, the data was available only starting with 2007 for three of the banks, which limits our time series to a span of 5 years of daily observations. Another consequence of data disclosure policies is that banks do not make public the actual composition of their portfolios. Hence, it is not possible to analyze the effect of individual risk factors on the overall returns of the portfolios. It would have been interesting to compute VaRs for each individual risk factor and aggregate these measures according to the correlations between risk factors in order to determine the VaR for the entire portfolio. This is the common practice for the banks we analyzed in computing Value-at-Risk through their in-

house models. The approach could have permitted to analyze the individual dynamics of the portfolio components in terms of risk which could shed light on the overall behavior of the market risk for the trading portfolio.

Secondly due to the time consuming computational burden required for estimating the regime-switching models as well as the limited availability of data, we only used a rolling window of 252 observations. An alternative would have been to also estimate the models on a longer rolling window, which could have lead to more stable forecasts and could have been able to better capture the regime switches in the volatility dynamics. A comparison of these two different estimates derived from two different rolling windows would have been most useful in the interpretation of the performance of the VaR models.

1.3. Structure of the paper

The paper is structured as follows. In the second part we present the theoretical background underlying the VaR measure. Section 3 presents the methodology used to conduct the study, and describes the volatility models used. The data is examined in section 4 and the empirical results are assessed in section 5. Finally, section 6 consists of the conclusion and we suggest further research venues in section 7.

2. Theoretical background

2.1. Value-at-Risk

In the past two decades Value-at-Risk (VaR) emerged as the primary tool to evaluate the market risk to which financial institutions are exposed. But what really is Value-at-Risk?

VaR represents the maximum amount of money that can be lost within a specified time horizon and with a given probability. As an example, if the one-day VaR has a value of \$1 million at a 99% confidence level then there is only a 1% probability that the bank will incur a loss larger than \$1 million over one day of trading.

A mathematical formulation of the above can be:

$$\Pr(R \leq -VaR) = 1 - \alpha \quad (1),$$

where α represents the confidence level for the VaR calculations, while R expresses the Profit or Loss (hereafter P/L) in the bank's portfolio.

The VaR estimates are usually compared to the bank's available capital to make sure that it will not suffer a liquidity shortage as a result of an extreme event, thus ensuring that the losses generated can be covered without risking the bank's entire operations. Thus, it can be stated that VaR is a risk measure whose focus lies clearly on the downside risk and more specifically on those events that have an extremely-low occurrence probability.

Dowd (2005) enumerates the following advantages in employing VaR as a risk measure. Firstly, it is a simple to grasp concept as it gives a maximum amount that can be lost with a certain probability. Secondly, it can be applied both at a firm and at a subsidiary level and management will be able to better track the risk targets. Thirdly, VaR has become the risk measure financial regulators, such as the Basel committee, have based their capital requirements framework. Last but not least, VaR takes into consideration all risk factors, while other risk measures only analyze risk one factor at a time.

As all risk measures VaR comes with its own limitations, the most important ones being listed by Damodaran (2007). A first critique would be the need in some models to make assumptions about the distribution of the returns, which if proved to be incorrect would result in miscalculated VaR estimates. Secondly, VaR is dependent upon the time period over which the historical data is collected. If the sample was gathered during a relatively calm period, from an economic point of view, then the calculated VaR would be smaller and hence it

would understate the risk exposure. Contrarily, if the historical data spans over a tumultuous time then VaR estimates will be too high, causing the financial institution to set aside unnecessary regulatory capital. Thirdly, VaR is criticized because it is mainly a short-term risk measure, as it is usually computed over a day, a week or ten days, but banks are also exposed to long-term risks. Fourthly, VaR is subject to manipulations from managers and thus creates an opening to agency problems. Finally, the fact that VaR only focuses on market risk might cause banks to disregard the other risks they face.

As a consequence, VaR should not be seen as the Holy Grail in risk management, but as a tool or a signal that allows a bank manager to have an apprehension of the risk exposures inherent in the trading activities the financial institution is engaged in. VaR should be placed in a more comprehensive context that includes both internal factors, specific to the particular entity in question, as well as external factors which influence the economic environment of the bank. Therefore, decisions need not be based only on VaR but on an entire set of determinants.

Joe Nocera wrote in a 2009 New York Times article: “Nothing ever happens until it happens for the first time. This didn’t mean you couldn’t use risk models to sniff out risks. You just had to know that there were risks they didn’t sniff out – and be ever vigilant for the dragons. When Wall Street stopped looking for dragons, nothing was going to save it. Not even VaR.”

2.1.1. Parametric VaR

The parametric method implies estimating VaR by using a distributional assumption for the returns and a series of mean and volatility estimates.

Normal distribution. The normal distribution is the most popular assumption for the distribution of returns. Its attractiveness lies in the fact that it is characterized by the first two moments, the mean and the variance (or the standard deviation). Parametric Value-at-Risk can be easily computed as an expression of these two parameters (Dowd (2005)):

$$VaR = -\mu - t_{\alpha}\sigma \quad (2)$$

where μ is the mean, σ is the standard deviation and t_{α} is the corresponding critical value to the assumed confidence level α .

The probability density function of the normal distribution is expressed as $\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}$.

Despite the fact that many models continue to be employed under the assumption of a normal distribution, financial data generally exhibits positive skewness and leptokurtosis which can lead to a bias in estimating of risk (Angelidis et al. (2004)). Moreover, using the assumption of a normal distribution permits a return to take on any value, thus increasing the probability of incurring losses so large that we stand to lose more than what we have invested (Dowd (2005)).

Student-t distribution. The Student-t distribution can accommodate a third moment, i.e. excess kurtosis, in addition to the mean and standard deviation. Value-at-Risk under the assumption of a student-t distribution can be expressed as:

$$VaR(\alpha) = -\mu - t(\alpha, \nu) \sqrt{\frac{\nu-2}{\nu}} \sigma \quad (3)$$

VaR based on student-t distribution is a function of three parameters: the mean (μ), the standard deviation (σ) and the degrees of freedom (ν), while t represents the corresponding statistic given a confidence level (α) and the degree of freedom (ν).

The probability density function of the Student-t distribution is: $\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} (1 + \frac{x^2}{\nu})^{-\frac{\nu+1}{2}}$.

If the number of degrees of freedom is very large, the Student-t distribution is a generalization of the normal distribution, but when ν is finite it has the ability to account for higher than normal kurtosis (Dowd(2005)). Furthermore, the Student-t distribution will provide better predictive densities (Danielsson (1997)). Nonetheless, among the issues that the use of this distribution raises we can identify the possibility of producing fallacious high risk estimates can be pointed out as well as its instability, which makes that VaR forecasts over long periods of time less reliable.

2.1.2. Non-parametric and semi-parametric methods

There are also a series of non-parametric methods to estimate Value-at-Risk, whose main strength lies in the fact that they do not require any distributional assumptions since the forecasts are based on previous returns. The most popular non-parametric method is the historical simulation, which most financial institutions currently use as a basis for their in-house VaR models, according to Perignon and Smith (2006).

The historical simulation (HS) VaR is determined by using a sample quantile estimate based on historical returns (Kuester et al.(2005)):

$$VaR_t = q(\alpha) \quad (4)$$

A major drawback of this method is that it does not capture volatility dynamics and consequently can lead to clustering in VaR violations (Christoffersen et al.(2008)). A solution to this would be using time-varying volatilities, such as those derived from GARCH models (Berkowitz (2006), Perignon and Smith (2006)).

Volatility Weighted Historical Simulation (VWHS). This semi-parametric approach entails a rescaling of the returns as to account for the recent changes in volatility.

$$y_t^* = \frac{y_t}{\sigma_t} \sigma_T \quad (5)$$

Therefore, the returns are scaled upwards during times of high volatility and downwards during periods of low volatility, which leads to risk estimates that display an accurate sensitivity to current volatility estimates (Dowd(2005)).

The present paper employs the VWHS method using volatility forecasts derived through a series of GARCH, EGARCH, GJR and Markov Regime Switching GARCH models, under both a normal and Student-t distributional assumption for the innovations.

HS-VIX. Another possibility to account for volatility dynamics is to use the implied volatility give by a volatility index¹. The HS-VIX model that was introduced by Nossman and Wilhelmsson (2011) is a filtered volatility weighted historical simulation method that uses the implied volatility given by the VIX index.

The portfolio returns are rescaled using the option implied volatility:

$$y_t^* = \frac{y_t}{VIX_{t-1}} VIX_{T-1} \quad (6)$$

The method is completely non-parametrical, forward-looking and is based on what is supposed to be the best variance forecast that exists.

The model was tested on the S&P500 index and its performance was proven to be superior compared to the basic Historical Simulation and the HS-GARCH model.²

In the present study the HS-VIX model is applied to the 5 trading portfolios. However, three different volatility indexes are used, according to the main market each bank operates in.

¹ Volatility indices measure the implied volatility for a basket of put and call options related to a specific index. For instance, the VIX index measures the implied volatility for a basket of out-of-the-money put and call options for the S&P 500 market index. More to the point, the VIX measures the expected 30-day volatility for the S&P 500. (www.cboe.com)

² For a more detailed description of the historical simulation and HS-VIX methods see Birtoiu and Dragu (2011)

The reason behind this approach is that each portfolio is more likely to be better correlated to a volatility index from its own market than the VIX index, which is mostly representative for the American market. Thus, the VIX index is used for the portfolio of Bank of America, VDAX for Deutsche Bank and VSTOXX for Danske Bank and Swedbank. The synthetic portfolio is tested against all three indices.

2.2. Backtesting

No matter how simple or ingenuous a VaR estimation model is its accuracy and overall performance need to be evaluated. In this paper several backtesting techniques are used to this purpose, namely the Christoffersen, GMM-duration and the Risk Map approaches. But what is backtesting more exactly and why is it needed?

J.P. Morgan Chase (2009) defines backtesting as a technique employed to establish the efficacy of a VaR model, i.e. to determine how many of the actual losses recorded by a financial institution exceeded those predicted by the model at the end of the selected time horizon (in this case, one day, but this horizon can also span over 10 days, 1 month, 1 year etc). Backtesting is needed as a bank's capital requirements will be computed based on the maximum loss predictions given by these models. Therefore it is extremely important for these estimates to be accurate, as it will otherwise result in an overestimation or underestimation of capital. An overestimation will incur an opportunity cost as these funds are set aside to cover risks that do not actually exist instead of being used to profit generating activities, whereas an underestimation would cause the bank to struggle since it would have insufficient funds to cover the resulted losses.

Backtesting has evolved in time along with the Value-at-Risk models. The various backtesting techniques are based on different concepts, some underlying the frequency of tail losses (Kupiec (1995) or Christoffersen (1998)) while others lean on their magnitude (Colletaz et al. (2011)). Furthermore, other methods are founded on multivariate Portmanteau test statistic (Hurlin and Tokpavi (2006)) or tests based on GMM duration (Candelon et al(2011)). Moreover, Engle and Manganelli (2004) propose a dynamic quantile backtesting procedure.

All these techniques have their own advantages and disadvantages, but since our subsequent inquiry will hinge on the Christoffersen, GMM-duration and Risk Map methods, some insights with respect to these tests will be presented in the subsections below.

The motivation for choosing these particular backtesting methods is given by the fact that they cover both the occurrence frequency of tail losses as well as the duration between exceptions. Moreover, by making use of the Risk Map it can be accounted not only for the number of violations but also for their magnitude, which means it accounts for those possible excessive losses that can cripple a bank's activity. Additionally, the GMM approach allows for developing a conditional coverage test in a duration-based framework and proved to have better power for sample sizes banks usually use when performing backtesting.

2.2.1. Christoffersen's approach to backtesting

This approach is a statistical method based on the frequency of tail losses, but as opposed to the Kupiec (1995) test, Christoffersen (1998) develops a conditional coverage test that accounts for the fact that the violations have the tendency to cluster over time if the model is misspecified.

The basic idea behind this method consists of separating the test hypotheses. Whereas the standard frequency test has the null hypothesis that the model generates a correct frequency of violations and that these violations are independent one from the other (Dowd(2005)), the Christoffersen approach breaks-down this hypothesis into two sub-hypotheses: one stipulating the correct frequency of exceptions, while the other assumes that violations are independent.

To test these hypotheses, firstly, Christoffersen and Pelletier (2004) define the following indicator variable:

$$I_t(\alpha') = \begin{cases} 1, & \text{if } R_t < -VaR_{t,\alpha'} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where R_t represents the observed loss at time t , and VaR_t is the predicted loss given by the model, at the same moment. This counts all the actual losses that were in excess to the maximum possible loss as predicted by the VaR-model.

The first test statistic, corresponding to the null hypothesis of generating the correct frequency of violations, is the same as the one used for the Kupiec (1995) test.

$$LR_{UC} = -2\ln[p^x(1-p)^{n-x}] + 2\ln\left[\left(\frac{x}{n}\right)^x \left(1-\frac{x}{n}\right)^{n-x}\right] \quad (8)$$

To account for the violations' clustering the independence test is then conducted by means of the following test-statistic (Christoffersen (1998)):

$$LR_{ind} = -2\ln\left\{\left(1 - \frac{x}{n}\right)^{n_{00}+n_{10}} \left(\frac{x}{n}\right)^{n_{01}+n_{11}}\right\} + 2\ln\{(1 - \pi_0)^{n_{00}}\pi_0^{n_{01}}(1 - \pi_1)^{n_{10}}\pi_1^{n_{11}}\} \quad (9)$$

where n_{ij} represents the number of observations in which state j occurred in a certain day while being in state i in the previous day, $\pi_0 = \frac{n_{01}}{n_{00}+n_{01}}$ (10), and $\pi_1 = \frac{n_{11}}{n_{10}+n_{11}}$ (11).

By combining the two test statistics yields a joint test of coverage and independence, thus a test of conditional coverage, as follows:

$$LR_{CC} = LR_{UC} + LR_{ind} \quad (12)$$

Both the unconditional coverage and the independence test statistics follow a chi-squared distribution with one degree of freedom, while the conditional coverage one it is also chi-squared distributed but has two degrees of freedom.

Applying this method allows not only to check if the model is appropriate or not, but also to identify the reasons why the model might not be appropriate (if this should be the case).

2.2.2.GMM-duration backtesting

Candelon et al. (2011) propose a new backtesting procedure that can be used to assess the precision of a VaR model forecasts. The basic idea behind this approach is that the duration-based test of Christoffersen and Pelletier (2004) can be applied in a GMM framework. This brings about several advantages that the previous duration tests did not have. Firstly, it allows for separate testing of the conditional coverage (CC), independence (IND) and unconditional coverage (UC) hypotheses. Secondly, it does not require a distributional assumption and it is easy to apply. Thirdly, the power of this backtest is larger than that of other validating procedures, performing very well even in realistic sample sizes (Candelon et al.(2011)).

As mentioned before the test relies on the duration between two sequential VaR exceptions. This can be expressed mathematically as:

$$d_i = t_i - t_{i-1} \quad (13)$$

where t_i signifies the time when the i th exception occurred.

Then, based on this duration and the confidence level α , the authors construct the following polynomial that is associated with a geometric distribution:

$$M_{j+1}(d, \alpha) = \frac{(1-\alpha)(2j+1)+\alpha(j-d+1)}{(j+1)\sqrt{1-\alpha}} M_j(d, \alpha) - \frac{j}{j+1} M_{j-1}(d, \alpha)^3 \quad (14)$$

Next, these polynomials are used in forming the moment conditions:

$$E[M_j(d, \alpha)] = 0 \quad (15)$$

The moment conditions will then be materialized into the null hypotheses of conditional coverage, unconditional coverage, and independence:

$$H_{0,CC}: E[M_j(d_i, \alpha)] = 0, \text{ with } j = 1, 2, \dots, p, \quad (16)$$

where p denotes the number of moment conditions.

$$H_{0,UC}: E[M_1(d_i, \alpha)] = 0 \quad (17)$$

$$H_{0,IND}: E[M_j(d_i, \beta)] = 0, \text{ with } j = 1, 2, \dots, p, \quad (18)$$

where β represents the success probability.

In order to determine the validity of these hypotheses or whether they can be rejected or not, Candelon et al. (2011) resort to test statistics similar to the J -statistic:

$$J_{CC}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M(d_i, \alpha) \right)^T \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M(d_i, \alpha) \right) \sim \chi^2(p) \quad (19)$$

$$J_{UC}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M_1(d_i, \alpha) \right)^2 \sim \chi^2(1) \quad (20)$$

$$J_{IND}(p) = \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M(d_i, \beta) \right)^T \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N M(d_i, \beta) \right) \sim \chi^2(p), \quad (21)$$

where N denotes the number of violations.

A rejection of the null hypotheses of correct conditional coverage, independence and unconditional coverage will lead to a rejection of the model as the violation sequence is characterized by the dependence between the probability of having an exception at the current time and the duration from the last violation.

³ $M_{-1}(d, \alpha) = 0$ and $M_0(d, \alpha) = 1$

2.2.3. Risk Map in backtesting

The techniques presented above only account either for the number of exceptions or their number and independence, but both are inefficient in ascertaining the magnitude of the losses that can occur in a bank's trading portfolio. To meet this shortcoming, Colletaz, Hurlin and Perignon (2011) developed a new procedure, called the Risk Map, which is based on the number of VaR exceptions as well as their amplitude.

This approach is constructed on the concept of "super exception", defined as a loss greater than an estimate of Value-at-Risk calculated at a very large confidence interval (i.e. 99.8%). The hit variable in this case takes the following form:

$$I_t(99.8\%) = \begin{cases} 1, & \text{if } R_t < -VaR_{t,99.8\%} \\ 0, & \text{otherwise} \end{cases} \quad (22)$$

Both $VaR_{99\%}$ and $VaR_{99.8\%}$ sequences are then backtested using a standard Kupiec test, using the following likelihood ratio test statistic:

$$LR_{UC}(\alpha) = -2 \ln[(1 - \alpha)^{T-N} \alpha^N] + 2 \ln \left[\left(1 - \frac{N}{T}\right)^{T-N} \left(\frac{N}{T}\right)^N \right] \quad (23)$$

where α represents the occurrence probability of a violation (and equals $1 - \text{confidence level}$), N represents the number of exceptions whereas T equals the total number of VaR forecasts. The same formula as the one above will also be used to validate the super exceptions, only that in this case $\alpha' = 99.8\%$ and N is replaced by N' which represents the number of super exceptions.

The results of these two unconditional coverage tests are then presented graphically, by placing the rejection and non-rejection zones of both $LR_{UC}(\alpha)$ and $LR_{UC}(\alpha')$ on a map. This will yield a figure with four quadrants, each corresponding to either a non-rejection, or a rejection of both tests, and two areas where one is rejected and the other is not.

The authors further argue that only employing these unconditional coverage tests independently can result in a rejection of a valid model and therefore they also suggest the use of a joint test of the number of exceptions and super exceptions, i.e. a multivariate unconditional coverage test, whose test statistic is given by the formula below:

$$LR_{MUC}(\alpha, \alpha') = 2 \left\{ N_0 \ln \left(\frac{N_0}{T} \right) + N_1 \ln \left(\frac{N_1}{T} \right) + N_2 \ln \left(\frac{N_2}{T} \right) - [N_0 \ln(1 - \alpha) + N_1 \ln(\alpha' - \alpha) + N_2 \ln(\alpha')] \right\} \quad (24)$$

where N_i gives the number of super exceptions, exceptions and of values lower than VaR, and they are calculated by means of the following hit variables:

$$J_{1,t} = \begin{cases} 1, & \text{if } -VaR_{t,99.8\%} < R_t < -VaR_{t,99\%} \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

$$J_{2,t} = \begin{cases} 1, & \text{if } R_t < -VaR_{t,99.8\%} \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

and

$$J_{0,t} = 1 - J_{1,t} - J_{2,t} \quad (27).$$

Regarding the distribution of the test statistics presented above, the 99% and the 99.8% likelihood ratios follow a chi-squared distribution with one degree of freedom. The multivariate unconditional coverage test statistic follows a chi-squared distribution and has two degrees of freedom.

There are several benefits that can arise from using this technique. Firstly, the Risk Map presents the backtesting results in a simple, effective and easy to understand display, while at the same time it also accounts for the severity of VaR violations. Secondly, it is very general as it can be applied to any VaR model and can be therefore used not only to assess the market risk but also default risk or systemic risk (Colletaz, Hurlin and Perignon, 2011). Moreover, Hurlin and Perignon (2011) employed the Risk Map to check the validity of a margining system in the derivatives market.

2.2. Previous research

The literature on Value-at-risk and backtesting techniques is very extensive and complex, with numerous articles and textbooks presenting, developing and assessing a diversity of models having been published in the past years. Hence, for more information on VaR models we recommend reviewing papers such as Kuester et al. (2005) which presents several of the most used VaR methods for financial data, or more detailed textbooks such as Dowd (2005) or Jorion (2001). For more details on the backtesting procedures, we recommend a series of articles such as Kupiec (1995), Christoffersen (1998), Campbell (2005) or Colletaz et al. (2011).

Previous research regarding the performance of VaR models in financial institutions is very limited in comparison to research on VaR in general. The majority of empirical studies employ either individual assets or simulated portfolios to assess these models, the main

problem being the financial institutions' non-disclosure policies with regard to their trading portfolios. Among the few papers that do investigate the use and accuracy of VaR models in financial institutions we must name Berkowitz and O'Brien (2002), Perignon and Smith (2008), Billinger and Eriksson (2009) and Birtoiu and Dragu (2011). These studies focus on comparing different VaR estimation methods and their performance using either real bank data or theoretical portfolios built to mimic real bank portfolios. All four papers conclude that models based on volatility forecasts modeled using GARCH processes give the best results.

The current paper has a very narrow focus on determining the performance of regime-switching GARCH (MRS-GARCH) models in computing Value-at-Risk in the context of real trading portfolio data from financial institutions. The literature on MRS-GARCH models and their accuracy in forecasting volatility is in a continuous development, yet no article to this date has investigated their efficiency in what concerns Value at Risk on real bank data. We chose to shortly review the most important studies that have set a benchmark in the development of Markov regime switching-GARCH models and that establish a basis for the current paper.

Markov Regime-switching processes were introduced by Hamilton (1989) with the purpose of capturing the periodic shifts from recessions to booms and vice-versa which characterized the US business cycle. Hamilton and Susmel (1994) and Cai (1994) were the first to combine the Markov Switching model of Hamilton (1989) with an ARCH specification in order to account for the possibility of structural breaks in the variance. To avoid the problem of path dependence, the authors restricted their investigation to ARCH models and did not choose a GARCH specification. Gray (1996) developed a generalization to the Markov regime switching GARCH models, which was later modified by Klaasen(2002). An MRS GARCH model employs an ARMA structure to describe volatility while adding the possibility of sudden jumps from a turbulent regime to a more stable one and vice versa. The Markov process is used to govern the shifts between regimes with different variances.

Klaasen's (2002) article introduces a generalized GARCH model by distinguishing two regimes with different volatility levels, and tests its efficiency by applying it on three major USD daily exchange rate series. This yields significantly better out-of-sample volatility forecasts compared to the high single-regime GARCH forecasts.

Gau and Tang (2004) analyze the application of the Markov-switching ARCH model (Hamilton and Susmel (1994)) in improving the forecasting power of VaR models. Their findings show that the VaR forecasts derived from the Markov-switching ARCH model are preferred to alternative parametric and nonparametric VaR models that only consider time-varying volatility.

Haas et al. (2004) propose a Markov regime-switching GARCH model based on mixtures of normal distributions which they apply to three exchange rate return series. They conclude that these methods provide better volatility forecasts than simple GARCH models. Moreover, allowing for skewness was also found to have an important role in determining accurate volatility and VaR forecasts for the data series that were employed.

Marcucci (2005) compares a series of standard GARCH models with a group of Markov Regime-Switching GARCH models (MRS-GARCH) as to their ability to forecast the volatility of the S&P100 index. In the latter models, all parameters switch between a low and a high volatility regime and several distributional assumptions are used for the residuals (normal, Student-t and GED). The empirical results show a superior performance of the MRS-GARCH models over the standard GARCH models in forecasting volatility for time horizons shorter than one week.

Ane and Ureche-Rangau (2006) extend the regime-switching model developed by Gray(1996) to an Asymmetric Power (AP) GARCH model and evaluate its performance on four Asian stock market indices. The study indicates that all the generalizations introduced by the MS-APGARCH model are statistically and economically significant.

Sajjad et al. (2008) apply an asymmetric Markov regime switching GARCH (MRS-GARCH) model to estimate Value-at-Risk for long and short positions of the FTSE100 and S&P500 indices. The study shows that MRS-GARCH under a Student-t distribution for the innovations outperforms other models in estimating the VaR for both long and short positions of the FTSE returns data. In the case of the S&P index, the MSGARCH-t and EGARCH-t models have the best performance, while the MRS-GARCH also performs quite accurately. Furthermore, the paper also concludes that ignoring skewness and regime changes has the effect of imposing larger than necessary conservative capital requirements.

Liu (2011) employs two types of regime-switching GARCH-jump models based on Chan and Maheu (2002)'s autoregressive jump intensity (ARJI) framework to model the non-linearity in return series. The first one is a Markov regime-switching model which generalizes

the GARCH model by distinguishing two regimes with different GARCH volatility and jump intensity levels while the second is a threshold GARCH-jump model with an exogenous threshold variable. The data consisted of daily observations on Japanese YEN-US Dollar exchange rate and IBM stock price and the results showed that the regime-switching models have a better performance than traditional GARCH models for the in-sample period.

In brief, the common findings of the papers reviewed above determine that volatility forecasts derived from regime-switching GARCH models outperform those derived from standard GARCH models. Our paper builds on the previous literature, since we employ the Markov Regime-Switching model introduced by Klaassen (2002). However, our contribution to current research materializes through the application of said model as well as a series of GARCH models without the regime-switching framework in the context of financial institutions, by using real bank trading portfolios. Furthermore, we apply the volatilities derived from these models as well as implied volatilities given by a series of volatility indices to compute volatility weighted historical simulation VaR estimates, in order to bypass the distributional assumptions regarding the returns series that are required by the parametric models. The use of real data allows us to evaluate the models in an empirically relevant context and adds a practical dimension to the study. Also original is the use of two novel backtesting methods, the GMM-duration approach and the RiskMap, in addition to the traditional Christoffersen test, to assess the performance of the VaR models.

3. Methodology

3.1. Estimating volatility using standard GARCH models

Bollerslev (1986) developed a generalized form of the autoregressive conditional heteroskedasticity model introduced by Engle (1982), which allows for a more flexible structure in modeling volatility.

In the basic GARCH(p, q) model, the conditional volatility is dependent upon previous own lags and on past returns and is described by the formula (Brooks (2008)):

$$\sigma_t^2 = \omega + \alpha_1 u_{t-1}^2 + \dots + \alpha_q u_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_p \sigma_{t-p}^2 \quad (28)$$

The most popular version in practice is the GARCH(1, 1) process. This has been proven to be a more parsimonious model that provides a good fit to the data, making it less likely to violate the non-negativity constraints. GARCH models also have the advantage that they account for volatility clusters and fat tails in the distribution (leptokurtosis).

The model specification for calculating VaR based on the volatilities estimated with a GARCH(1, 1) process and under a normal distributional assumption is the following (Perignon and Smith (2008)):

$$R_t = \mu + u_t \quad (29)$$

$$\sigma_t^2 = \omega + \alpha u_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (30)$$

$$VaR_{95\%t} = -\hat{R}_t - 1.96\hat{\sigma}_t \quad (31)$$

Although using the standard GARCH model leads to a lower probability of violating the non-negativity constraints, it does not eliminate this possibility altogether. Moreover the standard version does not account for leverage effects, i.e. the possibility that a negative shock on the market can cause the volatility to rise by more than a positive shock of the same magnitude would.

Extensions of the GARCH model that account for possible asymmetries, such as the Glosten, Jagannathan and Runkle (1993) model (GJR model) or the exponential GARCH proposed by Nelson (1991) have been developed as a response to these drawbacks.

The **EGARCH** model was introduced by Nelson in 1991. Its superiority in modeling volatility over a simple GARCH model is given by the fact that it allows for asymmetries under its formulation (Brooks (2008)).

Nelson's EGARCH model is defined as:

$$\ln(\sigma_t^2) = \omega + \beta \ln(\sigma_{t-1}^2) + \alpha \frac{u_{t-1}}{\sqrt{\sigma_{t-1}^2}} + \phi \left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} - E \left[\frac{|u_{t-1}|}{\sqrt{\sigma_{t-1}^2}} \right] \right] \quad (32)$$

Several empirical studies have determined the superior predictive performance of the EGARCH models compared to standard GARCH models in what concerns financial data. Alberg et al. (2008) applied several GARCH models on Tel Aviv Indexes and concluded that EGARCH skewed student-t model had the best fit in characterizing the dynamic behavior of the index returns. This model was able to capture the serial correlation, asymmetric volatility clustering and leptokurtic innovation. Su (2010) came to a similar conclusion after running a comparative study on the Chinese stock market. EGARCH models had a better performance than symmetric GARCH in modeling volatility of the Chinese stock returns.

The **GJR-GARCH** model was introduced by Glosten, Jagannathan and Runkle (1993). This is a simple extension of GARCH that includes an additional term to account for possible asymmetries.

The specification for a GJR-GARCH(1, 1) model is:

$$\sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \alpha \varepsilon_{t-1}^2 + \phi \varepsilon_{t-1}^2 I_{t-1} \quad (32)$$

$$\varepsilon_t = \sigma_t Z_t \quad \text{where } Z_t \text{ is iid}$$

$$I_{t-1} = \begin{cases} 0 & \text{if } \varepsilon_{t-1} \geq 0 \\ 1 & \text{otherwise} \end{cases} \quad (33)$$

Liu and Hung (2010) investigated the daily volatility forecasting for the S&P100 stock index series from 1997 to 2003. Their results indicated that the GJR-GARCH model achieved the most accurate volatility forecasts, closely followed by the EGARCH model. Su et al. (2009) also confirmed the superiority of the GJR models in a study that employed several symmetric and asymmetric GARCH models to determine Value at Risk forecasts, applied on the QQQQ index returns. Their findings determined that GJR-GARCH models outperformed both the symmetric GARCH models and the asymmetric NA-GARCH models used. In addition, the ARMA(1, 1)-GJR-GARCHM(1, 1) was the best market risk management tool for financial portfolios in terms of the smallest violation number and the smallest capital charge.

3.2. Markov regime-switching GARCH models (MRS-GARCH)

A problem with GARCH models is that they tend to overestimate volatility forecasts, particularly during periods of high volatility. The reason behind this is the high degree of persistence implied by the GARCH model. This spurious persistence could be derived from structural changes in the variance process (Hamilton and Susmel (1994)) and can lead to weak forecasts since the impact of shocks lasts for a shorter than accounted for. In order to overcome this issue, a Markov-Switching process can be employed to model changes in parameters.

Regime-switching models are characterized by the possibility for some or all of the parameters to change across several states of the world according to a Markov process, which is governed by a state variable S_t . The model actually draws the current value of the variable from a mixture of distributions, based on the more likely state that could have determined such an observation.

The transition probability represents the probability of switching from state i at $t-1$ to state j at t .

$$Pr(s_t = j | s_{t-1} = i) = p_{ij} \quad (34)$$

If we consider, for simplicity, that there are only 2 states, the transition matrix⁴ can be written as:

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p & (1-q) \\ (1-p) & q \end{pmatrix} \quad (35)$$

A general form of the MRS-GARCH models can be written as:

$$r_t | \Omega_{t-1} \sim \begin{cases} f(\theta_t^{(1)}) & \text{with probability } p_{1t} \\ f(\theta_t^{(2)}) & \text{with probability } (1 - p_{1t}) \end{cases} \quad (36)$$

Where $f(*)$ stands for one of the possible conditional distributions used in the model (normal, Student-t, GED), $\theta_t^{(i)}$ is the vector of parameters in the i -th regime that characterizes the distribution, $p_{1t} = Pr(s_t = 1 | \Omega_{t-1})$ is the ex-ante probability and Ω_{t-1} is the information set at $t-1$.

⁴ p and q are probabilities that the volatility remains in the same regime

The vector of time-varying parameters, θ , can be written as a function of three components:

$$\theta_t^{(i)} = (\mu_t^{(i)}, h_t^{(i)}, v_t^{(i)}) \quad (37)$$

$\mu_t^{(i)} = E(r_t | \Omega_{t-1})$ is the conditional mean, $h_t^{(i)} = Var(r_t | \Omega_{t-1})$ is the conditional variance and $v_t^{(i)}$ is the shape parameter of the conditional distribution.

Therefore, there are four main input elements in a MRS-GARCH model: the conditional mean, the conditional variance, the regime process and the conditional distribution.

Since the main purpose of this study is determining the volatility forecasts, the conditional mean is simply modeled as:

$$r_t = \mu_t^{(i)} + \varepsilon_t \quad (38)$$

where $i = 1, 2$ since we assumed two regimes, and $\varepsilon_t = \eta_t \sqrt{h_t}$ and $\eta_t \sim iid(0,1)$

The conditional variance, given the whole unobserved regime path $\tilde{s}_t = (s_t, s_{t-1}, \dots)$ is $h_t^{(i)} = V(\varepsilon_t | \tilde{s}_t, \Omega_{t-1})$. The conditional variance is expressed under a GARCH(1, 1) formulation:

$$h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)} \varepsilon_{t-1}^2 + \beta_1^{(i)} h_{t-1} \quad (39)$$

with h_{t-1} expressed as an average of past conditional variances. This simplification is due to the fact that the state-dependency of the past conditional variances of a GARCH model is infeasible a regime-switching context. The explanation is that the conditional variance would depend on the observable information set Ω_{t-1} , the current regime s_t which determines the parameters as well as on all past states \tilde{s}_{t-1} . The model would become very complex and most likely impossible to estimate since this multiple dependency would require the integration over a number of regime paths that would grow exponentially with the sample size (Marcucci (2005)).

In order to avoid this path-dependence problem, Gray (1996) and Klaassen (2002) advocate the use of the conditional expectation of the lagged variance as a proxy for the lagged variance. Gray (1996) uses the information available at $t - 2$ to integrate out the unobserved regimes:

$$h_{t-1} = E_{t-2} [h_{t-1}^{(i)}] = p_{1,t-1} [(\mu_{t-1}^{(1)})^2 + h_{t-1}^{(1)}] + (1 - p_{1,t-1}) [(\mu_{t-1}^{(2)})^2 + h_{t-1}^{(2)}] - [p_{1,t-1}\mu_{t-1}^{(1)} + (1 - p_{1,t-1})\mu_{t-1}^{(2)}]^2 \quad (40)$$

Nonetheless, this specification has one major drawback, in that multi-step ahead volatility forecasts are usually very complicated to determine.

Klaasen (2002) recommends using the conditional expectation of the lagged conditional variance with an extended information set. The model integrates out the past regimes by also taking into account the current ones using the following specification for the conditional variance: $h_t^{(i)} = \alpha_0^{(i)} + \alpha_1^{(i)}\varepsilon_{t-1}^2 + \beta_1^{(i)}E_{t-1}[h_{t-1}^{(i)}|s_t]$ (41) where the conditional expectation is defined as:

$$E_{t-1} [h_{t-1}^{(i)}|s_t] = \tilde{p}_{ii,t-1} [(\mu_{t-1}^{(i)})^2 + h_{t-1}^{(i)}] + \tilde{p}_{ji,t-1} [(\mu_{t-1}^{(j)})^2 + h_{t-1}^{(j)}] - [\tilde{p}_{ii,t-1}\mu_{t-1}^{(i)} + \tilde{p}_{ji,t-1}\mu_{t-1}^{(j)}]^2 \quad (42)$$

And the probabilities are computed as:

$$\tilde{p}_{ji,t} = \Pr(s_t = j|s_{t+1} = i, \Omega_{t-1}) = \frac{p_{ji}\Pr(s_t=j|\Omega_{t-1})}{\Pr(s_{t+1}=i|\Omega_{t-1})} \frac{p_{ji}p_{j,t}}{p_{i,t+1}} \quad (43) \quad \text{with } i, j = 1, 2$$

This regime-switching GARCH model proposed by Klaasen (2002) has two main advantages compared to previous models. Firstly, it allows for higher flexibility in capturing the persistence of shocks to volatility. Secondly, direct expressions for the multi-step ahead volatility forecasts exist, which can be determined recursively, as is the case with the standard GARCH models. For instance, the n -step ahead volatility forecast at $T-1$ can be computed as:

$$\hat{h}_{T,T+n} = \sum_{\tau=1}^n \hat{h}_{T,T+\tau} = \sum_{\tau=1}^n \sum_{i=1}^2 \Pr(s_{T+\tau} = i|\Omega_{T-1}) \hat{h}_{T,T+\tau}^{(i)} \quad (44)$$

where $\hat{h}_{T,T+\tau}$ is the time T aggregated volatility forecast for the next n steps and $\hat{h}_{T,T+\tau}^{(i)}$ is the τ -step ahead volatility forecast in regime i made at time t that can be computed recursively:

$$\hat{h}_{T,T+\tau}^{(i)} = \alpha_0^{(i)} + (\alpha_1^{(i)} + \beta_1^{(i)})E_{t-1} [h_{T,T+\tau-1}^{(i)}|s_{T+\tau}] \quad (45)$$

The multi-step ahead volatility forecasts are determined as a weighted average of the multi-step ahead volatility forecasts in each regime, using prediction probabilities as weights. A GARCH model is used to determine the volatility forecast for each regime, by weighting the volatilities from the previous regime with the probabilities in equation (43).

The estimation of the parameters for the regime switching models is generally done by employing the maximum likelihood method.

A necessary step is to determine the ex-ante probability⁵ $p_{1,t} = \Pr(s_t = 1 | \Omega_{t-1})$ using the following specification:

$$p_{1,t} = \Pr(s_t = 1 | \Omega_{t-1}) = (1 - q) \left[\frac{f(r_{t-1} | s_{t-1}=2)(1-p_{1,t-1})}{f(r_{t-1} | s_{t-1}=1)p_{1,t-1} + f(r_{t-1} | s_{t-1}=2)(1-p_{1,t-1})} \right] + p \left[\frac{f(r_{t-1} | s_{t-1}=1)p_{1,t-1}}{f(r_{t-1} | s_{t-1}=1)p_{1,t-1} + f(r_{t-1} | s_{t-1}=2)(1-p_{1,t-1})} \right] \quad (47)$$

where p and q are the transition probabilities in equation (35) and $f(*)$ is the likelihood function in equation (36).

The log-likelihood function can thus be rewritten as (Marcucci(2005)):

$$l = \sum_{t=-R+\omega+1}^{T+\omega} \log[p_{1,t}f(r_t | s_t = 1) + (1 - p_{1,t})f(r_t | s_t = 2)] \quad (48)$$

Where $\omega=0, 1, \dots, n$ and $f(* | s_t = i)$ denotes the conditional distribution considering that regime i occurs at time t .

The estimation of the MRS-GARCH models is done by numerical maximization of (48) using Matlab. For the forecast of the volatilities the coefficients were re-estimated as the 252 observations rolling window moved along with each observation. This lead to 987 sets of parameters which were further on used to obtain the volatility forecasts. The results of the estimation are sensitive to the starting values for the parameter optimization. The algorithm we used for the MRS-GARCH model estimates the starting values for the first rolling window by using a grid of reasonable values. The log-likelihood is evaluated on this grid and the best fit is then used. As the rolling window moves along with one observation, the starting values will be updated to the estimated parameters from the previous window. Estimating the starting values with the grid for each rolling window would significantly increase the time duration necessary for the estimation, which would make it very inefficient considering the time span allocated for the development of this paper.

The Markov Regime Switching GARCH model proposed by Klaasen (2002) is employed to derive volatility forecasts for the five banks trading portfolios, assuming the data follows two different volatility regimes and using both a normal and a fat-tailed Student-t distribution assumption for the innovations. These forecasts are further on used to compute both parametric and volatility-weighted historical simulation Value-at-Risk and their performance is assessed through three different backtesting methods.

⁵ Which is interpreted as the probability of being in the first regime at time t given the information set at $t-1$

4. Data

Sample. The data sample spans over a five-year period from January 2007 to December 2011 and consists of real, daily P/L figures from four commercial banks, more specifically Bank of America, Deutsche Bank, Swedbank and Danske Bank, and it contains a number of 1239 daily observations. The graphs disclosed by the above-mentioned banks in their annual and risk management reports were used to obtain the data. To this purpose a software application called GetData Graph Digitizer was employed.

The choice of these particular banks is motivated by several considerations. Firstly, we wanted to take into account both European and North American banks, since these markets tend to have different characteristics and behavior. Secondly, we chose those banks that had published P/L plots in their annual reports for at least five years, in order to obtain a large enough data set for our estimation. Bank of America was chosen since it is one of the largest banks in the US and thus representative for the US banking industry. Deutsche Bank is representative for the German banking system and it is a large multinational financial company, similar to Bank of America. Danske Bank and Swedbank are smaller banks and active mainly on the Scandinavian markets which have a somewhat different dynamics than Western European markets. Therefore, we aimed for diversity in location and size of the banks, which leads to a diversity in the structure of the trading portfolios. This permits us to get a broader view of the behavior of the analyzed VaR models in light of the particularities of each data set.

To further assess the validity of the VaR models applied, the analysis will extend to also encompass a mimicking portfolio for Bank of America with an unchanged composition. The synthetic portfolio was constructed based on the bank's average portfolio composition during the January 2007 – December 2011 interval.

To build the mimicking portfolio (following Billinger and Eriksson (2009)), Bank of America's average market risk VaR for each trading activity was used. These figures were then averaged across the entire five-year sample period to obtain the proportion of each asset in the trading portfolio. The portfolio is composed of six different assets, namely foreign exchange, interest rate, credit, real estate, equities and commodities. For each of the assets proxies were created.

As a proxy for foreign exchange, a basket of six exchange rates was constructed and it contained six most traded currencies in Bank of America's portfolio, i.e. the euro, the

Japanese yen, the British pound, the Australian dollar, the Canadian dollar and the Swiss franc. Equal weights were assigned to each of these currencies. Secondly, as an intermediary for the interest rate the three-month US treasury-bill was used, while credit was substituted by a corporate bond index, i.e. Dow Jones US Corporate Bond Index. When it comes to real estate this was proxied by the FTSE NAREIT United States Real Estate Index. Furthermore, the S&P500 stock index was used to proxy equities, while commodities were assigned to the Dow Jones - UBS commodity index.⁶

Daily data for the indexes, currencies and interest rate was extracted from the Thomson Reuters' Datastream available in the Finance Lab.

Even though keeping the bank's portfolio composition unchanged might seem counterintuitive as it is well known that banks' trading activities may fluctuate drastically across a time period and with it so will the portfolio composition, using a constant portfolio both at the beginning and at the end of the period will allow for a better assessment of a VaR model predictive ability (J.P. Morgan Chase (2009)). This is also in line with the new revisions to the Basel III market risk framework which give supervisors the possibility to require banks to carry out backtesting on either hypothetical (when end-of-day positions remain unchanged) or actual P/Ls (Basel (2011)).

The VaR models are estimated using a rolling window with an in-sample period comprised of 252 observations, while the other 987 observations will form the out-of-sample period used in the models' validation. The length of the estimation period is consistent with the regulations adopted by the Basel Committee. Furthermore, the out-of-sample is compliant with the current regulatory framework that states a backtesting period of at least three years.

Software. The P/L data corresponding to each of the four banks trading portfolio was extracted using GetData Graph Digitizer, from the daily trading-related revenues graph which are disclosed in the annual reports. To obtain the index, currency and interest rate data it was resorted to Thomson Reuters' Datastream.

Data analysis was conducted using Eviews 7, while Microsoft Excel 2007 was used to perform the models' backtesting with the Risk Map technique. Moreover, MatLab was employed to estimate the GARCH and regime switching processes, to calculate the VaR predictions, but also to backtest the models with the GMM-duration and Christoffersen's approaches and to perform the ARCH effects and Ljung-Box Q tests.

⁶ The portfolio components and their associated weights can be found in Appendix 1 - Table 1

The GARCH models were estimated using the Econometric Toolbox in MatLab, while for the MRS-GARCH models we used the code written by Marcucci (2005)⁷. The MatLab code written by Colletaz, Candelon, Hurlin and Tokpavi⁸ and the code by Hurlin and Perignon⁹ which were available online were employed to compute the GMM-duration and the Christoffersen (1998) backtesting statistics. All the codes were modified to be able to fit our data series.

Descriptive statistics. It is well known that financial returns have the characteristic of volatility clustering, that is large returns tend to be followed by large returns and small returns are more likely to come after small returns. This can also be defined as a positive correlation between the current level of volatility and those that preceded it (Brooks (2008)). This exact phenomenon can be observed in the P/L graphs below. The returns on the banks' portfolios seem to follow two different volatility regimes. Volatility bursts occur throughout the 2008-2009 financial crisis period when all four banks included in the analysis experienced a high level of volatility in their returns. Furthermore, the three European banks showed an increased unpredictability during selected periods in 2010 and 2011 when several European countries (e.g. Greece, Ireland, Portugal, Italy, Spain) were affected by the sovereign debt crisis. Bank of America's return volatility was not only affected by the debt crisis in Europe but also by the debt ceiling negotiations in the United States Congress, whose failure to reach an agreement in August 2011 led to an historic, first time rating downgrade of the US government debt.

When analyzing the autocorrelation and the partial autocorrelation functions of the five P/L series¹⁰ it can be ascertained that the data for all the four banks exhibits autocorrelation. This inference made from the P/L plots above and the correlograms was then tested in a more formal framework by applying the ARCH effects and Ljung-Box Q tests.¹¹ To conduct the tests the mean of each series was subtracted from the P/Ls. The null hypothesis of no serial correlation can be safely rejected in both cases, which supports the deduction made from the graphs, i.e. there exists volatility clustering in the returns.

⁷ <http://www.degruyter.com/view/j/sn-de.2005.9.4/sn-de.2005.9.4.1145/sn-de.2005.9.4.1145.xml?format=INT>

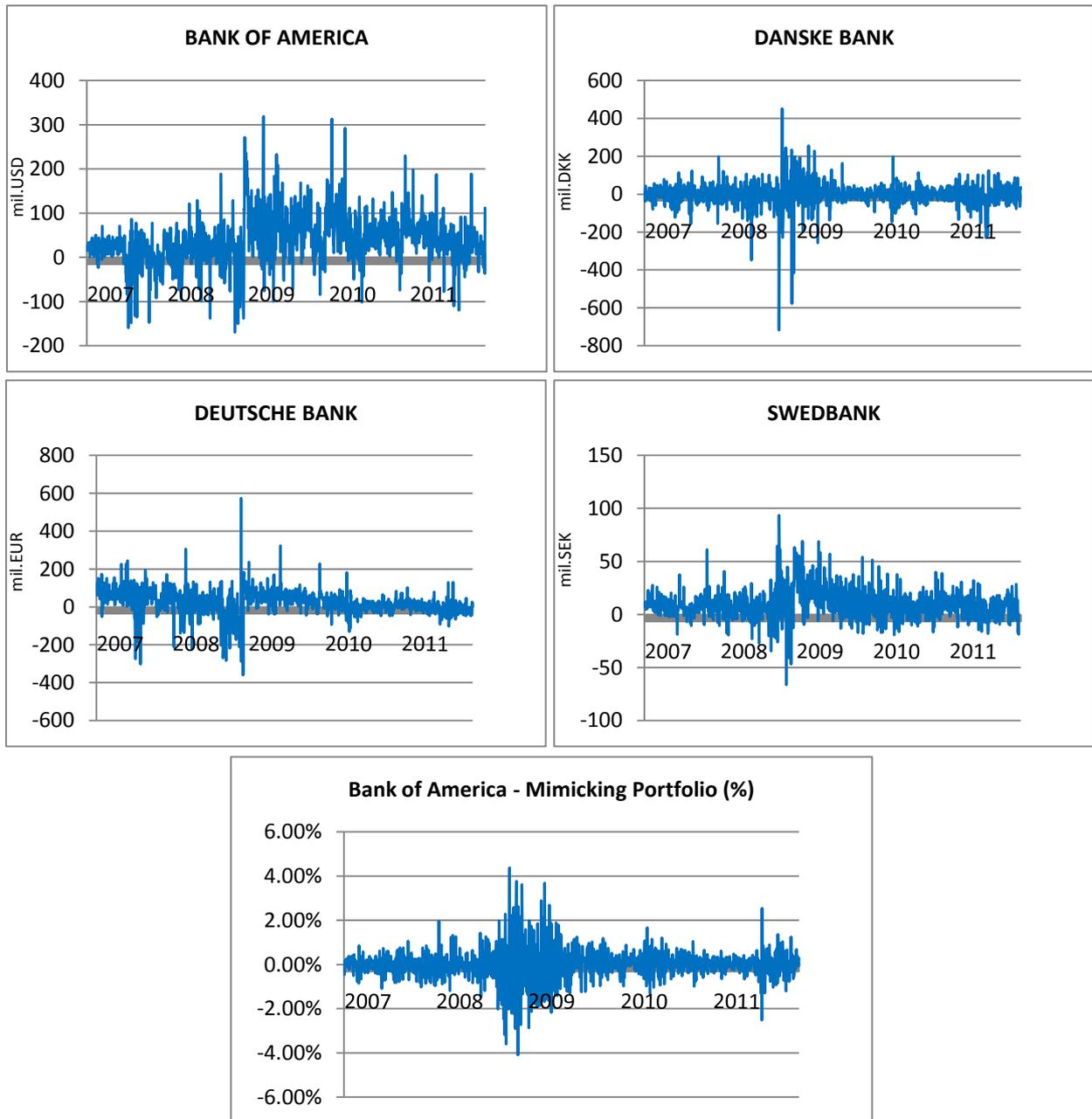
⁸ Candelon B., Colletaz G., Hurlin C. and Tokpavi S. (2011), "Backtesting Value-at-Risk : A GMM Duration-based Test" - RunMyCode Companion Website www.runmycode.fr/CompanionSite/Site3

⁹ Hurlin C. and Perignon C. (2012), "Backtesting Value-at-Risk : A Duration-Based Approach". RunMyCode Companion Website, www.runmycode.org/CompanionSite/Site68

¹⁰ See Appendix 1 – Figure 1

¹¹ See Appendix 1 – Tables 2 and 3

Figure 1: Daily Profit/Loss for Bank of America, Danske Bank, Deutsche Bank and Swedbank



The table below summarizes the descriptive statistics for the four real P/L series as well as for the percentage return series of Bank of America's mimicking portfolio. There can be seen a rather significant difference between the maximum and minimum values which can be attributed to the turbulent time period across which the analysis was conducted. Furthermore, all series present asymmetries from the mean. In case of Bank of America (both real and hypothetical P/Ls) and Swedbank the asymmetry is positive, while for the other two banks the series are negatively skewed. With respect to the kurtosis coefficient, all time series display a value that is larger than 3, which means that returns distributions are leptokurtic,

having fatter tails and a sharper peak. These figures for skewness and kurtosis imply that the series do not follow a normal distribution. This inference also backed by the Jarque-Bera test, whose null hypothesis of normality can be dismissed for all four banks and for the artificial portfolio. Moreover, regarding the stationarity of the series, after applying the Augmented Dickey-Fuller test it can be assumed that the series are stationary since the test null hypothesis of a unit root can be rejected in all cases.

Table 1: Descriptive statistics

	Bank of America (mil. USD)	Danske Bank (mil. DKK)	Deutsche Bank (mil. EUR)	Swedbank (mil. SEK)	Bank of America - mimicking portfolio (%)
Mean	37.28718	-1.74153	22.26597	9.60249	4.89E-05
Median	33.16062	-2.88462	21.17647	8.21530	0.000222
Maximum	318.49870	452.31020	573.68420	93.60272	0.043713
Minimum	-169.6185	-717.8774	-359.3301	-66.40219	-0.040936
Std. Dev.	54.70913	64.21122	66.52825	13.53415	0.007119
Skewness	0.37640	-1.62442	-0.17655	0.71685	0.143889
Kurtosis	5.69975	25.39471	10.27785	7.09594	9.015723
Jarque Bera	405.53370 <i>(0.00000)</i>	26,436.01 <i>(0.00000)</i>	2,740.86 <i>(0.00000)</i>	967.51 <i>(0.00000)</i>	1,872.528 <i>(0.00000)</i>
Augmented Dickey-Fuller	-8.071847 <i>(0.00000)</i>	-11.29807 <i>(0.00000)</i>	-5.214142 <i>(0.00000)</i>	-7.857297 <i>(0.00000)</i>	-42.01209 <i>(0.00000)</i>

In conclusion, given that the data analysis showed that P/L data exhibits volatility clustering and seems not to follow a normal distribution it can be expected that the empirical results will prove that the models based on time-varying volatility forecasts and a more fat-tailed distributional assumption will perform better.

Data reliability. The quality of the P/L data extracted with the digitizing software was verified by comparing the graph created with values given by the software to the original plot in the banks' annual reports. Moreover, the accuracy of the digitizing software is also supported by its use in various journal articles (Barron et al., 2010; Sadras et al., 2011) with topics ranging from medical research to physics, chemistry or ecology.

5. Empirical Results

The present study employs the Markov regime-switching GARCH model proposed by Klaasen (2002) under the assumption of two volatility regimes. Furthermore, the GARCH, EGARCH and GJR models are employed to forecast volatilities. For each of the models mentioned above two different distributional assumptions are applied, namely the normal and the Student-t distribution. The volatility forecasts are used to compute the parametric and volatility weighted historical simulation VaRs for each trading portfolio. In addition, we also use realized volatilities (VIX, VSTOXX, VDAX volatility indices) in order to calculate Value-at-Risk for the five portfolios. The performance of all the models is further on tested using three backtesting frameworks.

Both parametric and non-parametric models based on time varying volatilities were applied in the estimation of one day ahead Value-at Risk for all 5 portfolios. The performance of the models was assessed using the Christoffersen, the GMM-duration approach and the RiskMap backtesting methods. The backtesting window consists of 4 years of daily observations for the parametric methods and the HS-VIX and 3 years of daily observations for the volatility-weighted historical simulation method. We used a 252 days rolling window to forecast the one-day-ahead volatilities which were employed in computing the historical simulation with another 252 days rolling window and that left us with 3 years of out-of-sample observations for backtesting the VWHS method. While using a longer rolling window could offer more stability to the estimated volatilities and increase the probability of the models converging to a solution, we chose to employ the 252 days rolling window since it is consistent with the Basel recommendations for Value-at-Risk calculation.

For the forecast of the volatilities through GARCH and MRS-GARCH models, the coefficients were re-estimated as the rolling window moved along with each observation. This lead to 987 sets of parameters which were further on used to obtain the volatility forecasts.

The conditional mean and variance specification for the GARCH models was determined using the last 252 observation taken from each data series. For each portfolio and for each GARCH, EGARCH and GJR model several specifications were tested and a choice was made based on the statistical significance of the coefficients and the SIC information criteria. The model with the largest number of coefficients was chosen between two models with a different number of coefficients but all coefficients statistically significant. Although

less parsimonious, the larger models can provide a better description of the composition of the conditional variance.

The specifications and coefficient estimates for each portfolio and each GARCH model are centralized in Tables 1-5 in Appendix 2.

While it is possible for the most accurate specification to differ between time periods inside the same data series, we are interested in those models that best fit the data series at the present moment, and therefore we used the latest 252 observations to choose the model specification. However, we also tested the specifications on a series of random 252 days windows in each data series and the results were very similar.

The number of degrees of freedom for the computation of parametric VaR under a t -distribution was determined the closest integer to $(4k-6)/(k-3)$ where k is the kurtosis coefficient for the return distribution, as advocated by Dowd (2005). Moreover, the VaR estimates were determined employing volatilities derived from GARCH models that assume the innovations follow a t -distribution when estimating the conditional volatility.

The Markov Regime Switching GARCH model was estimated under the assumption of a random walk with drift for the conditional mean and GARCH(1,1)-like specification for the conditional variance, assuming two volatility regimes. This was due to the main focus on volatility forecasting, as well as the computational burden a more complex model would entail in the context of the time limit imposed for this study. The coefficient estimates for each portfolio are presented in Table 6 in Appendix 2.

When building the graphical representation of the Risk Map it is necessary to take into consideration the number of out-of-sample VaR forecasts. In this case, the Risk Maps are constructed based on the three years of observations for the volatility weighted historical simulation models, while for the other models the out-of-sample data contains four years of observations. The boundaries of the map will expand as the number of forecasts increases. Therefore, for the three years map the number of violations must not exceed 15 and be larger than 2, while the number of super exceptions must be strictly lower than 6. Conversely, for the four years map these limits will grow larger, placing the number of exceedances in the [3 18] interval, whereas the number of super exceptions has to be lower or equal to 6.

The results are centralized firstly by each bank portfolio and then an analysis centered on each model is conducted. The structure of each trading portfolio depends on the bank's

specific policies, purposes and particularities which this can influence the accuracy of the methodology used. Thus, we choose to first assess the fit of the models to each of the five data series and then we take an overall view of the behavior of each model across all five portfolios. The latter will allow us to determine whether these models are indeed sensitive to the composition of the portfolios or if there are models that have a consistent performance globally.

5.1. Separate portfolio analysis

In the following subchapters an analysis of the results will be undertaken in the context of each portfolio (for each bank), since it is necessary to highlight the way particularities of financial data influence these models.

The results for each bank will be aggregated based on the three backtesting techniques used to validate the models' forecasts and on the distributional assumption employed for a particular model. It is important to note that each of the three backtesting methods accounts for different critiques brought in time to these validating procedures. Thus, the Christoffersen test takes into consideration the frequency of tail losses as well as the violations' tendency to bundle in time. The GMM-duration approach accounts for the time interval between the individual exceptions and gives the possibility to assess the independence and the conditional coverage hypotheses. Lastly, the Risk Map not only verifies the unconditional coverage hypothesis as in Kupiec (1995) but also brings forward those models that have the capacity to capture excessive losses, i.e. losses whose occurrence will have an incommensurate effect on a bank's ability to run its operations normally.

This allows for the construction of a unified backtesting framework in which the VaR models will be evaluated from various perspectives.

In addition, to further assess the behavior of the estimated models, we compare their performance with the in-house VaR methodologies used by the banks to account for their market risk, in terms of the number of exceptions.

The four banks we analyze published the number of exceptions their model gave for each year as well as a short description of the VaR models they employ.

The common feature they share is that the in-house models use Monte-Carlo simulations applied to different periods of historical data. This choice is motivated by the fact that using historical data does not require any distributional assumptions for the return series.

The VaR is generally computed for each risk factor and then they are aggregated using the correlation between different risk factors to obtain the VaR measure for the entire portfolio. This allows for a better understanding of the way each factor impacts the total risk of the portfolio. Furthermore, in periods of market stress, the banks can choose to selectively reduce the market risk for some of their lines of business as Bank of America states in its annual reports. This is done that by selling out positions or executing macroeconomic hedges to reduce exposure where economically feasible.

The institutions also admit to the drawbacks of their models, in that the use of historical data may not give good predictions of potential future events, particularly of the ones that are extreme in nature. This backward-looking approach could lead to either an under-estimation or over-estimation of the market risk, none of which is beneficial for the banks' business. In addition, the correlations between different risk factors may not be accurate or may not hold during extreme market movements. The length of the historical data window used to estimate VaR can also influence the measure. Bank of America uses three years of data, which lead to an overall over-estimation of the risk during the 2009-2010 period when no exceptions occurred. The model did not manage to adapt to the lower volatility regime that occurred during those 3 years, since it was still affected by the turbulent market changes in 2008.

Due to these admitted flaws in the models, the banks supplement the VaR measures with additional methods of assessing market risk, such as stressed VaRs, Expected Shortfall measures or desk level limits.

The comparison between the methods applied in this paper and the banks' models is made difficult by the fact that the financial institutions reassess their models and assumptions after years when many exceptions occurred. This means that we are not comparing one single actual model over the years with the models we estimated, but a series of different models. For instance Danske Bank replaced their parametric model with a historical simulation based one in mid-2007, it updated from constant weights for the historical data to varying weights in the beginning of 2009, following the poor performance in predicting the turbulent market movements in 2008, and it expanded its model to include a series of new risk factors in 2011. Therefore, Danske Bank practically used four different models over this time span.

5.1.1. Bank of America

The backtesting results for Bank of America show that most of the models perform properly, except for the parametric EGARCH and GARCH under a normal distribution. Even

though the latter passes the unconditional and conditional coverage tests, due to the large number of super exceptions the model is rejected by the Risk Map, while the former does not pass the Christoffersen test as well. The only other rejection encountered in this case is that given by the GMM-duration approach for the t -distributed MRS GARCH parametric model.

Table 2: Backtesting results - Bank of America

Model		N	N'	Christoffersen			GMM duration			Risk Map			
				LR _{UC}	LR _{IND}	LR _{CC}	J _{UC}	J _{IND} (5)	J _{CC} (5)	LR _{99%}	LR _{99.8%}	LR _{MUC}	
	HS-VIX	9	3	0.0763 (0.7824)	0.1660 (0.6837)	0.2423 (0.8859)	0.0023 (0.9657)	2.0505 (0.2330)	2.0776 (0.4590)	0.0780 (0.7800)	0.4625 (0.4965)	0.9551 (0.6203)	
Normal	Volatility Weighted HS	GARCH	7	4	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.8492)	0.6207 (0.6492)	0.5857 (0.8803)	0.0162 (0.8989)	2.9639 (0.0851)	4.6698 (0.0968)
		EGARCH	7	4	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.8341)	0.6207 (0.6509)	0.5857 (0.8773)	0.0162 (0.8989)	2.9639 (0.0851)	4.6698 (0.0968)
		GJR	8	3	0.0601 (0.8064)	3.2715 (0.0705)	3.3315 (0.1890)	0.3130 (0.4573)	2.2928 (0.1860)	2.4802 (0.3688)	0.0582 (0.8093)	1.2275 (0.2679)	1.3613 (0.5063)
		GARCH	6	4	0.2598 (0.6102)	0.0990 (0.7530)	0.3589 (0.8357)	0.0157 (0.9252)	0.2074 (0.9608)	0.3482 (0.9537)	0.2635 (0.6077)	2.9639 (0.0851)	6.3934 (0.0409)
		EGARCH	6	3	0.2598 (0.6102)	0.0990 (0.7530)	0.3589 (0.8357)	0.0157 (0.9170)	0.2074 (0.9621)	0.3482 (0.9577)	0.2635 (0.6077)	1.2275 (0.2679)	2.9412 (0.2298)
		GJR	7	2	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.7114)	0.8681 (0.5114)	0.8054 (0.7928)	0.0162 (0.8989)	0.1734 (0.6771)	0.3096 (0.8566)
Normal		GARCH	16	9	3.2626 (0.0709)	0.5284 (0.4673)	3.7910 (0.1502)	3.0376 (0.0569)	1.1318 (0.4387)	3.3263 (0.2671)	3.2500 (0.0714)	13.3215 <i>(0.0003)</i>	13.4138 <i>(0.0012)</i>
		EGARCH	21	11	9.6241 <i>(0.0019)</i>	0.9150 (0.3388)	10.5391 <i>(0.0051)</i>	6.7779 (0.0122)	3.2290 (0.1143)	10.4251 (0.0362)	9.6011 <i>(0.0019)</i>	19.8417 <i>(0.0000)</i>	20.4071 <i>(0.0000)</i>
		GJR	11	4	0.1307 (0.7177)	2.5856 (0.1078)	2.7163 (0.2571)	0.3891 (0.5510)	35.0062 <i>(0.0002)</i>	22.2659 (0.0104)	0.1283 (0.7202)	1.6062 (0.2050)	1.7072 (0.4259)
		GARCH	11	3	0.1307 (0.7177)	0.2485 (0.6182)	0.3791 (0.8273)	0.3891 (0.5915)	0.7041 (0.6260)	0.8841 (0.7850)	0.1283 (0.7202)	0.4625 (0.4965)	0.4643 (0.7928)
		EGARCH	15	4	2.3446 (0.1257)	0.4639 (0.4958)	2.8085 (0.2455)	2.3878 (0.1219)	0.1460 (0.9861)	2.8365 (0.3209)	2.3340 (0.1266)	1.6062 (0.2050)	2.7212 (0.2565)
		GJR	9	2	0.0763 (0.7824)	0.1660 (0.6837)	0.2423 (0.8859)	0.0023 (0.9039)	16.3803 <i>(0.0012)</i>	22.1625 (0.0109)	0.0780 (0.7800)	0.0004 (0.9841)	0.1051 (0.9488)
Normal		MRS parametric	7	4	0.9265 (0.3358)	0.1002 (0.7516)	1.0267 (0.5985)	0.4321 (0.4212)	0.7167 (0.6142)	1.4641 (0.6070)	0.9323 (0.3343)	1.6062 (0.2050)	5.5860 (0.0612)
		MRS vwhs	6	3	0.2598 (0.6102)	0.0990 (0.7530)	0.3589 (0.8357)	0.0157 (0.9141)	0.2074 (0.9628)	0.3482 (0.9602)	0.2635 (0.6077)	1.2275 (0.2679)	2.9412 (0.2298)
		MRS parametric	4	2	4.5256 (0.0334)	0.0326 (0.8567)	4.5582 (0.1024)	4.7520 (0.0291)	3.2745 (0.1079)	67.2034 <i>(0.0022)</i>	4.5376 (0.0332)	0.0004 (0.9841)	6.3227 (0.0424)
		MRS vwhs	6	4	0.2598 (0.6102)	0.0990 (0.7530)	0.3589 (0.8357)	0.0157 (0.8986)	0.2074 (0.9621)	0.3482 (0.9540)	0.2635 (0.6077)	2.9639 (0.0851)	6.3934 (0.0409)

Note: N – number of violations, N' – number of super exceptions (expected N and N' , given the 99% confidence level are equal to 8 and 2, for a 3-year sample, and to 10 and 2 for a 4-year sample). Figures in parentheses represent the p -values associated with the test statistics. The probabilities in *Italic* denote a model rejection.

A possible explanation for rejecting the GARCH and EGARCH model is that these two models seems to underestimate the risk to which the bank is exposed, inference confirmed both by the considerable number of exceptions and by the fact the mentioned

models yield the smallest average VaR figures out of all models employed.¹² Conversely, the model with the most conservative VaR estimates is the parametric GJR- t , followed by the MRS-GARCH- t and GJR with a normal distribution.

Furthermore these results confirm the expectation that models working with the assumption of a t -distribution give more accurate predictions than those that operate under the normal distribution. However, the anticipation that models accounting for asymmetries in the financial data have a better forecasting ability is not corroborated by the results as the performance of the simple GARCH is similar to that of the GJR-GARCH or EGARCH both under the parametric and the volatility weighted frameworks.

The volatility weighted historical simulation seems to have the best performance when using volatility forecasts yielded by various GARCH specifications but also when Markov Regime Switching volatility predictions are being employed.

From the table above it can also be inferred that the non-parametric HS-VIX efficiency is adequate as it passes all three validating procedures. The number of exceptions is lower than those given by most parametric VaR methods, while the number of super exceptions is on par or even better than those yielded by other models.

Moreover, the models based on Markov's regime switching process give accurate forecasts in general, even though the t -distributed parametric appears to overestimate the risk to which the bank is exposed and thus resulted in a rejection from the GMM-duration backtest.¹³

The in-house VaR model that Bank of America consists of a historical simulation approach based on 3 years of historical data and an expected shortfall methodology equivalent to a 99 percent confidence level, as described in their annual reports. The total number of violations that occurred using the bank's VaR methodology was 2 for the 2008-2011 period. Out of these, 2 violations occurred in 2008, while none were registered during the next 3 year period. All the models we estimated on this portfolio gave a much larger number of exceptions. At first glance it may seem that the actual model used by the bank is more accurate. However, we must take into consideration the fact that such a small number of violations could also mean an overestimation of the VaR measures, which is not beneficial for the institution since it increases its capital charges. This reasoning is confirmed by the annual

¹² Table 1-2 in Appendix 5

¹³ See Figures 1 and 6 in Appendix 3 for Bank of America VaR forecasts graphs.

report of Bank of America, where they state that since they use 3 years of historical observations, this includes the volatile 4th quarter of 2008 for the VaR forecasts for 2009-2011. In the context of lower market volatility during these years the magnitude of the largest trading losses has been lower than would be expected based on the VaR measures. Although we do not have the actual VaR series but only the number of violations for the in-house model, we can infer that the backtesting techniques we used would probably either reject this model for overestimating the market risk or could not be applicable on such a small number of violations.

5.1.2. Mimicking portfolio

Unlike the actual trading portfolio of Bank of America, the mimicking portfolio is more sensitive to the VaR model specification. Thus all three parametric GARCH models under the normal distribution are rejected by the RiskMap as well as by the Christoffersen unconditional test. They do pass the GMM-duration test, but with very small probabilities.

Their poor performance in backtesting can also be explained through the small value of average VaR, which points toward an underestimation of risk by these models. In this case the VWHS models are the ones that return the highest average VaR.¹⁴

The models under the t-distribution seem to have better results since the parametric GARCH and asymmetric GARCH models are not rejected by any of the three backtesting techniques.

The parametric MRS is in the same situation in the sense that the MRS VaR under the normal distribution is rejected by all three backtesting statistics while for the t-distribution is only rejected by the risk map.

The non-parametric models seem to have a better performance in forecasting VaR for the synthetic portfolio. The Christoffersen independence and conditional coverage tests reject the VWHS GARCH models under both the normal and the t-distribution, but neither the GMM nor the RiskMap invalidate the model. The asymmetric GARCH VWHS models pass all the three backtests.

The MRS VWHS is rejected by the independence tests for both the GMM and the Christoffersen frameworks, but passes the conditional coverage tests, as well as the RiskMap.

¹⁴ Table 1 and 3 in Appendix 5

The HS-VIX method also yields suitable results, since it is not rejected by either backtest. We can also infer that the volatility of the portfolio is better correlated to the VIX index than the VSTOXX and VDAX, since the GMM independence test rejects the latter two VWHS models.

Table 3: Backtesting results for the synthetic portfolio

Model		N	N'	Christoffersen			GMM duration			Risk Map			
				LR _{UC}	LR _{IND}	LR _{CC}	J _{UC}	J _{IND} (5)	J _{CC} (5)	LR _{99%}	LR _{99.8%}	LR _{MUC}	
VIX - HS	VIX	10	5	0,0023 (0,9618)	0,2051 (0,6506)	0,2074 (0,9015)	0,1214 (0,6505)	4,4337 (0,0539)	4,3385 (0,1735)	0,0020 (0,9643)	3,2572 (0,0711)	4,4649 (0,1073)	
	VSTOXX	11	5	0,1307 (0,7177)	0,2485 (0,6182)	0,3791 (0,8273)	0,3891 (0,6037)	7,9320 <i>(0,0085)</i>	7,7030 (0,0605)	0,1283 (0,7202)	3,2572 (0,0711)	3,7422 (0,1540)	
	VDAX	14	5	1,5621 (0,2114)	0,4037 (0,5252)	1,9658 (0,3742)	1,7860 (0,1578)	21,0731 <i>(0,0002)</i>	16,5670 (0,0144)	1,5536 (0,2126)	3,2572 (0,0711)	3,4154 (0,1813)	
Normal	Volatility Weighted HS	GARCH	11	2	1,6088 (0,2047)	13,3743 <i>(0,0003)</i>	14,9831 <i>(0,0006)</i>	1,8358 (0,1805)	6,0339 (0,0160)	4,7435 (0,1313)	1,5987 (0,2061)	0,1734 (0,6771)	1,6220 (0,4444)
		EGARCH	9	4	0,3583 (0,5494)	0,2238 (0,6362)	0,5821 (0,7475)	0,7201 (0,3675)	0,2272 (0,9491)	0,7958 (0,8045)	0,3537 (0,5520)	2,9639 (0,0851)	3,0954 (0,2127)
		GJR	9	3	0,3583 (0,5494)	0,2238 (0,6362)	0,5821 (0,7475)	0,7201 (0,4034)	6,6305 (0,0127)	5,4281 (0,1043)	0,3537 (0,5520)	1,2275 (0,2679)	1,2308 (0,5404)
	t	GARCH	11	3	1,6088 (0,2047)	13,3743 <i>(0,0003)</i>	14,9831 <i>(0,0006)</i>	1,8358 (0,1738)	6,0339 (0,0198)	4,7435 (0,1324)	1,5987 (0,2061)	1,2275 (0,2679)	1,9347 (0,3801)
		EGARCH	9	5	0,3583 (0,5494)	0,2238 (0,6362)	0,5821 (0,7475)	0,7201 (0,3906)	0,2272 (0,9435)	0,7958 (0,7996)	0,3537 (0,5520)	5,2084 (0,0225)	5,8680 (0,0532)
		GJR	10	3	0,8820 (0,3476)	0,2766 (0,5989)	1,1587 (0,5603)	1,2368 (0,2443)	4,0494 (0,0554)	4,4688 (0,1496)	0,8747 (0,3497)	1,2275 (0,2679)	1,4380 (0,4872)
Normal	Volatility Weighted HS	GARCH	19	9	6,7509 <i>(0,0094)</i>	8,1367 <i>(0,0043)</i>	14,8876 <i>(0,0006)</i>	5,2032 (0,0139)	3,9899 (0,0697)	9,4496 (0,0395)	6,7321 <i>(0,0095)</i>	13,3215 <i>(0,0003)</i>	13,8779 <i>(0,0010)</i>
		EGARCH	22	16	11,2091 <i>(0,0008)</i>	0,4284 (0,5128)	11,6375 <i>(0,0030)</i>	7,5943 (0,0119)	2,5762 (0,1681)	12,0464 (0,0287)	11,1841 <i>(0,0008)</i>	39,1382 <i>(0,0000)</i>	39,5819 <i>(0,0000)</i>
		GJR	20	10	8,1364 <i>(0,0043)</i>	0,6513 (0,4197)	8,7877 (0,0124)	5,9799 (0,0118)	1,1888 (0,4264)	9,8124 (0,0373)	8,1155 <i>(0,0044)</i>	16,4804 <i>(0,0000)</i>	17,0413 <i>(0,0002)</i>
	t	GARCH	16	4	3,2626 (0,0709)	1,2860 (0,2568)	4,5486 (0,1029)	3,0376 (0,0577)	1,3764 (0,3696)	4,0027 (0,1956)	3,2500 (0,0714)	1,6062 (0,2050)	3,4862 (0,1750)
		EGARCH	18	6	5,4727 (0,0193)	0,9338 (0,3339)	6,4065 (0,0406)	4,4510 (0,0433)	3,0307 (0,1215)	6,6360 (0,0757)	5,4559 (0,0195)	5,3130 (0,0212)	7,2101 (0,0272)
		GJR	14	5	1,5621 (0,2114)	1,7225 (0,1894)	3,2846 (0,1935)	1,7860 (0,1626)	4,3795 (0,0565)	6,6840 (0,0779)	1,5536 (0,2126)	3,2572 (0,0711)	3,4154 (0,1813)
Normal	MRS parametric	27	13	20,4551 <i>(0,0000)</i>	7,8348 <i>(0,0051)</i>	28,2899 <i>(0,0000)</i>	11,8839 <i>(0,0035)</i>	5,3743 (0,0321)	26,6968 <i>(0,0097)</i>	20,4197 <i>(0,0000)</i>	27,1015 <i>(0,0000)</i>	31,1202 <i>(0,0000)</i>	
		10	5	0,8820 (0,3476)	7,7715 <i>(0,0053)</i>	8,6535 (0,0132)	1,2368 (0,2110)	8,5613 <i>(0,0041)</i>	9,1460 (0,0371)	0,8747 (0,3497)	5,2084 (0,0225)	5,3375 (0,0693)	
	t	MRS parametric	13	8	0,9246 (0,3363)	1,9787 (0,1595)	2,9033 (0,2342)	1,2426 (0,1966)	2,9537 (0,1334)	4,4613 (0,1701)	0,9182 (0,3380)	10,3873 <i>(0,0013)</i>	11,5774 <i>(0,0031)</i>
		MRS vwhs	10	4	0,8820 (0,3476)	7,7715 <i>(0,0053)</i>	8,6535 (0,0132)	1,2368 (0,2216)	8,6643 <i>(0,0048)</i>	9,2184 (0,0392)	0,8747 (0,3497)	2,9639 (0,0851)	2,9677 (0,2268)

Note: N – number of violations, N' – number of super exceptions (expected N and N', given the 99% confidence level are equal to 8 and 2, for a 3-year sample, and to 10 and 2 for a 4-year sample). Figures in parentheses represent the p-values associated with the test statistics. The probabilities in *Italic* denote a model rejection.

5.1.3. Danske Bank

When assessing the performance of the seventeen models used to compute Value-at-Risk for Danske Bank, it can be observed that the parametric models assuming a student- t distribution are more efficient than those working under a normal distribution, since the latter are rejected by both the Risk Map and the Christoffersen test, while for the GMM-duration procedure the probabilities associated with the test statistics are fairly close to the rejection barrier. Additionally, in case of the Markov Regime Switching (MRS) model with a normal distribution all three backtesting procedures reject the null hypothesis of an adequate model at the 99% confidence level, as this model produces the largest number of exceptions both for the 99% VaR and for the 99.8% VaR.

The volatility weighted historical simulation performs well under both distributional assumptions and with volatility calculations based on three different GARCH-model specifications (i.e. simple GARCH, E-GARCH and GJR-GARCH) but also on the MRS process. This group of models yields a low number of violations as well as super exceptions, assuring thus not only the unconditional coverage and independence of the exceptions but also providing a better account for the magnitude of a possible loss in the bank's trading portfolio.

It is important to mention though that this type of models gave the highest average VaR which could indicate that they are overestimating VaR predictions. Similarly to the previous two portfolio analysis, the models that seem to underestimate the VaR calculations are the parametric ones with a normal distribution.

Although it was expected that the GARCH specifications that take into account the leverage effects characteristic of the financial data to yield better results than the simple GARCH, the performance of these volatility models is quite similar and in some cases, for Danske Bank, the GARCH model even gives better forecasts than EGARCH or GJR-GARCH.

Moreover, the non-parametric HS-VIX returns adequate Value-at-Risk predictions, accommodating conditional coverage under both Christoffersen and GMM-duration tests even though the number of exceptions is close to the rejection limit. Out of these 18 violations only three of them are super exceptions, which mean that the model also takes into consideration the magnitude of losses the bank can incur.

The backtesting results are presented model by model and are aggregated by validating procedure in the table below:

Table 4: Danske Bank - Backtesting results

Model		N	N'	Christoffersen			GMM duration			Risk Map			
				LR _{UC}	LR _{IND}	LR _{CC}	J _{UC}	J _{IND} (5)	J _{CC} (5)	LR _{99%}	LR _{99.8%}	LR _{MUC}	
	HS-VIX (VSTOXX)	18	3	5.4727 (0.0193)	0.6702 (0.4130)	6.1429 (0.0464)	4.4510 (0.0432)	6.2423 (0.0184)	12.7495 (0.0232)	5.4559 (0.0195)	0.4625 (0.4965)	5.5867 (0.0612)	
Normal	Volatility Weighted HS	GARCH	7	1	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.7153)	0.4915 (0.7357)	0.5235 (0.8961)	0.0162 (0.8989)	0.1685 (0.6815)	0.1711 (0.9180)
		EGARCH	9	3	0.3583 (0.5494)	0.2238 (0.6362)	0.5821 (0.7475)	0.7201 (0.3577)	0.8841 (0.5061)	1.3534 (0.6224)	0.3537 (0.5520)	1.2275 (0.2679)	1.2308 (0.5404)
		GJR	6	1	0.2598 (0.6102)	4.4482 (0.0349)	4.7081 (0.0950)	0.0157 (0.9350)	3.4997 (0.0853)	6.2514 (0.0800)	0.2635 (0.6077)	0.1685 (0.6815)	0.3071 (0.8577)
t	Volatility Weighted HS	GARCH	9	1	0.3583 (0.5494)	0.2238 (0.6362)	0.5821 (0.7475)	0.7201 (0.3683)	2.9053 (0.1174)	1.2952 (0.5081)	0.3537 (0.5520)	0.1685 (0.6815)	0.8639 (0.6492)
		EGARCH	10	2	0.8820 (0.3476)	0.2766 (0.5989)	1.1587 (0.5603)	1.2368 (0.2101)	0.8841 (0.5151)	1.8552 (0.5016)	0.8747 (0.3497)	0.1734 (0.6771)	0.8747 (0.6458)
		GJR	8	1	0.0601 (0.8064)	0.1766 (0.6744)	0.2366 (0.8884)	0.3130 (0.5195)	2.2118 (0.1895)	1.9822 (0.4703)	0.0582 (0.8093)	0.1685 (0.6815)	0.3728 (0.8299)
Normal	GARCH	19	13	6.7509 <i>(0.0094)</i>	0.7475 (0.3873)	7.4984 (0.0235)	5.2032 (0.0152)	0.1310 (0.9933)	6.9678 (0.0693)	6.7321 <i>(0.0095)</i>	27.1015 <i>(0.0000)</i>	27.5563 <i>(0.0000)</i>	
		EGARCH	21	12	9.6241 <i>(0.0019)</i>	0.9150 (0.3388)	10.5391 <i>(0.0051)</i>	6.7779 (0.0139)	0.5976 (0.6881)	10.6624 (0.0350)	9.6011 <i>(0.0019)</i>	23.3871 <i>(0.0000)</i>	23.5621 <i>(0.0000)</i>
		GJR	26	10	18.4415 <i>(0.0000)</i>	0.1334 (0.7149)	18.5749 <i>(0.0001)</i>	11.0035 <i>(0.0027)</i>	4.5083 (0.0486)	19.4268 (0.0115)	18.4081 <i>(0.0000)</i>	16.4804 <i>(0.0000)</i>	23.0910 <i>(0.0000)</i>
t	GARCH	14	4	1.5621 (0.2114)	0.4037 (0.5252)	1.9658 (0.3742)	1.7860 (0.1900)	0.6934 (0.6334)	2.0891 (0.4571)	1.5536 (0.2126)	1.6062 (0.2050)	2.1404 (0.3429)	
		EGARCH	14	5	1.5621 (0.2114)	0.4037 (0.5252)	1.9658 (0.3742)	1.7860 (0.1701)	1.2441 (0.4055)	3.2967 (0.2689)	1.5536 (0.2126)	3.2572 (0.0711)	3.4154 (0.1813)
		GJR	16	5	3.2626 (0.0709)	0.5284 (0.4673)	3.7910 (0.1502)	3.0376 (0.0690)	0.4444 (0.7808)	3.6517 (0.2267)	3.2500 (0.0714)	3.2572 (0.0711)	4.3788 (0.1120)
Normal	MRS parametric	33	16	34.0504 <i>(0.0000)</i>	0.0111 (0.9160)	34.0615 <i>(0.0000)</i>	17.3269 <i>(0.0006)</i>	2.2724 (0.2037)	37.8912 <i>(0.0068)</i>	34.0023 <i>(0.0000)</i>	39.1382 <i>(0.0000)</i>	47.3738 <i>(0.0000)</i>	
		MRS vwhs	7	3	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.7293)	0.8600 (0.5248)	0.9037 (0.7514)	0.0162 (0.8989)	1.2275 (0.2679)	1.8972 (0.3873)
		MRS parametric	14	4	1.5621 (0.2114)	1.7225 (0.1894)	3.2846 (0.1935)	1.7860 (0.1558)	0.6706 (0.6439)	2.7951 (0.3258)	1.5536 (0.2126)	1.6062 (0.2050)	2.1404 (0.3429)
t	MRS vwhs	7	2	0.0152 (0.9018)	0.1350 (0.7133)	0.1502 (0.9276)	0.0567 (0.8129)	0.3091 (0.8829)	0.2991 (0.9635)	0.0162 (0.8989)	0.1734 (0.6771)	0.3096 (0.8566)	

Note: N – number of violations, N' – number of super exceptions (expected N and N', given the 99% confidence level are equal to 8 and 2, for a 3-year sample, and to 10 and 2 for a 4-year sample). Figures in parentheses represent the p-values associated with the test statistics. The probabilities in *italic* denote a model rejection.

A simple way to interpret the backtesting results for Danske Bank is provided by the Risk Map in the appendix¹⁵ where it can be easily noticed that 13 of the models lie in the non-rejection area of both likelihood ratio tests whereas the other 4 are situated in the red zone that signifies rejection of the two test statistics.

¹⁵ See Figure 1 and 2 in Appendix 4

Danske Bank uses its own VaR model that is based on Monte Carlo simulations and employs 2 years of historical market data. Each calculation is based on 1000 scenarios that represent future outcomes of the risk factors and which are then used to determine the empirical loss distribution that serves for the estimation of VaR. According to the bank's annual reports, the number of violations over the 2008-2011 time span¹⁶ has totaled 18, out of which 10 occurred in 2008 in the context of high market turbulence due to the financial crisis, and therefore the rest of 8 from 2009 to 2011. Comparing only by the number of violations, since the VaR series derived by Danske Bank's model is not available, we can conclude that all the semi-parametric models we estimated have given a similar number of exceptions. The regime-switching GARCH based VWHS gave 7 exceptions both under the normal and the t -distribution and also passed all 3 backtests, suggesting that the magnitude of the exceptions was not significantly large and also the independence of these exceptions. Therefore, these models are comparable in accuracy to the bank's actual model.

5.1.4. Deutsche Bank

A similar pattern as in the case of Danske Bank can also be observed when analyzing the results obtained for Deutsche Bank, although for this financial institution more model rejections are encountered.

HS-VIX still provides a good unconditional coverage as shown by the Christoffersen and GMM-duration tests and a good coverage of excessive losses, since Risk Map's test statistics cannot be rejected at the 99% confidence level. However, the independence hypothesis is not respected as the two independence tests conducted dismiss the assumption of uncorrelated violations in the hit sequence. Moreover, the Christoffersen test also rejects the conditional coverage hypothesis.

The Markov Regime Switching models' performance is confirmed by two of the backtesting techniques, namely the GMM-duration and the Risk Map, while the Christoffersen test asserts that the parametric MRS (normal distribution) and the t -distributed volatility weighted MRS do not give a proper conditional coverage.

When it comes to the parametric VaR, just as before, the efficiency of the models under the assumption of a normal distribution is rather substandard as the GARCH and the EGARCH fail to account for the number of extreme losses. Furthermore, the GJR-GARCH is

¹⁶ A disaggregation of the number of violation for each year is presented in Appendix 2 – Table 7

rejected by the Christoffersen method. As expected, the parametric models with a t -distribution give better results as they all pass the conditional coverage of all three backtesting techniques. Interestingly, this is not the case with the volatility weighted historical simulation where the roles seem to be reversed and now the normally distributed models have a better performance than those with a student- t distribution, since in the case of the latter only the GJR-GARCH model is not rejected by any of the three backtests.

When ranking the models based on the average VaR¹⁷ values it can be noticed that the t -distributed parametric GARCH and EGARCH models not only do they provide adequate coverage in backtesting but they also give a low average VaR value, while the normally distributed versions continue to miscalculate the risk exposure and yield too many exceptions. For this financial institution the model that returns the highest average VaR is the non-parametric HS-VIX.

Similarly to the previous bank the asymmetric GARCH models do not yield significantly better VaR forecasts than the simple GARCH, with a slight advantage for the GJR-GARCH which seems to have a better fit on the data than in the previous case.

The results above can be translated into a Risk Map¹⁸ where, after placing the coordinates of each model, it can be seen that six of the models used are situated in the rejection area while the rest are located in the lower left quadrant, non-rejection area of the map.

Also notable for this bank is that the independence test statistic is rejected in 13 of the 17 models, implying that the violations are not independently distributed. Furthermore the validity of a model as a whole is rejected in 8 of the cases by at least one of the backtesting techniques.

Deutsche Bank's in-house model lead to a number of 35 exceptions in 2008 and a total of 6 exceptions for the 2009-2011 time span¹⁹. None of the parametric models we estimated gave more than 18 violations for the 4 year period, not even the models that were rejected by the backtesting. In particular the MRS GARCH- t and the GJR- t parametric VaRs gave the smallest number of exceptions out of all the parametric and non-parametric models employed. This leads to the conclusion that Deutsche Bank's model, even if it was supposed to be

¹⁷ Table 5 in Appendix 5

¹⁸ Figures 1-2 in Appendix 4

¹⁹ Appendix 2 – Table 7

tailored to the characteristics of its trading portfolio, was far from able to correctly capture the extreme market movements that occurred in 2008. Therefore, using a model based on time-varying volatilities and possibly regime-switching could have prevented such a large underestimation of the bank's market risk.

Table 5: Backtesting results - Deutsche Bank

Model		N	N'	Christoffersen			GMM duration			Risk Map			
				LR _{UC}	LR _{IND}	LR _{CC}	J _{UC}	J _{IND} (5)	J _{CC} (5)	LR _{99%}	LR _{99.8%}	LR _{MUC}	
	HS-VIX (VDAX)	16	6	3.2626 (0.0709)	10.2189 <i>(0.0014)</i>	13.4815 <i>(0.0012)</i>	3.0376 (0.0737)	18.8654 <i>(0.0004)</i>	20.9975 (0.0138)	3.2500 (0.0714)	5.3130 (0.0212)	5.8561 (0.0535)	
Normal	Volatility Weighted HS	GARCH	14	4	4.8398 (0.0278)	0.5452 (0.4603)	5.3850 (0.0677)	3.9615 (0.0338)	11.9211 <i>(0.0011)</i>	16.1894 (0.0157)	4.8213 (0.0281)	2.9639 (0.0851)	5.4082 (0.0669)
		EGARCH	14	4	4.8398 (0.0278)	0.5452 (0.4603)	5.3850 (0.0677)	3.9615 (0.0345)	11.9198 <i>(0.0014)</i>	16.1644 (0.0146)	4.8213 (0.0281)	2.9639 (0.0851)	5.4082 (0.0669)
		GJR	12	7	2.5205 (0.1124)	6.2610 (0.0123)	8.7815 (0.0124)	2.4980 (0.1105)	8.4856 <i>(0.0046)</i>	11.1152 (0.0274)	2.5076 (0.1133)	10.8460 <i>(0.0010)</i>	10.9705 <i>(0.0041)</i>
t	Volatility Weighted HS	GARCH	16	4	7.7436 <i>(0.0054)</i>	4.0440 (0.0443)	11.7876 <i>(0.0028)</i>	5.5561 (0.0266)	28.4989 <i>(0.0001)</i>	30.4735 <i>(0.0068)</i>	7.7196 <i>(0.0055)</i>	2.9639 (0.0851)	7.9558 (0.0187)
		EGARCH	16	7	7.7436 <i>(0.0054)</i>	4.0440 (0.0443)	11.7876 <i>(0.0028)</i>	5.5561 (0.0277)	26.9575 <i>(0.0001)</i>	29.3816 <i>(0.0053)</i>	7.7196 <i>(0.0055)</i>	10.8460 <i>(0.0010)</i>	12.3383 <i>(0.0021)</i>
		GJR	14	5	4.8398 (0.0278)	1.2532 (0.2629)	6.0930 (0.0475)	3.9615 (0.0409)	14.5823 <i>(0.0006)</i>	18.6368 (0.0144)	4.8213 (0.0281)	5.2084 (0.0225)	6.6831 (0.0354)
Normal	Volatility Weighted HS	GARCH	16	8	3.2626 (0.0709)	1.2860 (0.2568)	4.5486 (0.1029)	3.0376 (0.0583)	22.9643 <i>(0.0002)</i>	22.0943 (0.0107)	3.2500 (0.0714)	10.3873 <i>(0.0013)</i>	10.3906 <i>(0.0055)</i>
		EGARCH	19	11	6.7509 <i>(0.0094)</i>	0.7845 (0.3758)	7.5354 (0.0231)	5.2032 (0.0165)	9.7153 <i>(0.0045)</i>	16.4596 (0.0172)	6.7321 <i>(0.0095)</i>	19.8417 <i>(0.0000)</i>	19.8461 <i>(0.0000)</i>
		GJR	16	7	3.2626 (0.0709)	10.2189 <i>(0.0014)</i>	13.4815 <i>(0.0012)</i>	3.0376 (0.0707)	13.1678 <i>(0.0015)</i>	16.2462 (0.0183)	3.2500 (0.0714)	7.7058 <i>(0.0055)</i>	7.8686 (0.0196)
t	Volatility Weighted HS	GARCH	11	4	0.1307 (0.7177)	2.5856 (0.1078)	2.7163 (0.2571)	0.3891 (0.5587)	11.6645 <i>(0.0025)</i>	11.7651 (0.0290)	0.1283 (0.7202)	1.6062 (0.2050)	1.7072 (0.4259)
		EGARCH	14	6	1.5621 (0.2114)	1.7225 (0.1894)	3.2846 (0.1935)	1.7860 (0.1939)	14.6634 <i>(0.0013)</i>	15.8585 (0.0203)	1.5536 (0.2126)	5.3130 (0.0212)	5.3157 (0.0701)
		GJR	7	1	0.9265 (0.3358)	0.1002 (0.7516)	1.0267 (0.5985)	0.4321 (0.4785)	2.6423 (0.1607)	3.9337 (0.2084)	0.9323 (0.3343)	0.5869 (0.4436)	1.0873 (0.5806)
Normal	MRS parametric	MRS parametric	18	5	5.4727 (0.0193)	4.2075 (0.0402)	9.6802 <i>(0.0079)</i>	4.4510 (0.0412)	4.7118 (0.0458)	11.4366 (0.0291)	5.4559 (0.0195)	3.2572 (0.0711)	6.0817 (0.0478)
		MRS vwhs	11	4	1.6088 (0.2047)	0.3352 (0.5626)	1.9440 (0.3783)	1.8358 (0.1561)	2.0302 (0.2176)	4.2336 (0.1674)	1.5987 (0.2061)	2.9639 (0.0851)	3.1776 (0.2042)
t	MRS parametric	MRS parametric	7	3	0.9265 (0.3358)	0.1002 (0.7516)	1.0267 (0.5985)	0.4321 (0.5051)	1.5741 (0.3147)	2.6603 (0.3605)	0.9323 (0.3343)	0.4625 (0.4965)	2.8134 (0.2450)
		MRS vwhs	14	4	4.8398 (0.0278)	5.0446 (0.0247)	9.8844 <i>(0.0071)</i>	3.9615 (0.0310)	20.8688 <i>(0.0002)</i>	22.9474 (0.0103)	4.8213 (0.0281)	2.9639 (0.0851)	5.4082 (0.0669)

Note: N – number of violations, N' – number of super exceptions (expected N and N', given the 99% confidence level are equal to 8 and 2, for a 3-year sample, and to 10 and 2 for a 4-year sample). Figures in parentheses represent the p-values associated with the test statistics. The probabilities in *Italic* denote a model rejection.

The semi-parametric and non-parametric methods give a larger number of exceptions over the 2009-2011 period. It is, however, difficult to compare their performance to that of the parametric models, since the backtesting period for the latter does not include the volatile

2008 year. Deutsche Bank's model is based on Monte Carlo simulations using on one year of historical data and thus uses the persistent effect of the 2008 volatile period to forecasts risk for 2009 which was a less volatile period and thus possibly led to an over-estimation of the VaRs. Compared to the bank's in-house model in terms of number of violations, it could be said that the semi-parametric models were less accurate in predicting risk. Nonetheless, taking this into consideration as well as the fact that the fact that some of these latter models were not rejected by the backtests, we can infer that the methods based on time varying volatilities have had an overall better performance than the model used by Deutsche Bank.

5.1.5. Swedbank

The most interesting observation that can be made when looking at the backtesting results for Swedbank displayed in the table below is that the parametric Markov Regime Switching model appears to overestimate the risk to which the bank is exposed, as both the total number of violations as well as the number of super exceptions are quite low. Thus, the parametric MRS under a t -distribution is dismissed by all three backtesting methods whereas the model assuming a normal distribution is only rejected by the GMM-duration approach.

The affirmation that the parametric MRS models over-estimate the market risk exposures is also backed by the average VaR values which are the highest for these models. Same as before the parametric GARCH and EGARCH under a normal distribution return the lowest average VaR values, the only difference being that this time the former is not rejected by any of the validating procedures.

Just as was the case with the previous two banks, the parametric VaR forecasts based on various GARCH models work better when assuming a t -distribution, since for these models all three backtesting procedures assess that they give a correct conditional coverage. Regarding the normally distributed parametric VaRs, the GJR-GARCH is rejected by the Risk Map as the model does not properly account for the size of losses, but not by the Christoffersen and GMM-duration tests. This is not the case however with the EGARCH model where all three backtesting methods clearly discard the model. Therefore, it can be concluded that for this family of models the simple GARCH models yield the best predictions both for normally and t -distributed values. This result is not in line with the expectation that asymmetric models should provide better outputs as it was clear from the data analysis that Swedbank's returns are characterized by volatility clustering.

Table 6: Backtesting results - Swedbank

Model		N	N'	Christoffersen			GMM duration			Risk Map		
				LR _{UC}	LR _{IND}	LR _{CC}	J _{UC}	J _{IND} (5)	J _{CC} (5)	LR _{99%}	LR _{99.8%}	LR _{MUC}
	HS-VIX (VSTOXX)	14	5	1.5621 (0.2114)	0.4037 (0.5252)	1.9658 (0.3742)	1.7860 (0.1779)	9.8882 <i>(0.0046)</i>	11.6567 (0.0278)	1.5536 (0.2126)	3.2572 (0.0711)	3.4154 (0.1813)
Normal	GARCH	14	3	4.9516 (0.0261)	5.3020 (0.0213)	10.2535 <i>(0.0059)</i>	3.2679 (0.0522)	5.7445 (0.0188)	9.9397 (0.0347)	4.9328 (0.0264)	1.2528 (0.2630)	4.9503 (0.0841)
	EGARCH	12	7	2.5988 (0.1069)	0.3690 (0.5436)	2.9678 (0.2268)	1.8832 (0.1580)	9.3388 <i>(0.0047)</i>	9.8599 (0.0359)	2.5856 (0.1078)	10.9375 <i>(0.0009)</i>	11.0485 <i>(0.0040)</i>
	GJR	9	3	0.3866 (0.5341)	0.2004 (0.6544)	0.5870 (0.7456)	0.3359 (0.5319)	0.7453 (0.5801)	0.5256 (0.8999)	0.3818 (0.5366)	1.2528 (0.2630)	1.2589 (0.5329)
t	GARCH	14	4	4.9516 (0.0261)	5.3020 (0.0213)	10.2535 <i>(0.0059)</i>	3.2679 (0.0521)	5.7445 (0.0193)	9.9397 (0.0328)	4.9328 (0.0264)	3.0057 (0.0830)	5.5196 (0.0633)
	EGARCH	13	5	3.6968 (0.0545)	0.4367 (0.5087)	4.1335 (0.1266)	2.5511 (0.1320)	5.3740 (0.0239)	8.1582 (0.0486)	3.6808 (0.0550)	5.2668 (0.0217)	6.0222 (0.0492)
	GJR	10	4	0.9269 (0.3357)	0.2509 (0.6165)	1.1778 (0.5549)	0.7528 (0.4247)	1.2145 (0.3930)	0.9743 (0.7391)	0.9194 (0.3376)	3.0057 (0.0830)	3.0124 (0.2218)
Normal	GARCH	12	4	0.5666 (0.4516)	0.3041 (0.5813)	0.8707 (0.6470)	0.9035 (0.4024)	1.2673 (0.3854)	1.5063 (0.5963)	0.5615 (0.4537)	1.7163 (0.1902)	1.7310 (0.4208)
	EGARCH	22	10	11.8771 <i>(0.0006)</i>	0.4003 (0.5270)	12.2773 <i>(0.0022)</i>	7.8976 <i>(0.0049)</i>	16.9061 <i>(0.0004)</i>	26.1649 <i>(0.0093)</i>	11.8508 <i>(0.0006)</i>	16.9135 <i>(0.0000)</i>	19.0786 <i>(0.0001)</i>
	GJR	15	9	2.6308 (0.1048)	1.4502 (0.2285)	4.0810 (0.1300)	2.5939 (0.0797)	4.7088 (0.0476)	8.0243 (0.0562)	2.6194 (0.1056)	13.7006 <i>(0.0002)</i>	14.0766 <i>(0.0009)</i>
t	GARCH	7	1	0.7797 (0.3772)	0.1029 (0.7483)	0.8827 (0.6432)	0.3192 (0.6168)	1.3934 (0.3582)	2.7408 (0.3438)	0.7852 (0.3756)	0.5362 (0.4640)	0.9401 (0.6250)
	EGARCH	12	4	0.5666 (0.4516)	0.3041 (0.5813)	0.8707 (0.6470)	0.9035 (0.3507)	6.2982 (0.0182)	7.1989 (0.0669)	0.5615 (0.4537)	1.7163 (0.1902)	1.7310 (0.4208)
	GJR	15	2	2.6308 (0.1048)	1.4502 (0.2285)	4.0810 (0.1300)	2.5939 (0.0795)	33.0281 <i>(0.0001)</i>	17.0280 (0.0149)	2.6194 (0.1056)	0.0033 (0.9542)	3.0786 (0.2145)
Normal	MRS parametric	4	2	4.4540 (0.0348)	0.0328 (0.8562)	4.4868 (0.1061)	4.6352 (0.0335)	5.0741 (0.0372)	94.8129 <i>(0.0014)</i>	4.4659 (0.0346)	0.0008 (0.9773)	6.2510 (0.0439)
	MRS vwhs	13	5	3.6968 (0.0545)	1.6110 (0.2044)	5.3078 (0.0704)	2.5511 (0.1329)	4.4934 (0.0429)	7.5241 (0.0570)	3.6808 (0.0550)	5.2668 (0.0217)	6.0222 (0.0492)
t	MRS parametric	2	1	9.2896 <i>(0.0023)</i>	0.0082 (0.9279)	9.2978 <i>(0.0096)</i>	15.5233 <i>(0.0024)</i>	0.7023 (0.6192)	147.5770 <i>(0.0014)</i>	9.3056 <i>(0.0023)</i>	0.5751 (0.4483)	10.1982 <i>(0.0061)</i>
	MRS vwhs	12	4	2.5988 (0.1069)	1.8979 (0.1683)	4.4967 (0.1056)	1.8832 (0.1754)	5.8569 (0.0194)	6.8057 (0.0668)	2.5856 (0.1078)	3.0057 (0.0830)	3.7551 (0.1530)

Note: N – number of violations, N' – number of super exceptions (expected N and N', given the 99% confidence level are equal to 8 and 2, for a 3-year sample, and to 10 and 2 for a 4-year sample). Figures in parentheses represent the p-values associated with the test statistics. The probabilities in *Italic* denote a model rejection.

The volatility weighted historical simulation models return quite adequate VaR forecasts as five of the eight models are not rejected by any of the three validating methods and the other three models are dismissed only by either the Christoffersen or the Risk Map tests. Surprisingly, the GARCH model is not supported by the backtesting results of the Christoffersen test for both distributional assumptions. Furthermore, the VaR predictions based on volatilities yielded by the Markov Regime Switching process are backed by the three methods used to attest the accuracy of the VaR models, when employing a normal distribution as well as a *t*-distribution.

The HS-VIX yields 14 violations of which 5 are super exceptions, which means that the model provides a consistent unconditional coverage and also accounts for the magnitude of losses. Moreover, both the Christoffersen and the GMM-duration tests results imply that the conditional coverage is adequate as well.

Only one of the models used for Swedbank is placed the rejection area of both 99% and 99.8% VaRs when looking at the Risk Map²⁰, whereas the other three model rejections either lie in the non-rejection zone of the exceptions test or the super exceptions test.

Similar to the previous banks we analyzed, the Swedbank employs a VaR model based on Monte Carlo simulations with one year of historical data to measure its market risk.

The in-house model of Swedbank gave 4 violations in 2008, zero in 2009 and 2010 and 2 in 2011 according to their annual reports²¹, a total of 6 exceptions for the 4 year time span. In comparison, out of the models we used, only the parametric MRS-GARCH models under both the normal and the t distribution give a smaller number of exceptions. However, these models have been rejected by some of the backtests, since they tend to overestimate risk, which would lead to larger than optimal capital charges for the institution. Hence, based on the similarity between the number of exceptions given by these models, we can infer that it is possible that the actual model used by Swedbank may also lead to an over-estimation of risk. Using an EGARCH under the t-distribution, either parametric or semi-parametric model, the MRS-GARCH-t historical simulation or the HS-VSTOXX model could provide more balanced risk forecasts than the in-house model for Swedbank.

5.2. Model analysis

5.2.1. Parametric models

The results concerning the performance of the parametric models are very different across banks and also differ at times between the backtesting methods used.

Normal distribution. Under the normal distribution assumption, most of the models were rejected by the backtesting frameworks for the five banks. For Bank of America, the EGARCH model was rejected by both the RiskMap and the Christoffersen test, while the GARCH and GJR models passed the test. However, in the case of the synthetic portfolio and Danske Bank, all three GARCH models were rejected by the RiskMap and the Christoffersen

²⁰ Figures 1 and 2 in Appendix 4

²¹ Appendix 2 – Table 7

method. While the GMM framework did not reject these models at a 99% confidence level, the probabilities for non rejection were very small.

The MRS-GARCH model is also rejected by all three backtesting frameworks in the case of the synthetic portfolio and Danske Bank. For Bank of America it is not rejected by either backtesting, while for Deutsche Bank and Swedbank it is rejected by the Christoffersen and the GMM respectively, while it passes the other tests.

T-distribution. As expected, these models show improved performance compared to the models under the normal distribution assumption. All three models pass all three backtesting frameworks. The GJR mode in the case of Bank of America and both the GARCH and EGARCH model for Deutsche Bank are very close to rejection by the GMM-duration statistics, but are still in the non-rejection area. Furthermore, neither the RiskMap nor the Christoffersen test reject the model. This comes to show that the Student-t distribution provides a better fit for the returns of the banks' portfolios, which was to be expected, since the descriptive statistics show that the data series exhibit significant excess kurtosis.

The MRS-GARCH passes the backtesting for Bank of America, Danske Bank and Deutsche Bank, is rejected by the RiskMap for the synthetic portfolio, and does not pass either test in the case of Swedbank.

5.2.2. Non-parametric and semi-parametric models

Semi-parametric models based on the volatility estimates seem to yield better results than the parametric ones. In the case of Bank of America and Danske Bank all three GARCH models under both the normal and t-distribution pass all the backtests. Nonetheless, Deutsche Bank's portfolio does not seem to be provided with a very good fit by the GARCH and EGARCH models under the t-distribution, these being rejected by the Christoffersen conditional and unconditional coverage statistics, the GMM duration test as well as the RiskMap for VaR at 99% level. The GMM independence test rejects all three VWHS models computed with GARCH volatilities derived under a normal distribution, while the VWHS-GJR-normal is also rejected by RiskMap. The RiskMap and the Christoffersen test, however, do not reject the VWHS GARCH and EGARCH models under the normal distribution.

For Swedbank all models pass the tests but for VWHS GARCH under both the normal and t-distribution, which fails the Christoffersen test and VWHS EGARCH-n which is rejected by both the RiskMap and the GMM independence test. Similarly, the VWHS

GARCH under both distributional assumptions is rejected by the Christoffersen test but not by the others.

As far as VWHS VaR computed with volatilities from MRS GARCH models go, their performance is improved compared to parametric VaRs under the same volatility model. Both VWHS MRS normal and t pass the backtests for Bank of America, Danske Bank and Swedbank. In the case of Deutsche Bank the GMM independence and the Christoffersen conditional coverage tests reject the VWHS MRS under the t-distribution. In the case of the synthetic portfolio, both the Christoffersen and the GMM independence tests reject the VWHS MRS model under both the normal and the t-distribution, but the models are not rejected by the conditional coverage statistics or the RiskMap. The latter comes to show that the VWHS models, while may lead to correlated violations, the magnitude of these exceptions is not very large, proven by the small number of super-exceptions.

The HS-VIX model seems to have the most consistent performance over all the banks' trading portfolios. For Bank of America (VIX) and Danske Bank (VSTOXX) the model passes all three backtesting techniques. The same goes for the synthetic portfolio VaR computed with implied volatilities given by the VIX index. Using the VSTOXX and VDAX index also lead to good results for this portfolio, since, although the VaRs computed with said volatility indexes were rejected by the GMM independence test, they passed the conditional coverage test as well as the Christoffersen and RiskMap backtests. These results come to show a good correlation between the volatilities of the portfolios and the volatility indices. As expected, both the real and synthetic portfolio for Bank of America have a good correlation with the VIX index, while Danske Bank's portfolio is better correlated with the European VSTOXX portfolio. Similarly, Swedbank's portfolio is also well correlated with the VSTOXX index, although the HS-VIX model is rejected under the GMM independence test. Although counterintuitive, using the VDAX volatility index in the case of Deutsche Bank did not pass the Christoffersen backtest or the GMM independence test. We can thus infer that the trading portfolio of Deutsche Bank may not be very well correlated with the VDAX index volatility-wise. One drawback we need to mention in the interpretation of these results for the four banks' real trading portfolios is that there is a possibility the P/L data and the indices returns may not be properly aligned. This is due to the procedure that was applied to extract the P/Ls from graphs, since there is no security regarding the number of days presented in each plot or the exact starting day for the observations. The graphs the data was extracted from were not dated for each observation, but only yearly, which means we made an

assumption about the number of observations for each year, which may have resulted in more or less observations than in the real portfolio. This could lead to a misalignment of the P/Ls in time compared to the volatility indices returns, which needs to be taken into account in assessing the accuracy of the VaR estimates. Nonetheless, there is no such problem in the case of the synthetic portfolio, since the data was extracted perfectly aligned with calendaristic dates and therefore the performance of the model is only dependent on the correlation with the volatility index.

The table below summarizes the backtesting results across all the models and all portfolios included in the analysis:

Table 7: Traffic light system to aggregate backtesting results

Model		Bank of America	Mimicking portfolio	Danske Bank	Deutsche Bank	Swedbank	
HS-VIX	VIX			N/A	N/A	N/A	
	VSTOXX	N/A	GMM _{ind}		N/A	GMM _{ind}	
	VDAX	N/A	GMM _{ind}	N/A	CHR GMM _{ind}	N/A	
Normal	Volatility Weighted HS	GARCH	CHR		GMM _{ind}	CHR	
		EGARCH			GMM _{ind}	MAP GMM _{ind}	
		GJR			MAP GMM _{ind}		
		GARCH	CHR			CHR	
		EGARCH					
		GJR			GMM _{ind}		
t	Volatility Weighted HS	GARCH	CHR			CHR	
		EGARCH					
		GJR			GMM _{ind}		
		GARCH	MAP	CHR MAP	CHR _{UC} MAP	MAP GMM _{ind}	
		EGARCH	CHR MAP	CHR MAP	CHR MAP	MAP CHR _{UC} GMM _{ind}	
		GJR	GMM _{ind}	CHR _{UC} MAP	CHR MAP GMM _{UC}	CHR MAP GMM _{ind}	MAP
Normal	Volatility Weighted HS	GARCH			GMM _{ind}		
		EGARCH			GMM _{ind}		
		GJR	GMM _{ind}			GMM _{ind}	
		MRS parametric				CHR	GMM
		MRS vwhs		CHR _{ind} GMM _{ind}			
		MRS parametric	GMM	MAP			
t	Volatility Weighted HS	MRS parametric	GMM	MAP			
		MRS vwhs		CHR _{ind} GMM _{ind}		CHR GMM _{ind}	

Note: N/A – Not Applicable, CHR – Christoffersen test, GMM – GMM-duration approach, MAP – Risk Map, ind – Independence, UC – unconditional coverage, Colors: Green – the model is not rejected by any of the

backtesting methods, Yellow – the model is rejected by one or two of the backtests, Red – the model is rejected by all three validating procedures.

The conclusion that can be drawn from this model analysis is that there is not one single model that has the best performance across all data series. Non-parametric and semi-parametric models based on time varying volatility do tend to outperform the parametric models. The explanation for this lies in the fact that the historical simulation does not require any distributional assumptions for the computing of VaR measures, unlike the parametric models. Therefore, the parametric models may be flawed since the data may not follow an exact normal or t-distribution which can lead to poor risk estimates. The superiority of VWHS is also confirmed by Johansson (2011) in a study on an equally-weighted portfolio of index bonds. The results of his paper showed that VWHS was the most suitable method compared to several other parametric models, due to its ability to capture the volatility dynamics, simplicity and straight-forwardness.

Nevertheless, it is worth mentioning that, unlike the parametric models, the semi-parametric methods have had a out-of-sample backtesting window of 3 years, and did not include the 2008 financial crisis which was characterized by extreme market movements that had a big impact on the trading portfolios of the banks. Thus, there is a possibility for the behavior of these models to change during such turbulent times in a way that would make their performance similar to the parametric models.

In addition, we must keep in mind that the performance of the MRS model can be sensitive to the length of the rolling window used. The 252 daily observations that were used in this study may not have been sufficient for the model to capture the movements from one regime to the other, therefore, reestimating the model on a larger window may have yielded different results. Nonetheless, the restricted availability of data made it impossible to have a very large window, since there is a tradeoff between the length of the in-sample data and the backtesting window. A larger window for estimating the MRS volatilities may have led to better forecasts but it also would have shortened the out-of-sample window which could affect the accuracy of the performance assessment through backtesting. Ideally, at least 10 years of data would have given more precision to the results, with a rolling window of more than 5 years to be able to capture the regimes in volatility dynamics.

6. Conclusions

Value-at-Risk is currently the most popular measure for market risk, its importance increasing dramatically since its introduction in Basel regulatory framework. The predictive accuracy of VaR models has been an ongoing and controversial issue with numerous empirical studies trying to shed light on the performance of this measure over the past years.

The current paper is one of the few studies employing real bank data to estimate the efficiency of VaR models. Furthermore, it is the only paper at this point that tests the performance of Value-at-Risk computed by using volatilities estimated through Markov Regime-Switching GARCH-type models on real bank trading portfolios.

The focus of our research was determining the accuracy of regime-switching time-varying volatility models in quantifying and managing market risk in the context of financial institution. To this purpose, Klaasen's (2002) MRS-GARCH model both under the normal and t-distribution assumptions, was applied in order to calculate VaR for four real trading portfolios, obtained from American and European banks, as well as a mimicking portfolio built with constant weights. VaR measures computed with standard GARCH volatilities were used as a benchmark for comparison of the MRS models. Parametric, non-parametric and semi-parametric approaches were applied to obtain the VaR forecasts. To test the validity of the estimations, three different backtesting techniques were employed, including the newly developed GMM duration and RiskMap methodologies, which account for frequency and the duration between violations and for the magnitude of the exceptions.

The results showed a high variation of the performance of these models depending on the data series they were applied to. Therefore, the characteristics of each portfolio has a large impact on the accuracy of the VaR models, which may be the reason why the Basel framework allows each institution to use its own in-house developed VaR methodology to assess their market risk exposure.

The MRS model has not had a consistent performance over the five data series, and, in many cases was outperformed by the more simplistic GARCH models according to the backtests.

The parametric MRS VaR under the t-distribution passed the backtesting for Bank of America, Danske Bank and Deutsche Bank, was rejected by the RiskMap for the synthetic portfolio, and did not pass either test for Swedbank. The semi-parametric MRS VaR behaved

better, however. Both VWHS MRS normal and t passed the backtests for Bank of America, Danske Bank and Swedbank. In the case of Deutsche Bank the GMM independence and the Christoffersen conditional coverage tests reject the VWHS MRS under the t-distribution. For the synthetic portfolio, both the Christoffersen and the GMM independence tests rejected the VWHS MRS model under both the normal and the t-distribution, but the models were not passed the conditional coverage statistics or the RiskMap.

However, we must keep in mind that the MRS model can be sensitive to the length of the rolling window used, and a larger in-sample period could have added more precision to the results.

Overall, the semi-parametric and non-parametric VaR models as well as the models that accommodated for excess kurtosis in the return series seemed to deliver a better overall performance compared to the parametric models and the models under the assumption of a normal distribution. This was to be expected, since the returns series tend to exhibit fat tails. Furthermore, the t-distribution may not perfectly fit the data series, which was partly bypassed by the employment of semi-parametric models. These use the distributional assumption only in the computation of the volatility forecasts and not in the actual VaR forecast.

Comparing the models estimated in this paper with the in-house models used by banks strictly in terms of the number of exceptions, we find that the MRS VaR gives similar results to the banks' models in most of the cases. However, the most of the in-house models seem to over-estimate the risk for most of the portfolios, with a very small number of violations occurring in the past three years. While the other time-varying volatility employed give larger number of exceptions, they do pass the backtests which suggests that they may have a better fit in what concerns the dynamics of market risk than the Monte Carlo scenario-based models of the financial institutions. In the case of a 4 year period, including the turbulent 2008 year, the parametric models give a much lower number of exceptions, for Danske Bank and particularly for Deutsche Bank. We can thus infer that the models based on time-varying volatility, including the regime-switching models have a more stable performance in adapting to the market dynamics than the in-house models used by these banks.

7. Future research

One possible venue for further research would be to apply a series of more complex GARCH models to forecast the time-varying volatilities, such as APARCH, NGARCH or NAGARCH.

Moreover, these models could be estimated under a series of different distributional assumptions, besides the normal and the Student-t, such as the GED or skewed-t. These distributions could provide a better fit for the returns series which could thus lead to more accurate VaR forecast models.

Furthermore, while the regime-switching GARCH model literature has had a series of breakthroughs, these models have never been applied to real bank data up to this point to our knowledge. Therefore, it would be interesting to apply a series of different regime switching GARCH models or GARCH models with an added jump component, other than Klaasen's (2002) models which was used in the present study, to the 5 bank portfolios and assess their performance. One suggestion would be the GARJI model proposed by Maheu and McCurdy (2004), and its variations such as proposed by Liu (2011) or the NIG-GARJI Nyberg and Wilhelmsson (2009).

In addition, it would be ideal if data series that covered a longer time-span would be employed in these investigations. At least ten years of daily observations would allow for more precise volatility estimates through regime-switching models, since they would be able to better capture the evolution of the regimes. It is very difficult to gain access to real bank data for such a long period, since very few banks have published their P/L plots for more than 5 years. Out of the bank sample analyzed, only Bank of America has published around 10 years of daily P/L plots. An analysis on this portfolio of several different regime-switching and GARCH jump models would be worth doing with a longer estimation window.

References

- Alberg, D.; Shalit, H; Yosef, R.(2008), “Estimating Stock Market Volatility Using Asymmetric GARCH models”, *Journal of Applied Financial Economics*, 18, 15, 1201-1208.
- Ane, T., Ureche-Rangau, L., 2006. Stock market dynamics in a regime-switching asymmetric power GARCH model. *International Review of Financial Analysis*. 15, 109-129.
- Angelidis, T., Benos, A., and Degiannakis, S. (2004), “The use of GARCH models in VaR estimation”, *Journal of Statistical Methodology*, 1, 1-2, 105-128.
- Bank of America Annual Reports and Risk Management Reports 2007-2010.
- Barron, D. G., Brawn, J. D., and Weatherhead, P. J. (2010), “Meta-analysis of transmitter effects on avian behavior and ecology”, *Methods in Ecology and Evolution*, Volume 1, Issue 2, pages 180–187.
- Basel Committee on Banking Supervision (2011), “Basel III: A global regulatory framework for more resilient banks and banking systems”, *Bank for International Settlements*.
- Berkowitz, J. and O’Brien, J. (2002), “How Accurate Are Value-at-Risk Models at Commercial Banks?”, *The Journal of Finance*, 57, 3, 1093-1111.
- Billinger, O.; Eriksson, B., (2009) , ”Star Vars: Finding the optimal Value-at-Risk approach for the banking industry”, *University essay from Lund University*, <http://www.essays.se/essay/ce32043378/>
- Birtoiu, A. and Dragu, F. G. (2011), “Market Risk Management: The Applicability and Accuracy of Value-at-Risk Models in Financial Institutions”, *Master Thesis*, Lund University.
- Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, 31, 3, 307–327.
- Campbell, S.D., (2005), ”A Review of Backtesting and Backtesting Procedures”, Board of Governors of the Federal Reserve System (U.S.) - [*Finance and Economics Discussion Series*](#), nr. 2005-21.
- Cai J, (1994),”A Markov model of unconditional variance in ARCH”, *Journal of Business and Economic Statistics*, 12(3):309-316.
- Candelon, B., Colletaz, G., Hurlin, C., and Tokpavi, S. (2011), “Backtesting Value-at-Risk: A GMM Duration-Based Test”, *Journal of Financial Econometrics*, Vol. 9, No. 2, 314–343.
- Chan, W. H.; Maheu, J. M. (2002), ”Conditional jump dynamics in stock market returns”, *Journal of Business and Economic Statistics*, 20:377–389
- Chan, W.H.;Young, D. (2009), “A New Look at Copper Markets: A Regime-Switching Jump Model”, No 2009-13, *Working Papers from University of Alberta, Department of Economics*
- Christoffersen, P. (1998), “Evaluating Interval Forecasts”, *International Economic Review*, 39, 4, 841–862.

Christoffersen, P. and Pelletier, D. (2004), "Backtesting Value-at-Risk: A Duration-Based Approach", *Journal of Financial Econometrics*, 2, 84-108.

Colletaz, G., Hurlin, C., and Perignon, C. (2011), "The Risk Map: A New Tool for Backtesting Value-at-Risk Models", *Working Papers Series (SSRN)*.

Damodaran, A. (2007), "Strategic risk taking. A framework for risk management", 1st edition, *Wharton School Publishing*

Danielsson, J. and de Vries, C. G. (1997), "Value-at-Risk and Extreme Returns", *Annales d'Economie et de Statistique*, 60, 239–270.

Danske Bank Annual Reports and Risk Management Reports 2007-2010

Deutsche Bank Annual Reports and Risk Management Reports 2007-2010

Deng, Z-H., (2010) –"Volatility Forecasting Using APARCH with Skewed Conditional distributions", 2010 *International Conference on E-Business and E-Government*

Ding, Z., Granger, C. W. J., Engle, R. F., (1993), "A long memory property of stock market returns and a new model", *Journal of Empirical Finance* 1, 83-106.

Dowd, K. (2005), "Measuring Market Risk", 2nd edition, *John Wiley & Sons, Ltd.*

Engle, R.F.(1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica* Vol. 50, pg 987–1007.

Engle, R.F. and Manganelli, S. (2004), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles", *Journal of Business & Economic Statistics*, Volume 22, Issue 4, pp. 367-381.

Gau, Y-F; Tang, W-T.(2004), "Forecasting Value-at-Risk Using the Markov-Switching ARCH Model", No 715, *Econometric Society 2004 Far Eastern Meetings from Econometric Society*

Giot, P.; Laurent, S., (2003) – "Value-at-risk for long and short trading positions", *Journal of Applied Econometrics*, Volume 18, Issue 6, pages 641–663, November/December 2003

Glosten, L. R., Jaganathan, Runkle, D.,(1993) – "On the relation between the expected value and the volatility of the nominal excess return on stocks", *Journal of Finance* 48, 1779-1801.

Gray, S.F. (1996), "Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process," *Journal of Financial Economics*, 42, 27-62.

Haas, M., Mittnik, S., Paolletta, M. S.,(2004), "A new approach in Markov switching GARCH models", *Journal of Financial Econometrics*, 2(4), 493-530.

Hamilton, J. D., R. Susmel, (1994), "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307--333.

Hurlin, C. and Perignon, C. (2011), "Margin Backtesting", *Working Papers Series (SSRN)*.

Hurlin, C. and Tokpavi, S. (2006), "Backtesting VaR Accuracy: A New Simple Test", *Universite d'Orleans*, <http://www.univ-orleans.fr/deg/masters>.

- Johansson, M. (2011), "VaR for a portfolio of Swedish Index-bonds - An empirical evaluation", *Dissertation Thesis, Lund University*,
<http://www.lu.se/o.o.i.s?id=19463&postid=2167277>
- Jorion, P. (2001), "Value at Risk: The New Benchmark for Managing Financial Risk", 2nd edition, *McGraw-Hill*.
- J.P. Morgan Chase (2009), "Backtesting Value-at-Risk", *J.P. Morgan Investment Analytics and Consulting*, September, pp. 5-6
- Kim, C-J.; Piger, J.M.; Startz, R., (2003) – "Estimation of Markov Regime-Switching Regression Models with Endogenous Switching", *Working Paper, Federal Reserve Bank of St. Louis*
- Klassen, F., (2002), "Improving GARCH volatility forecasts", *Empirical Economics* 27, 363-394
- Kuester, K., Mittnik, S., and Paolella, M.S., (2005), "Value-at-Risk Prediction: A Comparison of Alternative Strategies", *Journal of Financial Econometrics*, 4, 1, 53-89.
- Kupiec, P. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models" *Journal of Derivatives*, 3, 2, 73-84.
- Lamoureux, C.G.; Lastrapes, W.D., (1990), "Persistence in Variance, Structural Change, and the GARCH Model", *Journal of Business & Economic Statistics, American Statistical Association*, vol. 8(2), pages 225-34
- Liu, P., (2011), "Regime-Switching GARCH-Jump Models with Autoregressive Jump Intensity", http://publish.uwo.ca/~pliu46/Pujun_paper2.pdf
- Liu, H-C; Hung, J-C, (2010) – "Forecasting S&P-100 stock index volatility: The role of volatility asymmetry and distributional assumption in GARCH models", *Expert Systems with Applications*, Volume 37, Issue 7, July 2010, Pages 4928–493
- Maheu, J. M.; McCurdy, T.H., (2004) "News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns", *Journal of Finance, American Finance Association*, vol. 59(2), pages 755-793
- Marcucci, J. (2005), "Forecasting Stock Market Volatility with Regime-Switching GARCH Models", *Studies in Nonlinear Dynamics & Econometrics*, 2005, vol. 9, issue 4, pages 6
- McMillan, D.G.; Kambouroudis, D., (2009) "Are RiskMetrics forecasts good enough? Evidence from 31 stock markets", *International Review of Financial Analysis, Volume 18, Issue 3, June 2009, Pages 117–124*.
- Nocera, J. (2009), "Risk Mismanagement – What Led to the Financial Meltdown", *New York Times*, January 2nd 2009, <http://www.nytimes.com/2009/01/04/magazine/04risk-t.html?pagewanted=1&r=2>
- Nyberg, P. M.; Wilhelmsson, A., (2009) – "Measuring Event Risk", *Journal of Financial Econometrics*, 2009, 7, 3, 265-287

Pérignon, C. and Smith, D.R. (2008), "A New Approach to Comparing VaR Estimation Methods", *Working Papers Series (SSRN)*.

Sadras, V. O., Trentacoste, E. R., and Meinzer, F. (2011), "Phenotypic plasticity of stem water potential correlates with crop load in horticultural trees", *Tree Physiol*, 31, 5, 494-499.

Sajjad, R.; Coakley, J.; Nankervis, J.C. (2008), "Markov-Switching GARCH Modelling of Value-at-Risk", *Studies in Nonlinear Dynamics & Econometrics*, 2008, vol. 12, issue 3, pages 7

Su, C. (2010), "Application of EGARCH Model to Estimate Financial Volatility of Daily Returns: The empirical case of China", *Gothenburg University, Master Thesis*.

Su, Y.; Lin, C.; Chen, P. (2009) – Asymmetric GARCH Value at Risk for QQQQ, *Working Paper Series (SSRN)*.

Swedbank Annual Reports and Risk Management Reports 2007-2010.

Yu, J-S.; Daal, E., (2005) – "A Comparison of Mixed GARCH-Jump Models with Skewed t-Distribution for Asset Returns" - http://papers.ssrn.com/sol3/papers.cfm?abstract_id=670502

Appendix

Appendix 1. Portfolio composition and statistic tests

Table 1: Portfolio components and weights

Component	Weight
Foreign Exchange	5.67%
Interest Rate	16.60%
Credit	42.27%
Real Estate	18.01%
Equities	12.53%
Commodities	4.92%

Figure 2: Correlograms

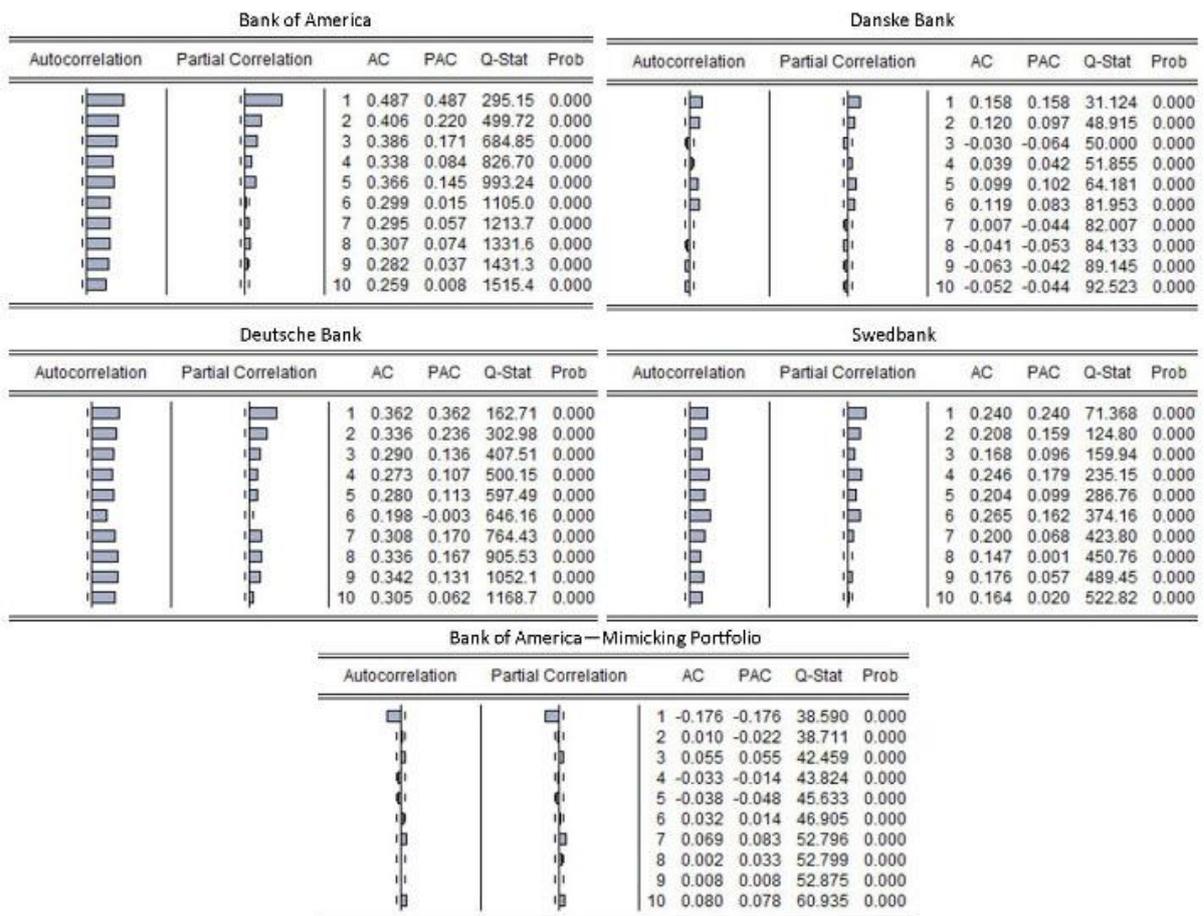


Table 2: ARCH effects test results

	Lags	H	p-Value	Test Statistic	Critical Value
Bank of America	10	1	0	112.9821	23.2093
	15	1	0	113.9266	30.5779
	20	1	2.11E-15	115.2435	37.5662
Danske Bank	10	1	0	127.5364	23.2093
	15	1	0	128.7644	30.5779
	20	1	0	131.7283	37.5662
Deutsche Bank	10	1	0	235.9108	23.2093
	15	1	0	239.9806	30.5779
	20	1	0	251.6346	37.5662
Swedbank	10	1	0	158.2647	23.2093
	15	1	0	177.7808	30.5779
	20	1	0	178.5177	37.5662
Bank of America - mimicking portfolio	10	1	0	328.4267	23.2093
	15	1	0	367.1451	30.5779
	20	1	0	388.0687	37.5662

Table 3: Ljung-Box Q-test results

	Lags	H	p-Value	Test Statistic	Critical Value
Bank of America	10	1	0	1515.3550	23.2093
	15	1	0	1821.6945	30.5779
	20	1	0	2055.5241	37.5662
Danske Bank	10	1	0	92.5227	23.2093
	15	1	0	103.4151	30.5779
	20	1	0	109.3856	37.5662
Deutsche Bank	10	1	0	1,168.7379	23.2093
	15	1	0	1,469.3276	30.5779
	20	1	0	1,762.8954	37.5662
Swedbank	10	1	0	522.8205	23.2093
	15	1	0	670.6688	30.5779
	20	1	0	749.2253	37.5662
Bank of America - mimicking portfolio	10	1	2.41E-09	60.9353	23.2093
	15	1	3.78E-10	75.9729	30.5779
	20	1	4.34E-12	96.9893	37.5662

Appendix 2. Coefficient estimates

Table 1: GARCH Coefficient estimates for Bank of America

BANK OF AMERICA	GARCH		EGARCH		GJR	
	n	t	n	t	n	t
	ARMA(0,0) - GARCH(2,1)		AR(1)-EGARCH(1,1,1)		AR(1)-GJR(2,1,1)	
Mean equation						
C	37.6111 (4.519870)	40.71 (2.283373)	40.34411 (4.878483)	40.17897 (4.040135)	38.80889 (2.476618)	40.24862 (2.272921)
AR(1)	0.549603 (0.033299)		0.519569 (0.066334) 0.487321 (0.059813)			
Variance equation						
ω	991.0106 (144.079000)	1519.476 (441.318700)	4.835796 (0.914032)	5.562608 (1.806263)	1032.013 (209.224000)	1186.654 (346.711400)
α_1	0.592048 (0.125681)	0.538943 (0.172641)	0.859424 (0.139498)	0.684214 (0.176050)	1.03606 (0.195386)	0.708491 (0.235272)
α_2	0.38023 (0.134772)				0.279454 (0.036986)	0.259958 (0.058029)
β	-0.056696 (0.028512)	-0.527115 (0.264964)	0.239576 (0.121806)	0.163446 (0.243860)	-0.252916 (0.060530)	-0.291377 (0.147662)
ϕ			0.262379 (0.087394)	0.155113 (0.111708)	-0.716164 (0.233603)	-0.376379 (0.285280)
Log likelihood	-1265.168	-1284.82	-1260.523	-1255.776	-1287.927	-1284.518

**The values in brackets are standard errors*

Table 2: GARCH Coefficient estimates for the synthetic portfolio

Portfolio	GARCH		EGARCH		GJR	
	n	t	n	t	n	t
	AR(1)-GARCH(1,1)		AR(1)-EGARCH(1,1)		AR(1)-GJR(1,1)	
Mean equation						
C	0.000223 (0.000190)	0.000246 (0.000190)	-0.000003 (0.000199)	0.000072 (0.000201)	0.000093 (0.000198)	0.000151 (0.000199)
AR(1)	-0.067277 (0.072043)	-0.050693 (0.065919)	-0.031964 (0.067452)	-0.029078 (0.064206)	-0.049859 (0.069761)	-0.044740 (0.065985)
Variance equation						
ω	0.000000 (0.000000)	0.000001 (0.000000)	-0.577939 (0.216783)	-0.574308 (0.250199)	0.000000 (0.000000)	0.000001 (0.000000)
α	0.171774 (0.032520)	0.155851 (0.047729)	0.240000 (0.068678)	0.228439 (0.083069)	0.048752 (0.044685)	0.047790 (0.053147)
β	0.816339 (0.032807)	0.821370 (0.047221)	0.964893 (0.018330)	0.964602 (0.020614)	0.849994 (0.035153)	0.851911 (0.043951)
ϕ			-0.136242 (0.034532)	-0.123927 (0.044977)	0.154947 (0.049583)	0.139803 (0.063673)
Log likelihood	1093.992	1096.086	1095.546	1096.574	1097.343	1098.244

**The values in brackets are standard errors*

Table 3: GARCH Coefficient estimates for Danske Bank

DANSKE BANK	GARCH		EGARCH		GJR	
	n	t	n	t	n	t
	ARMA(1,1)-GARCH(1,1)		ARMA(1,1)-EGARCH(1,2,1)		ARMA(1,1)-GJR(1,2,1)	
Mean equation						
C	-0.061959 (2.556104)	-0.209848 (2.562590)	-2.812967 (2.598202)	-2.800476 (2.671025)	-2.584156 (2.653740)	-2.528658 (2.709538)
AR(1)	-0.813003 (0.153692)	-0.799001 (0.166768)	-0.808166 (0.154736)	-0.806762 (0.160695)	-0.816453 (0.151063)	-0.812927 (0.157196)
MA(1)	0.700495 (0.184997)	0.690591 (0.200866)	0.697403 (0.181607)	0.695914 (0.188982)	0.71252 (0.179709)	0.708563 (0.187174)
Variance equation						
ω	358.6526 (219.767900)	380.8095 (249.445000)	0.909212 (0.503489)	0.909435 (0.505707)	409.6152 (198.648200)	408.9254 (201.852900)
α	0.229745 (0.068483)	0.234724 (0.085487)	0.265129 (0.103320)	0.265345 (0.105867)	0.022704 (0.066583)	0.024333 (0.072140)
β	0.637231 (0.122360)	0.623822 (0.143654)	0.855025 (0.066103)	0.854969 (0.066412)	0.642113 (0.112979)	0.341872 (0.140062)
ϕ			-0.214323 (0.070432)	-0.213942 (0.070813)	0.344084 (0.137443)	0.641612 (0.117388)
Log likelihood	-1325.041	-1324.336	-1320.731	-1320.733	-1320.264	-1320.25

*The values in brackets are standard errors

Table 4: GARCH Coefficient estimates for Deutsche Bank

DEUTSCHE BANK	GARCH		EGARCH		GJR	
	n	t	n	t	n	t
	ARMA(1,1)-GARCH(1,1)	ARMA(1,1)-GARCH(2,1)	ARMA(1,1)-EGARCH(1,1,1)		ARMA(1,1)-GJR(2,1,1)	ARMA(0,0)-GJR(1,1,1)
Mean equation						
C	-14.12441 (4.938976)	-4.528575 (2.136822)	-23.53403 (13.934110)	-27.93896 (17.304560)	-17.66863 (8.664441)	-2.681273 (1.647515)
AR(1)	0.989002 (0.004506)	0.637949 (0.103069)	0.991668 (0.005900)	0.993278 (0.004807)	0.991652 (0.004622)	
MA(1)	-0.995104 (0.001904)	-0.514402 (0.118415)	-0.989381 (0.009286)	-0.987032 (0.014358)	-0.994146 (0.004120)	
Variance equation						
ω	40.24942 (22.92486)	96.90312 (75.40379)	12.40331 (0.22663)	12.47548 (0.46435)	93.31754 (22.10035)	33.74253 (22.60906)
α_1	0.053041 (0.02651)	-0.034679 (0.04105)	0.135283 (0.03396)	0.119359 (0.06012)	-0.04837 (0.01311)	-0.001228 (0.02508)
α_2		0.114917 (0.06997)			0.081981 (0.02977)	
β	0.898291 (0.04689)	0.823045 (0.10222)	-0.907117 (0.02402)	-0.900348 (0.04774)	0.799189 (0.05221)	0.930863 (0.03563)
ϕ			-0.126803 (0.03725)	-0.117685 (0.06580)	0.09948 (0.05682)	0.074697 (0.03251)
LL	-1186.658	-1181.413	-1178.106	-1169.631	-1184.813	-1199.963

*The values in brackets are standard errors

Table 5: GARCH coefficient estimates for Swedbank

SWEDBANK	GARCH		EGARCH		GJR		
	n	t	n	t	n	t	
	AR(1)-GARCH(1,2)		ARMA(1,1)-GARCH(1,2)		AR(1)-EGARCH(1,2,1)		AR(1)-GJR(1,2,1)
Mean equation							
C	6.867238 (0.73311)	8.124101 (0.87512)	7.116954 (0.76692)	7.004453 (0.79807)	6.933534 (0.76878)	7.288653 (0.76202)	
AR(1)	0.230424 (0.06423)	-0.634515 (0.05488)	0.270361 (0.06067)	0.24991 (0.06379)	0.242867 (0.04654)	0.232136 (0.06769)	
MA(1)		0.984887 0.011571					
Variance equation							
ω	71.29465 (15.32335)	3.109484 (2.14735)	3.394465 (0.41706)	3.60924 (3.48753)	196.8133 (21.64071)	71.81496 (17.25450)	
α	0.164919 (0.07782)	-0.094047 (0.04259)	0.235052 (0.11253)	0.298498 (0.21614)	0.11412 (0.05605)	0.187104 (0.09306)	
β_1	0.774354 (0.12992)	0.361574 (0.44729)	1.038066 (0.07320)	0.102025 (0.55636)	-0.300575 (0.04410)	0.755561 (0.13284)	
B_2	-0.676603 (0.20449)	0.706415 (0.45588)	-0.838855 (0.10184)	0.045469 (0.52077)	-0.877551 (0.07531)	-0.664059 (0.21437)	
ϕ			0.052945 (0.06750)	0.006859 (0.11652)	-0.162472 (0.04208)	-0.050633 (0.08515)	
LL	-928.9011	-914.4234	-926.4913	-929.658	-924.0159	-928.5939	

*The values in brackets are standard errors

Table 6: Markov Regime-Switching GARCH coefficient estimates

	Bank of America		Danske Bank		Deutsche Bank		Swedbank		Portfolio	
	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-N	MRS-GARCH-t	MRS-GARCH-N	MRS-GARCH-t
$\delta^{(1)}$	22.4274 (3.85223)	-20.8655 (-3.58395)	-37.887 (4.24132)	39.6759 (4.78103)	-12.0329 (2.77565)	7.6116 (2.17005)	11.7625 (1.07812)	-1.9957 (1.34946)	4.928 (7.26541)	6.5041 (11.96613)
$\delta^{(2)}$	-23.7945 (2.32291)	22.191 (3.15807)	5.9867 (2.37449)	-7.212 (5.40395)	8.2806 (2.08149)	-1.6771 (0.25415)	3.6722 (1.10454)	10.9614 (0.84481)	-42.4448 (2.22813)	-42.0515 (2.83641)
$\alpha_0^{(1)}$	9.5537 (2.00395)	8.9631 (1.10071)	0 (0.28503)	0 (5.38673)	2.2145 (27.5449)	1.9258 (13.89756)	0.7495 (4.41310)	9.02 (22.79986)	0.6625 (22.36529)	7.8168 (12.27677)
$\alpha_0^{(2)}$	0 (22.4919)	8.657 (0.24330)	7.0565 (46.5223)	0.7102 (77.40917)	0 (9.89117)	9.1982 (0.11683)	0.0791 (3.31911)	7.5858 (10.65316)	0 (19.74605)	0 (24.09018)
$\alpha_1^{(1)}$	0.0809 (0.01265)	0 (0.43388)	0.6674 (0.48750)	0.5214 (0.19955)	0.0831 (0.05141)	0.0918 (0.06221)	0.0784 (0.05953)	0.3535 (0.21706)	0.0397 (0.43062)	0.0287 (0.26740)
$\alpha_1^{(2)}$	0.0614 (0.07557)	0.0753 (0.09662)	0.0851 (0.06553)	0.1445 (0.11151)	0 (0.02337)	0.243 (0.06750)	0.1118 (0.05393)	0.2 (0.12485)	0.5276 (0.02872)	0.9974 (0.07430)
$B_1^{(1)}$	0.8778 (0.01270)	0.965 (0.01438)	0.0045 (0.00855)	0.0365 (0.17724)	0.9068 (0.05986)	0.9081 (0.04093)	0.9041 (0.09805)	0.1337 (0.26212)	0.9203 (0.24105)	0.9139 (0.28299)
$\beta_1^{(2)}$	0.9178 (0.07557)	0.8897 (0.05538)	0.9059 (0.04521)	0.8535 (0.11124)	0.9983 (0.02359)	0.757 (0.01029)	0.7322 (0.05441)	0.537 (0.15320)	0 (0.02874)	0 (0.06709)
p	0.989 (0.01381)	0.9505 (0.01512)	0.1341 (0.25967)	0 (0.23186)	0.9928 (0.00968)	0.9868 (0.01640)	0.9646 (0.02460)	0.5481 (0.10505)	0.9454 (0.18217)	0.9653 (0.27996)
q	0.9608 (0.01792)	0.9848 (0.04843)	0.9277 (0.04242)	0.7538 (0.12883)	0.9915 (0.00752)	0.9456 (0.00694)	0.9504 (0.03355)	0.7978 (0.06308)	0.4116 (0.02279)	0.3554 (0.04784)
v		5.5411 (33.3711)		18.4579 (93.82248)		4.5276 (10.13071)		16.84 (29.92268)		269.776 (93.82248)
LL	-1182.61	-1175.73	-1325.24	-1319.31	-1200.44	-1187.28	-929.48	-927.88	-1259.32	-1256.67

*The values in brackets are standard errors

Table 7: Number of violations for banks' in-house VaR models

	Bank of America	Danske Bank	Deutsche Bank	Swedbank
2011	0	2	3	2
2010	0	4	2	0
2009	0	2	1	0
2008	2	10	35	4
2007	14	2	12	1

Appendix 3. VaR forecasts – graphical representation

Figure 3: Bank of America - Comparison of VaR methods (4 years)

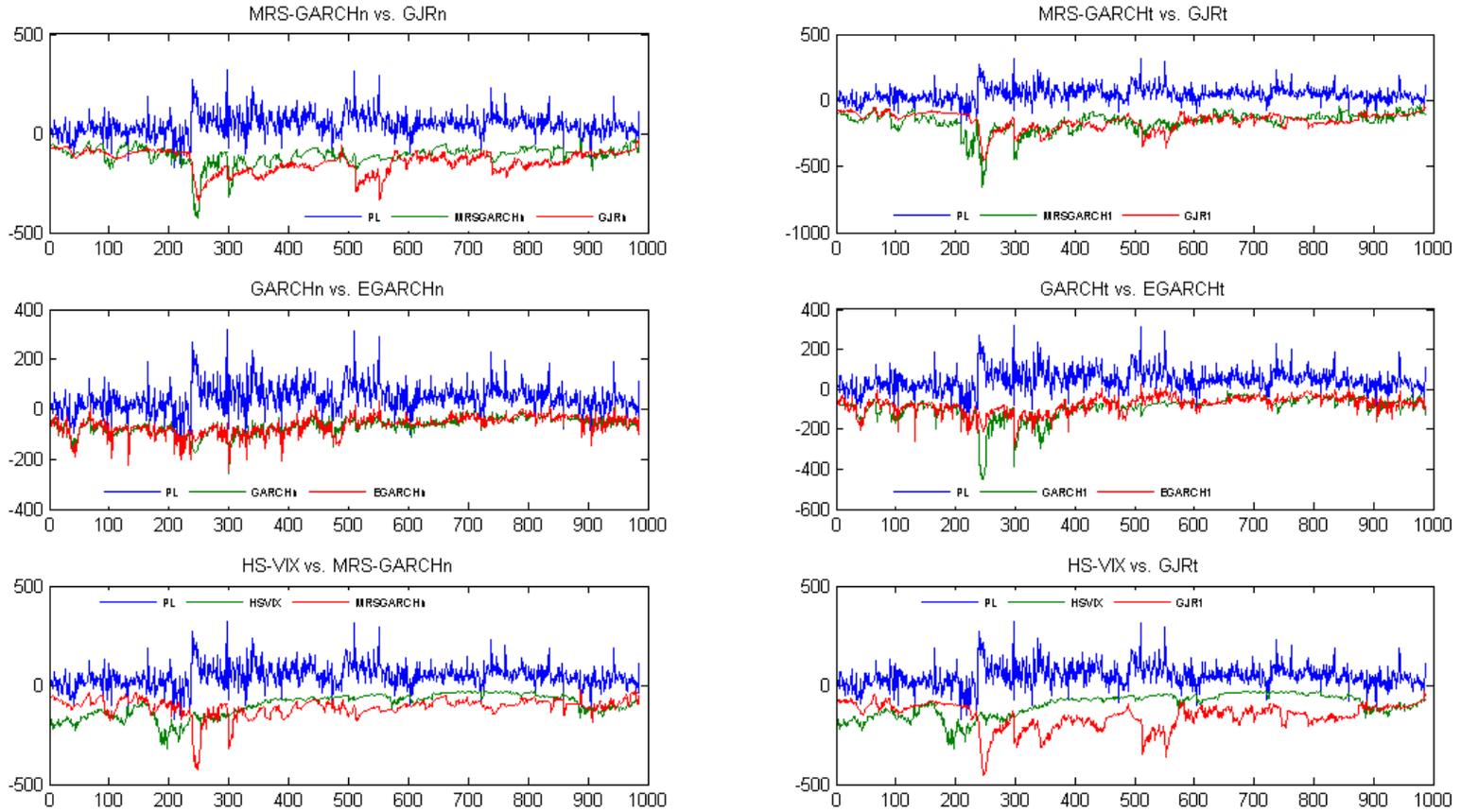


Figure 4: Mimicking portfolio - Comparison of VaR methods (4 years)

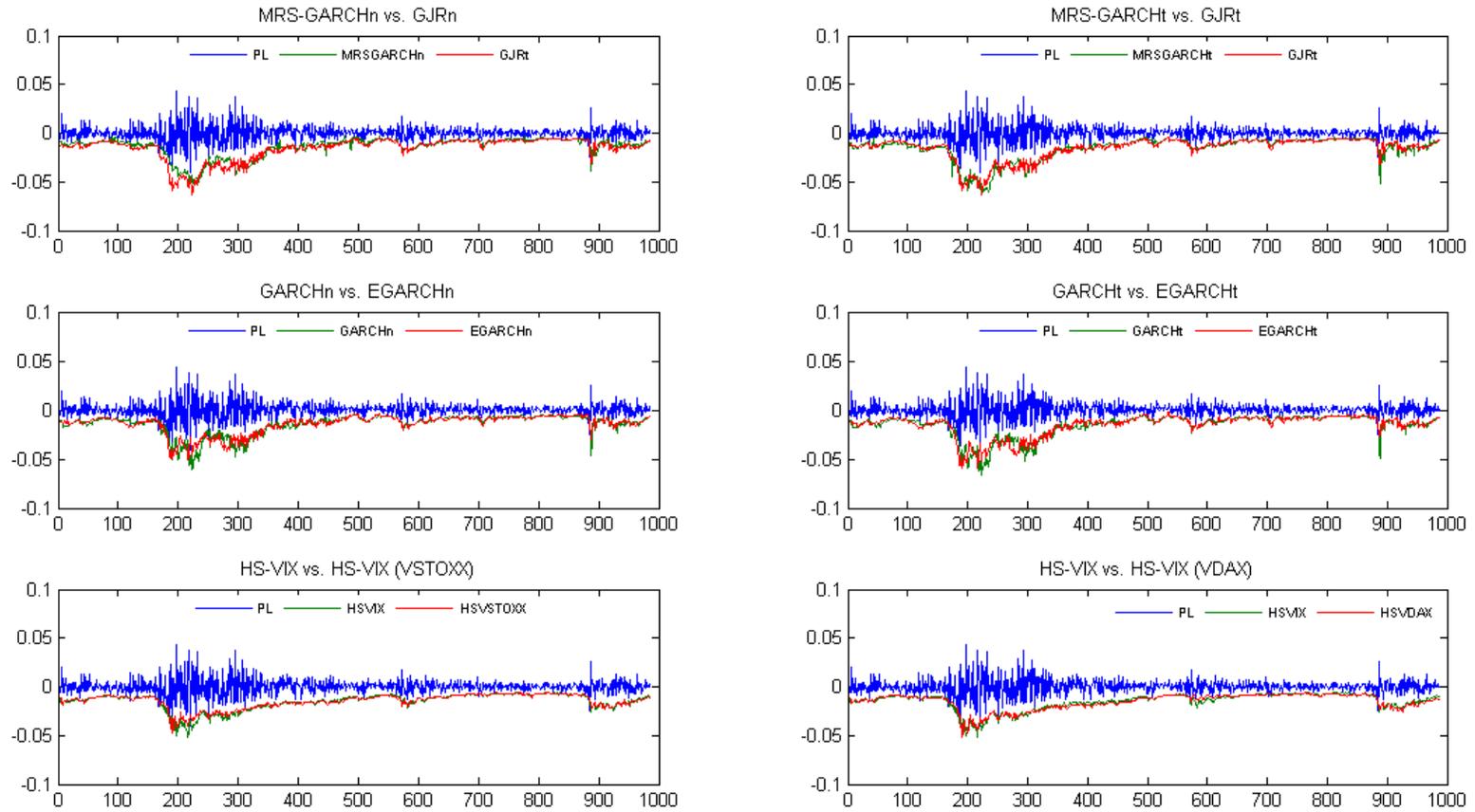


Figure 5: Danske Bank - Comparison of VaR methods (4 years)

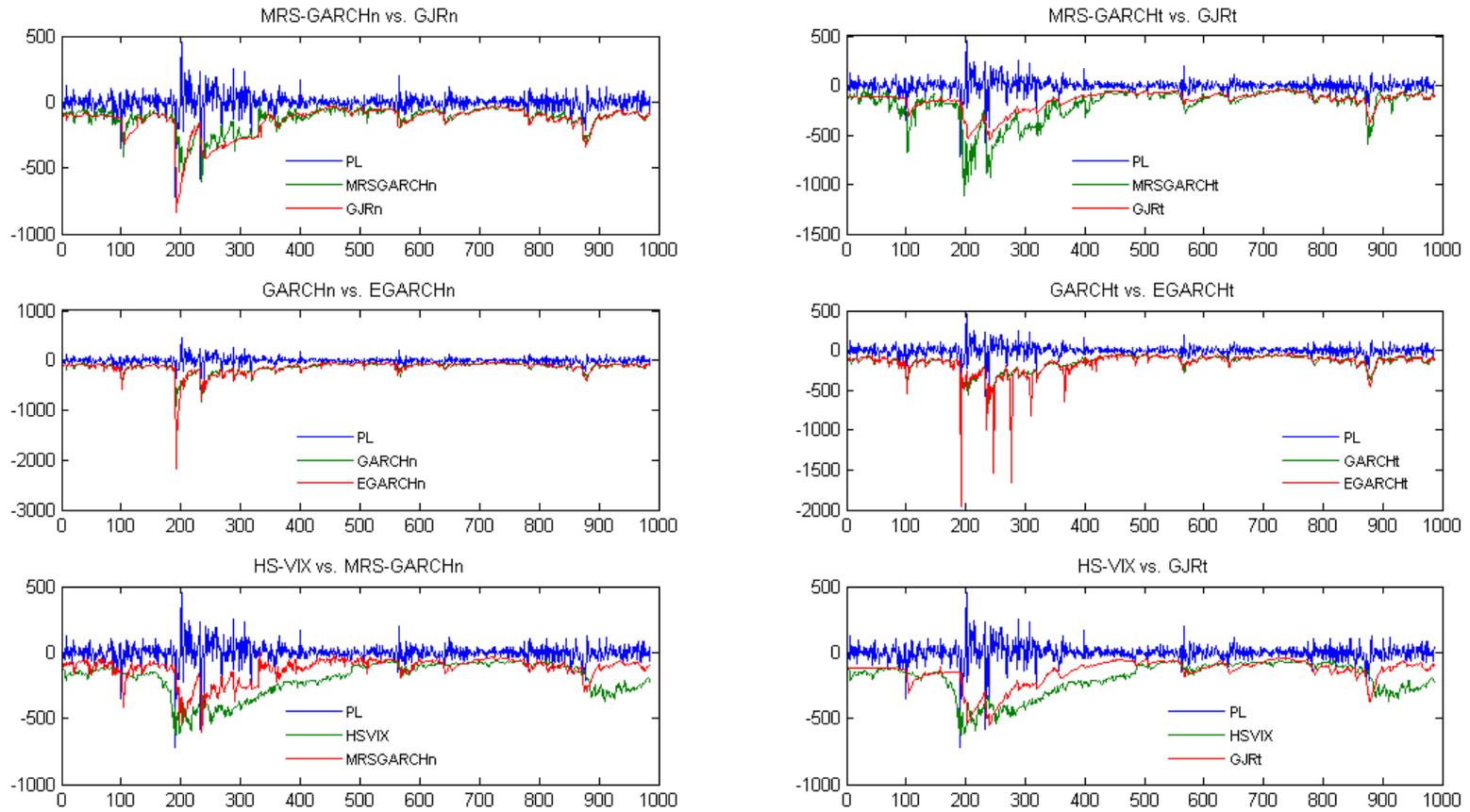


Figure 6: Deutsche Bank - Comparison of VaR methods (4 years)

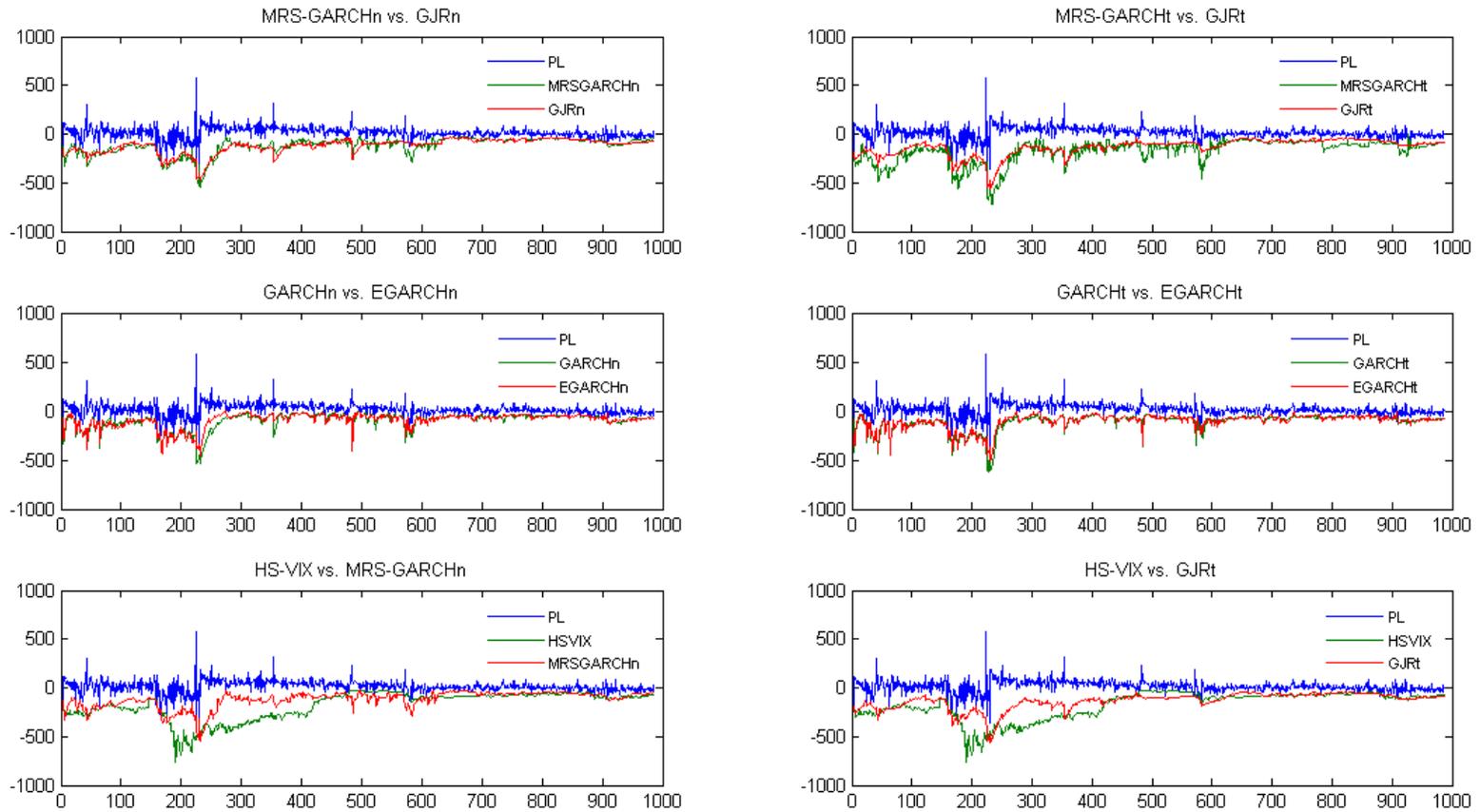


Figure 7: Swedbank - Comparison of VaR methods (4 years)

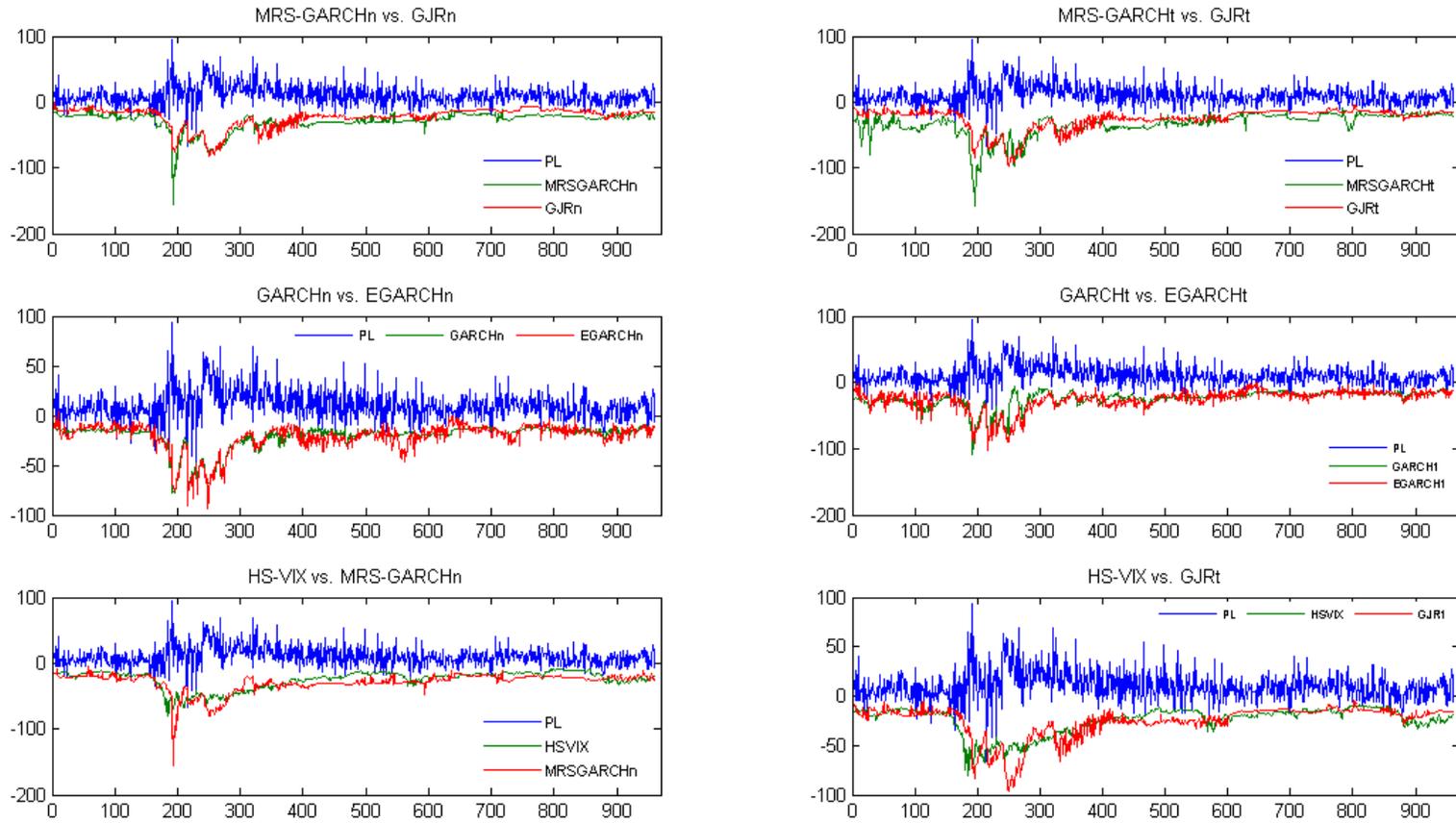


Figure 8: Bank of America - Comparison of VaR methods (3 years)

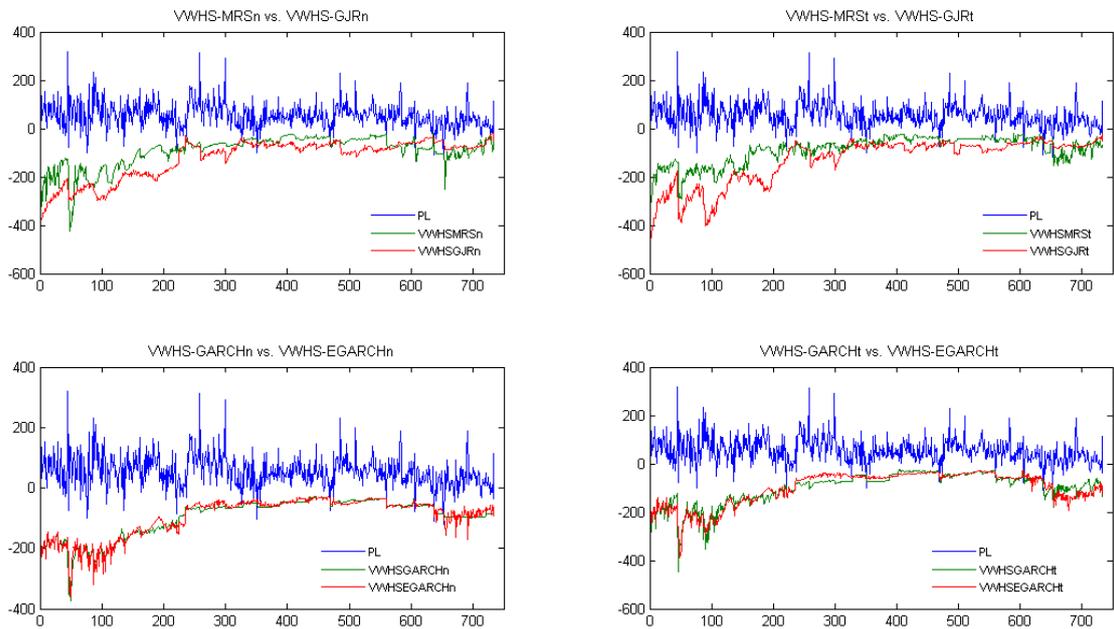


Figure 9: Mimicking Portfolio - Comparison of VaR methods (3 years)

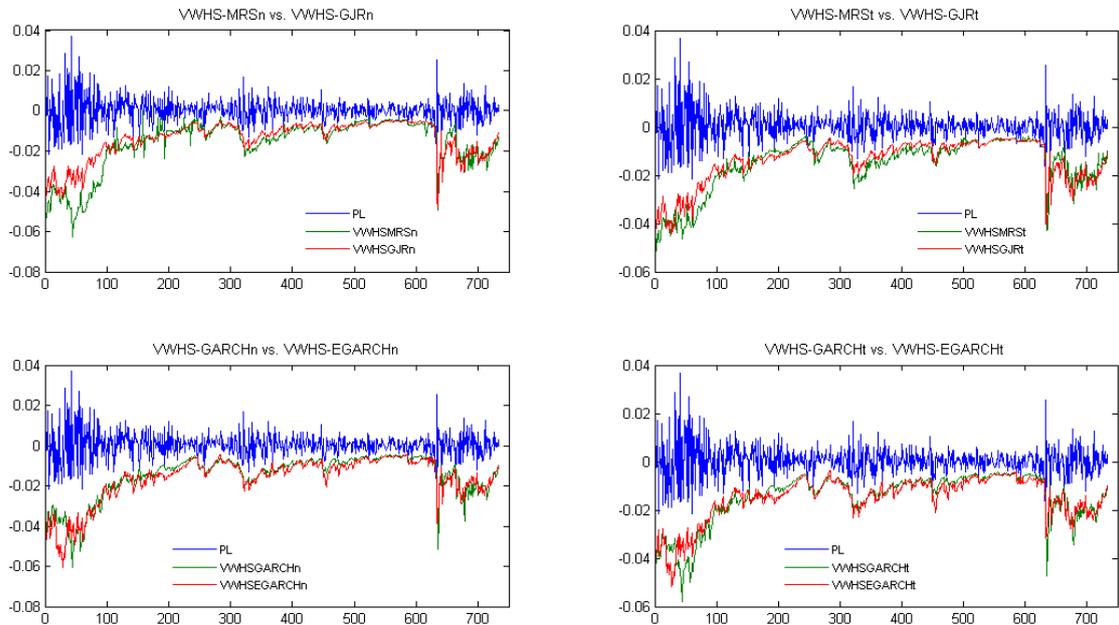


Figure 10: Danske Bank - Comparison of VaR methods (3 years)

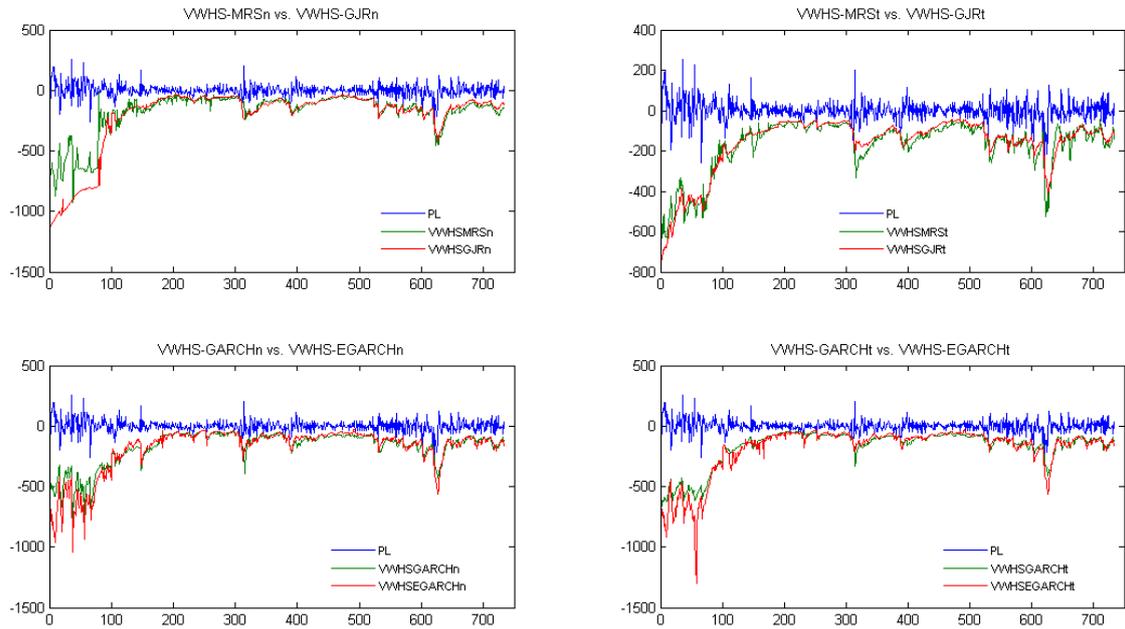


Figure 11: Deutsche Bank - Comparison of VaR methods (3 years)

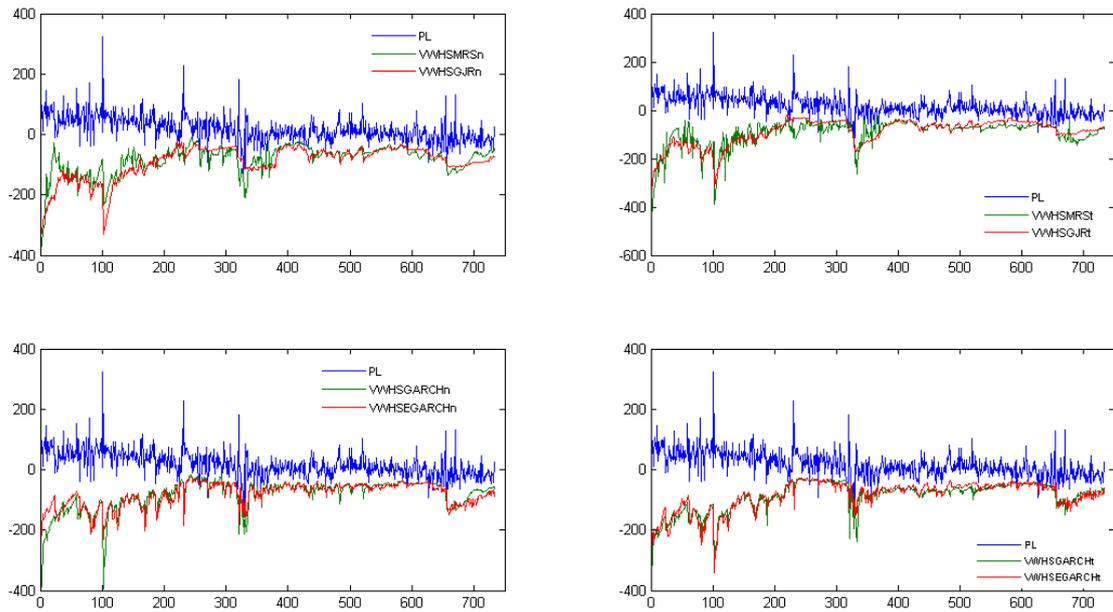
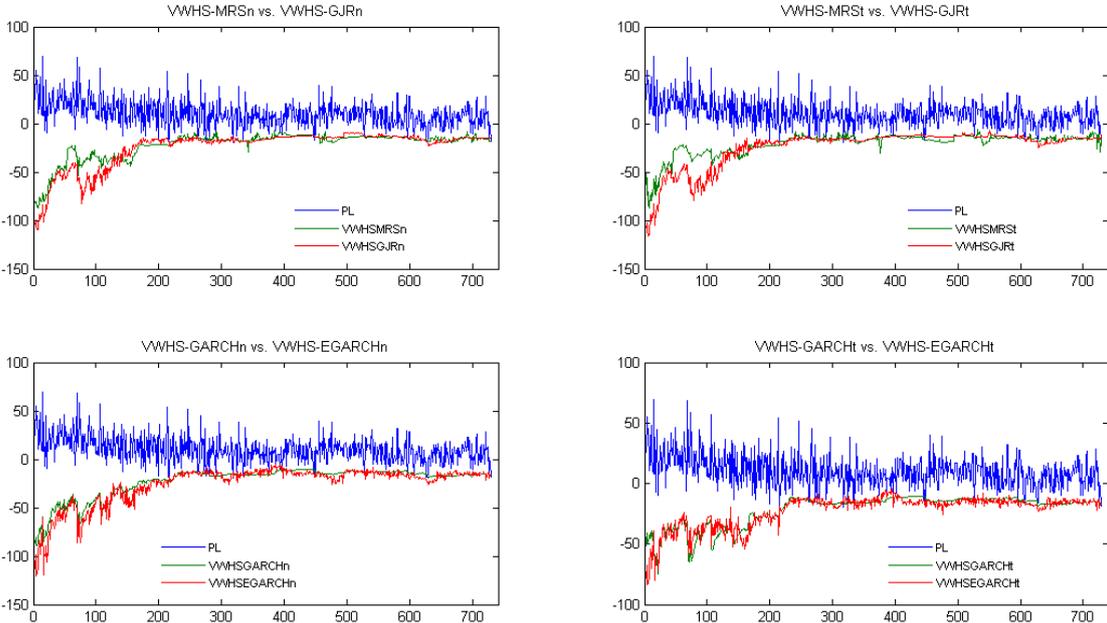


Figure 12: Swedbank - Comparison of VaR methods (3 years)



Appendix 4. Risk Map graphical representation

Figure 1: Risk Map - 4 years of observations

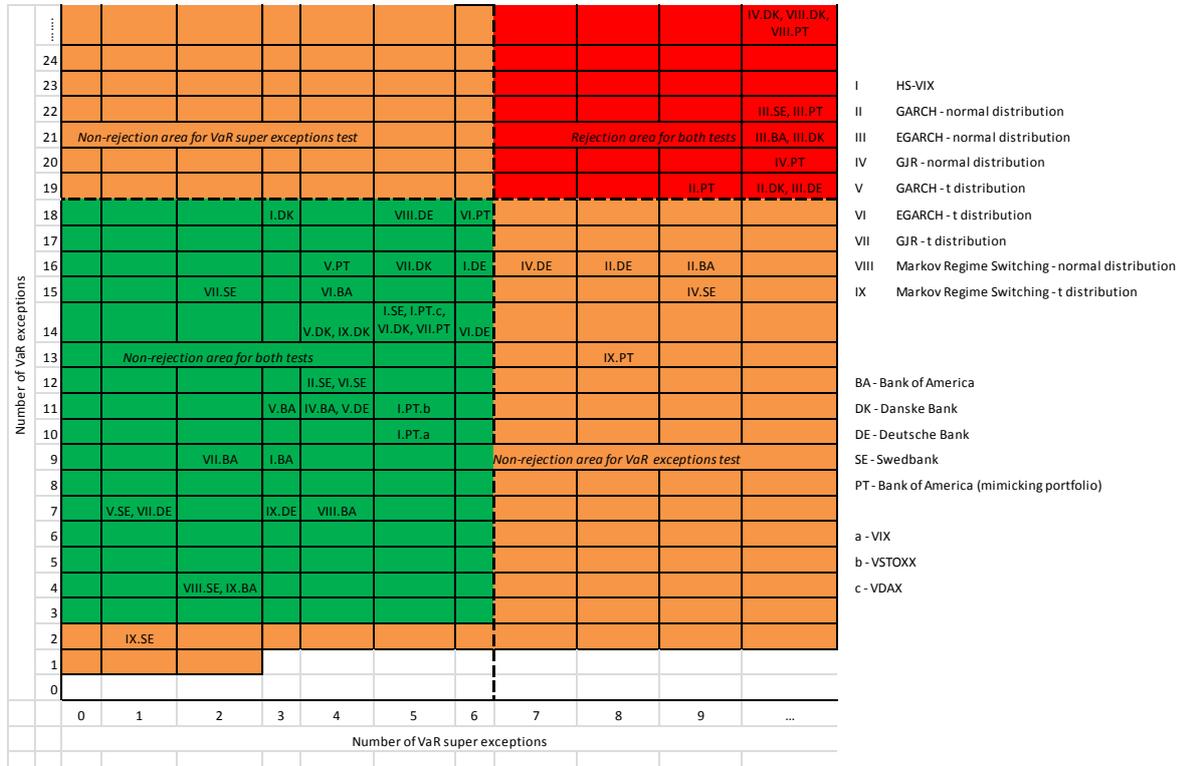
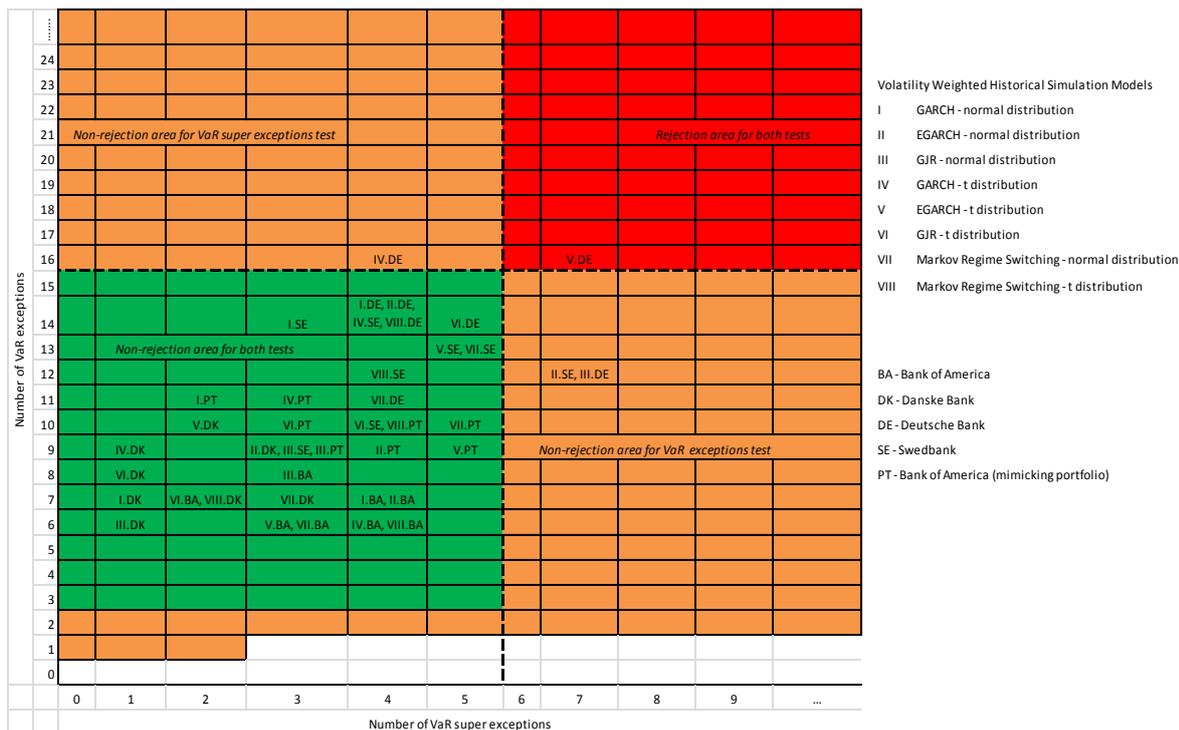


Figure 2: Risk Map - 3 years of observations



Appendix 5. Average VaR

Table 2: Average VaR across models and portfolios (3 years)

Model		Bank of America (mil. USD)	Mimicking portfolio (%)	Danske Bank (mil. DKK)	Deutsche Bank (mil. EUR)	Swedbank (mil. SEK)	
HS-VIX	VIX	-76.3043	-1.3662%	N/A	N/A	N/A	
	VSTOXX	N/A	-1.3687%	-181.6940	N/A	-22.8882	
	VDAX	N/A	-1.4179%	N/A	-133.5823	N/A	
Normal	Volatility Weighted HS	GARCH	-97.1802	-1.5079%	-171.1154	-83.0017	-22.6566
		EGARCH	-95.1895	-1.6223%	-183.7851	-81.6160	-25.5772
		GJR	-122.5744	-1.4081%	-211.8743	-91.4429	-24.5348
t	Volatility Weighted HS	GARCH	-100.1255	-1.4829%	-168.1141	-88.0066	-21.7025
		EGARCH	-100.4689	-1.4995%	-185.8669	-85.6237	-22.3916
		GJR	-130.3250	-1.3949%	-163.3026	-89.5659	-24.5818
Normal		GARCH	-59.0544	-1.2701%	-118.0307	-67.2984	-19.0132
		EGARCH	-56.5737	-1.2506%	-118.2306	-64.5839	-19.5793
		GJR	-155.1534	-1.2507%	-119.7522	-92.9037	-22.0427
t		GARCH	-84.4522	-1.3599%	-129.3623	-78.4628	-19.7971
		EGARCH	-70.5003	-1.3416%	-140.0398	-76.4460	-22.1039
		GJR	-168.4042	-1.3345%	-127.3538	-103.2142	-24.3368
Normal	MRS GARCH	-110.2587	-1.2440%	-110.0616	-84.5381	-27.7046	
	MRS vwhs	-93.4427	-1.6162%	-181.2922	-82.7443	-22.1600	
t	MRS GARCH	-156.7638	-1.3258%	-162.3145	-127.0810	-28.8578	
	MRS vwhs	-87.7311	-1.5243%	-171.4239	-93.9406	-20.8581	

Table 3: Bank of America - Model ranking on 3-years average VaR (mil. USD)

Model	Average
EGARCH-n	-56.5737
GARCH-n	-59.0544
EGARCH-t	-70.5003
HS-VIX	-76.3043
GARCH-t	-84.4522
MRS vwht-t	-87.7311
MRS vwht-n	-93.4427
VWHS-EGARCH-n	-95.1895
VWHS-GARCH-n	-97.1802
VWHS-GARCH-t	-100.1255
VWHS-EGARCH-t	-100.4689
MRS GARCH-n	-110.2587
VWHS-GJR-n	-122.5744
VWHS-GJR-t	-130.3250
GJR-n	-155.1534
MRS GARCH-t	-156.7638
GJR-t	-168.4042

Table 4: Mimicking portfolio - Model ranking on 3-years average VaR (%)

Model	Average
MRS GARCH-n	-1.2440%
EGARCH-n	-1.2506%
GJR-n	-1.2507%
GARCH-n	-1.2701%
MRS GARCH-t	-1.3258%
GJR-t	-1.3345%
EGARCH-t	-1.3416%
GARCH-t	-1.3599%
HS-VIX	-1.3662%
HS-VIX (VSTOXX)	-1.3687%
VWHS-GJR-t	-1.3949%
VWHS-GJR-n	-1.4081%
HS-VIX (VDAX)	-1.4179%
VWHS-GARCH-t	-1.4829%
VWHS-EGARCH-t	-1.4995%
VWHS-GARCH-n	-1.5079%
MRS vwht-t	-1.5243%
MRS vwht-n	-1.6162%
VWHS-EGARCH-n	-1.6223%

Table 5: Danske Bank - Model ranking on 3-years average VaR (mil. DKK)

Model	Average
MRS GARCH-n	-110.0616
GARCH-n	-118.0307
EGARCH-n	-118.2306
GJR-n	-119.7522
GJR-t	-127.3538
GARCH-t	-129.3623
EGARCH-t	-140.0398
MRS GARCH-t	-162.3145
VWHS-GJR-t	-163.3026
VWHS-GARCH-t	-168.1141
VWHS-GARCH-n	-171.1154
MRS vwhs-t	-171.4239
MRS vwhs-n	-181.2922
HS-VIX	-181.6940
VWHS-EGARCH-n	-183.7851
VWHS-EGARCH-t	-185.8669
VWHS-GJR-n	-211.8743

Table 6: Deutsche Bank - Model ranking on 3-years average VaR (mil. EUR)

Model	Average
EGARCH-n	-64.5839
GARCH-n	-67.2984
EGARCH-t	-76.4460
GARCH-t	-78.4628
VWHS-EGARCH-n	-81.6160
MRS vwhs-n	-82.7443
VWHS-GARCH-n	-83.0017
MRS GARCH-n	-84.5381
VWHS-EGARCH-t	-85.6237
VWHS-GARCH-t	-88.0066
VWHS-GJR-t	-89.5659
VWHS-GJR-n	-91.4429
GJR-n	-92.9037
MRS vwhs-t	-93.9406
GJR-t	-103.2142
MRS GARCH-t	-127.0810
HS-VIX	-133.5823

Table 7: Swedbank - Model ranking on 3-years average VaR (mil. SEK)

Model	Average
GARCH-n	-19.0132
EGARCH-n	-19.5793
GARCH-t	-19.7971
MRS vwht	-20.8581
VWHS-GARCH-t	-21.7025
GJR-n	-22.0427
EGARCH-t	-22.1039
MRS vwht-n	-22.1600
VWHS-EGARCH-t	-22.3916
VWHS-GARCH-n	-22.6566
HS-VIX	-22.8882
GJR-t	-24.3368
VWHS-GJR-n	-24.5348
VWHS-GJR-t	-24.5818
VWHS-EGARCH-n	-25.5772
MRS GARCH-n	-27.7046
MRS GARCH-t	-28.8578

Appendix 6: GMM-duration backtesting results

Table 1: Bank of America

Model		GMM duration						
		J _{UC} (1)	J _{CC} (3)	J _{CC} (5)	J _{IND} (1)	J _{IND} (3)	J _{IND} (5)	
VIX - HS		0.0023 (0.9657)	1.5131 (0.4459)	2.0776 (0.4590)	0.1028 (0.3615)	1.2961 (0.2148)	2.0505 (0.2330)	
Normal	Volatility Weighted HS	GARCH	0.0567 (0.8492)	0.4784 (0.8152)	0.5857 (0.8803)	0.1286 (0.4470)	0.5240 (0.4969)	0.6207 (0.6492)
		EGARCH	0.0567 (0.8341)	0.4784 (0.8116)	0.5857 (0.8773)	0.1286 (0.4764)	0.5240 (0.4974)	0.6207 (0.6509)
		GJR	0.3130 (0.4573)	1.3627 (0.4762)	2.4802 (0.3688)	0.1148 (0.6732)	1.3322 (0.1749)	2.2928 (0.1860)
t		GARCH	0.0157 (0.9252)	0.0663 (0.9950)	0.3482 (0.9537)	0.1464 (0.2975)	0.1726 (0.8946)	0.2074 (0.9608)
		EGARCH	0.0157 (0.9170)	0.0663 (0.9962)	0.3482 (0.9577)	0.1464 (0.3320)	0.1726 (0.8944)	0.2074 (0.9621)
		GJR	0.0567 (0.7114)	0.4150 (0.8341)	0.8054 (0.7928)	0.1286 (0.5404)	0.4487 (0.5501)	0.8681 (0.5114)
Normal	GARCH	GARCH	3.0376 (0.0569)	3.2731 (0.1699)	3.3263 (0.2671)	0.0618 (0.9537)	0.6473 (0.4363)	1.1318 (0.4387)
		EGARCH	6.7779 (0.0122)	9.6151 (0.0256)	10.4251 (0.0362)	0.0484 (0.9822)	0.5562 (0.4876)	3.2290 (0.1143)
		GJR	0.3891 (0.5510)	21.3363 (0.0069)	22.2659 (0.0104)	0.0862 (0.6300)	29.5197 (0.0002)	35.0062 (0.0002)
t	GARCH	GARCH	0.3891 (0.5915)	0.7470 (0.7149)	0.8841 (0.7850)	0.0862 (0.6701)	0.6402 (0.4456)	0.7041 (0.6260)
		EGARCH	2.3878 (0.1219)	2.7803 (0.2036)	2.8365 (0.3209)	0.0654 (0.9390)	0.0981 (0.9887)	0.1460 (0.9861)
		GJR	0.0023 (0.9039)	20.0839 (0.0072)	22.1625 (0.0109)	0.1028 (0.3527)	15.9372 (0.0008)	16.3803 (0.0012)
Normal	MRS parametric	MRS parametric	0.4321 (0.4212)	1.1270 (0.5693)	1.4641 (0.6070)	0.1277 (0.1579)	0.2382 (0.7475)	0.7167 (0.6142)
		MRS vwhs	0.0157 (0.9141)	0.0663 (0.9948)	0.3482 (0.9602)	0.1464 (0.3494)	0.1726 (0.8917)	0.2074 (0.9628)
t	MRS parametric	MRS parametric	4.7520 (0.0291)	47.2090 (0.0024)	67.2034 (0.0022)	0.2024 (0.0276)	1.8340 (0.1207)	3.2745 (0.1079)
		MRS vwhs	0.0157 (0.8986)	0.0663 (0.9945)	0.3482 (0.9540)	0.1464 (0.2832)	0.1726 (0.8955)	0.2074 (0.9621)

Table 2: Bank of America - mimicking portfolio

Model		GMM duration						
		J _{UC} (1)	J _{CC} (3)	J _{CC} (5)	J _{IND} (1)	J _{IND} (3)	J _{IND} (5)	
VIX - HS	VIX	0.1214 (0.6505)	4.0109 (0.1165)	4.3385 (0.1735)	0.0937 (0.4849)	4.0555 (0.0172)	4.4337 (0.0539)	
	VSTOXX	0.3891 (0.6037)	7.1437 (0.0423)	7.7030 (0.0605)	0.0862 (0.6339)	7.8882 (0.0041)	7.9320 (0.0085)	
	VDAX	1.7860 (0.1578)	13.5649 (0.0112)	16.5670 (0.0144)	0.0696 (0.8583)	18.6378 (0.0002)	21.0731 (0.0002)	
Normal	Volatility Weighted HS	GARCH	1.8358 (0.1805)	2.7030 (0.2138)	4.7435 (0.1313)	0.0872 (0.8629)	1.2838 (0.1808)	6.0339 (0.0160)
		EGARCH	0.7201 (0.3675)	0.7437 (0.7028)	0.7958 (0.8045)	0.1037 (0.6999)	0.1185 (0.9855)	0.2272 (0.9491)
		GJR	0.7201 (0.4034)	1.7812 (0.3636)	5.4281 (0.1043)	0.1037 (0.6727)	1.7893 (0.0973)	6.6305 (0.0127)
		GARCH	1.8358 (0.1738)	2.7030 (0.2182)	4.7435 (0.1324)	0.0872 (0.8648)	1.2838 (0.1863)	6.0339 (0.0198)
		EGARCH	0.7201 (0.3906)	0.7437 (0.7000)	0.7958 (0.7996)	0.1037 (0.7087)	0.1185 (0.9859)	0.2272 (0.9435)
		GJR	1.2368 (0.2443)	2.3584 (0.2482)	4.4688 (0.1496)	0.0947 (0.8368)	1.7544 (0.1015)	4.0494 (0.0554)
t	Volatility Weighted HS	GARCH	5.2032 (0.0139)	7.7916 (0.0326)	9.4496 (0.0395)	0.0530 (0.9746)	1.6917 (0.1418)	3.9899 (0.0697)
		EGARCH	7.5943 (0.0119)	11.1083 (0.0201)	12.0464 (0.0287)	0.0465 (0.9829)	0.5736 (0.4766)	2.5762 (0.1681)
		GJR	5.9799 (0.0118)	8.8241 (0.0269)	9.8124 (0.0373)	0.0506 (0.9777)	0.8579 (0.3564)	1.1888 (0.4264)
		GARCH	3.0376 (0.0577)	3.6686 (0.1325)	4.0027 (0.1956)	0.0618 (0.9466)	0.2431 (0.7400)	1.3764 (0.3696)
		EGARCH	4.4510 (0.0433)	5.9058 (0.0582)	6.6360 (0.0757)	0.0556 (0.9742)	0.5650 (0.4776)	3.0307 (0.1215)
		GJR	1.7860 (0.1626)	4.6897 (0.0945)	6.6840 (0.0779)	0.0696 (0.9023)	4.0334 (0.0182)	4.3795 (0.0565)
Normal	MRS parametric	MRS parametric	11.8839 (0.0035)	21.6919 (0.0086)	26.6968 (0.0097)	0.0388 (0.9818)	4.9155 (0.0109)	5.3743 (0.0321)
		MRS vwhs	1.2368 (0.2110)	4.4959 (0.0865)	9.1460 (0.0371)	0.0947 (0.8146)	5.6520 (0.0060)	8.5613 (0.0041)
	MRS vwhs	MRS parametric	1.2426 (0.1966)	3.1898 (0.1810)	4.4613 (0.1701)	0.0743 (0.8160)	2.5999 (0.0561)	2.9537 (0.1334)
		MRS vwhs	1.2368 (0.2216)	4.5118 (0.0842)	9.2184 (0.0392)	0.0947 (0.8318)	5.6896 (0.0068)	8.6643 (0.0048)

Table 3: Danske Bank

Model		GMM duration						
		J _{UC} (1)	J _{CC} (3)	J _{CC} (5)	J _{IND} (1)	J _{IND} (3)	J _{IND} (5)	
VIX - HS		4.4510 (0.0432)	9.7479 (0.0229)	12.7495 (0.0232)	0.0556 (0.9688)	6.1520 (0.0061)	6.2423 (0.0184)	
Normal	Volatility Weighted HS	GARCH	0.0567 (0.7153)	0.1065 (0.9789)	0.5235 (0.8961)	0.1286 (0.5106)	0.1586 (0.9261)	0.4915 (0.7357)
		EGARCH	0.7201 (0.3577)	0.8985 (0.6287)	1.3534 (0.6224)	0.1037 (0.7751)	0.2946 (0.6943)	0.8841 (0.5061)
		GJR	0.0157 (0.9350)	0.7785 (0.6771)	6.2514 (0.0800)	0.1464 (0.2869)	0.3841 (0.6048)	3.4997 (0.0853)
t		GARCH	0.7201 (0.3683)	0.9012 (0.4627)	1.2952 (0.5081)	0.1037 (0.7671)	0.9322 (0.2834)	2.9053 (0.1174)
		EGARCH	1.2368 (0.2101)	1.4822 (0.4352)	1.8552 (0.5016)	0.0947 (0.8113)	0.2648 (0.7315)	0.8841 (0.5151)
		GJR	0.3130 (0.5195)	0.4960 (0.8096)	1.9822 (0.4703)	0.1148 (0.5351)	0.3679 (0.6184)	2.2118 (0.1895)
Normal	GARCH	5.2032 (0.0152)	6.8308 (0.0417)	6.9678 (0.0693)	0.0530 (0.9760)	0.0864 (0.9944)	0.1310 (0.9933)	
	EGARCH	6.7779 (0.0139)	9.9106 (0.0261)	10.6624 (0.0350)	0.0484 (0.9787)	0.4335 (0.5722)	0.5976 (0.6881)	
	GJR	11.0035 (0.0027)	17.5762 (0.0088)	19.4268 (0.0115)	0.0401 (0.9804)	0.9173 (0.3203)	4.5083 (0.0486)	
t	GARCH	1.7860 (0.1900)	2.0401 (0.3212)	2.0891 (0.4571)	0.0696 (0.9010)	0.0715 (0.9981)	0.6934 (0.6334)	
	EGARCH	1.7860 (0.1701)	2.7774 (0.2101)	3.2967 (0.2689)	0.0696 (0.8765)	1.0061 (0.2909)	1.2441 (0.4055)	
	GJR	3.0376 (0.0690)	3.5582 (0.1414)	3.6517 (0.2267)	0.0618 (0.9507)	0.1293 (0.9469)	0.4444 (0.7808)	
Normal	MRS parametric	17.3269 (0.0006)	32.1297 (0.0051)	37.8912 (0.0068)	0.0325 (0.9829)	2.1159 (0.0945)	2.2724 (0.2037)	
	MRS vwhs	0.0567 (0.7293)	0.1390 (0.9585)	0.9037 (0.7514)	0.1286 (0.4243)	0.1847 (0.8617)	0.8600 (0.5248)	
t	MRS parametric	1.7860 (0.1558)	2.4244 (0.2562)	2.7951 (0.3258)	0.0696 (0.9066)	0.5183 (0.5011)	0.6706 (0.6439)	
	MRS vwhs	0.0567 (0.8129)	0.1329 (0.9581)	0.2991 (0.9635)	0.1286 (0.5124)	0.1831 (0.8628)	0.3091 (0.8829)	

Table 4: Deutsche Bank

Model		GMM duration						
		J _{UC} (1)	J _{CC} (3)	J _{CC} (5)	J _{IND} (1)	J _{IND} (3)	J _{IND} (5)	
VIX - HS		3.0376 (0.0737)	13.1285 (0.0159)	20.9975 (0.0138)	0.0618 (0.9556)	18.5817 (0.0002)	18.8654 (0.0004)	
Normal	Volatility Weighted HS	GARCH	3.9615 (0.0338)	10.1248 (0.0193)	16.1894 (0.0157)	0.0706 (0.9657)	11.1452 (0.0008)	11.9211 (0.0011)
		EGARCH	3.9615 (0.0345)	10.1210 (0.0185)	16.1644 (0.0146)	0.0706 (0.9660)	11.1242 (0.0009)	11.9198 (0.0014)
		GJR	2.4980 (0.1105)	6.4926 (0.0465)	11.1152 (0.0274)	0.0808 (0.9146)	7.1544 (0.0035)	8.4856 (0.0046)
t		GARCH	5.5561 (0.0266)	17.5751 (0.0104)	30.4735 (0.0068)	0.0628 (0.9785)	27.8288 (0.0001)	28.4989 (0.0001)
		EGARCH	5.5561 (0.0277)	17.1029 (0.0088)	29.3816 (0.0053)	0.0628 (0.9762)	26.1853 (0.0001)	26.9575 (0.0001)
		GJR	3.9615 (0.0409)	10.9940 (0.0177)	18.6368 (0.0144)	0.0706 (0.9720)	13.9434 (0.0004)	14.5823 (0.0006)
Normal		GARCH	3.0376 (0.0583)	14.8777 (0.0110)	22.0943 (0.0107)	0.0618 (0.9520)	21.5752 (0.0002)	22.9643 (0.0002)
		EGARCH	5.2032 (0.0165)	12.0076 (0.0168)	16.4596 (0.0172)	0.0530 (0.9776)	9.1475 (0.0026)	9.7153 (0.0045)
		GJR	3.0376 (0.0707)	9.2679 (0.0283)	16.2462 (0.0183)	0.0618 (0.9588)	11.4601 (0.0010)	13.1678 (0.0015)
t	GARCH	0.3891 (0.5587)	7.9129 (0.0332)	11.7651 (0.0290)	0.0862 (0.6385)	9.4149 (0.0020)	11.6645 (0.0025)	
	EGARCH	1.7860 (0.1939)	10.2765 (0.0237)	15.8585 (0.0203)	0.0696 (0.8566)	14.0668 (0.0005)	14.6634 (0.0013)	
	GJR	0.4321 (0.4785)	3.2749 (0.1725)	3.9337 (0.2084)	0.1277 (0.1422)	1.0060 (0.2993)	2.6423 (0.1607)	
Normal	MRS parametric	4.4510 (0.0412)	8.6355 (0.0284)	11.4366 (0.0291)	0.0556 (0.9704)	4.4680 (0.0137)	4.7118 (0.0458)	
	MRS vwhs	1.8358 (0.1561)	3.0861 (0.1730)	4.2336 (0.1674)	0.0872 (0.8632)	1.4571 (0.1470)	2.0302 (0.2176)	
t	MRS parametric	0.4321 (0.5051)	1.8587 (0.3738)	2.6603 (0.3605)	0.1277 (0.1668)	0.4455 (0.5624)	1.5741 (0.3147)	
	MRS vwhs	3.9615 (0.0310)	13.2900 (0.0130)	22.9474 (0.0103)	0.0706 (0.9646)	19.9708 (0.0002)	20.8688 (0.0002)	

Table 5: Swedbank

Model		GMM duration						
		J _{UC} (1)	J _{CC} (3)	J _{CC} (5)	J _{IND} (1)	J _{IND} (3)	J _{IND} (5)	
VIX - HS		1.7860 (0.1779)	7.2217 (0.0386)	11.6567 (0.0278)	0.0696 (0.8601)	8.9880 (0.0022)	9.8882 (0.0046)	
Normal	Volatility Weighted HS	GARCH	3.2679 (0.0522)	6.7368 (0.0456)	9.9397 (0.0347)	0.0000 (0.9807)	5.4062 (0.0065)	5.7445 (0.0188)
		EGARCH	1.8832 (0.1580)	5.0245 (0.0734)	9.8599 (0.0359)	0.0000 (0.9814)	7.0283 (0.0043)	9.3388 (0.0047)
		GJR	0.3359 (0.5319)	0.3872 (0.8562)	0.5256 (0.8999)	0.0000 (0.9879)	0.1807 (0.8767)	0.7453 (0.5801)
		GARCH	3.2679 (0.0521)	6.7368 (0.0436)	9.9397 (0.0328)	0.0000 (0.9791)	5.4062 (0.0066)	5.7445 (0.0193)
		EGARCH	2.5511 (0.1320)	5.2410 (0.0691)	8.1582 (0.0486)	0.0000 (0.9797)	4.5731 (0.0103)	5.3740 (0.0239)
		GJR	0.7528 (0.4247)	0.7573 (0.7223)	0.9743 (0.7391)	0.0000 (0.9863)	0.1346 (0.9667)	1.2145 (0.3930)
t	Volatility Weighted HS	GARCH	0.9035 (0.4024)	0.9748 (0.6186)	1.5063 (0.5963)	0.0799 (0.7316)	0.1956 (0.8117)	1.2673 (0.3854)
		EGARCH	7.8976 (0.0049)	17.0692 (0.0111)	26.1649 (0.0093)	0.0466 (0.9820)	15.3988 (0.0003)	16.9061 (0.0004)
		GJR	2.5939 (0.0797)	6.4483 (0.0509)	8.0243 (0.0562)	0.0655 (0.9256)	4.6814 (0.0137)	4.7088 (0.0476)
		GARCH	0.3192 (0.6168)	2.2507 (0.2814)	2.7408 (0.3438)	0.1278 (0.1913)	1.0048 (0.2961)	1.3934 (0.3582)
		EGARCH	0.9035 (0.3507)	5.4735 (0.0685)	7.1989 (0.0669)	0.0799 (0.7458)	6.0369 (0.0067)	6.2982 (0.0182)
		GJR	2.5939 (0.0795)	15.0645 (0.0118)	17.0280 (0.0149)	0.0655 (0.9365)	26.2262 (0.0001)	33.0281 (0.0001)
Normal	MRS parametric	4.6352 (0.0335)	70.9674 (0.0014)	94.8129 (0.0014)	0.2025 (0.0308)	2.8925 (0.0423)	5.0741 (0.0372)	
	MRS vwhs	2.5511 (0.1329)	5.0376 (0.0740)	7.5241 (0.0570)	0.0000 (0.9826)	3.9002 (0.0144)	4.4934 (0.0429)	
	MRS parametric	15.5233 (0.0024)	111.6150 (0.0014)	147.5770 (0.0014)	0.3354 (0.0008)	0.3945 (0.5819)	0.7023 (0.6192)	
	MRS vwhs	1.8832 (0.1754)	3.8843 (0.1172)	6.8057 (0.0668)	0.0000 (0.9804)	3.8002 (0.0165)	5.8569 (0.0194)	

Appendix 7. Selective Matlab Code

Matlab function for the forecast of mean and variance through a GARCH process

```
function [c] = forecastrw(series, rw, spec)
%forecasts conditional mean and variance on a rolling window using the
functions from the Econometric toolbox
%series = the P/L data series
%rw = the length of the rolling window
%spec = the GARCH specification structure for each model (needs to be set
%before applying the forecast, using the garchset function
c=zeros(988,1);
for i=1:(numel(series)-rw)
    s=series(i:(i+rw));
    [Coeff] = garchfit(spec, s);
    [sigmaforecast, meanforecast]=garchpred(Coeff,s,1);
    c(i,1)=sigmaforecast;
    c(i,2)=meanforecast;
end
end
```

Matlab function for the estimation of parametric VaR

```
function [varfor]=paramvar(series,vol_mean,conf,dist)
% computes parametric VaR for normal and t-distribution
%series - the return series to be estimated
%vol_mean - volatility and mean forecasts matrix through garch models
%conf- the confidence level for estimation
%dist - distribution 0 Normal 1- Student t
for i=1:numel(vol_mean(:,1))
    if dist==0
        q=norminv(conf,0,1);
        varfor(i,1)=-(-vol_mean(i,2)+q*vol_mean(i,1));
    else
        k=kurtosis(series(i:i+251,1));
        df=(4*k-6)/(k-3);
        q=tinv(conf,df);
        varfor(i,1)=-(-vol_mean(i,2)+q*sqrt((df-2)/df)*vol_mean(i,1));
    end
end
end
end
```

Matlab function for the estimation of volatility weighted historical simulation VaR

```
function [vwhs]=varvwhs(series,volforecast)
%computes the Volatility Weighted Historical Simulation VaR
%series - data series of returns
%volforecast - vector of variance(std deviations) forecasts from a GARCH
process in our case

for i=1:(numel(volforecast(:,1))-252)
    s = series(i+252:i+252+251,1);%we actually used a rw of 253 to estimate
the volatility forecasts
    vf=volforecast(i:i+251+1,1);% we need the volatility forecast for the
last day as well=>253 observations
```

```

    for k=1:252;
        resc(k,1)=s(k,1)*sqrt(vf(253,1))/sqrt(vf(k,1));
    end
    vwhs(i,1)= prctile(resc(:,1),0.2);
end
end

```

Matlab function for the estimation of VaR using the HS-VIX model

```

function [vixvol]=vixhs(series,volindex)
%computes the Volatilitt Weighted Historical Simulation VaR
%series - data series of returns
%volforecast - vector of variance(std deviations) forecasts from a GARCH
process in our case

for i=1:(numel(volindex)-252)%i from 1 for all banks and 7 for swedabnk
since we have less observations
    s = series(i:i+252,1);%we actually used a rw of 253 to estimate the
volatililtiy forecasts
    % vf=volindex(i+6:i+252+6,1);% for swedbank
    vf=volindex(i:i+252,1);
    for k=2:253;%still want a 252 rw-equivalent to volatility for day 254
        resc(k-1,1)=s(k,1)*vf(253,1)/vf(k-1,1);
    end

    vixvol(i,1)= prctile(resc(:,1),0.2);
end
end

```