Erasmus University Rotterdam

MSc. Quantitative Finance

Master Thesis

Forecasting and Evaluating Portfolio Value-at-Risk: A Multivariate Approach

Shell Asset Management Company

Supervisors:
Prof. Dr. Dick van Dijk
dr. Kaj Martensen
Co-reader:
dr. Wing Wah Tham

Author:
Ilayda Ozyildiz

Abstract

This paper compares RiskMetrics and several multivariate GARCH models that are used to forecast Value-at-Risk. We consider a data set that includes the benchmarks that SAMCo uses, the data consists of both fixed income and equity indices. We evaluate the forecasting power of the Value-at-Risk of these models by using the backtesting test and the Comparative Predictive Ability test (CPA). Also the economic significance of time-varying, predictable volatility is examined by using both the minimum-variance and the mean-variance asset allocation rules. We find that incorporating the asymmetry in the correlation between assets in the models to forecast Value-at-Risk bears fruit, as they have a better performance for the statistical tests. In general, assigning a Student-t distribution to the error terms leads to an improvement of the model according to the CPA-test. Finally, we find that it is unlikely that the gains to volatility-timing are due to chance.

Acknowledgements

All things come to an end and the time has come to write the final words of this thesis. Many people deserve my most sincere thanks for their support during my thesis. First of all, I would like to thank Dick van Dijk for his supervision and contribution in this thesis. His practical comments are gratefully acknowledged. In his busy agenda, there was always time to discuss my research. Furthermore, I would like to thank Wing Wah Tham for reviewing the earlier draft of this thesis, his valuable feedback and for creating awareness on certain topics. Your feedback was very useful and highly appreciated. This thesis would not have been possible without SAMCo and Kaj Martensen. I thank Kaj for providing me the opportunity to have a glimpse behind the scenes. It was a very valuable experience to be at the firm every day and have challenging and inspiring chats with the bright people that work there. I would like to take this opportunity to thank Rabia Ibrahim. We have had very nice conversations during many lunches. I am grateful for the time you spent with me during my internship at SAMCo and I wish you great success with your new job in the Philippines. I would also like to thank Tomasz Katzur for his feedback during the early stage of my thesis. Of course I would like to thank all other people that work at SAMCo for making sure that I have a pleasant time at the office. I am lucky to have many friends and family around that supported me not only during my thesis but throughout my life. I am grateful to my parents for their unconditional trust and support. A big thank you to my friends and sister for the fun outside the university life and for making me realize that there is much more in life than my computer screen and econometric models.

Ilayda Ozyildiz
Rotterdam, March 2013

Contents

Chapter 1

Introduction

Managing portfolio risk is of paramount importance to portfolio managers due to the financial regulations imposed by the Basel Accords. Risk managers have to come up with models that can take into account all kinds of risk and they use several risk measures to manage the risk. Value-at-Risk is one of the widely used risk measures to control and manage market risks, which are risks to a financial portfolio from movements in market prices such as equity prices, FX rates and interest rates. Value-at-Risk is the maximum loss that could occur over a given holding period or the minimum return that could occur over a given holding period with a specified confidence level under adverse market conditions. Until now no agreement has been reached on which procedure provides the most accurate VaR estimates. Inaccurate VaR-estimates can lead to disastrous consequences: Underestimating the VaR means that the investor is not aware of the maximum loss that could be made when investing on a specific manner. On the other side, overestimating VaR means that the investor will lose opportunity costs, because setting money aside to absorb losses can reduce the money they have to make bets.

A potential problem with estimating the VaR of a portfolio lies in inadequate covariance matrix estimation. Specifically if we assume constant correlations between asset prices. The goal of this paper is to compare the performance of the RiskMetrics model with multivariate GARCH models that can incorporate the dynamic changes in the correlations between asset prices and the asymmetry in the correlations between asset prices when implemented to forecast Value-at-Risk. The findings of this paper will have several important implications from practitioner points of view. The Value-at-Risk that is calculated by a RiskMetrics model could lead to an underestimation in the risk of a portfolio, this can have tremendous effects for financial institutions. Furthermore, assuming conditional correlations to be time-varying (rather than constant) and allowing for assymetric effects can lead to VaR forecast improvements. In this paper we will investigate whether this believes hold for our portfolio, which will be explained in detail in the latter of this section and in the Data section of this research.

Past research has shown that assuming constant correlations between asset classes is not a realistic assumption. Especially in some contexts, which are listed below, it is shown that correlations vary over time, hence it is of great importance to model the correlations dynamically: First, in crisis periods asset prices move more closely together, which is usually called contagion. Negative returns move more together than positive returns, especially in equity markets. This asymmetric dependence has been reported by many previous studies including Erb et al. (1994), Longin and Solnik (2001), Ang and Bekaert (2002, Ang and Chen (2002), Das and Uppal (2003) and Patton (2004). Second, in a stock-bond setting correlations have been shown to change over time (Connolly et al. (2005)). Finally, in studies related to global equity market integration it is proven that correlations between markets change over time. The dependence between international stock markets can change because of increasing economic and financial integration (Bekaert and Harvey (1995), Longin and Solnik (1995), Goetzmann et al. (2005), Cappiello et al. (2006), Patton (2006b), Bekaert et al. (2009), for example.), macroeconomic conditions (Bracker and Koch (1999)) and market liquidity (Baela et al. (2009)). Christoffersen et al. (2010) show that correlations have been significantly trending upward for both developed markets and emerging markets. Hence it seems like it is plausible that models that can incorporate the dynamics and/or the asymmetry in the correlation between assets produce more accurate Value-at-Risk forecasts.

This research is performed for Shell Asset Management Company (SAMCo). SAMCo is an asset management company established in 2006 to provide investment advice and asset management services to pension funds associated with Shell worldwide. Total assets under management are in the order of 35 billion EUR at mid 2010. It is SAMCO's aim to be a relatively small, internationally oriented and dynamic organisation. Continuous innovation in financial products, changes in industry supervisory regulations and client needs make it a professionaly challenging and dynamic environment.

The data used in this research include the benchmarks that SAMCo uses. The SSPS fund, which is the largest fund that SAMCo uses, consists of Financial Investments and Treasury. The Financial Investments consist of Liability Hedge and Return Seeking Assets. The Return Seeking Assets can be split into Cash, Equity, Fixed Income, Hedge Funds, Other, Private Equity and Real Estate. In our research we will focus on the benchmarks of equity and fixed income. The dataset consists of six equity indices (MSCI World, MSCI World Small Cap, MSCI Emerging Markets, MSCI North America, MSCI Europe and MSCI Japan) and seven fixed income indices (Merrill EMU Direct Governments, iBoxx Euro Financials, iBoxx Euro Non-Financials, iBoxx USD Treasuries Total Return Index, Merrill Euro High Yield, Merrill US High Yield and Merrill Emerging Market Governments). We use weekly data for the period starting at May 7th, 1999 and ending at May 18th, 2012, which results in 681 observations.

Currently, SAMCo uses the RiskMetrics approach, hence we will use this model as a natural benchmark. The Dynamic Conditional Correlation models are used to construct the covariance forecasts in order to compare their performance with the performance of the RiskMetrics Model. We also consider multivariate models with asymmetric time-varying correlations, alternative distributions for the innovations and we use multivariate DCC models with different specifications for the univariate volatilities, such as GARCH and the Threshold-GARCH model. The forecasts are evaluated by using the backtesting tests based on coverage/independence criteria proposed by Christoffersen (1998). Also the comparative predictive ability (CPA) test, which is a statistical test designed to evaluate the comparative predictive performance among candidate models is used to enhance the backtesting analysis. Besides using statistical tests to compare the different models with each other we will also look at their economic significance. More specifically, we examine the economic value of volatility timing. This will be done by investigating the ex-post portfolio returns, Sharpe Ratios and the Manipulation Proof Performance Measure (MPPM) that emerge from the minimum-variance and mean-variance asset allocation rules. Some of our dynamic models are used to produce the covariance forecasts. Finally, we investigate the statistical significance of the volatility timing results by conducting simulations. For this analysis we use a simple portfolio that consists of a stock, bond and a risk-free rate.

We find that the ADCC-GARCH-t and ADCC-TGARCH-t models pass the conditional coverage test of Christoffersen (1998). These models take the asymmetry between asset classes into account. Obviously, incorporating this asymmetry bears fruit as the statistical tests indicate that the ADCC-GARCH-t and ADCC-TGARCH-t are the best performing models. Furthermore, RiskMetrics is outperformed by eight other models according to the CPA-test. In general, assigning a Student-t distribution to the error terms leads to an improvement of the model according to the CPA-test. For the economic significance results we find that it is unlikely that the gains to volatility timing are due to chance, because the frequency (in %) with which the simulation beat the portfolio according to the Sharpe Ratio lie between 2.6% and 8.7% and according to the MPPM between 3.5% and 8.9% . For the minimum-variance analysis RiskMetrics performs best, because the ex-post portfolio volatility is the lowest. However, for the mean-variance analysis the ADCC-TGARCH-t model performs best, according to the the frequency with which the portfolio outperformed the simulation based on the Sharpe Ratio and the MPPM.

The remainder of this paper is organized as follows. Section 2 describes the literature review and Section 3 describes the data set used in this research. In section 4 we elaborate on the methodology. Section 5 brings the results and discussion. Finally, Section 6 concludes, gives specific insights for SAMCo and brings suggestions for further research. Appendix A.1 shows the time series of the returns. Appendix A.2 shows the scatter plots of the return series. Appendix A.3. shows the table of the covariance matrix of the return series, Appendix A.4. presents the derivation of the variance-covariance matrix Ω that is a necessary input for the test developed by Hong, Tu and Zhou (2007) and Appendix A.5. shows the empirical autocorrelation functions of the return series.

Chapter 2

Literature Review

Understanding and predicting the dependence between asset returns is important for many issues in finance. Over the past few years there has been done a lot of research about this topic, because it opens the door to better decision tools in various areas, such as asset pricing, risk management and portfolio management. Engle (1982) introduced the Generalized Autoregressive Conditional Heteroskedasticity model (GARCH), which is a powerful model that is able to capture volatility clustering. This model is now widely used to describe and forecast changes in the volatility of financial time series. However, it still has several limitations; the parameters of the model are constrained by non-negativity to ensure that the volatilities remain larger than zero at all times and it is unable to account for different regimes in volatility. For this purpose, several extensions have been made in order to incorporate these limitations, for example the Threshold-GARCH model (T-GARCH), the Non-Linear GARCH model (N-GARCH), the exponential GARCH model (E-GARCH) and the quadratic GARCH model (Q-GARCH).

Engle (2002) proposes a new class of multivariate models called dynamic conditional correlation models. These have the flexibility of univariate GARCH models coupled with parsimonious parametric models for the correlations. It has three advantages over other estimation methods. First, the DCC-model estimates correlation coefficients of the standardized residuals and thus accounts for heteroskedasticity separately. Second, the model allows to include additional explanatory variables in the mean equation to ensure that the model is well specified. Third, the multivariate GARCH model can be used to examine multiple asset returns without adding too many parameters.[1]. However, the DCC model has been criticized because the DCC-model is an assumed rather than derived model, because it models the conditional covariances of the standardized residuals and hence does not yield conditional correlations between asset prices.

[1]The full vec and the BEKK model (Enle and Kroner, 1995) would become computationally very intensive if expanded to three asset returns

Santos, Nogales and Ruiz (2013) compare multivariate and univariate GARCH models to forecast portfolio Value-at-Risk. They find that even in large systems, it could be worth to predict the VaR of a portfolio by fitting multivariate modes. Our research is in line with their research but goes beyond statistical evaluations and does not focus on the comparison between multivariate and univariate GARCH models solely. Furthermore, Santos et al. (2013) include the Constant Conditional Correlation model (CCC) in their investigation, we only focus on the Dynamic Conditional Correlation model (DCC). Our goal is to investigate if RiskMetrics can be outperformed by other models that can incorporate the dynamic changes in the correlation and the asymmetry in the correlation between asset prices.

As in line with Santos et al. (2013) we use the backtesting test developed by Christoffersen (1998) and the CPA-test developed by Giacomini and White (2006) to assess our Value-at-Risk forecasts. To extend the comparison we also compare the economic significance of our models as in line with Fleming, Kirby and Ostdiek (2001). We examine if our models have economic value by using the minimum-variance and mean-variance asset allocation rules for a simple stock, bond and cash portfolio. Fleming et al. (2001) estimate the Conditional Covariance Matrix by using rolling estimators that are constructed in an asymptotically manner as in line with Foster and Nelson (1996). We use the RiskMetrics approach and the best performing DCC-model to estimate the covariance matrix, which is then used to construct the optimal portfolio.

Besides the symmetric DCC-model we also use the asymmetric DCC model to forecast VaR. In the past couple of years researchers have developed various extensions on the DCC-model: Hafner and Franses (2003) introduce a model that allows for asset-specific correlation sensitivities, which is useful in particular if one aims to summarize a large number of asset returns. Billio and Caporin (2006) introduce a block structure in parameter matrices that allows for interdependence with a reduced number of parameters. Their model nests the Flexible DCC-model of Billio et al. (2006) and is named Quadratic Flexible DCC Multivariate GARCH. Bauwens et al. (2006) present an overview of the literature on multivariate GARCH models.

Chapter 3

Data

3.1 Data description

The data set used in this research covers several types of stock and bond indices. The data set presented in Table 3.1, includes the benchmarks that SAMCo uses and are weekly end-of-week prices (total return indices, hence including dividend and coupon payments) expressed in USD. The SSPF fund, which is the largest fund that SAMCo uses consists of Financial Investments and Treasury. The Financial Investments consist of Liability Hedge and Return Seeking Assets. The Return Seeking Assets can be split into Cash, Equity, Fixed Income, Hedge Funds, Other, Private Equity and Real Estate. In our research we will focus on the equity and fixed income benchmarks. The data were downloaded from Bloomberg and cover the range May 7th, 1999 - May 18th, 2012 at the weekly frequence. Hence, we have 681 observations.

In a first step the price data has been converted into historical week-to-week excess return data given by equation 3.1:

$$Return_t = log(\frac{Price_t}{Price_{t-1}}) \tag{3.1}$$

The time series of the returns are presented in Figure A.1.1, A.1.2 and A.1.3. All indices behave more or less the same during the crisis period, with return values that can decrease down to 22% for the MSCI Europe index. The time series pattern of some series are similar so that we can expect correlations to be high. It is obvious that equities have a much more negative return value during the crisis and are more volatile than the fixed income indices. The largest negative value for the fixed income index is -8.2% for the Merrill US High Yield index.

Table 3.1: **Description of the data set**

Equity	Fixed Income
World Indexes (Large and Small cap) MSCI World (MW) *The MSCI World index is a free-float weighted equity index. It includes developed markets and does not include emerging markets*	**Governments** Merrill EMU Direct Governments (MDG) *MDG tracks the performance of EUR denominated sovereign debt publicly issued by Euro member countries in either the eurobond market or the issuer's own domestic market.*
MSCI World Small Cap (MSC) *MSC captures small cap representation across 24 developed markets. With 4272 constituents, the index covers 14% of the free float-adjusted market capitalization in each country.*	Merrill Emerging Market Governments (MEMG) *MEMG contains all securities in The BofA Merrill Lynch Emerging Markets Corporate Plus Index that are rated investment grade based on an average of Moody's, S&P and Fitch. Index constituents are capitalization-weighted, based on their current amount outstanding.*
Developed Markets MSCI North America (MNA) *MNA is designed to measure the performance of the large and mid cap segments of the US and Canada markets. With 701 constituents, the index covers 85% of the free float-adjusted market capitalization in the US and Canada.*	**Corporates (Financial and Non-Financial)** iBoxx Euro Financials (IBF) *Bonds secured by a floating charge over some or all assets of the issuer are considered corporate bonds. Corporate bonds are further classified into Financials and Non-Financials. Financials focus on the following market sectors: Banks, Life/Nonlife insurance, Financial Services (also Real Estate) and Insurance-Wrapped.*
MSCI Europe (ME) *ME captures large and mid cap representation across 16 developed markets countries in Europe. With 436 constituents, the index covers 85% of the free float-adjusted market capitalization across the European Developed Markets equity universe.*	iBoxx Euro Non-Financials (IBNF) *Non-Financials focus on Oil & Gas, Basic Materials, Industrials, Consumer Goods, Health Care, Consumer Services, Telecommunications, Utilities and Technology.*
MSCI Japan (MJ) *MJ is designed to measure the performance of the large and mid cap segments of the Japan market. With 317 constituents, the index covers 85% of the free float-adjusted market capitalization in Japan.*	Merrill Euro High Yield (MHY) *MHY contains all non-financial securities in the BofA Merrill Lynch Euro High Yield Index that are rated BB1 through BB3, based on an average of Moody's, S&P and Fitch. Index constituents are capitalization-weighted, based on their current amount outstanding.*

	Merrill US High Yield (MUHY) *MHY contains all non-financial securities in the BofA Merrill Lynch US High Yield Index that are rated BB1 through BB3, based on an average of Moody's, S&P and Fitch. Index constituents are capitalization-weighted, based on their current amount outstanding.*
Emerging Markets MSCI Emerging Markets (MEM) *MEM captures large and mid cap representation across 21 Emerging Markets countries. With 821 constituents, the index covers 85% of the free float-adjusted market capitalization in each country.*	**Treasuries** iBoxx USD Treasuries Total Return Index (IBT) *IBT represents the investment grade fixed income market for USD denominated bonds.*

This table shows the composition of the equity and fixed income indices with their abbreviations in brackets used in this research

3.2 Descriptive Statistics and Stylized Facts

In this section we will further analyze the data set and we will do this by investigating the descriptive statistics of the return series and the stylized facts which are the properties that many asset returns have. This section will focus on the univariate characteristics, whereas the next sections will focus on multivariate characteristics such as the correlations between the assets.

Empirical research has shown that many asset returns have the same properties. Taylor (2005) proposes that these properties can be summarized by the stylized facts, which are as follows:

◇ Distribution of returns is not normal.
◇ No significant autocorrelations in returns.
◇ Small, but slowly declining autocorrelations in squared and absolute returns.

In the remainder of this section we will investigate if these stylized facts also hold for the equity and fixed income benchmarks that SAMCo uses. Table 3.2 presents the summary statistics of the return series. The first stylized fact is that the distribution of returns is not normal. From Table 3.2 we can see that all returns show excess kurtosis and are negatively skewed. The excess kurtosis is due to some large shocks on the data inducing fat tails. MUHY has the highest kurtosis value (29.605) and MDG has the lowest kurtosis value (3.458). None of the return series satisfy or even come close to normality as the p-value of the Jarque-Bera test statistic for all return series are almost equal to zero. The average returns do not deviate much from each other, whereas the standard deviations do deviate from each other.

The second stylized fact implies that there are no significant autocorrelations in the returns. This is in line with the returns used in this research. A graphical representation of the empirical autocorrelation function can be found in Figure A.3.1, A.3.2 and A.3.3 of the Appendix. The third stylized fact is that the autocorrelations in squared and absolute returns are slowly declining. Most of the returns exhibit this feature. However, for some of the returns this is less obvious, for example for the IBF and IBNF indices. The presence of the third stylized fact means that most returns suffer from volatility clustering.

Table 3.2: **Descriptive Statistics for international equity and fixed income indices.**

	Mean	St.Dev	Min	Max	Skewness	Excess Kurtosis	P-value
MW	-0.007	2.647	-0.224	0.116	-1.149	8.881	0.001
MSC	0.109	2.806	-0.198	0.118	-1.067	5.683	0.001
MEM	0.120	3.372	-0.226	0.185	-0.763	5.994	0.001
MNA	-0.002	2.763	-0.206	0.120	-0.793	6.296	0.001
ME	-0.041	2.864	-0.244	0.124	-1.055	9.106	0.001
MJ	-0.100	2.867	-0.223	0.096	-0.896	5.302	0.001
MDG	0.109	1.619	-0.061	0.06	-0.134	0.458	0.038
IBF	0.079	0.600	-0.051	0.019	-1.464	9.164	0.001
IBNF	0.091	0.470	-0.021	0.018	-0.487	1.311	0.001
IBT	0.111	0.646	-0.030	0.024	-0.405	0.965	0.001
MHY	0.131	0.937	-0.080	0.052	-1.506	14.423	0.001
MUHY	0.135	0.8	-0.086	0.051	-2.395	26.605	0.001
MEMG	0.153	0.698	-0.066	0.034	-2.201	18.781	0.001

This table presents the descriptive statistics for the return indices. Returns are weekly, denominated in USD, include dividends and are excess returns. The sample period is May 7th, 1999 - May 18th, 2012. The P-value evolves out of the Jarque-Bera test, which is a test for normality based on the skewness and kurtosis.

Table 3.3 shows the correlation matrix for the return series for the whole sample period. In this table we find positive and in some cases very high correlations between the equity assets, which means that there is an opportunity in incorporating correlations. The equity and the fixed income assets are in some cases negatively correlated with magnitudes up to 0.51. The fixed income assets are in most cases postively correlated with each other with magnitudes up to 0.78. In order to check the presence of perfect correlation between the MSCI Equity indexes and the MSCI world index, we add up MNA, ME and MJ up and check its correlation with the rest. It follows that the last mentioned series is almost perfectly correlated with MW with a correlation value of 0.95. Furthermore, this return series is highly correlated with the rest of the equity indices. On the contrary, it does not correlate a lot with the fixed income indices. We have to take these findings into account when we start modelling the asset returns. Table A.2.1 in the Appendix shows the covariance matrix with the variances of the indices on the diagonal.

Table 3.3: **Correlation Matrix**

	MW	MSC	MEM	MNA	ME	MJ	MDG	IBF	IBNF	IBT	MHY	MUHY	MEMG	MNA+ME+MJ
MW	1	0.93	0.81	0.96	0.90	0.66	0.20	0.12	-0.03	-0.31	0.41	0.47	0.15	0.95
MSC	-	1	0.84	0.87	0.81	0.68	0.23	0.13	-0.03	-0.32	0.46	0.53	0.19	0.89
MEM	-	-	1	0.72	0.73	0.65	0.23	0.13	0.03	-0.27	0.44	0.51	0.26	0.80
MNA	-	-	-	1	0.84	0.55	0.06	0.07	-0.06	-0.32	0.37	0.41	0.07	0.90
ME	-	-	-	-	1	0.59	-0.08	0.06	-0.09	-0.38	0.39	0.41	0.01	0.92
MJ	-	-	-	-	-	1	0.03	0.10	-0.04	-0.25	0.35	0.39	0.12	0.81
MDG	-	-	-	-	-	-	1	0.38	0.35	0.33	0.08	0.20	0.55	0.00
IBF	-	-	-	-	-	-	-	1	0.78	0.43	0.40	0.47	0.59	0.09
IBNF	-	-	-	-	-	-	-	-	1	0.59	0.35	0.42	0.59	-0.07
IBT	-	-	-	-	-	-	-	-	-	1	-0.13	0.06	0.53	-0.36
MHY	-	-	-	-	-	-	-	-	-	-	1	0.77	0.41	0.42
MUHY	-	-	-	-	-	-	-	-	-	-	-	1	0.57	0.46
MEMG	-	-	-	-	-	-	-	-	-	-	-	-	1	0.07
MNA+ME+MJ	-	-	-	-	-	-	-	-	-	-	-	-	-	1

This table presents the Correlation Matrix of the indices we use in our research. In the last column we count up MNA, ME and MJ to investigate if they are perfectly correlated with MW. The correlation between MNA+ME+MJ and MW is large but the series are not perfectly correlated with each other.

3.3 Rolling Variance and Correlation

As already mentioned, in many realisting applications we have to deal with multiple assets. Then, in addition to their volatilities, we also have to consider their correlations. It is not reasonable to assume that correlations and variances are constant over time, for this reason we illustrate the historical variances and correlations based on a 52-week moving window, which boils down to a computation of the variance and correlation on yearly basis.

Figure A.5.1 shows a graphical representation of the time-varying variance of the equity and fixed income indices. We can immediately see that the variances considerably increase during the global financial crisis between 2008 and 2010. As there is a large deviation in the variance, we can conclude that heteroskedasticity is present both for the equity and the fixed income return series. Figures A.5.3 and A.5.4 show the time-varying historical correlations based on a 52-week moving window. For both equity and fixed income we see a lot of fluctuations in the time-varying correlation. Hence, it seems plausible that a model that incorporates this time-variability of the correlation will perform better in comparison with a model that does not take it into account. It is obvious that correlations increase during the global financial crisis in 2008. This evidences the finding that negative returns are more dependent than positive returns and that correlations are higher during a bear market than during a bull market. Concluding, this graphical analysis shows that correlations are not stable over time and, more interestingly that correlations show similar patterns between return series.

To investigate the effect of the moving window length, we will perform the same analysis but now on quarterly basis which is a 12-week moving window. The reason that we do not perform this analysis on a more frequent basis is because the figures will be too unclear due to the large amount of fluctuations. Figures A.5.5 and A.5.6 show the time-varying correlations corresponding to a moving window of 12 weeks. We can immediately see that in these figures there are even more fluctuations and the correlations are more time-varying. However, this is a logical consequence of using a smaller window. When using a smaller window we observe that the variances during the financial crisis are larger than the variances during the financial crisis that are obtained with a larger moving window. Figure A.5.2 shows the time-varying variance based on a 12-week moving window. Here again we can see that there are more fluctuations in comparison with the 52-week moving window variance.

3.4 Asymmetry in the Correlations

Besides the existence of time-varying correlations, we also investigate if there is asymmetry in the correlation. In other words, does correlation increase (decrease) when we have negative (positive) returns? In order to show this we present the threshold correlations computed on returns. Examining asymmetric correlations is very important: the first reason is that hedging relies on the correlations between the assets hedged and the financial instruments used. The second reason is that the value of portfolio diversification might be questionable if all assets tend to fall as the market falls. The threshold correlations computed on returns which are standardized by their unconditional means and variances are calculated as follows:

$$\tilde{\rho}_\gamma(z_i, z_j) = \begin{cases} Corr(z_i, z_j | z_i \leq \gamma, z_j \leq \gamma) & \text{if} \quad \gamma \leq 0 \\ Corr(z_i, z_j | z_i > \gamma, z_j > \gamma) & \text{if} \quad \gamma > 0 \end{cases} \tag{3.2}$$

where z_i are error terms for index i and γ denotes a vector with grid points from -1 till 1. The returns are standardized to have zero mean and unit variance so that the mean and variance do not appear explicitly in the right-hand side of the definition, making both the computation and statistical analysis easy.

For the sake of convenience we only present the threshold correlation between the following indices: MW - MSC, MW - MDG, MW - IBF, MW - IBNF, IBF - IBNF and MEM-MNA. We choose for these pairs because they are economically interesting as some of them are each others opposite as regards the asset class, such as MEM-MNA (developed market vs. emerging market), IBF-IBNF (financial vs. non-financial), MW-MDG (equity vs. fixed income). Furthermore it is useful to investigate the correlations within an asset-class such as MW-MSC (equity vs. equity) and IBF-IBNF (fixed income vs. fixed income). First we present the threshold correlations from May 1999 till May 2012 in Figure 3.1. Then in Figure 3.2

we present the threshold correlations during the crisis period 2008 till 2011 to explore if the lower tail dependence during the crisis period is larger than the lower tail dependence for the whole sample.

At one glance we can see that there are indeed asymmetric correlations between the different benchmarks. Especially for the following pairs: MW-MSC, MW-IBNF and MW-IBF. Another remarkable result is that there is not much difference between the lower tail dependence and upper tail dependence for the correlation between IBF and IBNF. A possible explanation for this result is that these two indices are not very different from each other. In Table 3.3 we can observe that the correlation between these two indices is 0.78, which is quite large as a priori expected.

When looking at the differences between the upper and the lower tail dependences in Figure 3.1 and in Figure 3.2 we can conclude that there are not extremely large differences between the lower/upper tail dependences. However, to be able to make a final conclusion about the latter we have to perfom statistical tests to evaluate the difference between upper and lower tail dependences, this is done in the end of this section. An interesting fact that emerges from Figure 3.2 is that the upper tail dependence between MW and IBNF suddenly increases from 0.9 onwards. Obviously, these two indices move together for extremely large positive returns. This is interesting because usually very negative returns move more closely.

As mentioned before, we have to test formally if there is a difference between the lower and the upper tail dependence. We will use the test developed by Hong, Tu and Zhou (2007), which has the advantage of being a model-free test. The null hypothesis of symmetric correlation is

$$H_0 : \rho^+(\gamma) = \rho^-(\gamma) \quad \text{for all} \quad \gamma \geq 0 \tag{3.3}$$

where $\rho^+(\gamma)$ is the upper tail dependence and $\rho^-(\gamma)$ is the lower tail dependence. The formula above tests if the correlation between the positive returns of the two assets is the same as that between their negative returns. The alternative hypothesis is

$$H_A : \rho^+(\gamma) \neq \rho^-(\gamma) \quad \text{for some} \quad \gamma \geq 0 \tag{3.4}$$

If the null hypothesis is true, the following $m \times 1$ vector

$$\hat{\rho}^+ - \hat{\rho}^- = [\hat{\rho}^+(\gamma_1) - \hat{\rho}^-(\gamma_1), ..., \hat{\rho}^+(\gamma_m) - \hat{\rho}^-(\gamma_m)]' \tag{3.5}$$

must be close to zero. The statistic for testing the null hypothesis can be presented as follows:

$$J_p = T(\hat{\rho}^+ - \hat{\rho}^-)'\hat{\Omega}^{-1}(\hat{\rho}^+ - \hat{\rho}^-) \tag{3.6}$$

where T is the number of observations and $\hat{\Omega}$ is a positive definite variance-covariance matrix for all possible true distributions of the data satisfying some regularity conditions. The derivation of $\hat{\Omega}$ can be found in Appendix A.4.

Table 3.4 presents the results of this test. The pairs which are investigated are presented in the rows. The test is performed for a threshold value 0 and for a set of threshold values varying from 0 up to 1.5. The table presents both the p-values of the test for every pair and for every set of threshold values. Furthermore, the difference between the threshold correlations $\hat{\rho}^{+}(\gamma)$-$\hat{\rho}^{-}(\gamma)$ is presented for every threshold value and for every pair.

Full sample

We observe that in none of the cases we can reject the null hypothesis, hence according to this test there are no significant differences between the lower and the upper tail dependences. However the differences of the threshold correlations are negative in almost all cases. This means that the lower tail dependence is larger than the upper tail dependences. Hence, we can conclude that although the figures present an asymmetric difference between the threshold correlations, the differences are not significant for these pairs.

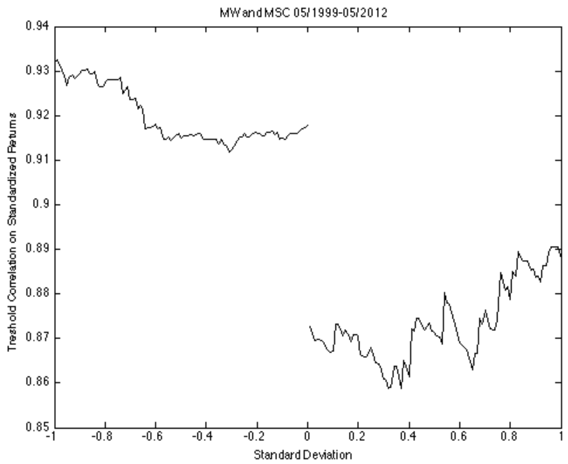
Crisis period

The results are more or less similar as the previous results as opposed to our expectations. We observe that we are still not able to reject our null hypothesis of equality. However, some values for the differences between the upper and the lower tail dependence become even larger. The reason is the fact that we now use the crisis period, in which correlations between asset returns usually increase. Some values are missing as the sample does not contain asset returns that are larger than 1.0 or 1.5.

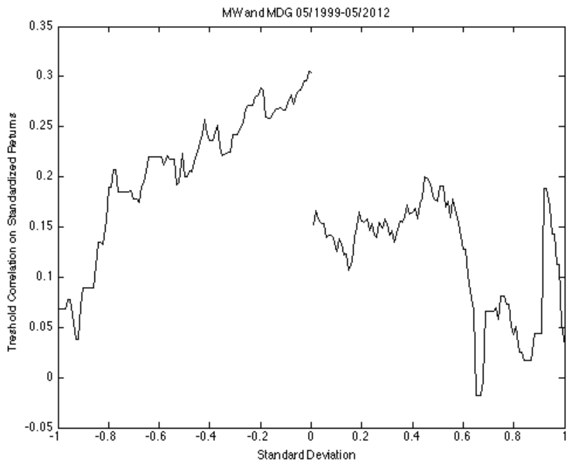
Note that the asymmetry in the correlations of few pairs are investigated, there could be pairs for which the difference between the upper and the lower tail dependence is significant. For this reason, we still include asymmetric terms in our models.

Figure 3.1: Empirical Threshold Correlations. The figures report pairwise threshold correlations computed on returns for different benchmarks that SAMCo uses for the period May 14,1999 to May 18, 2012. The left line denotes the lower tail dependence. The right line denotes the upper tail dependence. The Treshold Correlations are computed as in equation 3.2

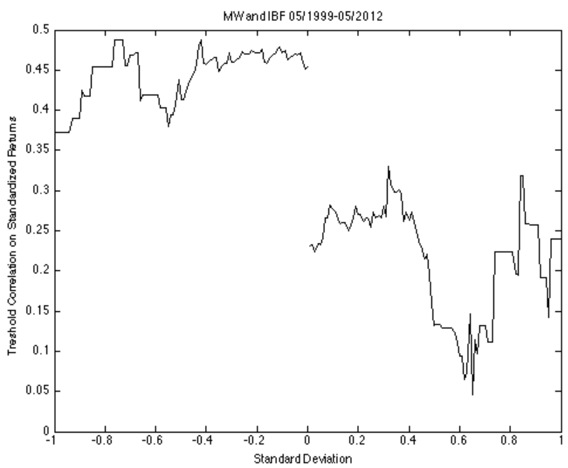
(a) MW-MSC



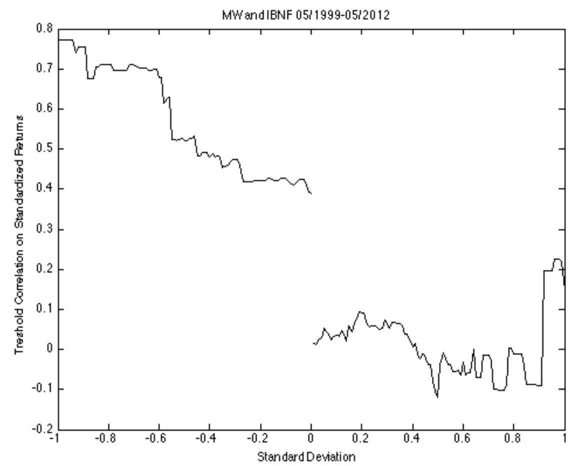
(b) MW-MDG



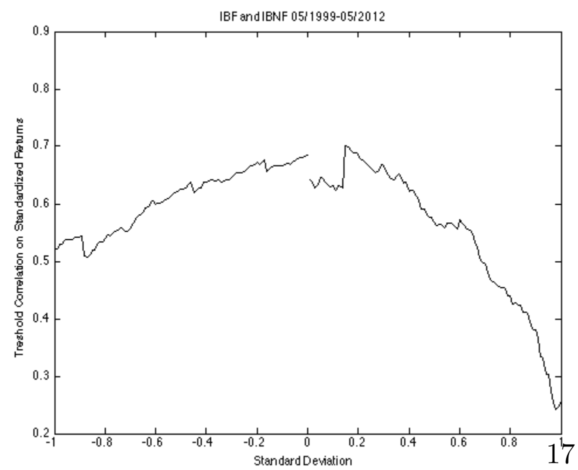
(c) MW-IBF



(d) MW-IBNF



(e) IBF-IBNF



(f) MEM-MNA

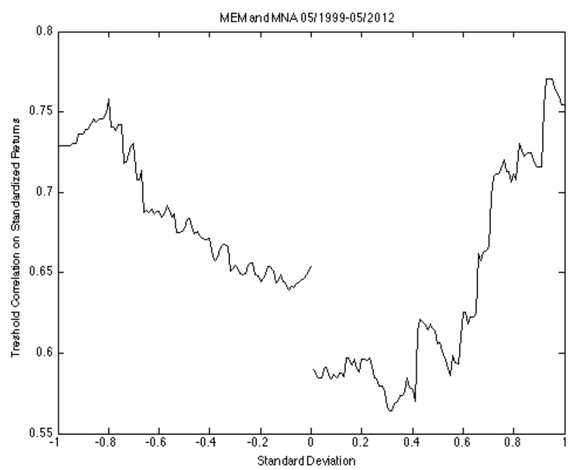
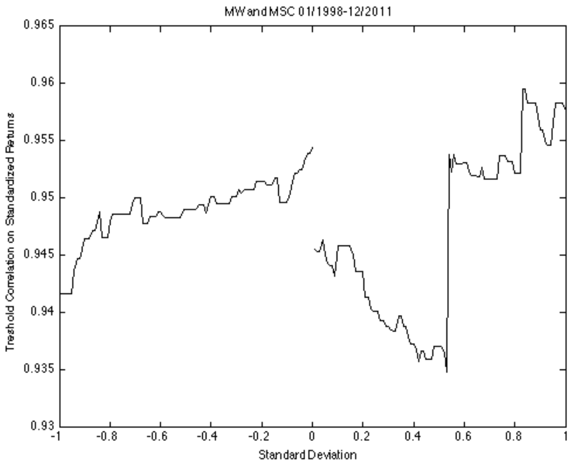
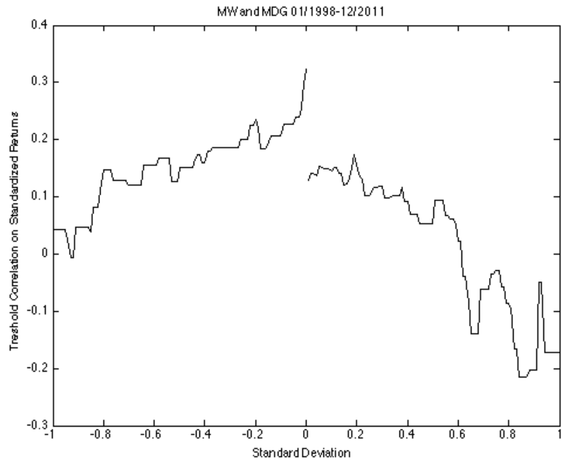


Figure 3.2: Empirical Threshold Correlations. The figures report pairwise threshold correlations computed on returns for different benchmarks that SAMCo uses for the period January 2008 to December 2011. The left line denotes the lower tail dependence. The right line denotes the upper tail dependence. The Treshold Correlations are computed as in equation 3.2.

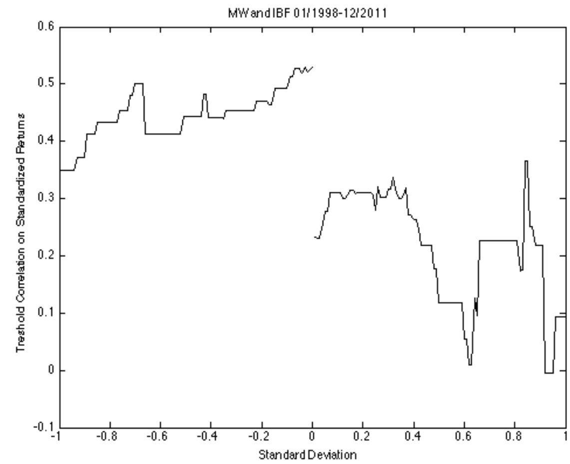
(a) MW-MSC



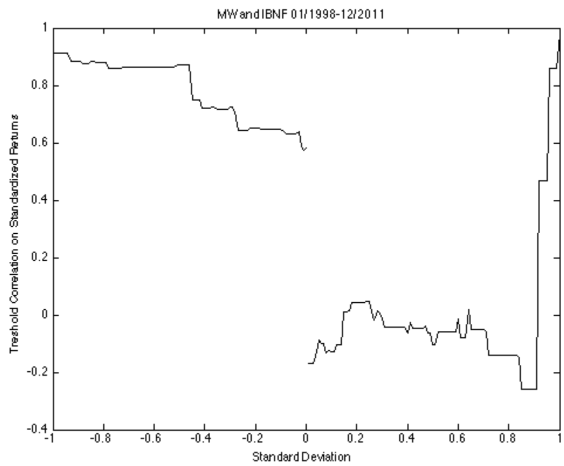
(b) MW-MDG



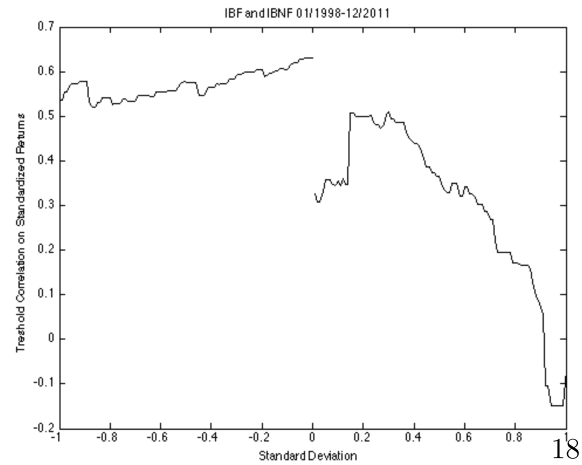
(c) MW-IBF



(d) MW-IBNF



(e) IBF-IBNF



(f) MEM-MNA

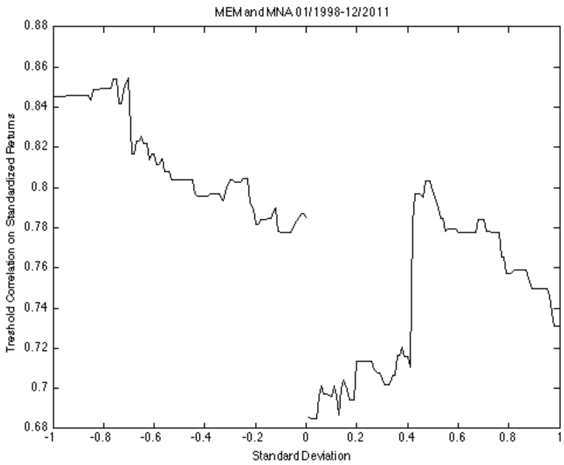


Table 3.4: **Exceedance Test**

	γ=0 P-value	$\hat{\rho}^{+}(\gamma)$-$\hat{\rho}^{-}(\gamma)$ $\gamma = 0$	γ=[0,0.5,1.0,1.5] P-value	$\hat{\rho}^{+}(\gamma)$-$\hat{\rho}^{-}(\gamma)$ $\gamma = 0$	$\gamma = 0.5$	$\gamma = 1.0$	$\gamma = 1.5$
Full Sample							
MW-MSC	0.89	-0.045	1.0	-0.045	-0.044	-0.044	-0.098
MW-MDG	0.27	-0.162	0.47	-0.162	-0.024	-0.036	0.320
MW-IBF	0.31	-0.222	0.59	-0.222	-0.280	-0.133	-0.371
MW-IBNF	0.21	-0.379	0.59	-0.379	-0.640	-0.627	
IBF-IBNF	0.82	-0.040	0.60	-0.040	-0.055	-0.265	-0.486
MEM-MNA	0.80	-0.067	0.95	-0.067	-0.072	0.025	-0.048
Crisis Period							
MW-MSC	0.98	-0.009	1.0	-0.009	0.004	0.014	0.027
MW-MDG	0.37	-0.158	0.66	-0.158	-0.176	-0.140	
MW-IBF	0.17	-0.350	0.41	-0.350	-0.226	-0.403	
MW-IBNF	0.24	-0.570	0.43	-0.567	-0.878		
IBF-IBNF	0.38	-0.258	0.14	-0.258	-0.236	-0.567	
MEM-MNA	0.83	-0.086	0.84	-0.086	-0.082	-0.187	-0.354

This table presents the P-values of the exceedance test developed by Hong, Tu and Zhou (2007) and the differences between the upper and the lower tail dependence for the full sample and for the crisis period. The test is performed for a threshold value 0 and for a set of threshold values varying from 0 up to 1.5. For the sake of convenience we only present the threshold correlation between some pairs. We choose for these pairs because they are economically interesting as some of them are each others opposite as regards the asset class, such as MEM-MNA (developed market vs. emerging market) or they are within the same asset-class such as MW-MSC (equity vs. equity). The crisis period starts in January 2008 and ends in December 2011.

Chapter 4

Methodology

In this chapter we elaborate on the methodology used in this research. We provide an explanation of the Riskmetrics and the DCC-model. For each model the definition, estimation procedure and the forecasting procedure is presented. Subsequently, we elaborate on the evaluation techniques used to assess our results. First, we explain the backtesting test developed by Christoffersen (1998) in section 4.3.1, then we explain the comparative predictive ability (CPA) test developed by Giacomini and White (2006) in section 4.3.2. In the last section of this chapter we will elaborate on the methodology of the economic value of volatility timing. In other words, this chapter explains all the methods and techniques that are used in the rest of this research. It can be quite technical in comparison with the rest of the chapters, though it provides a necessary theoretical foundation to understand the results of the research.

4.1 RiskMetrics Approach

4.1.1 Definition

RiskMetrics is a widely used model in forecasting volatilities and in calculating the Value-at-Risk. It was first introduced by J.P. Morgan in 1995. This model is currently used by SAMCo to obtain Value-at-Risk. For this reason we will use the RiskMetrics approach as a natural benchmark. The time varying covariance matrix is used as the dependence measure between the different assets. We can define the RiskMetrics covariance marix between two assets at time t, as follows:

$$\Sigma_t = \lambda\Sigma_{t-1} + (1-\lambda)(r_{t-1}r_{t-1}^{'}) \tag{4.1}$$

where Σ_t is the covariance matrix with dimension $N \times N$ (N is the number of assets in the portfolio) at time t, λ is a predefined decay factor that lies between 0 and 1. SAMCo uses 0.99 for the decay factor. Furthermore, r_t is a vector of the returns at time t. The index i defines asset i at time t.

4.1.2 Estimation

The only parameter that the RiskMetrics model uses, is the decay factor λ. In the standard RiskMetrics approach the decay factor λ is fixed at 0.94. However, SAMCo uses 0.99 for the decay factor, hence we will not put effort in estimating this parameter and we will keep it fixed through the research. The value of the decay factor is usually related to the frequence of the returns. SAMCo uses returns at the weekly frequence, hence for this reason they use 0.99.

4.1.3 Forecasting Procedure

This section will explain the algorithm that forecasts the covariance matrix with the RiskMetrics model. Before we start with the forecasting procedure we have to define the size of the window, given by the following parameter: $windowSize$, the decay factor λ, and the total sample length T. The length of the window should be large enough to avoid strange features and parameter instability. This is especially relevant for DCC-models. For these reasons we choose a window length that contains 200 observations.

Start at time $t = windowSize$

1. We calculate the forecasted RiskMetrics covariance matrix $\hat{H}_t$ over the sample $\tau - windowSize + 1$ until $\tau = t$.

2. The $t + K$ forecast of the covariance is given by the following equation:

$$h_{1,2,t+K|t} = Kh_{1,2,t+1|t} \tag{4.2}$$

here, 1 and 2 denote the row-index and the column-index, respectively of the covariance matrix $\hat{H}_{t+K|t}$.

3. Move to $t = t + 1$ and repeat this procedure until time $t = T$.

4.2 Dynamic Conditional Correlation Model

4.2.1 Definition

As already mentioned in the Literature Review, Engle (2002) introduced a new class of multivariate GARCH estimators that can be viewed as a generalization of the Bollerslev (1990) Constant Conditional Correlation (CCC) estimator. In Bollerslev's model the covariance matrix H_t at time t is denoted as follows:

$$H_t = D_t R D_t \tag{4.3}$$

where here again just as the RiskMetrics model $D_t = diag(\sqrt{h_{i,t}})$, where $h_{i,t}$ is the volatility estimated with an univariate volatility model, $diag(.)$ is the operator that transforms a $N \times 1$ vector into a $N \times N$ diagonal matrix. The volatilities can be estimated with any type of a volatility model, which makes the DCC-model very flexible. R is a symmetric positive definite conditional correlation matrix with elements $\rho_{ij,t}$, where $\rho_{ii,t} = 1$, we can directly see that R contains the conditional correlations by rewriting equation 4.3 as:

$$\mathbb{E}_{t-1}(\epsilon_t \epsilon_t^{'}) = D_t^{-1} H_t D_t^{-1} = R \tag{4.4}$$

since $\epsilon_t = D_t^{-1} r_t$, which are standardized disturbances that have mean zero and variance one for each series.
The dynamic conditional correlation model differs only in allowing R to be time varying:

$$H_t = D_t R_t D_t \tag{4.5}$$

We consider two different specifications for R_t: the Dynamic Conditional Correlation (DCC) model of Engle (2002) and the Asymmetric DCC (ADCC) model of Cappiello et al. (2006). In the DCC model, R_t remains the correlation matrix, except for the fact that it is now a time-varying $N \times N$ correlation matrix with diagonal elements equal to 1. R_t is now defined as follows:

$$R_t = diag(Q_t^{-1/2}) Q_t diag(Q_t^{-1/2}) \tag{4.6}$$

where $diag(Q_t)$ is a diagonal matrix containing the diagonal elements of the $N \times N$ positive definite matrix Q_t given by:

$$Q_t = (1 - \alpha - \beta)\overline{Q} + \alpha(\epsilon_{t-1}\epsilon_{t-1}^{'}) + \beta Q_{t-1} \tag{4.7}$$

where $\overline{Q}$ is the $N \times N$ unconditional covariance matrix of ϵ_t and α and β are non-negative scalar parameters. The model is covariance-stationary if $\alpha + \beta < 1$. Moreover, Q_t is guaranteed to be positive definite if $(1 - \alpha - \beta)\overline{Q}$ and Q_0 are themselves positive definite.

Finally, the ADCC model incorporates the leverage effect into the conditional correlations. The leverage effect refers to the phenomenon that volatility increases when stock price falls. The ADCC model is given by:

$$Q_t = (1 - \alpha - \beta)\overline{Q} - \delta\overline{\Gamma} + \alpha(\epsilon_{t-1}\epsilon_{t-1}^{'}) + \beta Q_{t-1} + \delta n_{t-1} n_{t-1}^{'} \tag{4.8}$$

where $n_t = I(\epsilon_t < 0) \circ \epsilon_t$ ($\circ$ is the Hadamard product: elementwise matrix multiplication) and $\overline{\Gamma} = E[n_t n_t']$. A necessary condition for Q_t to be covariance-stationary and positive definite is that $(1-\alpha-\beta)\overline{Q} - \delta\overline{\Gamma}$ and Q_0 are positive definite and $\alpha + \beta + \lambda\delta < 1$, where λ is the maximum eigenvalue of $\overline{Q}^{-1/2}\overline{\Gamma Q}^{-1/2}$.

The multivariate models are implemented using alternative specifications for the univariate conditional variances. To model the conditional variances, we use the GARCH-model and the TGARCH-model. Furthermore we assume two different types of error distributions, on the one hand the Normal distribution and on the other hand the Student-t distribution. Hence, we have 8 different kind of specifications, to facilitate the presentation of the results, the models are denoted by the following abbreviations: DCC-GARCH, DCC-GARCH-t (Student-t distribution for the error terms), DCC-TGARCH (Threshold-GARCH), DCC-TGARCH-t, ADCC-GARCH, ADCC-GARCH-t, ADDC-TGARCH and finally ADCC-TGARCH-t.

4.2.2 Estimation

In the first step of the DCC-model the univariate conditional volatilities for each return series are constructed. Consequently, in the second step these conditional volatilities are used to construct the covariance matrix. As already mentioned in the previous section, we use two specifications for the univariate conditional volatility model: the GARCH and the Threshold-GARCH (TGARCH) model. In this section, we will clarify the estimation procedure during the first and the second step of the DCC-model.

First we will start with clarifying the first step in the estimation procedure. The intuitive idea of the GARCH model is that the volatility changes only gradually over time such that h_t will be "close(ly related)" to h_{t-1}. Also, the squared (unexpected) return $(r_{t-1} - \mu)^2$ is an 'ex post' measure of volatility during period $t-1$ (given that $h_{t-1} = \mathbb{E}[\epsilon_{t-1}^2 | I_{t-2}]$). So, this may give useful information on how volatility is changing. The GARCH(1,1) model (Bollerslev, 1986) is given by:

$$h_t = \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} \tag{4.9}$$

where, ϵ_{t-1}^2 is a series of squared residuals of the returns, h_{t-1} is the one period lagged conditional variance. To guarantee that $h_t \geq 0$ for all t, ω should be larger than 0 and α and β should be larger than or equal to zero. The model is covariance stationary if $\alpha + \beta < 1$. Then $\mathbb{E}[\epsilon_t^2] = \sigma^2 = \frac{\omega}{1-\alpha-\beta}$. [1]

The GARCH model is symmetric in the sense that positive and negative values of ϵ_{t-1} have the same effect on h_t. In practice, period of high volatility often start with large negative

[1]Note that setting $\omega = 0$ and $1 - \alpha = \beta = \lambda$ gives RiskMetrics.

returns. The Threshold GARCH model allows for such asymmetry:

$$h_t = \omega + \alpha\epsilon_{t-1}^2 I[\epsilon_{t-1} \leq] + \gamma\epsilon_{t-1}^2 I[\epsilon_{t-1} > 0] + \beta h_{t-1} \tag{4.10}$$

where $I[A] = 1$ if A occurs and 0 otherwise. $\omega > 0$, $\alpha > 0$, $\gamma > 0$ and $\beta \geq 0$ are required for $h_t \geq 0$ for all t. $(\alpha + \gamma)/2 + \beta < 1$ is required for covariance stationarity.

When the parameters $\theta = (\omega, \alpha, \beta)'$ or $\theta = (\omega, \alpha, \beta, \gamma)'$ of the GARCH model 4.9 or 4.10 are estimated by Maximum Likelihood estimation and the standardized residuals $z_t = D_t^{-1}\epsilon_t$ are constructed, we use these to estimate $D_t = diag(\sqrt{h_{i,t}})$. The parameters $\phi = (\alpha, \beta, (\delta))$ of the DCC models in equations 4.3 and 4.8 are estimated by Maximum Likelihood estimation. $\overline{Q}$ is replaced by the unconditional correlation matrix of $\hat{z}_t$: $\overline{Q} = \frac{1}{T}\sum_{t=1}^{T} \hat{z}_t \hat{z}_t'$. This is called correlation targeting, in this manner we have $N(N-2)$ less parameters to estimate.

The quasi-loglikelihood function that is maximized by numerical optimization ($fmincon$ in Matlab) [2] to find the parameters in vector ϕ is defined as follows:

$$log(L(\phi; \hat{\theta}, r_t)) = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log|D_t| + r_t' D_t^{-1} D_t^{-1} r_t - \epsilon_t'\epsilon_t + log|R_t| + \epsilon_t' R_t^{-1}\epsilon_t) \tag{4.11}$$

The above formula denotes a Gaussian loglikelihood function. For models where we use a Student-t distribution for the error terms, we use a different loglikelihood function. Now ϕ consists of $\phi = (\alpha, \beta, (\delta), v)$, as the degrees of freedom v also need to be estimated. The loglikelihood function to estimate the parameters of models with a t-distribution is presented as follows:

$$log(L(\phi; \hat{\theta}, r_t)) = \sum_{t=1}^{T}(log[\Gamma(\frac{v+n}{2})] - log[\Gamma(\frac{v}{2})] - \frac{n}{2}log[\pi(v-2)] - \frac{1}{2}log[|R_t|] - \frac{v+n}{v}log[1 + \frac{\epsilon_t' R_t^{-1}\epsilon_t}{v-2}]) \tag{4.12}$$

The estimated parameters are used to calculate the correlation matrix Q_t. Then this is used to obtain R_t. Once we have R_t and D_t (which is obtained by an univariate volatility model), we can calculate H_t.

4.2.3 Forecasting Procedure

First, we should define the length of the moving window before we start with making K-step ahead forecasts. Let's call the length of the moving window $windowSize$ and the total sample length T.

[2] The symmetric DCC models are estimated by quasi maximum likelihood (QML) using the UCSD GARCH Matlab toolbox developed by Kevin Sheppard: http://www.kevinsheppard.com/wiki/UCSDGARCH.

Start at time $t = windowSize + 1$

1. Estimate the parameters $\theta = (\omega, \alpha, \beta, (\gamma))$ of the univariate volatility model over the sample $t - windowSize$ until t for every asset.

2. When we have these parameters we can forecast every assets volatility $\sqrt{h_{i,t+K}}$ and we can construct $D_{i,t+K}$

3. Now we can calculate the conditional correlation matrix $Q_{t+K|t}$ as defined in equation 4.7 or in the case of the Asymmetric DCC-model as in equation 4.8.

4. Construct $R_{t+K|t}$ as in equation 4.6.

5. Finally, construct the K-step ($t + K$ forecast) ahead covariance matrix as follows: $\hat{H}_{t+K} = D_{t+K} R_{t+K} D_{t+K|t}$.

6. Move to $t + 1$ and repeat this procedure until T.

4.3 Forecast Evaluation

In order to evaluate the performance of our models we use the following two main criteria, that are explained in this section. The first one is the backtesting test developed by Christoffersen (1998) and the second evaluation criterion is the CPA-test developed by Giacomini & White (2006).

4.3.1 Value-at-Risk

Financial risk model evaluation or backtesting is a key part of the internal models approach to market risk management as laid out by the Basel Commitee on Banking Supervision (1996). As risk exposures are typically quantified in terms of a Value-at-Risk (VaR) estimate, we will forecast the Value-at-Risk for different models and assess and quantify the accuracy of these estimates. The results obtained from this evaluation criteria have implications for SAMCo that wishes to assess the accuracy of their internal risk measurement model. Before we start with clarifying the testing procedure, we will explain the concept of Value-at-Risk.

VaR is the maximum loss that could occur over a given holding period or the minimum return that could occur over a given holding period with a specified confidence level q. For any $0 < q < 1$, the VaR at $100 \times (1 - q)\%$ is the return that is expected to be exceeded with probability $1 - q$. Furthermore, $VaR_{t+1}(1 - q)$ is the q-th quantile of the conditional distribution of the return r_{t+1}:

$$P[r_{t+1} \leq VaR_{t+1}(1-q)|I_t] = F_{t+1|t}(VaR_{t+1}(1-q)) = q \tag{4.13}$$

where $F_{t+1|t}(.)$ is the cumulative distribution function of r_{t+1} conditional on I_t.

If we have a GARCH(1,1) model for instance as defined in 4.9 then the $VaR_{t+1}(1-q)$ will be equal to $VaR_{t+1}(1-q) = \mu + z_q\sqrt{h_{t+1}}$, where z_q is the q-th quantile of the standard normal distribution when using a model with Gaussian distributions. If we assume an univariate volatility model with Student-t distributed error terms, z_q denotes the q-th quantile of the Student-t distribution. h_{t+1} is the volatility that is forecasted with the help of the univariate GARCH(1,1) model. For our multivariate RiskMetrics and GARCH-models we have a different specification for Value-at-Risk:

$$VaR_{t+k,(1-q)} = w'\mu_k + z_q\sqrt{w'H_{k,t}w} \tag{4.14}$$

where w denotes the portfolio weights, we assume an equally weighted portfolio with weights $w = 1/N$. Furthermore, $w'\mu_k$ is the portfolio mean, z_q denotes the q-th quantile of the standard normal distribution when using a model with Gaussian distributions. If we assume an univariate volatility model with Student-t distributed error terms, z_q denotes the q-th quantile of the Student-t distribution. Finally, $w'H_{k,t}w$ is the portfolio variance and $H_{k,t}$ is the covariance matrix with dimension $N \times N$ of the k-day asset returns.

In general, $(100 \times q)\%$ interval forecast for r_{t+1} are of the form $(L_{t+1|t}(q), U_{t+1|t}(q))$, where $L_{t+1|t}(q)$ is the lower bound and $U_{t+1|t}(q)$ is the upper bound. $(L_{t+1|t}(q), U_{t+1|t}(q))$ is constructed in such a way that:

$$P[L_{t+1|t}(q) \leq r_{t+1} \leq U_{t+1|t}q|I_t] = q \tag{4.15}$$

VaR is obtained by setting $L_{t+1|t}(q) = -\infty$ and $U_{t+1|t}(q) = VaR_{t+1}(1-q)$. We will make use of techniques developed for evaluating interval forecasts to evaluate VaR estimates which is explained in the remainder of this section.

Given the VaR estimates $(VaR_{t+1}(1-q))_{t=0}^{T-1}$, how should we determine whether these VaR estimates are "good" or "accurate"? A "good" interval forecast suffices the following points:

1. Fraction of observations inside the interval should be equal to the nominal coverage probability. This is called correct unconditional coverage, which means that the fraction of VaR violations should be equal to the nominal coverage probability. This test is introduced by Christoffersen (1998). We will elaborate on this point as follows:

Define the indicator function I_{t+1} as:

$$I_{t+1} = \begin{cases} 1 & : r_{t+1} \in (L_{t+1|t}(q), U_{t+1|t}(q)) \\ 0 & : r_{t+1} \notin (L_{t+1|t}(q), U_{t+1|t}(q)) \end{cases} \tag{4.16}$$

Recall that the VaR context is retrieved by setting $L_{t+1|t}(q) = -\infty$ and $U_{t+1|t}(q) = VaR_{t+1}(1-q)$ such that the indicator function I_{t+1} would be defined as:

$$I_{t+1} = \begin{cases} 1 & : r_{t+1} < VaR_{t+1}(1-q) \\ 0 & : r_{t+1} > VaR_{t+1}(1-q) \end{cases} \tag{4.17}$$

Hence, I_{t+1} indicates "violations" of the VaR. Testing correct unconditional coverage boils down to testing the following null hypothesis:

$$H_0 : E[I_{t+1}] = q \tag{4.18}$$

(the number of expected VaR violations is equal to the nominal coverage probability), given that $I_{t+1}, I_t, ...$ are independent. This null hypothesis may be tested using a Likelihood Ratio test. Given independence, the likelihood function for interval forecasts with coverage probability $\pi = P[I_{t+1} = 1]$ is given by:

$$\mathcal{L}(\pi; I_T, I_{T-1}, ..., I_1) = P[I_T = i_T, I_{T-1} = i_{T-1}, ..., I_1 = i_1] \tag{4.19}$$

$$= P[I_T = i_T]P[I_{T-1} = i_{T-1}]...P[I_1 = i_1] \tag{4.20}$$

$$= (1-\pi)^{T_0}\pi^{T_1} \tag{4.21}$$

where, $T_1 = \sum_{t=1}^{T} i_t$, $T_0 = T - T_1$, T_0 is the number of returns that lie within the interval of VaR and T_1 is the number of returns that exceed the VaR.
The Likelihood Ratio test compares the likelihood under the null with the likelihood under the alternative. Under the null hypothesis of correct unconditional coverage we have the following likelihood function:

$$\mathcal{L}(\pi_{exp}; I_T, I_{T-1}, ..., I_1) = (1-\pi_{exp})^{T_0}\pi_{exp}^{T_1} \tag{4.22}$$

where π_{exp} is the expected proportion of returns that lie within the prescribed interval of the distribution, which is equal to the normal coverage probability q. Under the alternative hypothesis, we have the following likelihood function:

$$\mathcal{L}(\pi_{obs}; I_T, I_{T-1}, ..., I_1) = (1-\pi_{obs})^{T_0}\pi_{obs}^{T_1} \tag{4.23}$$

where π_{obs} is the observed proportion of returns that are within the prescribed interval of the distribution. The maximum likelihood estimate of π_{obs} is equal to:

$$\hat{\pi}_{obs} = \hat{P}[I_{t+1} = 1] = \frac{T_1}{T_0 + T_1} \tag{4.24}$$

Finally, the likelihood ratio test of correct unconditional coverage is then computed as:

$$LR_{uc} = -2\log(\frac{\mathcal{L}(\pi_{exp})}{\mathcal{L}(\pi_{obs})}) \sim \chi^2(1) \tag{4.25}$$

where $\chi^2(1)$ is the χ^2-distribution with 1 degree of freedom.

Furthermore, a "good "VaR-forecast should also satisfy the following point:

2. Occurrences of observations outside the interval should be spread out over the sample and not come in clusters, this is called independence and it is especially relevant in the presence of time-dependent heteroskedasticity, such as "volatility clustering". This will be tested with the Likelihood Ratio Test of independence. Independence will be tested against an explicit first-order Markov alternative. Consider a binary first-order Markov Chain, (I_t), with transition probability matrix

$$\Pi_1 = \begin{bmatrix} 1-\pi_{01} & \pi_{01} \\ 1-\pi_{11} & \pi_{11} \end{bmatrix} \tag{4.26}$$

where $\pi_{ij} = P(I_{t+1} = j | I_t = i)$, the transition probability from state i to state j. Both $i, j \in (0, 1)$. Here 0 denotes a non-exceedance, whereas 1 denotes an exceedance. The approximate likelihood function for this process is

$$\mathcal{L}(\Pi_1; I_1, I_2, ..., I_T) = (1-\pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1-\pi_{11})^{n_{10}} \pi_{11}^{n_{11}} \tag{4.27}$$

where n_{ij} is the number of observations with value i followed by j. As is standard, we condition on the first observation everywhere. It is then easy to maximize the log-likelihood function and solve for the parameters, which are simply ratios of the counts of the appropriate cells:

$$\hat{\Pi}_1 = \begin{bmatrix} \frac{n_{00}}{n_{00}+n_{01}} & \frac{n_{01}}{n_{00}+n_{01}} \\ \frac{n_{10}}{n_{10}+n_{11}} & \frac{n_{11}}{n_{10}+n_{11}} \end{bmatrix} \tag{4.28}$$

Consider now the output sequence, I_t, from an interval model. We estimate a first-order Markov chain model on the sequence, and test the hypothesis that the sequence is independent by noting that

$$\Pi_2 = \begin{bmatrix} 1-\pi_2 & \pi_2 \\ 1-\pi_2 & \pi_2 \end{bmatrix} \tag{4.29}$$

corresponds to independence. The likelihood under the null hypothesis then becomes

$$\mathcal{L}(\Pi_2; I_1, I_2, ..., I_T) = (1-\pi_2)^{(n_{00}+n_{10})}\pi_2^{(n_{01}+n_{11})} \tag{4.30}$$

the Maximum Likelihood estimate is $\hat{\Pi}_2 = \hat{\pi}_2 = \frac{n_{01}+n_{11}}{n_{00}+n_{10}+n_{01}+n_{11}}$. The Likelihood Ratio test of independence is asymptotically distributed as a χ^2 distribution with $(s-1)^2$ degrees of freedom, as we are working with a binary sequence (so $s = 2$), that is,

$$LR_{ind} = -2log\frac{\mathcal{L}(\hat{\Pi}_2)}{\mathcal{L}(\hat{\Pi}_1)} \sim \chi^2(1) \tag{4.31}$$

Note that this test does not depend on the true coverage p, hence it only tests the independence part of the hypothesis.

Taken together, 1. and 2. can be called correct conditional coverage. A "good" interval forecast should perform well in both the independence test and the unconditional coverage test, which can be tested simultaneously with the correct conditional coverage test. Hence, we can combine the above tests for unconditional coverage and independence to form a complete test of conditional coverage. In this case, the null hypothesis of the unconditional coverage test will be tested against the alternative of the independence test. We have now the following Likelihood Ratio test:

$$LR_{cc} = -2log\frac{\mathcal{L}(\pi_{exp})}{\mathcal{L}(\hat{\Pi}_1)} \sim \chi^2(s(s-1)) = \chi^2(2) \tag{4.32}$$

As we are forecasting 4-step ahead VaR estimates, we modify the above procedure a little bit to account for the fact that optimal forecasts at horizon K are characterized by autocorrelations of order $K-1$. We have to do this because the indicator variables used to construct the χ^2 statistics will also exhibit autocorrelation of order $K-1$ when the forecasts are optimal. We use the procedure based on Bonferroni bounds suggested by Diebold et al. (1998) to overcome this problem. This procedure divides the indicator variable series into K sub-groups that are dependent under the null hypothesis. The sub-groups will look as follows: $(I_{t+K|t}, I_{t+2K|t+K}, \dots)$, $(I_{t+1+K|t+1}, I_{t+1+2K|t+1+K}, \dots)$, $\dots$, $(I_{t+(K-1)+K|t+(K-1)}, I_{t+(K-1)+2K|t+(K-1)+K}, \dots)$. As K equals four, we have four sub-groups. We then apply the conditional coverage, independence and the conditional coverage test to each of the K subgroups and reject the relevant null hypothesis for a given sub-group at the significance level of $\frac{\alpha}{K}$. Furthermore, we choose to stick to the 'normal' backtesting approach in the sense that we base inferences on the asymptotic distributions of the 'normal' backtesting test statistics. Some researchers prefer to follow Wallis (2003) and calculate exact p-values based on the observed and expected outcomes. This is an advantage if the number of out-of-sample forecasts is not large, this is the case if the forecast horizon is large. As our forecast horizon is not very large ($K = 4$) we will use the 'normal' backtesting test.

4.3.2 CPA test

Eventhough the backtesting tests based on coverage/independence criteria explained in the previous section are appropriate to evaluate the accuracy of a single model, it can provide an ambiguous decision for ranking alternative estimates of the VaR; As in line with Santos, Nogales and Ruiz (2013) we use a statistical test designed to evaluate the comparative predictive performance among candidate models. The comparative predictive ability test proposed by Giacomini and White (2006) is a good alternative for this purpose. In the realistic situation, models are possibly misspecified, due to unmodeled dynamics, unmodeled heterogeneity, incorrect functional form or any combination of these. Specifically, the error terms (the realized value minus the forecasted value) are usually generated from parametric models that have to be recursively estimated over time. This means that the error terms will be polluted by errors caused by estimation uncertainty concerning the parameters of the underlyig models. In relation to our study, our models might have an incorrect functional form or omission of lags.

Giacomini and White derive their tests in an environment where the finite sample properties of the estimators on which the forecasts may depend are preserved asymptotically. This test has several advantages: it captures the effect of estimation uncertainty on relative forecast performance, it can handle forecasts based on both nested and nonnested models, it allows the forecasts to be produced by general estimation methods and they are easy to compute.

The null hypothesis of the test claims that both models have equal predictive ability. In formula, we can present this as follows:

$$H_0 : E[L_1^{\zeta}(e_{t,1}) - L_2^{\zeta}(e_{t,2})] = 0 \tag{4.33}$$

where $L^{\zeta}(e_t)$ is the asymmetric linear loss function described below. $e_{t,1}$ and $e_{t,2}$ denote the difference between the portfolio return and forecasted VaR obtained by model 1 and 2, respectively, in formula this will look as follows: $e_t = y_{p,t} - VaR_t^{\zeta}$.

The test is performed using the following asymmetric linear loss function [3] of order ζ :

$$L^{\zeta}(e_t) = (\zeta - I(e_t < 0))e_t \tag{4.34}$$

Furthermore, finding the model that minimizes 4.34 is an intuitive and appealing criterion to compare predictive ability. A Wald-type test is performed as follows:

$$CPA^{\zeta} = T(T^{-1}\sum_{t=1}^{T-1} I_t LD_{t+1}^{\zeta})'\hat{\Omega}^{-1}(T^{-1}\sum_{t=1}^{T-1} I_t LD_{t+1}^{\zeta}) \tag{4.35}$$

where T is the sample size, LD_t^{ζ} is the loss difference between the two models. As was done by Giacomini and White (2006), we set $I_t = (1, LD_t^{\zeta})$ and $\hat{\Omega}$ is a covariance matrix that consistently estimates the variance of $I_t LD_{t+1}^{\zeta}$, $\hat{\Omega}$ has dimensions 2×2. The null hypothesis of equal predictive ability is rejected for a size ξ when $CPA^{\zeta} > \chi^2_{2,1-\xi}$.

4.4 The Economic Value of Volatility Timing

Past literature has proven that volatility models deliver reasonably accurate volatility/covariance forecasts. For instance, Andersen and Bollerslev (1998) show that volatility models produce strikingly accurate interdaily forecasts. They find that GARCH models explain approximately 50 percent of the variation in their performance meansure of ex post volatility. The question of economic significance however remains unanswered. Besides the statistical evaluation of the volatility/covariance models, we also assess the economic significance of time-varying, predictable volatilities in order to examine the economic value of volatility timing to investors. By doing this, we will not only obtain further insights in the economic significance but we will also evaluate our portfolio VaRs with a different point of view. We will do this by evaluating the impact of predictable changes in the covariance/volatility on the performance of short-horizon asset-allocation strategies. The framework for our analysis is the minimum-variance asset allocation rule and the mean-variance asset allocation rule, that are explained in the next sections.

4.4.1 Minimum-Variance

In this section we will elaborate on the minimum-variance asset allocation rule, that we use to obtain the portfolio weights. We consider an investor who uses a minimum-variance optimization rule to allocate funds across three asset classes: an equity index, a bond index

[3] We will clarify the loss function with a simple example: The portfolio return at time t is -4% and the two VaR forecasts obtained by the two different models at time t, calculated at time $t-1$ are -2% and -4%, respectively. The first model has a VaR violation, the second model does not have a VaR violation. The value of the loss function for the first VaR model is: (0.01-1)(-2)≅2, for the second VaR model it is (0.01-0)(2)=0.02. A model is more penalized when a VaR violation is observed. The greater the magnitude of the violation, the greater the penalization.

and cash. The minimum-variance optimization rule facilitates several aspects of our analysis. First, we avoid the estimation of the conditional expected returns, as these will probably have estimation errors. Second, the input parameter that is necessary for this strategy is the covariance matrix, which can be estimated with greater precision (Merton (1980)). We will use the RiskMetrics approach, and the best performing multivariate GARCH model for the estimation of the covariance matrix. The objective function of the minimum-variance optimization rule is as follows:

$$\min_{x_t} = x_t' \Sigma_t x_t \tag{4.36}$$

The function above minimizes the portfolio covariance matrix for each date t, x_t denotes a $N \times 1$ (N = number of assets) vector of portfolio weights and Σ_t is the forecasted covariance matrix at time t. We also include transaction costs in this analysis as transaction costs are a source of concern for portfolio managers and ignoring transaction costs can result in inefficient portfolios. Let $x_{i,t}^I$ be the amount by which a proportion in the i-th security is increased and $x_{i,t}^D$ be the amount by which a proportion in the i-th security is decreased at time t, then we have:

$$x_t = x_{t-1} - x_t^D + x_t^I \tag{4.37}$$

$$x_t^D \geq 0 \tag{4.38}$$

$$x_t^I \geq 0 \tag{4.39}$$

here $x_t^D = (x_{1,t}^D, x_{2,t}^D, x_{3,t}^D, \ldots, x_{n,t}^D)'$ and $x_t^I = (x_{1,t}^I, x_{2,t}^I, x_{3,t}^I, \ldots, x_{n,t}^I)'$, where N denotes the number of assets. The transaction cost at time t, $c_{i,t}$ of security i is assumed to be a V-shaped function of a difference between a given existing portfolio x_{t-1} and a new portoflio x_t and formulated explicitly into the portfolio return:

$$c_{i,t} = k_i(d_{i,t}^+ + d_{i,t}^-), \forall i \tag{4.40}$$

where k_i is a constant cost per change in a proportion of the i-th security which is assumed to be 1% per change in proportion of a security. By doing this we keep the analysis simple but still incorporate transaction costs. Note that the goal of this research is not optimizing the asset allocation. However to investigate whether volatility timing makes sense. Furthermore $c_t = (c_{1,t}, c_{2,t}, c_{3,t}, \ldots, c_{n,t})$ and

$$x_{i,t} - x_{i,t-1} = d_{i,t}^+ - d_{i,t}^-, \forall i \tag{4.41}$$

$$d_{i,t}^+ d_{i,t}^- = 0, \forall i \tag{4.42}$$

$$d_{i,t}^+, d_{i,t}^- \geq 0, \forall i \tag{4.43}$$

If the difference $x_{i,t} - x_{i,t-1}$ is positive (negative), $d^-_{i,t}(d^+_{i,t})$ becomes zero and $d^-_{i,t}(d^+_{i,t})$ becomes the difference. Because of Equation 4.42, both of d's cannot be positive at the same time. We also include the restriction $x'_t\iota = 1$, on the assets at time t, where ι is a vector of ones. The restriction makes sure that the weights sum up to one. Furthermore we have the restriction $x_t \geq 0$, which prohibits short sales and borrowings. Summarizing, we have the following restrictions:

$$
\begin{aligned}
&c_{i,t} = k_i(d^+_{i,t} + d^-_{i,t}), \forall i \\
&x_{i,t} - x_{i,t-1} = d^+_{i,t} - d^-_{i,t}, \forall i \\
&d^+_{i,t} d^-_{i,t} = 0, \forall i \\
&d^+_{i,t}, d^-_{i,t} \geq 0, \forall i \\
&x'_t\iota = 1 \\
&x_t \geq 0
\end{aligned}
$$

4.4.2 Mean-Variance

Mean variance optimization is a rather straightforward way to come to an "optimal" return portfolio, i.e. portfolio that gives the largest Sharpe Ratio. Central in this is the Markowitz theory. Here again we use the same restrictions as in the last section. However, now we have the following objective function:

$$\max_x U(E(r_c), Var(r_c)) = \max_x [r_f + x_t^{'}\mu_t^e - c_t^{'}\iota - \frac{x_t^{'}\Sigma_t x_t}{2}] \tag{4.44}$$

Where r_f is the risk free return, x_t contains the weights at time t, Σ_t is the forecasted covariance matrix at time t, c_t contains the transaction costs at time t and μ^e is the vector containing the excess returns over the risk free rate. The function above maximizes an investors mean-variance utility function subject to the restrictions

$$
\begin{aligned}
&c_{i,t} = k_i(d^+_{i,t} + d^-_{i,t}), \forall i \\
&x_{i,t} - x_{i,t-1} = d^+_{i,t} - d^-_{i,t}, \forall i \\
&d^+_{i,t} d^-_{i,t} = 0, \forall i \\
&d^+_{i,t}, d^-_{i,t} \geq 0, \forall i \\
&x'_t\iota = 1 \\
&x_t \geq 0
\end{aligned}
$$

the purpose of the formulas and the explanation of the symbols are presented in the previous section.

4.4.3 Performance Measures

To evaluate the performance of the volatility timing results we have to use performance measures, which are explained in this section.

Sharpe Ratio
The Sharpe Ratio is a benchmark that measures the ratio between returns and volatility. As quadratic utility functions are based on the first and second moments of the returns, the Sharpe Ratio is an adequate performance measure for this type of utility.

Manipulation Proof Performance Measure (MPPM)
The Sharpe Ratio is not an adequate performance measure as it is not able to take higher moments such as skewness and kurtosis into account. It is only an adequate measure of performance evaluation when investors believe that risk can be properly measured by standard deviation or in a world where returns are normally distributed. Eling and Schuhmacher (2007) analyze and compare 13 different performance measures for a data set of hedge fund returns and conclude that all these performance measures produce similar rankings. They used the following measures: Sharpe Ratio, Treynor, Jensens Alfa, Omega Ratio, Sortino Ratio, Kappa 3, the upside potential ratio, the Calmer Ratio, the Sterling Ratio, the Burke Ratio, the excess return on Value at Risk, the conditional Sharpe Ratio and the modified Sharpe Ratio. They showed that it seems like the Sharpe Ratio does not perform better than other measures they have included in their research. However, they did not include the Manipulation Proof Performance Measure (MPPM). For these reasons we use this performance measure that goes beyond the mean-variance world. The MPPM has been developed by Goetzmann et al. (2007). Their article describes three general strategies for manipulating a performance measure. The first is the manipulation of the underlying distribution to influence the measure. The second is the dynamic manipulation that induces time variation into the return distribution in order to influence measures that assume stationarity. The third type encompasses dynamic manipulation strategies that focus on inducing estimation error. Furthermore their article defines a MPPM as one that has four properties: i) produce a single valued score with which to rank each subject; ii) score's value should not depend on portfolio's size; iii) an uninformed investor cannot expect to enhance his estimated score by deviating from the benchmark and at the same time informed investors should be able to produce higher scoring portfolios by using arbitrage; and iv) measure should be consistent with standard financial market equilibrium conditions. It turns out that these four requirements are enough to uniquely identify a manipulation-proof measure. Another advantage is that the MPPM is easy to calculate. The MPPM they derive looks as follows:

$$\hat{\Theta} = \frac{1}{(1-\rho)\Delta t} \ln\left(\frac{1}{T}\sum_{t=1}^{T}[(1+r_t)/(1+r_{ft})]^{1-p}\right) \tag{4.45}$$

where $\hat{\Theta}$ is an estimate of the portfolio's premium return after adjusting for risk. That is, the portfolio has the same score as does a risk-free asset whose continuously-compounded return exceeds the interest-rate by $\hat{\Theta}$. It can be interpreted as the annualized continuously compounded excess return certainty equivalent of the portfolio. Hence, a risk-free portfolio earning $\exp[\ln(1+r_{ft})+\hat{\Theta}\Delta t]$ each period would have a measured performance of $\hat{\Theta}$. Furthermore T is the total number of observations and Δt is the length of time between observations, as we have quarterly returns for this analysis the latter is equal to 1/4. These two variables serve to annualize the measure. The portfolio's (un-annualized) rate of return at time t is r_t and the risk-free rate is r_{ft}. The coefficient ρ is a function of the relative risk aversion of the investor, which we set equal to 2 as is usually done in many empirical exercises.

Chapter 5

Results

Since the methodology is extensively explained in the previous chapter, we present our obtained results. We will start by giving the in-sample parameter estimates to obtain a certain sense of how the models/model parameters look like. In order to obtain an insight in the behaviour of the covariances that are forecasted by the different models, we present graphs that provide the evolution of the forecasted covariances through time. Subsequently, the forecasted covariances are used to obtain the forecasted 99% Value-at-Risk. We only consider 99% Value-at-Risk forecasts, as SAMCo is obligated to use 99% VaR. We compare these forecasted VaRs by using different testing procedures, that are all explained in the methodology section. We start with the backtesting test developed by Christoffersen (1998) to explore if the forecasted Value-at-Risks inside the interval are equal to the nominal coverage probability and if they are spread out over the sample. Then we use the pairwise (CPA) test in an environment where the finite sample properties of the estimators on which the forecasts may depend are preserved asymptotically. Obviously, we also want to explore the robustness of our results. This is done by performing the backtesting test for different samples. A robust estimator should be consistent and will approximately give the same results for different samples/subsamples. We also assess the economic value of our models, hence we show these results in the last section of this chapter.

5.1 In-Sample Parameter Estimates

This section presents the estimated parameters of the models we use. Table 5.1 contains all estimated parameters with their corresponding p-values in parenthesis. The parameters are estimated based on the whole in-sample and based on the restrictions to satisfy the positivity and the covariance-stationarity conditions. [1].

[1]Note that the positivity condition is $(1-\alpha-\beta)\overline{Q} > 0$ and $Q_0 > 0$ for the DCC-model and $(1-\alpha-\beta)\overline{Q}-\delta\overline{\Gamma} > 0$ and $Q_0 > 0$ for the Asymmetric DCC-model. The covariance-stationarity condition is $\alpha + \beta < 1$ for the DCC-model and $\alpha + \beta + \lambda\delta < 1$ for the Asymmetric DCC-model.

Full sample:
We observe that the news parameter α is significant for all models, both for the Gaussian distribution and the Student-t distribution. The parameter δ, which includes an asymmetric effect in the DCC-model, is significant and positive which means (according to equation 4.8) that negative returns have a positive effect on the correlation, whereas $\overline{\Gamma} = E[n_t n_t^{'}]$ has a negative effect on the correlation.

We perform the same analysis for the first half of the sample and the second half of the sample to investigate if the parameters remain significant.

First half of the sample:
The news parameter α of the DCC-GARCH, DCC-TGARCH and ADCC-GARCH is no longer significant. We observe that the asymmetric term δ remains significant and positive for the Asymmetric-DCC models.

Second half of the sample:
We observe that the news parameter α of DCC-GARCH, DCC-TGARCH and ADCC-GARCH is not significant. This is the same result as we obtained from the first half of the sample. The parameter β is not significant for ADCC-GARCH and ADCC-TGARCH. The asymmetric term δ is no longer significant for the ADCC-GARCH and the ADCC-TGARCH model.

Table 5.1: **Estimated parameters**

	ω	α	β	δ
Full sample				
DCC-GARCH		0.017	0.968	
		(0.001)*	(0.009)*	
DCC-GARCH-t		0.019	0.965	
		(0.003)*	(0.007)*	
DCC-TGARCH		0.019	0.964	
		(0.003)*	(0.012)*	
DCC-TGARCH-t		0.019	0.961	
		(0.000)*	(0.000)*	
ADCC-GARCH		0.003	0.963	0.033
		(0.004)*	(0.003)*	(0.000)*
ADCC-GARCH-t		0.045	0.890	0.021
		(0.001)*	(0.008)*	(0.000)*
ADCC- TGARCH		0.025	0.910	0.051
		(0.001)*	(0.007)*	(0.002)*
ADCC-TGARCH-t		0.043	0.897	0.042
		(0.010)*	(0.001)*	(0.003)*

This table presents the estimated parameters for each model with corresponding p-values in parenthesis, the full sample is used for this purpose. * Denotes significance at the 1% level. ** Denotes significance at the 5% level.

Table 5.2: **Estimated parameters**

	ω	α	β	δ
First half of the sample				
DCC-GARCH		0.014	0.951	
		(0.062)	(0.017)**	
DCC-GARCH-t		0.019	0.938	
		(0.002)*	(0.043)**	
DCC-TGARCH		0.016	0.947	
		(0.063)	(0.007)*	
DCC-TGARCH-t		0.031	0.952	
		(0.041)*	(0.039)*	
ADCC-GARCH		0.014	0.951	0.000
		(0.061)	(0.047)**	(0.001)*
ADCC-GARCH-t		0.071	0.891	0.093
		(0.031)**	(0.027)**	(0.038)**
ADCC- TGARCH		0.015	0.947	0.003
		(0.003)*	(0.042)**	(0.020)**
ADCC-TGARCH-t		0.073	0.839	0.031
		(0.020)**	(0.007)*	(0.004)*
Second half of the sample				
DCC-GARCH		0.003	0.977	
		(0.051)	(0.042)**	
DCC-GARCH-t		0.041	0.910	
		(0.035)**	(0.009)*	
DCC-TGARCH		0.005	0.971	
		(0.071)	(0.008)*	
DCC-TGARCH-t		0.081	0.914	
		(0.008)*	(0.047)**	
ADCC-GARCH		0.003	0.977	0.000
		(0.074)	(0.102)	(0.083)
ADCC-GARCH-t		0.061	0.894	0.031
		(0.046)**	(0.006)*	(0.004)*
ADCC- TGARCH		0.005	0.971	0.000
		(0.039)**	(0.077)	(0.064)
ADCC-TGARCH-t		0.071	0.910	0.032
		(0.021)**	(0.004)*	(0.033)**

This table presents the estimated parameters for each model with corresponding p-values in parenthesis, the first half of the sample (14/05/1999 -11/11/2005) and the second half of the sample (18/11/2005 - 18/05/2012) are used for this purpose. * Denotes significance at the 1% level. ** Denotes significance at the 5% level.

5.2 Forecasts of the Covariances

In order to evaluate the evolution of the forecasted covariances through time, we present the covariance in Figure 5.1. Figure 5.1a shows the 4-week ahead forecasted covariance between MSC and MEM that are obtained by both the RiskMetrics approach and the DCC models with different specifications for the volatility. Figure 5.1b presents the 4-week ahead forecasted covariance between MSC and MEM, obtained again by the RiskMetrics approach and the Asymmetric DCC models together with all different specifications for the univariate volatility. A window with 200 observations is used for this purpose. Using a rolling window that estimates the parameters every week and then forecasts the VaRs takes a lot of time. Since the DCC-models we use are computationally quite intensive, this would even take more time when using a rolling window that estimates the parameters every week. For this reason we make us of a different strategy, that uses a certain fixed window to estimate the parameters of the models and then re-estimates the parameters once a quarter of a year. The main advantage of this procedure is that it entails less computational effort since all parameters are estimated less frequently.

In general, the magnitudes of the covariances are not very informative, as they depend on the scale. Besides, it is not clear whether the differences in the covariances is due to differences in the volatility or differences in the correlations between the assets. For this reason, we will also discuss the results in terms of volatilities/correlations. Note that a large covariance means a large correlation between two assets and/or low individual volatilities of two assets.

Our first observation is that both the DCC and ADCC models with the TGARCH-specification for the univariate volatilities give the largest covariance in comparison with the rest of the covariances. This is due to the fact that the asymmetric term in the TGARCH model has a large positive value, which causes a large fluctuation in the covariance during the crisis period as the covariance depends on the correlation and the individual volatilities of the two indices.

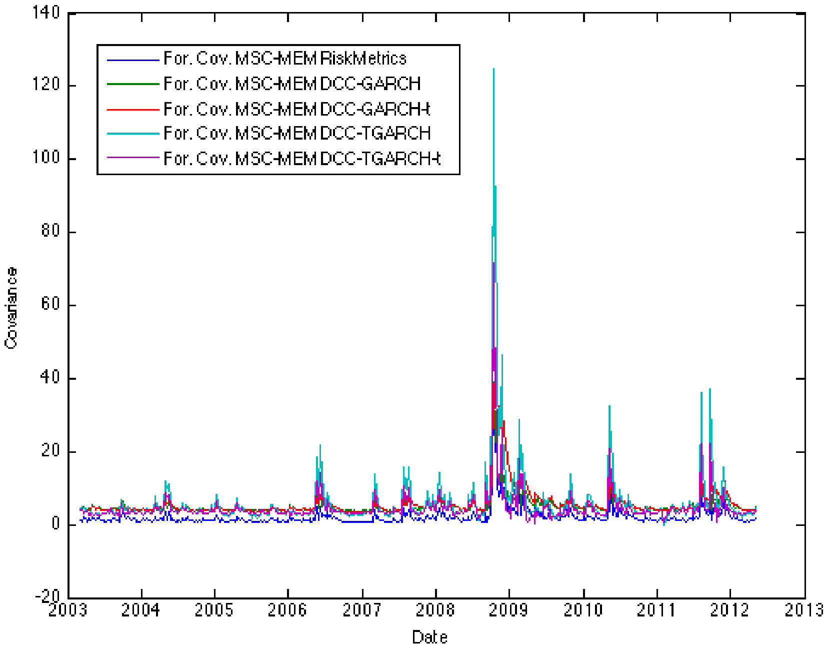
Furthermore, we can observe that the DCC model with the TGARCH specification produces a very large covariance between MSC and MEM. The covariances for the rest of the models have magnitudes up to 100, while the DCC-TGARCH model produces a covariance of approximately 120. Hence, MSC and MEM are possibly highly correlated. We can verify this by looking at Table 3.3 which shows us that MSC and MEM have a correlation value that is equal to 0.84.

If we take a look at fluctuations other than the fluctuation during the crisis, we can observe again that covariances of models with the TGARCH specification make larger jumps and react more extremely.

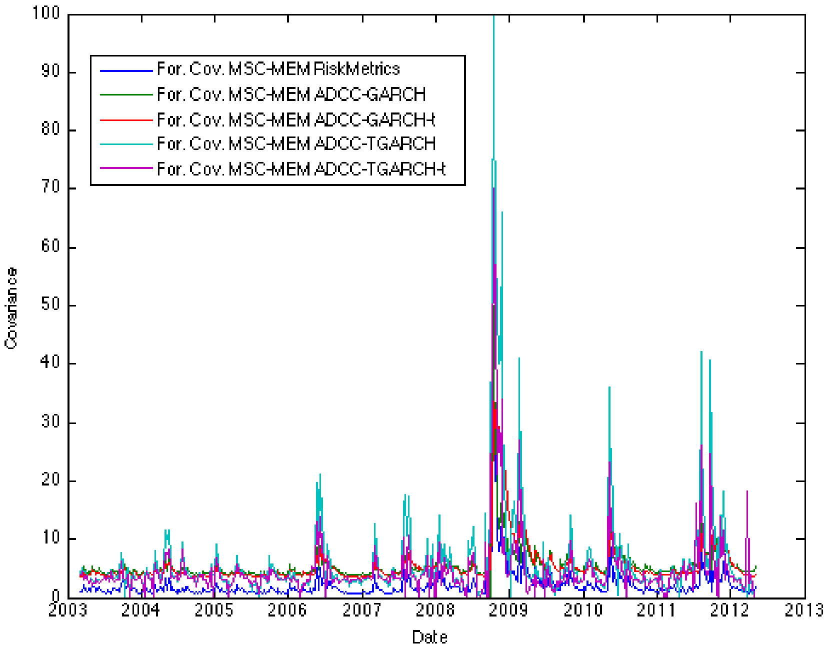
The forecasted covariance that is calculated with the help of the RiskMetrics approach is in comparison with the other forecasted covariances quite low. This can be explained by the decay factor that is used for the RiskMetrics approach, which is fixed at 0.99.

Figure 5.1: **Covariance forecasts**

(a) DCC



(b) ADCC



These figures present four-step ahead out-of-sample forecasts of the covariance between MSC and MEM obtained by RiskMetrics, all DCC-models and ADCC-models based on a fixed window of 200 observations. The parameters of the DCC-models are estimated once a quarter of a year to reduce computational effort.

5.3 Comparison of the Performances

5.3.1 Value-at-Risk

After estimating the in-sample parameter estimates, we obtain out-of-sample forecasts of the VaR. Table 5.3 presents the results of the backtesting test of Christoffersen (1998) based on out-of-sample 99%-VaR estimates. The table reports the p-values of the unconditional coverage, independence and conditional coverage test for each of the sub-groups.

We can observe that the null hypothesis of correct unconditional coverage is rejected for all models, except for the ADCC-GARCH-t and ADCC-TGARCH-t model at the $\frac{\alpha}{K} = \frac{0.10}{4} = 0.025$ significance level. Obviously the forementioned models produce a VaR violation that is not significantly different from the nominal coverage probability, whereas the other models are not able to reproduce this. The models that do not perform well according to the correct unconditional coverage test might underestimate risk during a period of stress that lies ahead, resulting in an increase in the number of VaR violations. When looking at the independence test, we can observe another interesting fact. The null hypothesis can not be rejected for all models, meaning that the VaR forecasts (of all models) outside the interval are spread out over the sample and do not come in clusters. A model should perform well both for the correct unconditional coverage test and the independence test. Taken these two criteria into account we can perform a correct conditional coverage test. Only the ADCC-GARCH-t and the ADCC-TGARCH-t models are able to perform well when taking the two criteria into account.

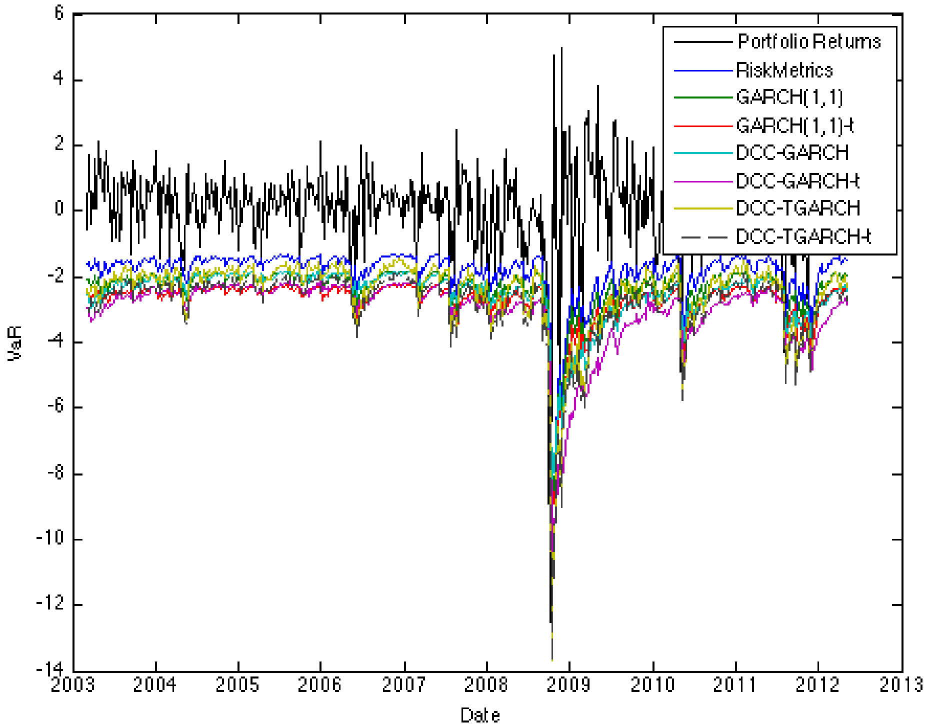
To summarize, only the ADCC-GARCH-t and ADCC-TGARCH-t are able to pass the correct unconditional coverage test meaning that the VaR violations that these models have is not significantly different from the nominal coverage probability. Clearly, the RiskMetrics model has VaR violations which are significantly different from the nominal coverage probability. All models pass the independence test. Apparently, all models are able to capture heteroskedasticity. Finally, only the ADCC-GARCH-t and the ADCC-TGARCH-t perform well for the conditional coverage test.

We can observe the forecasted VaRs through time for all models in Figure 5.2. The RiskMetrics model is not able to quickly respond to periods of stress and its pattern is very similar to the pattern of the GARCH(1,1) VaR forecasts. The reason is very simple: RiskMetrics is a restricted GARCH(1,1) model. We can also observe that the RiskMetrics VaR forecasts are quite large in comparison with the rest. Hence RiskMetrics might underestimate risk. The model that has the lowest peaks is the GARCH-t model, as this model gives the highest line in the graph in comparison with the rest. The DCC-TGARCH and the ADCC-TGARCH models produce the highest peak in the VaR forecasts. It might be the case that these models overestimate the risk. Both the DCC and the ADCC models move in a similar way and show the same pattern. Compared to the RiskMetrics model they are able to respond to periods of stress and do not underestimate the risk.

To be able to easily compare the DCC and the Asymmetric DCC models with each other we also present them in Figure 5.3. From Figure 5.3a we can observe that the models have the same pattern and are different in their lags. The DCC-GARCH-t model has the lowest VaR values, while the ADCC-GARCH model has the largest VaR values. Figure 5.3b shows that the ADCC-TGARCH and ADCC-TGARCH-t models have more fluctuations in comparison with the other models. This can be explained by the asymmetric terms of these models that take the leverage effect into account and therefore reacts more extreme in periods where the returns are negative.

Figure 5.2: Four-step ahead 99% Value-at-Risk forecasts obtained by all models based on a window of 200 observations. The parameters are re-estimated once a quarter of a year. The portfolio returns are based on an equally weighted portfolio.

(a) DCC



(b) ADCC

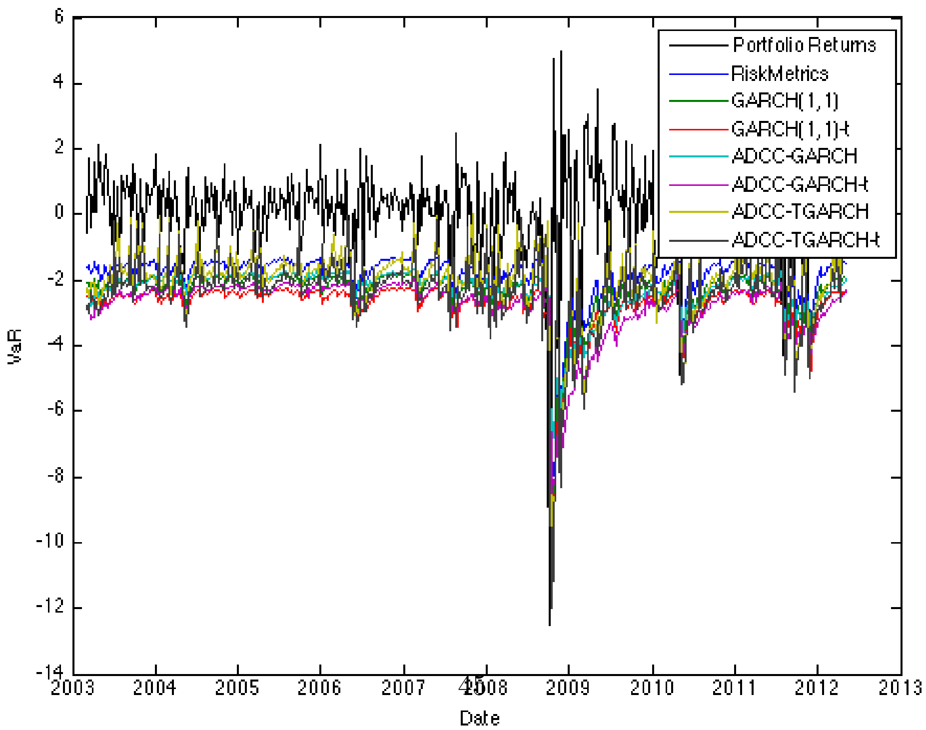
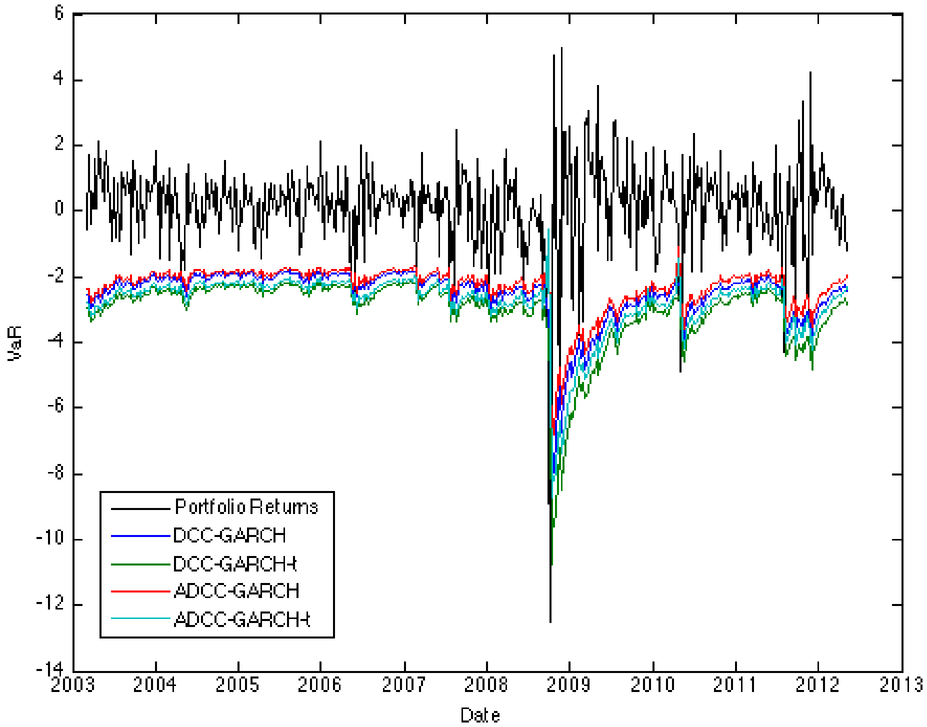


Figure 5.3: Four-step ahead 99% Value-at-Risk forecasts of the (A)DCC models based on a window of 200 observations. The parameters are re-estimated once a quarter of a year. The portfolio returns are based on an equally weighted portfolio.

(a) (A)DCC with GARCH-(t) specifications



(b) (A)DCC with TGARCH-(t) specifications

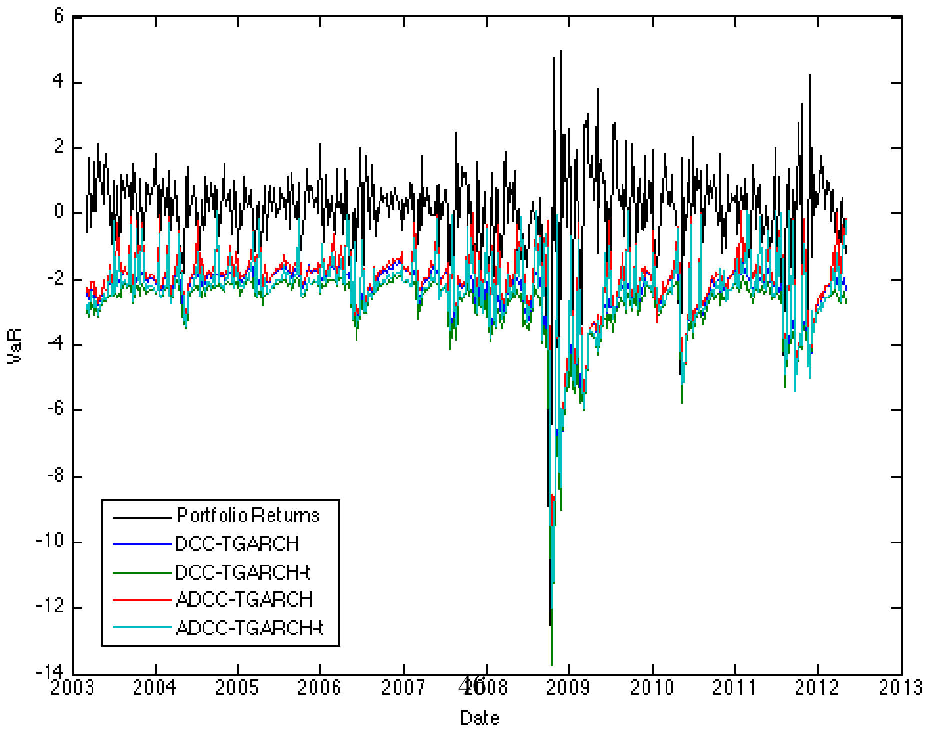


Table 5.3: **Value-at-Risk**

$K = 4$ Model	$\frac{\alpha}{K} = \frac{0.10}{4} = 0.025$ LR_{cuc}	LR_{ind}	LR_{cc}
RiskMetrics	[**0.00**], [**0.00**], [**0.00**], [**0.00**]	[0.10], [0.13], [0.10], [0.10]	[**0.00**], [**0.00**], [**0.00**], [**0.00**]
GARCH(1,1)	[**0.00**], [**0.00**], [**0.00**], [**0.00**]	[0.27], [0.09], [0.27], [0.15]	[**0.00**], [**0.00**], [**0.00**], [**0.00**]
GARCH(1,1)-t	[**0.00**], [**0.00**], [**0.00**], [**0.00**]	[0.43], [0.36], [0.36], [0.40]	[**0.00**], [**0.00**], [**0.00**], [**0.00**]
DCC-GARCH	[**0.00**], [**0.00**] [**0.00**], [**0.00**]	[0.46], [0.40], [0.46], [0.42]	[**0.00**], [**0.00**], [**0.00**], [**0.00**]
DCC-GARCH-t	[**0.00**], [**0.00**], [**0.00**], [**0.00**]	[0.05], [0.05], [0.06], [0.05]	[**0.02**], [**0.02**] [**0.02**], [**0.02**]
DCC-TGARCH	[**0.00**], [**0.00**], [**0.00**], [**0.00**]	[0.09], [0.15], [0.09], [0.12]	[**0.00**], [**0.00**], [**0.00**],[**0.00**]
DCC-TGARCH-t	[**0.00**],[**0.00**] [**0.00**],[**0.00**],	[0.52], [0.68], [0.48], [0.55]	[**0.00**],[**0.00**], [**0.00**],[**0.00**]
ADCC-GARCH	[**0.00**],[**0.00**], [**0.00**],[**0.00**],	[0.08], [0.09] [0.08], [0.07]	[**0.00**],[**0.00**], [**0.00**],[**0.00**]
ADCC-GARCH-t	[0.09], [0.09] [0.11], [0.09]	[0.21], [0.19] [0.21], [0.19]	[0.12], [0.12] [0.12], [0.12]
ADCC- TGARCH	[**0.00**],[**0.00**], [**0.00**],[**0.00**],	[0.46], [0.50], [0.48], [0.51]	[**0.00**],[**0.00**], [**0.00**],[**0.00**]
ADCC-TGARCH-t	[0.08], [0.08] [0.09], [0.08]	[0.72], [0.72] [0.66], [0.70]	[0.20], [0.20] [0.20], [0.20]

This table presents four-step ahead out-of-sample backtesting p-values for each of the four sub-groups. The likelihood ratio test statistics of correct unconditional coverage (LR_{cuc}), independence (LR_{ind}) and correct conditional coverage (LR_{cc}) are presented. Bold statistic indicate significance at the 0.10/K level. A coverage probability q of 0.01 and a window of 200 observations are used to obtain the Value-at-Risk forecasts and the p-values.

5.3.2 The Comparative Predictive Ability (CPA) test

Another method to evaluate the accuracy of the VaR forecast is by using the CPA-test developed by Giacomini & White (2006). It is a pairwise test, hence it compares the VaR forecasts of two models with each other. A model is more penalized when a VaR violation is observed. The greater the magnitude of the violation, the greater the penalization.

Table 5.5 presents the test-statistics, the p-values denoted in parenthesis, an upper arrow and a left arrow. An upper arrow means that the model in the column outperforms the model on the row. A left arrow means that the model in the row outperforms the model in the column.

RiskMetrics model is outperformed by eight models and it is only able to outperform the DCC-GARCH and the ADCC-GARCH model. If we compare the GARCH models with the multivariate models, we observe that the GARCH model outperforms the multivariate models four out of eight times. The GARCH-t model also outperfoms its multivariate counterparts four out of eight times. However, it is important to mention that a GARCH model is not very interesting for a long-term investor. A long-term investor does not care about sudden fluctuations in the volatility and volatility clustering, especially if he or she assumes the presence of mean reversion. However, it is of great importance for short-term investors.

In general, assigning a Student-t distribution to the error terms leads to an improvement of the model, as most models that have t-distributed error terms perform better. Table 5.5 shows that the ADCC-TGARCH-t model outperforms all other models. This is in line with the backtesting test in the previous section, where we found that the ADCC-TGARCH-t model was able to pass all tests. Furthermore, it is important to note that the results are mixed when estimating models with Gaussian errors. This result corroborates the evidence of the research of Santos et al. (2013). Table 5.4 presents the ranking of all models based on the CPA test. The larger the number of models that a specific model outperforms, the higher the ranking. We observe that indeed the ADCC-TGARCH-t model is able to outperform all other models based on the CPA test, that ADCC-GARCH-t is the second best performing model and that RiskMetrics is one of the worst performing models. Obviously, RiskMetrics produces many VaR violations in the VaR forecasts compared with other models. Here again we can clearly see that models with a Student-t distribution are performing better than the same model with a Normal distribution for the error terms.

To summarize the results of this section, assigning a Student-t distribution to the error terms leads to an improvement of the model. The RiskMetrics Approach is outperformed by eight other models. Finally, the ADCC-TGARCH-t model performs best based on the CPA-test as this model is able to outperform all other models.

Table 5.4: **Comparative Predictive Ability**

Model	Ranking	Model	Ranking
ADCC-TGARCH-t	1 (10)	ADCC-GARCH	5 (4)
ADCC-GARCH-t	2 (8)	ADCC-TGARCH	6 (3)
DCC-TGARCH-t	2 (8)	DCC-GARCH-t	7 (2)
ADCC-TGARCH	3 (6)	RiskMetrics	7 (2)
GARCH-t	3 (6)	DCC-GARCH	8 (1)
GARCH	4 (5)		

This table presents the Comparative Predictive Ability-test ranking from best performing model till worst performing model based on the number of outperformed models. The number in parenthesis denotes the number of models that are outperformed by the model on the row.

Table 5.5: **Comparative Predictive Ability Test Statistics**

	RM	GARCH(1,1)	GARCH(1,1)-t	DCC-GARCH	DCC-GARCH-t	DCC-TGARCH	DCC-TGARCH-t	ADCC-GARCH	ADCC-GARCH-t	ADCC-TGARCH	ADCC-TGARCH-t
RM	-	7.574 ↑	12.696 ↑	8.751 ←	10.939 ↑	13.567 ↑	12.367 ↑	15.704 ←	12.681 ↑	11.498 ↑	15.394 ↑
		(0.023)	(0.002)	(0.013)	(0.004)	(0.001)	(0.002)	(0.000)	(0.002)	(0.003)	(0.001)
GARCH(1,1)	-	-	7.019 ↑	10.425 ←	11.066 ←	15.992 ↑	18.757 ↑	14.287 ←	13.560 ↑	10.560 ←	16.533 ↑
			(0.029)	(0.005)	(0.004)	(0.000)	(0.000)	(0.001)	(0.001)	(0.005)	(0.000)
GARCH(1,1)-t	-	-	-	11.599 ←	11.220 ←	16.540 ↑	19.084 ↑	14.005 ←	13.385 ↑	10.154 ←	16.407 ↑
				(0.003)	(0.004)	(0.000)	(0.000)	(0.001)	(0.001)	(0.006)	(0.000)
DCC-GARCH	-	-	-	-	0.472↑	15.744 ↑	21.025 ↑	8.123 ↑	16.208 ↑	11.381 ←	16.152 ↑
					(0.790)	(0.000)	(0.000)	(0.017)	(0.000)	(0.003)	(0.000)
DCC-GARCH-t	-	-	-	-	-	15.060 ↑	20.119 ↑	8.306 ↑	6.463 ↑	11.218 ↑	16.077 ↑
						(0.001)	(0.000)	(0.016)	(0.040)	(0.004)	(0.000)
DCC-TGARCH	-	-	-	-	-	-	4.663	11.308 ↑	19.121 ↑	11.299 ←	17.102 ↑
							(0.097)	(0.004)	(0.000)	(0.004)	(0.000)
DCC-TGARCH-t	-	-	-	-	-	-	-	11.553 ←	9.184 ←	11.975 ↑	6.884 ↑
								(0.003)	(0.010)	(0.003)	(0.032)
ADCC-GARCH	-	-	-	-	-	-	-	-	12.997 ↑	8.962 ←	16.423 ↑
									(0.002)	(0.011)	(0.000)
ADCC-GARCH-t	-	-	-	-	-	-	-	-	-	7.783 ←	4.29 ↑6
										(0.020)	(0.117)
ADCC-TGARCH	-	-	-	-	-	-	-	-	-	-	4.055 ↑
											(0.132)
ADCC-TGARCH-t	-	-	-	-	-	-	-	-	-	-	-

Comparative Predictive Ability Test Statistics with associated p-values in parenthesis. An up-arrow denotes that the model in the column outperforms the model on the row, a left-arrow denotes that the model on the row outperforms the model in the column.

5.3.3 Robustness

A 'good' model is also able to perform well and is consistent for all different subsamples, crisis periods and less turbulent periods. We evaluate the robustness of our results by forecasting the VaRs for all models on different subsamples. The first one is the pre-crisis period that starts at May 14th, 1999 and ends at June 27th, 2008. The crisis period starts at July 4th, 2008 and ends at December 25th, 2009. Finally the post-crisis period that starts at January 1st, 2010 and ends at May 18th, 2012. Here again, we evaluate the VaR forecasts with the backtesting test, by performing an unconditional coverage test, independence test and conditional coverage test. Table 5.6 presents the number of rejections of the tests of correct unconditional coverage, independence and correct conditional coverage using the 5% significance level for all different sub-samples. We choose for this presentation to prevent an inconvenient table with many numbers.

Earlier, we found that only the ADCC-GARCH-t and the ADCC-TGARCH-t models are able to pass the correct unconditional coverage test, that all models pass the independence test and that only the ADCC-GARCH-t and the ADCC-TGARCH-t perform well for the conditional coverage test. We will investigate whether these findings remain the same for the different subsamples.

We start with the pre-crisis period. Here again, we see that it is hard for most models to pass the correct unconditional coverage test. If we make use of a 5% significance level, only the ADCC-GARCH-t and ADCC-TGARCH-t model pass this test. Also for this subsample we can conclude that all models pass the independence test. The ADCC-GARCH-t and the ADCC-TGARCH-t model are only able to pass the conditional coverage test at the 5% significance level.

When looking at the crisis period we see different results. None of the models are able to pass the unconditional coverage test. The performance for the independence test is similar, none of the models are able to pass this test. Hence all models are rejected for the conditional coverage test.

The results of the post-crisis period are similar to the results of the pre-crisis period. Obviously, the ADCC-GARCH-t and the ADCC-TGARCH-t model are quite consistent as they have similar results for different subsamples. For our third subsample, they can again pass the unconditional coverage test, the independence test and the correct conditional coverage test at the 5% significance level. The other models are not able to pass the unconditional coverage test, they are able to pass the independence test and finally they are not able to pass the conditional coverage test. We can conclude that the results are quite consistent especially for the pre-crisis and post-crisis period. These results corrobarate the findings of the backtesting test results earlier in this research. Obviously, the ADCC-GARCH-t and the ADCC-TGARCH-t are able to produce a VaR violation that is significantly not different from the nominal coverage probability and the VaR forecasts (of these two models) outside the interval are spread out over the sample and do not come in clusters for different subsamples (except for the crisis period).

Table 5.6: **Robustness results for the backtesting test**

q=0.01	LR_{cuc}	LR_{ind}	LR_{cc}
Before the crisis:			
RiskMetrics	4	0	4
GARCH(1,1)	4	0	4
GARCH(1,1)-t	4	0	4
DCC-GARCH	4	0	4
DCC-GARCH-t	4	0	4
DCC-TGARCH	4	0	4
DCC-TGARCH-t	4	0	4
ADCC-GARCH	4	0	4
ADCC-GARCH-t	0	0	0
ADCC- TGARCH	4	0	4
ADCC-TGARCH-t	0	0	0
During the crisis:			
RiskMetrics	4	3	4
GARCH(1,1)	4	3	4
GARCH(1,1)-t	4	3	4
DCC-GARCH	4	3	4
DCC-GARCH-t	4	3	4
DCC-TGARCH	4	3	4
DCC-TGARCH-t	4	3	4
ADCC-GARCH	4	3	4
ADCC-GARCH-t	4	3	4
ADCC- TGARCH	4	3	4
ADCC-TGARCH-t	4	3	4
After the crisis:			
RiskMetrics	4	0	4
GARCH(1,1)	4	0	4
GARCH(1,1)-t	4	0	4
DCC-GARCH	4	0	4
DCC-GARCH-t	4	0	4
DCC-TGARCH	4	0	4
DCC-TGARCH-t	4	0	4
ADCC-GARCH	4	0	4
ADCC-GARCH-t	0	0	0
ADCC- TGARCH	4	0	4
ADCC-TGARCH-t	0	0	0

The values denote the number of rejections of the tests of correct unconditional coverage (LR_{cuc}), independence (LR_{ind}) and correct conditional coverage (LR_{cc}) using the 5% significance level. A coverage probability q of 0.01 is used to obtain the Value-at-Risk forecasts and the test statistics. The pre-crisis period starts at May 14th, 1999 and ends at June 27th, 2008. The crisis period starts at July 4th, 2008 and ends at December 25th, 2009. Finally the post-crisis period that starts at January 1st, 2010 and ends at May 18th, 2012.

5.4 The Economic Value of Volatility Timing

This section will elaborate on the economic significance of volatility timing. In the previous sections we presented the results of the statistical tests. As discussed in the methodology, in this chapter we will evaluate the economic value of volatility timing for short-horizon asset-allocation strategies. We use the minimum-variance and the mean-variance asset-allocation rules to calculate the portfolio weights. Constructing the optimal portfolios requires estimates of the conditional covariance matrix. We use three different methods to forecast the Covariance matrix, which we then use as an input in the minimum-variance and the mean-variance asset allocation rules. The portfolio weights are used to calculate the ex-post portfolio returns. Consequently, the portfolio returns are used to calculate the Sharpe Ratio and the MPPM. To be able to assess the statistical significance of the volatility timing results we conduct a simulation exercise which is explained in the results section. We use a stock index, a bond index and a risk-free rate for this analysis. The first part of this section explains the data used in this section. Consequently, we will present the results.

5.4.1 Data

Our data is in line with Goyal & Welch (2008) and Rapach, Strauss and Zhou (2010). We use their data, because their data consists of a simple stock index, a risk free rate and a government bond return. This makes the portfolio very simple, straightforward and diversified. Furthermore their data is often used in portfolio management and entails less computational effort as the dimension of the portfolio is reduced. It is also SAMCo's preference to use a simple portfolio for this analysis. The data consist of 332 quarterly observations of 3 assets from 1926 to 2008. Below all types are discussed in detail.

Real risk-free rate: The risk-free rate is the Treasury bill rate. We model the real risk-free rate as $ln(1+R_f)-ln(1+INFL)$, where R_f is the risk-free rate and $INFL$ is the Consumer Price Index from the Bureau of Labor Statistics.
Excess stock return: We use S&P 500 index returns from the Center for Research in Security Press (CRSP) month end values. As Goyal and Welch (2008) mention, the stock returns are the continuously compounded returns on the S&P 500 index including dividends. The excess stock return is modeled as $ln(1+R_s)-ln(1+R_f)$, where R_s is the stock return and R_f is the risk-free rate.
Excess bond return: We use long-term government bond returns R_b to model the excess bond return as $ln(1+R_b)-ln(1+R_f)$.

Table 5.7 presents the descriptive statistics of these three assets and the correlations between them. As a priori expected, the stock has the largest mean return and the highest volatility. All assets show excess kurtosis and are non-normally distributed. The correlations between the assets are not very large. The largest correlation is between the risk-free rate and the long-term government bond. The risk-free rate and the stock are negatively correlated with each other, however the magnitude of this correlation is not very large.

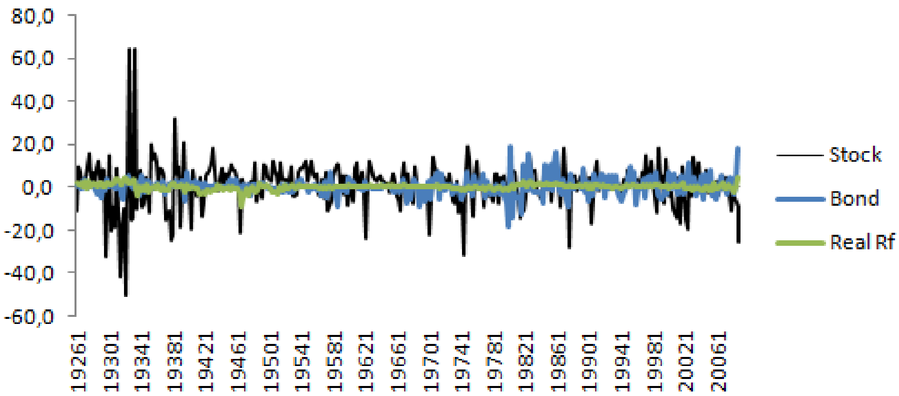
Figure 5.4 presents the time series of the three returns. The stock index is quite volatile whereas the real risk-free rate is stable through time. The bond becomes more volatile after the end of the 70's with a lot of fluctuations.

Table 5.7: **Descriptive Statistics**

	Risk-free rate	Stock	Bond
Mean	0.002	0.013	0.005
Median	0.003	0.029	0.003
Std.Dev	0.013	0.108	0.042
Min	-0.088	-0.499	-0.186
Max	0.045	0.640	0.185
Skewness	-1.049	0.218	0.349
Excess kurtosis	7.056	8.302	3.758
P-value	0.001	0.001	0.001
Correlation			
Risk-free rate	1	-0.056	0.214
Stock	-	1	0.076
Bond	-	-	1

This table presents the descriptive statistics of the assets in the simple portfolio we use and the correlation between the assets. Returns are quarterly, denominated in USD, include dividends and are NOT excess returns. The sample period is 1926Q1 - 2008Q4. The P-value evolves out of the Jarque-Bera test, which is a test for normality based on the skewness and kurtosis.

Figure 5.4: Time series of the excess stock return, excess bond return and the real risk-free rate. The first four numbers of the values on the x-axis denote the year, and the last number denotes the quarter. The returns on the y-axis are in percentage (%).



5.4.2 Results

The variation in the covariance matrix, which is proved in section 5.2 suggests that there is a (possible) role for volatility timing in asset-allocation decisions. In this section we will elaborate on the latter. We will do this by examining the mean portfolio return, the Sharpe Ratio's, the Sample Volatilities and the simulation results. To estimate the conditional covariance matrix we use different methodologies:

1. The RiskMetrics approach.
2. The DCC-GARCH model
3. The ADCC-TGARCH-t model.

The conditional covariance matrix is used to compute the optimal portfolio weights. Then we apply these portfolio weights to the actual returns to calculate the ex post portfolio return.

As already mentioned in the methodology, the portfolio weights are constructed using the minimum-variance and the mean-variance asset allocation rules along with the one-step-ahead estimates of the conditional covariance matrix. To perform the optimization we use *fmincon* in Matlab. Furthermore, we use 100 observations for the in-sample. This means that we have 231 observations for the out-of-sample. An expanding window is used to obtain the out-of-sample volatility and covariance estimates. In this section we demonstrate the ex-post mean portfolio return which is calculated by multiplying the portfolio weights with the observed next-day returns on stock, bond and cash. We also show the sample portfolio volatility and the estimated Sharpe Ratio.

We are aware of the fact that the objective of every asset allocation strategy is different and that in portfolio managament this would usually cause problems for the comparability. However note that the goal of this analysis is not comparing different asset allocation rules with each other, however we would like to investigate whether volatility timing makes sense. Using different asset allocation rules will serve as some kind of robustness analysis as we obtain insights for different asset allocation strategies.

Table 5.8 shows a number of results. It shows the mean quarterly portfolio return, sam-

Table 5.8: **Economic Significance results**

	Mean QR	SV	SR	Freq $SV_{port} > SV_{sim}$ (%)	Freq $SR_{sim} > SR_{port}$ (%)	Freq $\Theta_{sim} > \Theta_{port}$ (%)
Minimum-Variance						
RiskMetrics	1.92%	2.58%	0.74			
Simulation	0.59%	1.44%	0.41	5.1%	7.3%	7.9%
DCC-GARCH	2.14%	2.58%	0.83			
Simulation	1.43%	3.05%	0.47	7.4%	8.7%	8.9%
ADCC-TGARCH-t	2.52%	3.01%	0.84			
Simulation	1.87%	3.08%	0.61	8.1%	6.9%	6.2%
Mean-Variance						
RiskMetrics	2.05%	2.80%	0.74			
Simulation	1.41%	2.65%	0.39	7.4%	5.4%	5.9%
DCC-GARCH	2.88%	3.42%	0.84			
Simulation	2.34%	4.18%	0.56	9.3%	3.9%	4.3%
ADCC-TGARCH-t	2.95%	3.22%	0.92			
Simulation	2.49%	3.60%	0.69	8.4%	2.6%	3.5%

This table presents the Mean Quarterly Return (Mean QR), Sample Volatility (SV) and the Sharpe Ratio (SR) of the portfolio that consists of a simple stock, bond and risk free rate. RiskMetrics, DCC-GARCH and ADCC-TGARCH are used to calculate the time-varying volatilties. The simulation is conducted by randomly rearranging the returns and applying the actual weights to the randomly rearranged return series to compute portfolio returns. We use 10.000 trials for the simulations. The frequencies with which the simulation beat the portfolio according to the Sharpe Ratio and the MPPM are reported in %. Also the frequencies with which the portfolio beat the simulation according to the Sample Volatiliy is reported in %. The results are presented for the mean-variance rule and the minimum-variance rule. The MPPM is calculated as follows: $\hat{\Theta} = \frac{1}{(1-\rho)\Delta t} \ln(\frac{1}{T} \sum_{t=1}^{T} [(1 + r_t)/(1 + r_{ft})]^{1-p})$ (see equation 4.45).

ple volatility and the estimated Sharpe Ratio that are obtained using the minimum-variance and mean-variance asset allocation rules by using three different methodologies to estimate the conditional covariance matrix. Furthermore, to be able to assess the statistical significance of the volatility timing results, we also conduct simulations where the asset returns are generated independently of the portfolio weights as in line with Fleming, Kirby and Ostdiek (2001). First the actual return series are rearranged randomly and then we apply the actual weights to the randomly rearranged return series to compute portfolio returns. We use 10.000 trials for the simulations. If volatility-timing makes sense then we expect that the strategies should perform better using the actual data than in the simulations. The following facts emerge from Table 5.8:

Minimum-Variance results

Using the ADCC-TGARCH-t method to construct the covariance matrix yields the largest Mean Portfolio Return (2.52%), whereas using the RiskMetrics approach yields the lowest Mean Portfolio Return (1.92%). Using the ADCC-TGARCH-t model to construct the covariance matrix yields the largest Sharpe Ratio (0.84%), using RiskMetrics yields the lowest Sharpe Ratio (0.74%). Although we use the minimum-variance asset allocation rule, it is nog completely 'fair' to use the Sharpe Ratio as a performance measure, because the objective of the minimum-variance asset allocation rule is to minimize risk not to maximize the Sharpe Ratio. However it is interesting to show these results as we are wondering if the simulation exercise yields larger Sharpe Ratio's. As mentioned before, the objective of the minimum-variance asset allocation rule is minimizing risk. If we look at the volatilities then we observe that although Riskmetrics gives the lowest Sharpe Ratio, it is able to give the lowest volatility compared with the Sharpe Ratio that evolve out of other covariance-constructing models. Hence, for this performance measure it does a better job than DCC-GARCH and ADCC-TGARCH-t. Also the frequencies with which the portfolio beat the simulation according to the Sample Volatility is most favourable for the Riskmetrics approach. All simulations perform worse than the actual data as the sample volatilities for the actual data are only in a few cases larger than the sample volatility of the simulations. For all methodologies a very low percentage of the total number of trials (10.000) yield a higher Sharpe Ratio and a higher MPPM. For example for the RiskMetrics, only 7.3% of the trials yield a higher Sharpe Ratio and 7.9% of the trials yield a higher MPPM. Hence, we can conclude that these findings indicate that the volatility-gains are significant and that it is unlikely that the gains to volatility-timing are due to chance. We observe that ADCC-GARCH-t provides the lowest frequency with which the simulation beats the portfolio according to the MPPM and the Sharpe Ratio. Hence, we can conclude that ADCC-TGARCH-t is the best performing model when looking at the Sharpe Ratio and the MPPM.

Mean-Variance results

The results are quite consistent as here again using the ADCC-TGARH-t model gives the largest Mean Portfolio Return and the RiskMetrics approach provides the lowest Mean Portfolio Return. If we look at the volatility then the RiskMetrics approach is able to provide the lowest portfolio volatility in comparison with the other covariance-constructing methods. DCC-GARCH gives the largest portfolio volatility. As we are observing the Mean-Variance results, it is now more 'fair' to use the Sharpe Ratio as a performance measure. We observe that ADCC-TGARCH-t gives the largest Sharpe Ratio and the RiskMetrics the lowest Sharpe Ratio. Here again, the volatility timing gains are definitely not due to chance. Now only 5.4% of the 10.000 trials yield a higher Sharpe Ratio for the RiskMetrics approach, 3.9% of the trials yield a higher Sharpe Ratio for the DCC-GARCH approach and finally only 2.6% of the trials yield a higher Sharpe Ratio for the ADCC-TGARCH-t model. In general, the frequencies with which the simulation beat the portfolio according to the MPPM are larger than the frequencies based on the Sharpe Ratio. However, we can still observe that ADCC-GARCH-t provides the lowest frequency with which the simulation beats the portfolio according to the MPPM and the Sharpe Ratio. Based on these results, we can conclude that ADCC-TGARCH-t is the best performing model when looking at the Sharpe Ratio and the MPPM.

5.5 Summary of the Results

Now that we have performed all statistical tests and now that we have evaluated the economic significance of the covariance-constructing methods we will give an overview of the results we have found in this research.

Table 5.9 presents the summary of the results. In this research we have performed both statistical tests and economic significance tests to choose the best performing covariance-constructing model. The statistical tests test the accuracy of the VaR forecasts and hence solely focuses on the Value-at-Risk estimates. According to both tests the best performing model is the ADCC-TGARCH-t model. This model takes the asymmetry between asset classes into account. Obviously, incorporating this asymmetry bears fruit as the statistical tests indicate that the ADCC-TGARCH-t is the best performing model.

When looking at the volatility timing results, we first wanted to know whether it makes sense at all to use covariance-constructing models in the world of portfolio management. This is done by conducting simulations and comparing them with the actual data. The conclusion is that it is highly unlikely that gains from volatility timing are due to chance, because the percentages higher Sharpe Ratio that evolve out of the simulations lie between 2.6% and 8.7% and the percentages higher MPPM lie between 3.5% and 8.9%. As the objective of the minimum-variance asset allocation rule is to minimize risk, we have used the Sample Volatility and the frequency with which the actual data beats the simulation exercise according to the Sample Volatility as a performance measure to evaluate the covariance-constructing models. We observe that RiskMetrics is able to give the lowest sample volatility and ADCC-TGARCH-t gives the largest Sample Volatility. Furthermore, based on the frequency here again RiskMetrics performs best. We have also used the mean-variance asset allocation rule to obtain insights in the economic significance. Here again we have looked what the objective of mean-variance is. As its objective is maximizing the Sharpe Ratio, we will use the latter and the frequency with which the simulation beat the portfolio based on the Sharpe Ratio and the MPPM as evaluation criteria. In this case, the ADCC-TGARCH-t is able to give the largest Sharpe Ratio. Besides, the frequency with which the simulations beat the actual data based on the Sharpe Ratio and the MPPM is lowest for the ADCC-GARCH-t model. Hence

ADCC-TGARCH-t is the best performing model according to this criteria.

Table 5.9: **Summary of the Results**

Performance Measure	Best Performing Model
Statistical Tests	
Backtesting Test	ADCC-GARCH-t & ADCC-TGARCH-t
CPA-test	ADCC-TGARCH-t
Volatility Timing (Economic Significance)	
Minimum Variance	RiskMetrics
Mean Variance	ADCC-TGARCH-t

This table presents the best performing model for all evaluation criteria we have used in this reserach: the backtesting test, the comparative predictive ability test and the volatility timing results.

Chapter 6

Conclusion, Recommendation & Further Research

This paper investigates the forecasting power for the Value-at-Risk of several covariance constructing models. On the one hand we use the RiskMetrics approach. On the other hand we use multivariate GARCH models, such as the DCC and the Asymmetric-DCC model. We compare these models in the context of a portfolio that is used as a benchmark for the Return Seeking Assets of the SSPF fund that SAMCo uses. The models are compared by using the backtesting test developed by Christoffersen (1998) and the Comparative Predictive Ability test developed by Giacomini and White (2006). The results of the backtesting test obtained in this paper indicate that only the ADCC-GARCH-t and the ADCC-TGARCH-t models are able to pass the correct conditional coverage test. Hence, the VaR violations that these models have is significantly not different from the nominal coverage probability. Furthermore, their VaR forecasts outside the interval are spread out over the sample and do not come in clusters. This result is robust as the same result is obtained when using different samples, except for the crisis period where none of the models were able to pass the correct conditional coverage test. In this research we have performed the same test for the sample before the global financial crisis, during the global financial crisis and after the global financial crisis.

The results of the comparative predictive ability test (CPA) in this paper indicate that the RiskMetrics model is outperformed by eight other models and outperforms the DCC-GARCH and the ADCC-GARCH model. Furthermore, we now know that assigning a Student-t distribution to the error terms leads to an improvement of the model. According to the CPA-test the ADCC-TGARCH-t model is able to outperform all other models.

Past literature has proven that volatility models deliver reasonably accurate volatility/covariance forecasts. Besides the statistical evaluation of the covariance models, we also assess the economic significance of time-varying, predictable volatilities/covariances. We perform this analysis on a dataset that consists of three assets: A risk-free rate, excess stock return and excess bond return. The minimum-variance and the mean-variance asset-allocation strategies are used to obtain the portfolio weights, ex-post returns, the Sharpe Ratio and the MPPM for three covariance constructing models: RiskMetrics, DCC-GARCH model and the ADCC-TGARCH-t model. Furthermore to assess the statistical significance of the volatility timing results, we also conduct simulations where the asset returns are generated independently of the portfolio weights as in line with Fleming et al. (2001). The results of this simulation exercise indicate that it is unlikely that the gains to volatility-timing are due to chance.

Our main recommendation for SAMCo will be to take the advantages of other multivariate Value-at-Risk forecasting models into account. RiskMetrics is not able to take the asymmetry in the correlations between the asset returns into account. In this research we have seen that RiskMetrics is outperformed in most cases both for statistical tests and for economic significance tests. The RiskMetrics was only able to perform better when using the minimum-variance asset allocation rule, as the volatility was the lowest. Which we already expected, as we have seen that RiskMetrics underestimates risk in this research. We are aware of the fact that models that perform better than RiskMetrics in this research can be computationally quite intensive, but the benefits are very large. If VaR is forecasted with RiskMetrics, it is likely that the VaR will be underestimated, especially during high volatility periods. This can have tremendous effects for SAMCo, as they will be not aware of the losses they are faced to. They could choose to reestimate the parameters in the models once in a while, as is done in this research. In this way it will not be computationally intensive but the downside is that estimation errors will be increased.

Our work suggests a number of possible directions for future research. For further research we suggest to update the parameters of the DCC-models more often. We have done this on a quarterly basis but this might produce estimation errors and parameter uncertainty. The performance of the DCC models could be better when updating the parameters more often. Hence, the effect of the number of times the parameters are estimated on the results can be investigated for future research. In this research we have used a constant cost per change in a proportion of a security for the economic significance analysis. Another line of research could attempt to consider a variable cost per change of a security. For the volatility timing analysis we have used the minimum-variance and the mean-variance analysis as our framework. For further research we suggest to also use asset allocation rules that take higher moments into consideration by plugging in the skewness and the kurtosis in the objective directly and using a higher order Taylor series approximation. The approach of Guidolin & Timmermann (2006) could be used for this purpose.

Bibliography

Bollerslev, T., (1986). *Generalized autoregressive conditional heteroskedasticity.* Journal of Econometrics 31, 307-327.

Cappiello, L., Engle, R.F. and Sheppard, K. (2006). *Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns.* Journal of Financial Econometrics **4**(4), 537-572

Christoffersen, P. (1998). *Evaluating Interval Forecasts.* International Economic review, **39**(4).

Christoffersen, P., Errunza, V., Jacobs, K. and Langlois, H. (2012). *Is the Potential for International Diversification Disappearing? A Dynamic Copula Approach.* The Review of Financial Studies.

Diebold, F.X., Gunther, T.A. and Tay, A.S. (1998). *Evaluating Density Forecasts with Applications to Financial Risk Management.* International Economic Review. **39**(4), 863-883.

Engle, R. (2002). *Dynamic Conditional Correlation.* Journal of Business & Economic Statistics. **20**(3), 339-350.

Flemming, J., Kirby, C. and Ostdiek, B. (2001). *The Economic Value of Volatility Timing.* Journal of Finance **57**(1), 329-352.

Giacomini, R. and White, H. (2006). *Tests of Conditional Predictive Ability.* Econometrica. **74**(6), 1545-1578.

Goetzmann, W.N., Li, L. and Rouwenhorst, K.G. (2004). *Long-Term Global Market Correlations*, working paper, Yale School of Management and Yale university.

Goetzmann, W.N., Ingersoll, J. and Spiegel M. (2007). *Portfolio Performance Manipulation and Manipulation-proof Performance Measures.* The Review of Financial Studies. **20**(5), 1503-1546.

Hong, Y., Tu, J. and Zhou, G. (2007). *Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation.* The Review of Financial Studies. **20**(5), 1547-1581.

J.P. Morgan. (1996). *RiskMetrics.* Technical Document, 4th edition.

Kritzman, M., Li, Y., Page, S. and Rigobon, R. (2010). *Principal Components as a Measure of Systemic Risk.* working paper 4785-10, MIT Sloan School of Management.

Santos, A.A.P., Nogales and F.J., Ruiz, E. (2013). *Comparing univariate and multivariate models to forecast porfolio Value-at-Risk.* Journal of Financial Econometrics. **11**(2), 400-441.

Solnik, B. and Roulet, J. (2000). *Dispersion as Cross-Sectional Correlation.* Financial Analysts Journal. **56**(1), 54-61.

Wallis, K.F. (2002), *Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts.* International Journal of Forecasting. **19**, 165-175.

Yoshimoto, A. (1996). *The Mean-Variance Approach to Portfolio Optimization subject to Transaction Costs.* Journal of the Operations Research. **39**(1), 99-117.

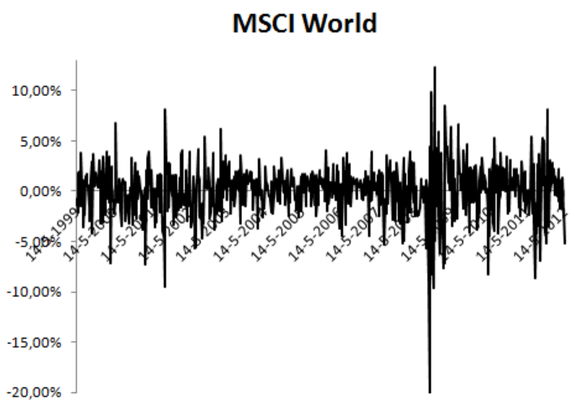
Appendix A

Appendix

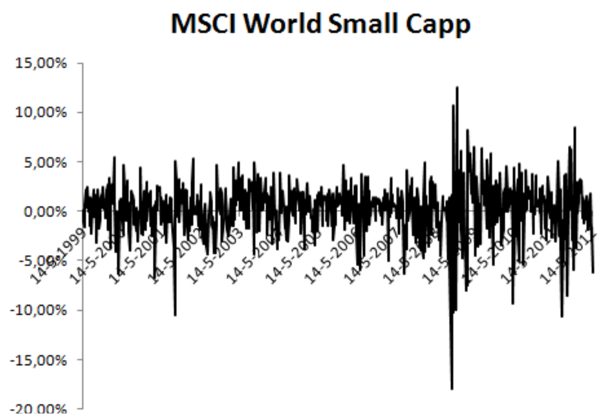
A.1 Time series of the returns

Figure A.1.1: The time series for the returns of the indices.

(a) MSCI World



(b) MSCI World Small Capp



(c) MSCI Emerging Markets

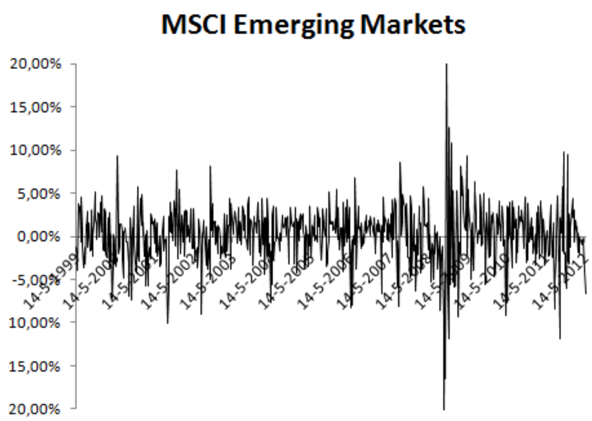
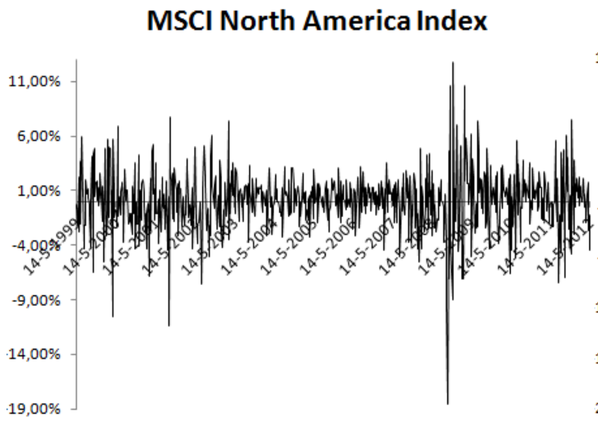
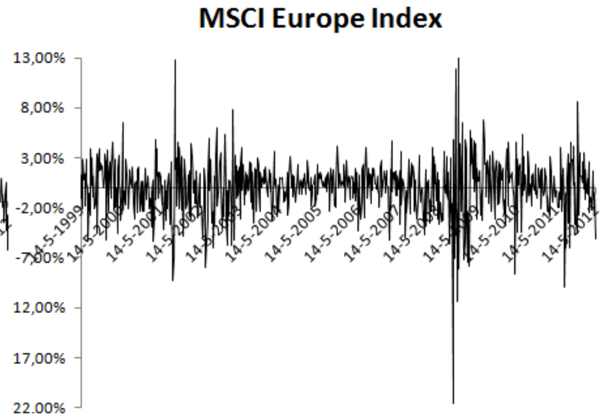


Figure A.1.2: The time series for the returns of the indices.

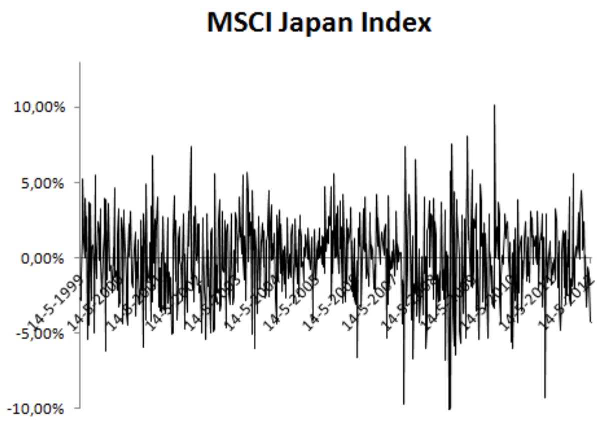
(a) MSCI North America



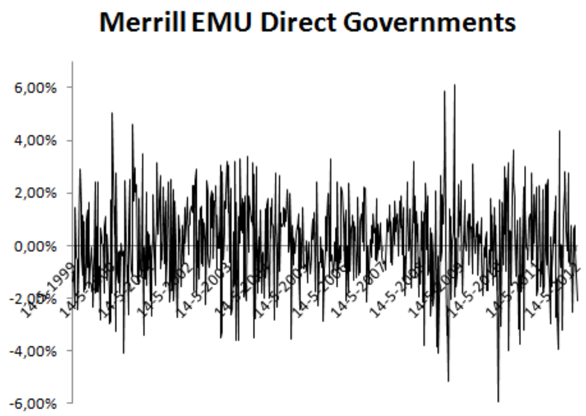
(b) MSCI Europe



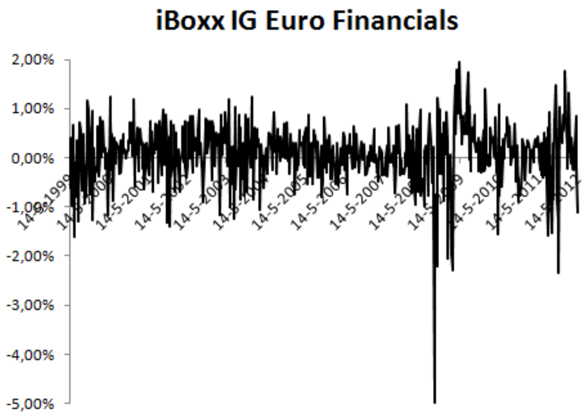
(c) MSCI Japan



(d) Merrill EMU Direct Governments



(e) iBoxx IG Euro Financials



(f) iBoxx IG Euro Non-Financials

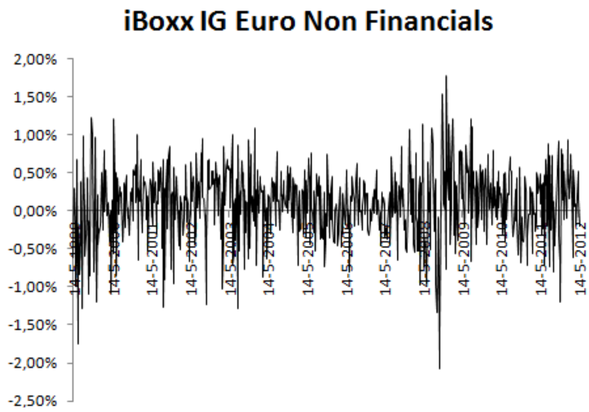
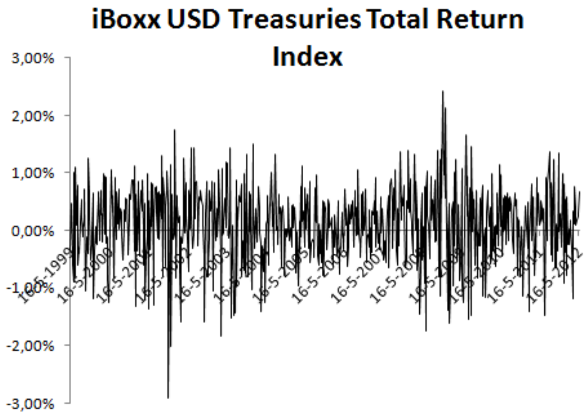
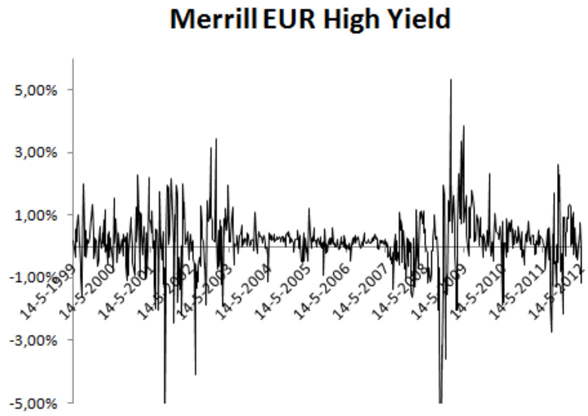


Figure A.1.3: The time series for the returns of the indices.

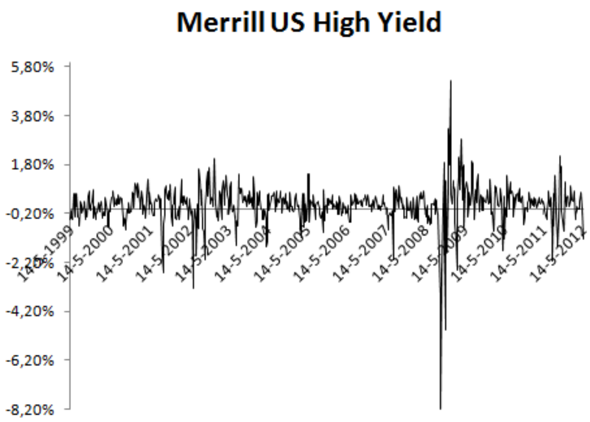
(a) iBoxx USD Treasuries Total Return Index



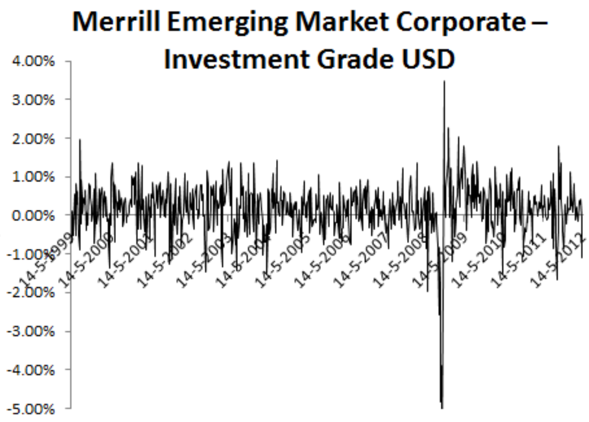
(b) Merill EUR High Yield



(c) Merrill US High Yield



(d) Merill Emerging Market Corporate



A.2 Table of the Covariance Matrix of the Return Series

Table A.2.1: Covariance Matrix

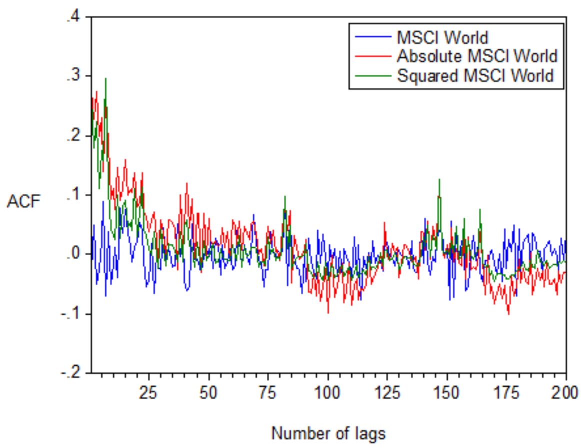
This table presents the covariances between the indices, the diagonal of the matrix shows the variances of the indices.

	MW	MSC	MEM	MNA	ME	MJ	MDG	IBF	IBNF	IBT	MHY	MUHY	MEMG
MW	6.82E-04	6.73E-04	7.10E-04	6.85E-04	6.63E-04	4.88E-04	8.60E-05	1.85E-05	-4.27E-06	-5.21E-05	1.01E-04	9.75E-05	2.79E-05
MSC	-	7.68E-04	7.77E-04	6.57E-04	6.38E-04	5.31E-04	1.02E-04	2.10E-05	-3.32E-06	-5.71E-05	1.19E-04	1.16E-04	3.60E-05
MEM	-	-	1.12E-03	6.58E-04	6.94E-04	6.12E-04	1.23E-04	2.68E-05	4.20E-06	-5.84E-05	1.38E-04	1.36E-04	5.95E-05
MNA	-	-	-	7.50E-04	6.52E-04	4.24E-04	2.71E-05	1.19E-05	-7.38E-06	-5.66E-05	9.40E-05	8.87E-05	1.34E-05
ME	-	-	-	-	7.99E-04	4.75E-04	-3.69E-05	9.97E-06	-1.14E-05	-6.94E-05	1.02E-04	9.19E-05	1.03E-06
MJ	-	-	-	-	-	8.02E-04	1.55E-05	1.69E-05	-4.75E-06	-4.62E-05	9.15E-05	8.76E-05	2.28E-05
MDG	-	-	-	-	-	-	2.62E-04	3.68E-05	2.65E-05	3.49E-05	1.23E-05	2.62E-05	6.21E-05
IBF	-	-	-	-	-	-	-	3.58E-05	2.19E-05	1.67E-05	2.24E-05	2.22E-05	2.43E-05
IBNF	-	-	-	-	-	-	-	-	2.21E-05	1.79E-05	1.52E-05	1.57E-05	1.92E-05
IBT	-	-	-	-	-	-	-	-	-	4.17E-05	-7.87E-06	2.93E-06	2.39E-05
MHY	-	-	-	-	-	-	-	-	-	-	8.68E-05	5.73E-05	2.66E-05
MUHY	-	-	-	-	-	-	-	-	-	-	-	6.30E-05	3.16E-05
MEMG	-	-	-	-	-	-	-	-	-	-	-	-	4.81E-05

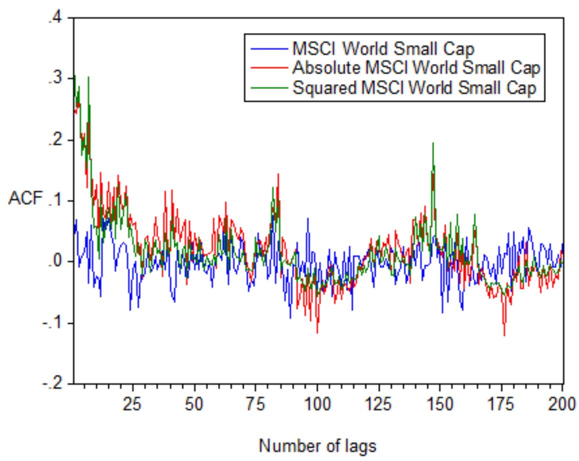
A.3 Empirical autocorrelation functions of the return series

Figure A.3.1: The autocorrelations of the returns, the absolute returns and the squared returns.

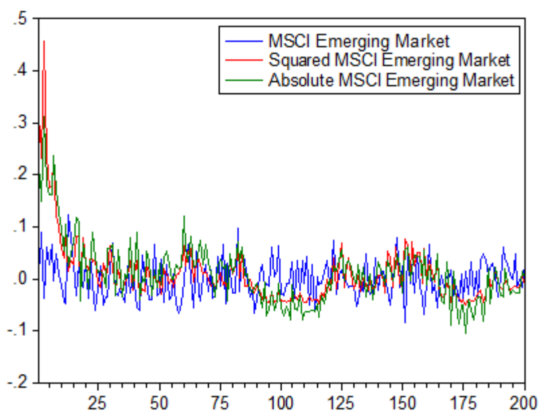
(a) MSCI World



(b) MSCI World Small Capp



(c) MSCI Emerging Markets



(d) MSCI World Minimum Volatility Index

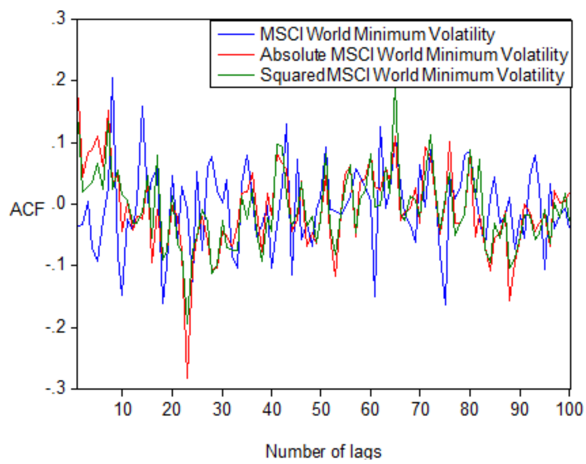
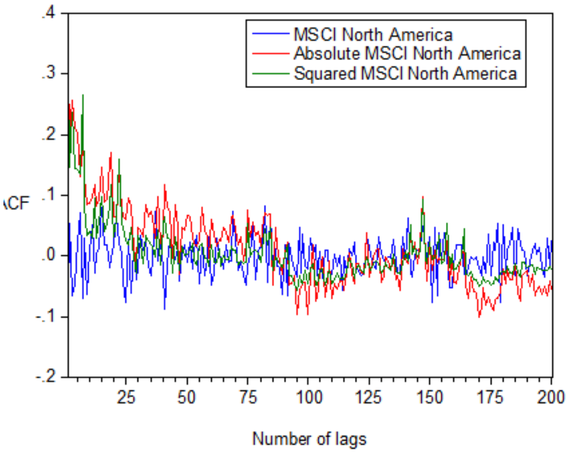
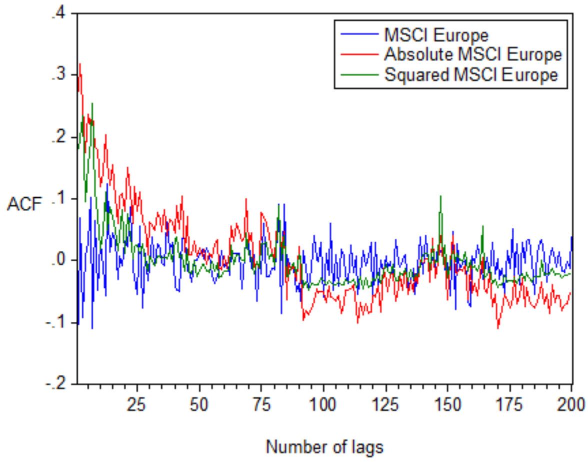


Figure A.3.2: The autocorrelations of the returns, the absolute returns and the squared returns.

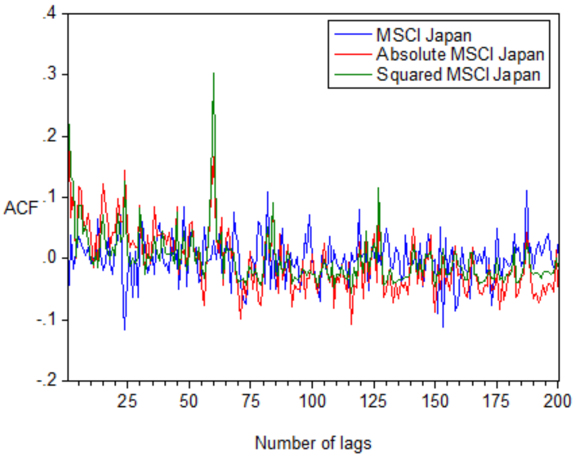
(a) MSCI North America



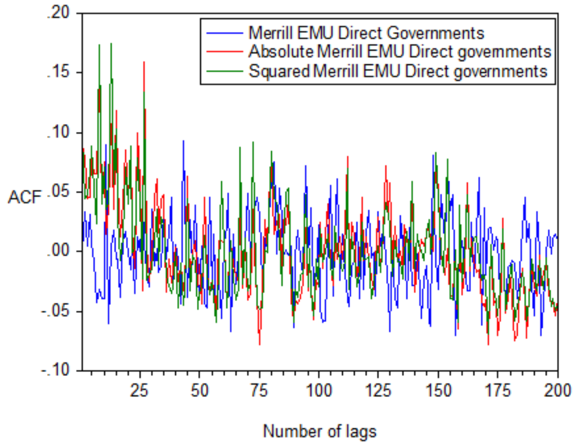
(b) MSCI Europe



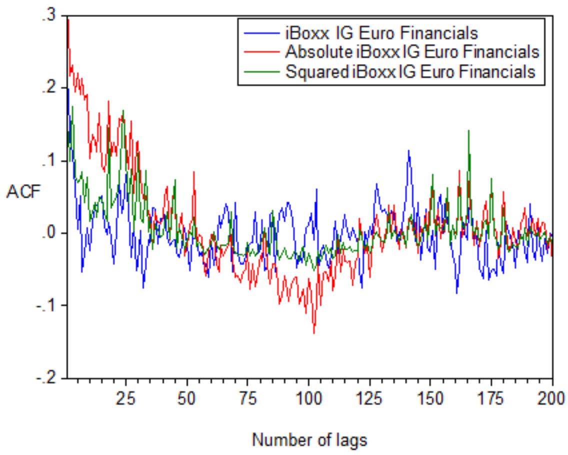
(c) MSCI Japan



(d) Merrill EMU Direct Governments



(e) iBoxx IG Euro Financials



(f) iBoxx IG Euro Non-Financials

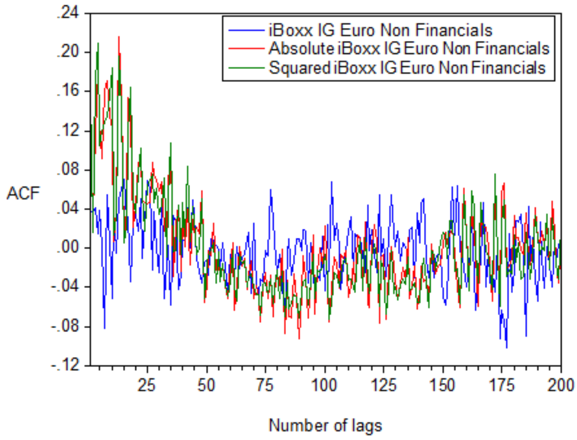
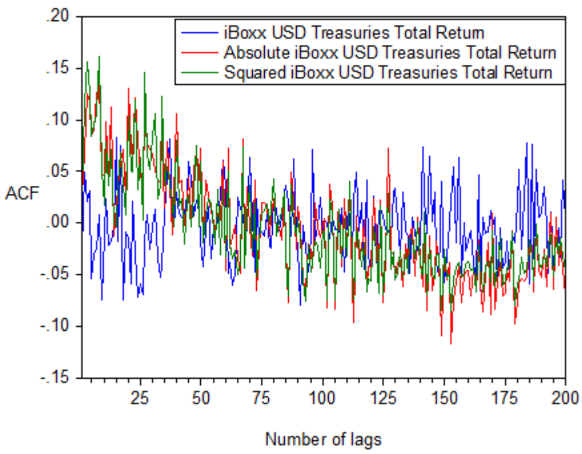
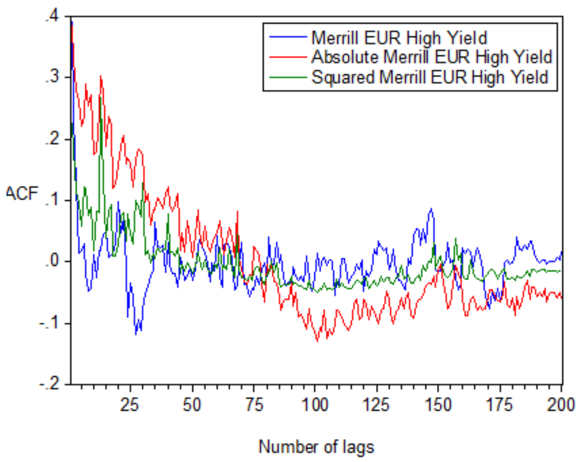


Figure A.3.3: The autocorrelations of the returns, the absolute returns and the squared returns.

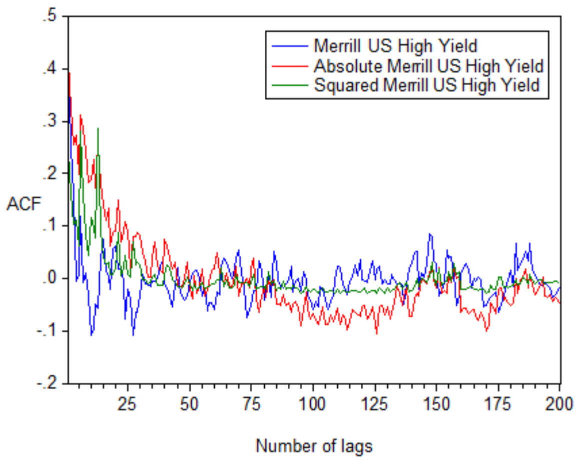
(a) iBoxx USD Treasuries Total Return Index



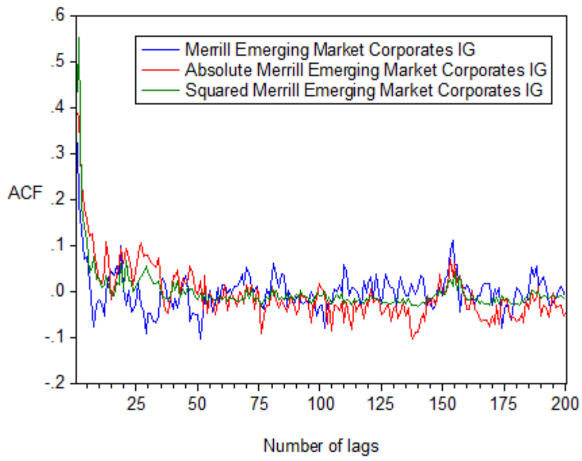
(b) Merill EUR High Yield



(c) Merrill US High Yield



(d) Merill Emerging Market Corporate



A.4 Derivation of $\hat{\Omega}$

To construct a feasible test statistic to test symmetry in the correlations we have to estimate the positive-definite variance-covariance matrix Ω for all possible true distributions of the data satisfying some regularity conditions. The sample means and variances of the two conditional threshold correlation series are computed as follows:

$$\hat{\mu}_1^+(\gamma) = \frac{1}{T_c^+}\sum_{t=1}^{T} R_{1t}(R_{1t} > \gamma, R_{2t} > \gamma), \tag{A.1}$$

$$\hat{\mu}_2^+(\gamma) = \frac{1}{T_c^+}\sum_{t=1}^{T} R_{2t}(R_{2t} > \gamma, R_{2t} > \gamma), \tag{A.2}$$

$$\hat{\sigma}_1^+(\gamma)^2 = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T}[R_{1t} - \hat{\mu}_1^+(\gamma)]^2 1(R_{1t} > \gamma, R_{2t} > \gamma) \tag{A.3}$$

$$\hat{\sigma}_2^+(\gamma)^2 = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T}[R_{2t} - \hat{\mu}_2^+(\gamma)]^2 1(R_{1t} > \gamma, R_{2t} > \gamma) \tag{A.4}$$

where $1(.)$ is the indicator function. The sample conditional correlation $\hat{\rho}^+(\gamma)$ can be expressed as follows:

$$\hat{\rho}^+(\gamma) = \frac{1}{T_c^+ - 1}\sum_{t=1}^{T} \hat{X}_{1t}^+(\gamma) - \hat{X}_{2t}^+(\gamma) 1(R_{1t} > \gamma, R_{2t} > \gamma) \tag{A.5}$$

where

$$\hat{X}_{1t}^+(\gamma) = \frac{R_{1t} - \hat{\mu}_1^+(\gamma)}{\hat{\sigma}_1^+(\gamma)} \tag{A.6}$$

and

$$\hat{X}_{2t}^+(\gamma) = \frac{R_{2t} - \hat{\mu}_2^+(\gamma)}{\hat{\sigma}_2^+(\gamma)} \tag{A.7}$$

we have a similar expression for $\hat{\rho}^-(\gamma)$. A consistent estimator of Ω is given by the following almost positive definite matrix:

$$\hat{\Omega} = \sum_{l=1-T}^{T-1} k(l/p)\hat{\gamma}_l \tag{A.8}$$

where $\hat{\gamma}_l$ is an $N \times N$ matrix with (i, j)-th element

$$\hat{\gamma}_l(\gamma_i, \gamma_j) = \frac{1}{T} \sum_{t=|l|+1}^{T} \hat{\xi}_t(\gamma_i)\hat{\xi}_{t-|l|}(\gamma_j) \tag{A.9}$$

and

$$\hat{\xi}_t(\gamma) = \frac{T}{T_c^+}[\hat{X}_{1t}^+(\gamma)\hat{X}_{2t}^+(\gamma) - \hat{\rho}^+(\gamma)]1(R_{1t} > \gamma, R_{2t} > \gamma) - \frac{T}{T_c^-}[\hat{X}_{1t}^-(\gamma)\hat{X}_{2t}^-(\gamma) - \hat{\rho}^-(\gamma)]1(R_{1t} < -\gamma, R_{2t} < -\gamma) \tag{A.10}$$

where $k(.)$ is a kernel function that assigns a suitable weight to each lag of order l and p is the smoothing parameter or lag truncation order. We will use the Bartlett kernel:

$$k(z) = (1 - |z|)1(|z| < 1) \tag{A.11}$$

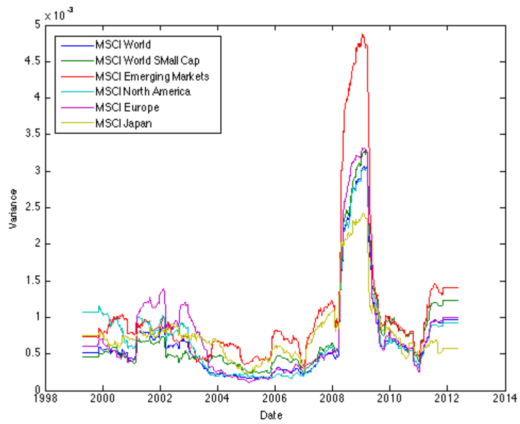
With these formulas we can define the test statistic as follows:

$$J_p = T(\hat{\rho}^+ - \hat{\rho}^-)'\hat{\Omega}^{-1}(\hat{\rho}^+ - \hat{\rho}^-) \tag{A.12}$$

A.5 Time-varying Variance and Correlations

Figure A.5.1: Time-varying variances of the Equity and Fixed Income return series based on a 52-week moving window.

(a) Equity



(b) Fixed Income

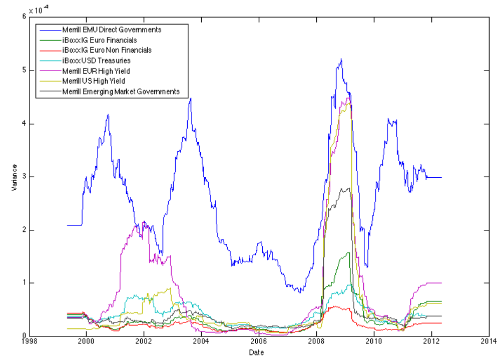
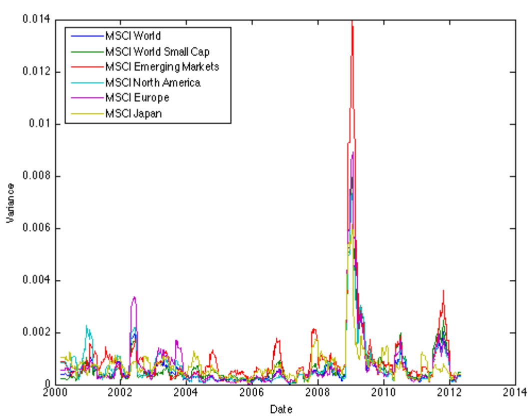


Figure A.5.2: Time-varying variances of the Equity and Fixed Income return series based on a 12-week moving window.

(a) Equity



(b) Fixed Income

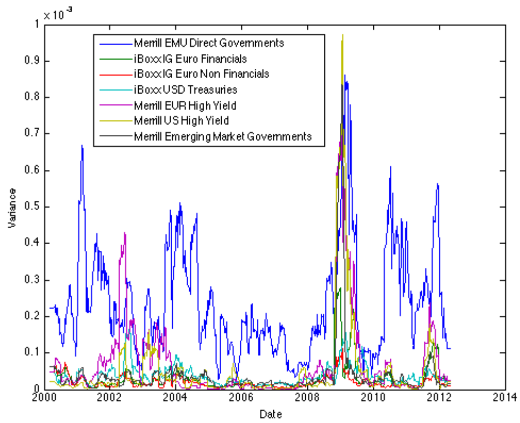
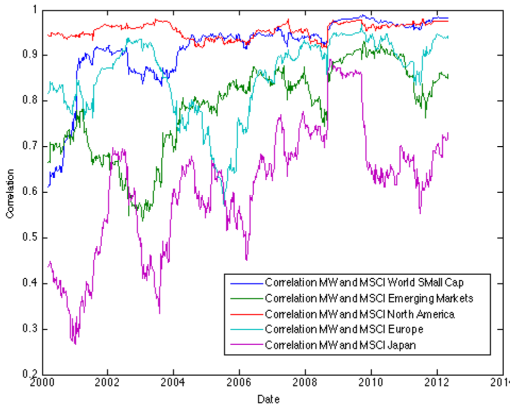
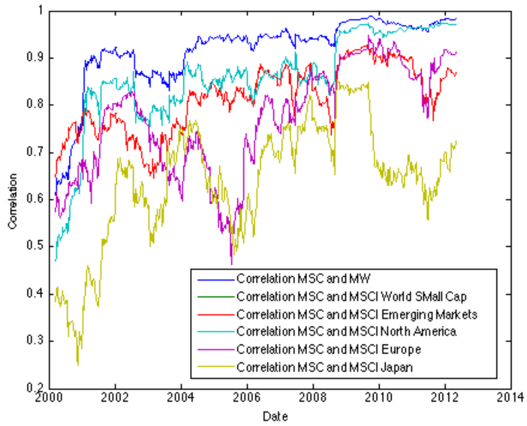


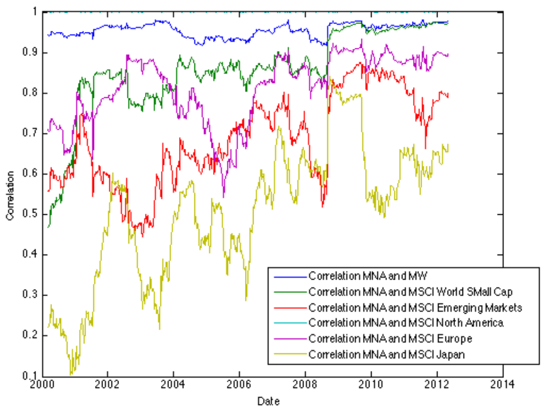
Figure A.5.3: Time-varying correlation of the Equity return series based on a 52-week moving window.



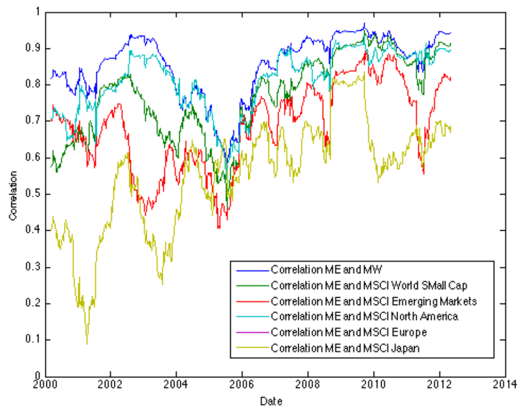
(a) Correlation between MW and the rest of the Equity series



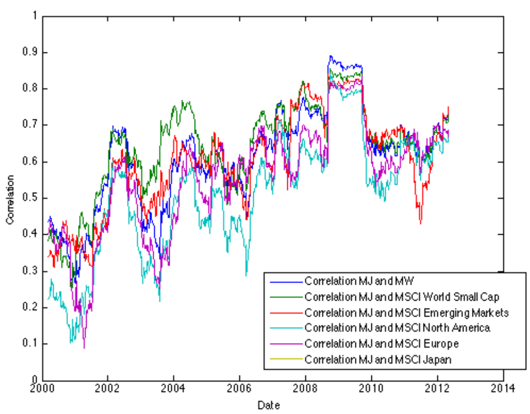
(b) Correlation between MSC and the rest of the Equity series



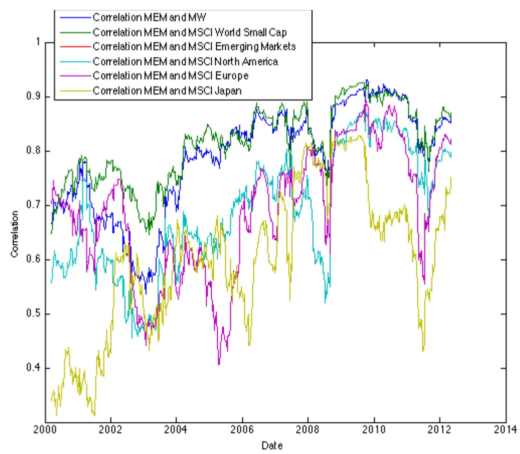
(c) Correlation between MNA and the rest of the Equity series



(d) Correlation between ME and the rest of the Equity series



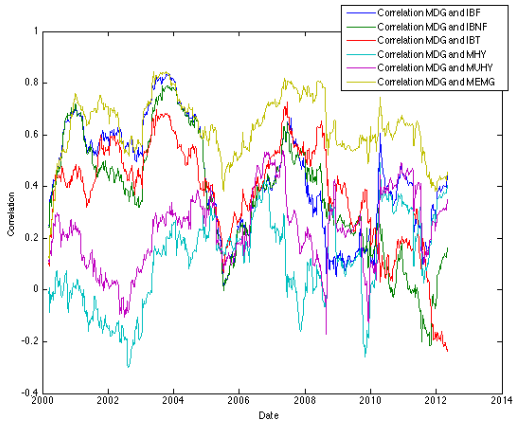
(e) Correlation between MJ and the rest of the Equity series



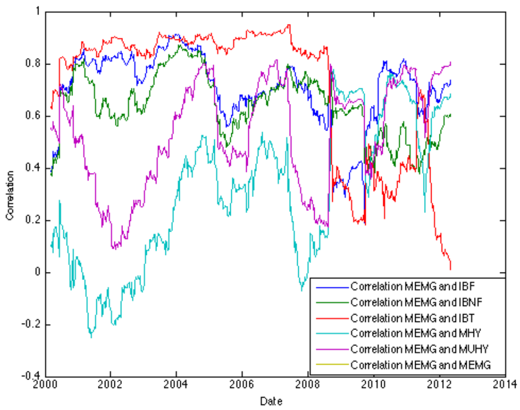
(f) Correlation MEM and the rest of the Equity series

Figure A.5.4: Time-varying correlation of the Fixed Income return series based on a 52-week moving window.

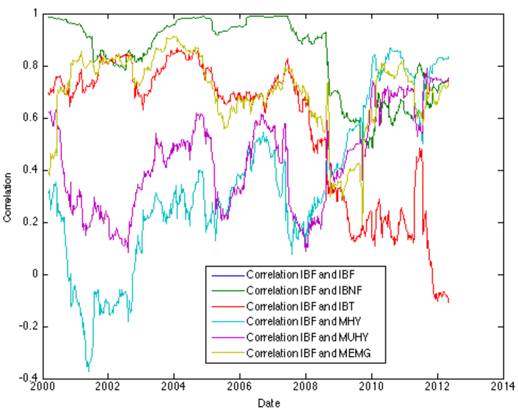
(a) Correlation between MDG and the rest of the Fixed Income series



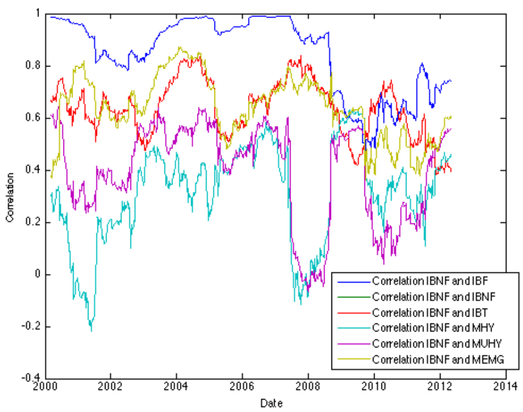
(b) Correlation between MEMG and the rest of the Fixed Income series



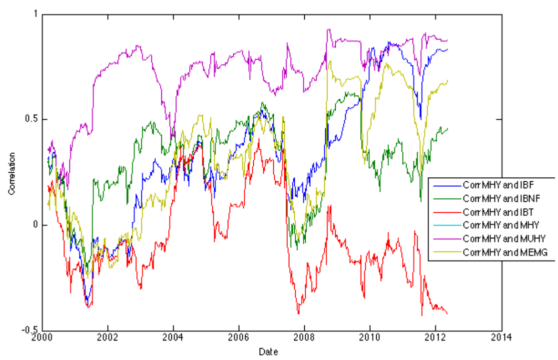
(c) Correlation between IBF and the rest of the Fixed Income series



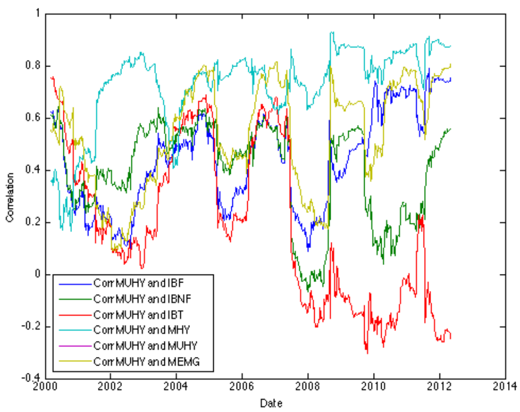
(d) Correlation between IBNF and the rest of the Fixed Income series



(e) Correlation between MHY and the rest of the Fixed Income series



(f) Correlation MUHY and the rest of the Fixed Income series



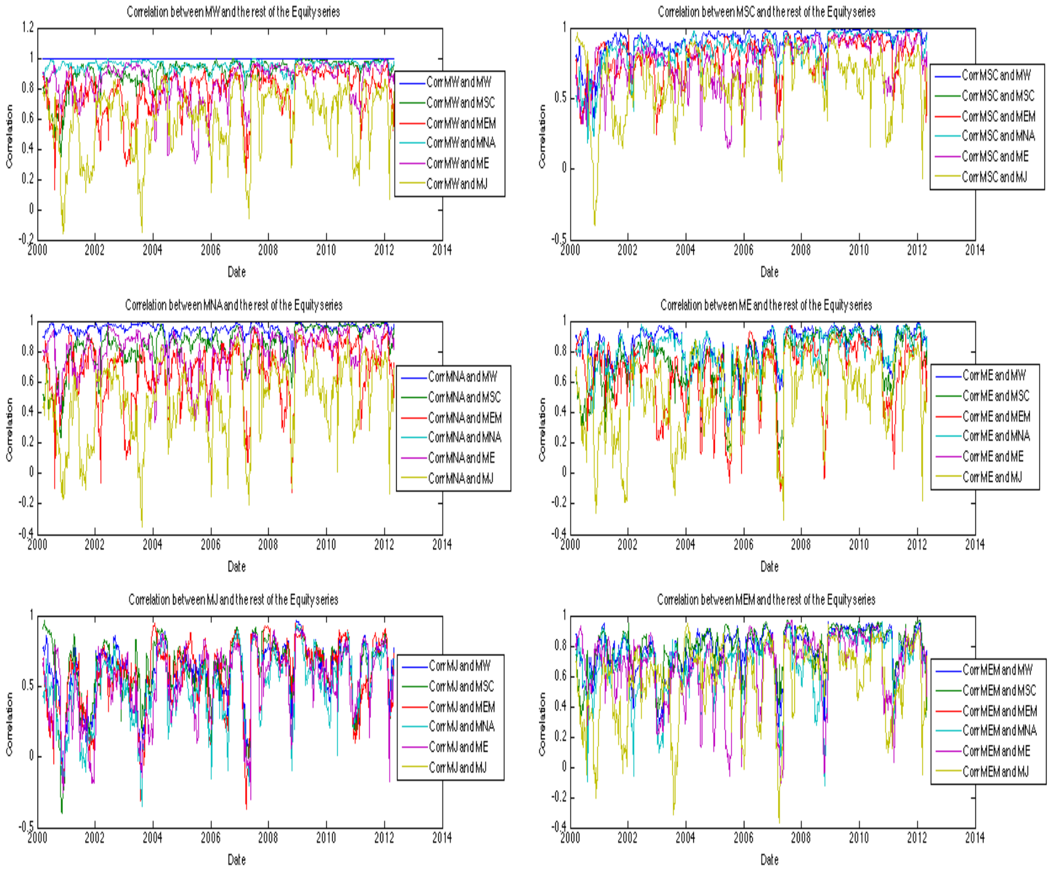


Figure A.5.5: Time-varying correlations of the Equity Series based on a 12-week moving window (quarterly)

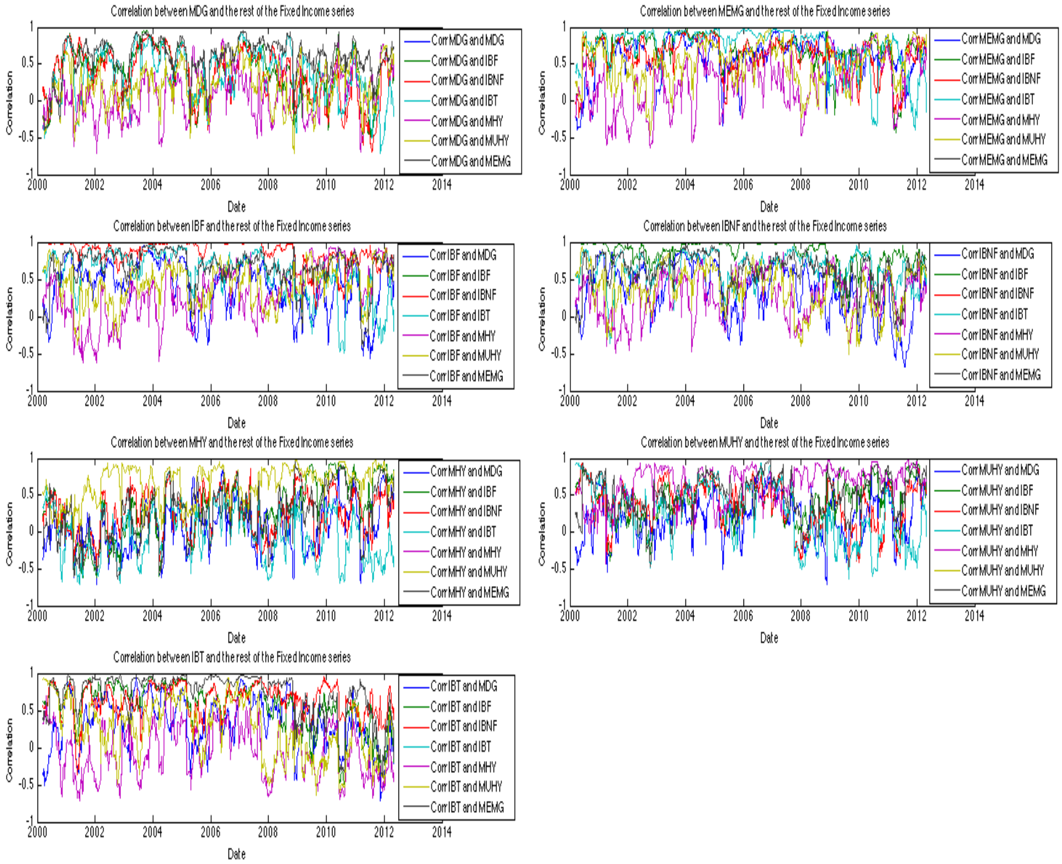


Figure A.5.6: Time-varying correlations of the Fixed Income Series based on a 12-week moving window (quarterly)