DEVELOPING A MODEL FOR DETERMINING AN OPTIMAL ASSET ALLOCATION

USING A TIME-VARIANT OPTIMIZATION FRAMEWORK FOR ESTIMATING FUTURE RETURNS, VOLATILITIES AND CORRELATIONS

MASTER THESIS

SUPERVISED BY

DR. A. MARKIEWICZ

EDWIN HAUWERT

August 2011

Abstract

During the last few years many inventions have been made in the area of financial economics, during the time of Fisher, 100 years ago, there were no appropriate quantitative methods available to estimate dynamic conditional variances and correlations. These methods have been developed in the last 20 years. In this research paper we combined some of the important advances. Many of the firms active in the financial sector still make use of the MPT framework as developed by Markowitz, which imply that they all make use of linear optimization techniques. These firms differ in the method in which they estimate or gather these inputs. Goal of this research was to develop a model and especially one which delivers an optimal asset allocation. This optimal asset allocation is defined in portfolio theory as the portfolio which delivers the best return for each unit of risk or has the lowest amount of risk relative to return. This portfolio is known as the tangency portfolio. This model improves the average Sharpe ratio over time up to 50% on a 2 year time frame, but the outperformance seems time-invariant. Our results with the 2 year and 10 year GARCH-DCC MPT framework are convenient, but it is questionable if the benefits out weights the costs of complexity. We must also see the fact that a 1-month switch of 50 percent of the portfolio from corporate bonds to other assets classes requires an extreme amount of flexibility. And let's not forget that we did not take into account the costs of switching, we would incur up till 0.5% in costs and possibly costs resulting from illiquidity. The results are too weak to claim that the GARCH-DCC framework is considerably stronger than using historical covariance matrix. Improvement of the return expectations could do the trick. This thesis shows a framework which is simple, easy to adjust and which can be tailored to your preferences.

MASTER FINANCIAL ECONOMICS
ERASMUS SCHOOL OF ECONOMICS

ERASMUS UNIVERSITY OF ROTTERDAM

1 Introduction

"To stop short in any research which bids fair to widen the gates of knowledge, to recoil from fear of difficulties or adverse criticism, is to bring reproach upon science."
Sir William Crookes, Chemist and Physicist (1832-1919)

One of the most important problems in asset management is to make decisions concerning the optimal asset allocation (Sharpe, 1990). The goal of asset managers is to realize the highest possible return corrected for risk, with the lowest value at risk (VaR) possible. The prices of different asset classes change in response to news in anticipation of future performance. In portfolio choice and in determining aggregate risk, the variance and correlation structures across assets are extremely important. This paper is to explore possible methods and to determine the quantitative elements to use for developing a model, which output is the best possible asset allocation.

Kaplan and Garick (1980) wrote about the quantitative definition of risk, which is used to discuss notions of "relative risk", "relativity of risk" and "acceptability of risk", this to reach agreement among academics. Physics has many linkages with economics since phenomenon in behavior often occur in matter, like the Hermetic saying "as above so below". Against the earlier materialist definition of economic science, Robbins (1932) propounded the, now well-known, scarcity definition: "Economics is the science which studies human behavior as a relationship between ends and scarce means which have alternative uses". He would later show to be an important colleague of Hicks, one of the inspirators of Markowitz. "The brain at rest produces random activity," Cowan said, or what physicists call "Brownian motion." "It so happens that there is an analogy between the behavior of chemical reaction networks and neural networks." (Cowan, 2008). Cowan is one of the scientist who uses applied physical knowledge to explain phenomena in other fields of science. He is among many scientists drawing similarities between different fields of science for better understanding of phenomena. Here I would like to refer to one of the earliest researchers of financial markets, W.D. Gann (1909), who wrote books about the behavior of markets and registered many panics and price patterns of the past and produced forecasts between 1909 and 1955. He wrote "I soon began to note the periodical recurrence of the rise and fall in stocks and commodities. This led me to conclude that natural law was the basis of market movements. I then decided to devote ten years of my life to the study of natural law as applicable to the speculative markets and to devote my best energies toward making speculation a profitable profession." Economist Fisher wrote his first book before 1907, here he makes notion of chance and uncertainty in his analysis:
"Up to this point we have ignored the element of chance, by assuming that the entire future income-stream, or at any rate, such portions of it as needed to influence present choice, are foreknown and mapped out in advance . . . This assumption, like the assumption that bodies fall in vacuo, in the ordinary presentation of the theory of gravitation, has enabled us to complete our formal statement of the theory more easily, although at the expenses of exact conformity to actual historical fact; for, in the concrete world, the most conspicuous characteristic of the future is its uncertainty" (Fisher, 1907).
Uncertainty about future asset prices or risk is often measured by the average deviation from the mean, Fisher was also one of the first economists to note this measurement of risk.

Risk and the reward for bearing risk or the risk premium are important concepts in investment theory. A goal of portfolio management is to produce efficient portfolios in which the optimal amount of premium for an amount of risk is chosen or vice versa the optimal risk for a required premium. The premium for risk is often seen as the explanation for return in asset pricing theory(Scholes, 2005). One potential source of misspecification of existing models is that the structural form of conditional means and variances is relatively inflexible and is held fixed throughout the entire sample period. These models are single-regime models in the sense that they effectively make the assumption that a single structure for the conditional mean and variance is present (Gray, 1996). In this research the focus is on the structure of the mean, variance and correlation coefficients which are used as the only three inputs in the Mean-Variance framework, which is used to determine an optimal and efficient allocation of resources. This framework is agreed upon by the Nobel laureates of Economics by rewarding Markowitz, Miller and Sharpe with the Nobel prize in 1990. Markowitz developed a theory of portfolio decisions of households and firms under conditions of uncertainty and described how the problem of investing in a large number of assets, each with different characteristics, may be reduced to the issue of a trade-off between expected return and the variance of the return of the portfolio. He also showed to solve this problem as a quadric programming problem. He won the price together with Sharpe, who was rewarded for the contribution to the understanding of how these asset prices are determined and with Miller, who together with earlier Nobel prize winner Modigliani (1985) founded the modern theory of corporate finance. Partly continuing on the work of Sharpe, Scholes and Merton won a Nobel prize in 1997 for their contribution to the understanding of asset pricing and especially their method of developing a formula for derivative pricing. Their interesting and controversial finding was that in pricing derivatives, no relative risk premium is used in the evaluation. One only needs to have the correct estimate for volatility. In my days at the Amsterdam derivative trading floor Beursplein 5, where I stayed during one of my internships, I learned that there is only one variable important in the pricing mechanism which was continually adjusted by computer inputs and especially by expert human input. This expert human input largely depends on implied volatility of the underlying security and market, the social interaction among those experts and the news or information developments. Scholes (2005) describes in his book 'Asset Pricing', which book won the Paul A. Samuelson award in 2001 for scholarly writing, that no problems are solved by the pure extremes. The CAPM and its successor factor models are paradigms of the absolute approach. In applications, they price assets "relative" to the market or other risk factors, without answering what determines the market or factor risk premia and betas. On the other end of the spectrum, even the most practical financial engineering questions usually involve assumptions beyond pure lack of arbitrage, assumptions about equilibrium "market prices of risk". And so Scholes determines the central and unfinished task of absolute asset pricing is to understand and measure the sources of aggregate or macroeconomic risk that drive asset prices. He points at the fact that expected returns vary across time and across assets in ways that are linked to macroeconomic variables, or variables that also forecast macroeconomic events. "A wide class of models suggests that a 'recession' or 'financial distress' factor lies behind many asset prices". At the time he wrote his book he explained that theory lags behind and there is no model which explains these interesting correlations. Standard macroeconomic models predict that people really do not care about business cycles (Lucas, 1987). Scholes view is that asset prices reveal they do, that asset prices correct for risk or forego risk premia to avoid assets that fall in recessions, this tells us something about the magnitude of environmental changes on the price of assets. This is the reason why we include a regime-switching model in this research to include environmental economic information like production growth(GDP) and price change(Inflation).

"Today in modern society, every day of our lives the media informs us with numbers on key economic variables: inflation, unemployment, interest rates, stock prices and much more. Most readers or viewers regard this information merely as quotations, more or less similar to weather reports. But researchers regard them as important integral parts of a massive flow of data which, over time, gives rise to economic time series. Such time series are important, not only for basic researchers who develop and test economic theory, but also for practitioners who require clear understanding and reliable forecasts as a basis for public policy or private decisions."

This was the introduction of the Nobel prize Award Ceremony Speech of 2003, when Engle and Granger won the Nobel prize for Economics. A prerequisite for capturing the special features of economic time series are statistical methods developed in the borderland between statistics and economics. Granger solved the difficult problem concerning the complex interplay among macroeconomic variables over time. It is difficult to distinguish adjustments towards a long-run relationship from short-run fluctuations in data. As with most macroeconomic variables, financial time series do not fluctuate around given values over time, but around stochastic trends. When traditional statistical methods are applied to such nonstationary time-series, the resulting relationships are often misleading. During the same period Engle, inspired by Laureates Friedman and Lucas, worked on uncertainty, volatility, which could shape economic relationships. Engle interpreted volatility as the size of the random term in a statistical model, and devised the so-called AutoRegressive Conditional Heteroskedasticity (ARCH) method to trace systematic variations in volatility over time. Application of his theory is especially conducted in the field of financial economics. Investors who choose between stocks and bonds and banks which want to limit the risk of large capital losses, all need proper measurement and forecasts on the riskiness of returns or their volatility. Engle was also one of the first to come up with a practical solution for multivariate correlation analysis. Engle modeled the behavior of volatility and correlation and paved the way for the new field of financial econometrics.

An extension of the Asset Liability Management (ALM) theory, heavily used in the pension sector and the asset management industry today, is the use of the research of Steehouwer (2005), who wrote the book 'Macroeconomic scenarios and reality' in which he uses a frequency domain approach for analyzing historical time series in generating scenarios for the future and warns that simply generating an average return value with "some volatility" around it is not enough. He uses advanced techniques beyond the scope of this thesis to develop a set of economic scenarios, which incorporate the non-linear properties of the data. The features the returns, volatilities and correlation are extremely important in the problem of finding an optimal asset allocation.

1.1 Asset allocation problem

In portfolio choice and in determining aggregate risk, the variance and correlation structures across assets are extremely important. It goes beyond this thesis to fully explain the causes of risk and to elaborate on the changes in valuations of specific risks which aggregate in portfolio theory. When academic students get assigned to determine the optimal asset allocation for a one million euro portfolio limited by stocks and bonds, and are only able to take long positions, they start with Modern Portfolio Theory (MPT). This is the framework they choose for optimizing a selection of stocks and bonds and from there they develop their strategies. Only some are capable of making the models estimated variance time-variant by using Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) series. The computational complexity of the model and time constraint of the students soon starts to limit the opportunities and only the most sophisticated future portfolio managers succeed in progressing from there. Stylized anomalies or predictive models are applied to select up till 100 assets used to construct a portfolio by using the mean-variance framework. Since these students will represent an important part of the financial sector of the future it is likely that many financial firms are encountering the same difficulties in constructing their portfolios and so there is a need for practical translation of complex theory into useful models, which incorporate the advanced features of the financial time-series.

The focus of this thesis is on efficient asset allocation, by using better inputs in the mean-variance framework one will reduce the realized variance by increasing the reliability of the diversification effects and will increase its return by using the knowledge of which economic conditions and so risks to expect. Key in this approach are the predictions for return, variance and correlation coefficients by use of the historical price series. In addition to the methods currently used state-of-the-art methods, like the use of GARCH in combination with Dynamic Conditional Correlation (DCC), for portfolio optimization are developed. DCC has the flexibility of univariate GARCH but not the complexity of conventional multivariate GARCH. An important facet of this paper is to reduce the complexity of the model, without losing too much performance and its practical application. Often academic models suffer of parsimony; the amounts of parameters that need to be estimated are large and take more than hours to calculate. The most important question that I will try to answer is if it possible to develop a model which outperforms an asset selection framework with naïve or historical variances and returns and which is not too complex in terms of understandability and parsimony? Secondly It is interesting to answer the question if the model can be improved with a factor model. Besides creating a model that uses univariate GARCH and DCC, it is valuable to determine quantitative macroeconomic factors to use for an estimation or forecast model, which output is the best possible mean input for the Mean-Variance(MV) optimization process. This factor model makes use of macro-economic factors to determine regimes or states of the economy. The Regime-Switching factor model will be used to forecast the means while the DCC-GARCH model will be used to forecast the variances. The model will be assessed with performance test statistics like the Sharpe-ratio and forecasting performance statistics like Root Mean Squared Error, and the Diebold and Mariano statistic. By assessing the model in terms of forecasting errors, Sharpe-ratios the goal of this research is to improve and select the best possible model.

The paper is organized as follows. The second section outlines the methodology in detail and discusses the estimation procedures used. Section three elaborates on some of the assumptions made. Section four is about the data and software used to apply the research methodology. Section five focuses on the model itself and explains the model specifications used for generating testing data. The sixth section shows the results. Concluding remarks are in the final section.

1.2 Literature overview

1.2.1 Developments in portfolio theory

Considering portfolio management we have to deal with risks, premiums for bearing those risks and the risk reduction effect on the portfolio from diversification. We can go back far in history to find the first signs of the notion of diversification in investments. Markowitz (1999) gives his research of the past, he found the following in the Merchant of Venice(1597), Shakespeare wrote

"My ventures are not in one bottom rusted, Nor to one place; nor is my whole estate Upon the fortune of this present year; Therefore, my merchandise makes me not sad" Act I, Scene I

One of the earliest written notions of diversification and intuition of covariance. It would take 300 years before mathematical theory was developed and another 50 years to link this to the behavior of risk and its risk premia and to quantify the behavior of the combined risk premia for portfolios.

In 1999 Markowitz wrote about the 'early' history of portfolio theory. He was inspired by the work of Hicks (1935), Nobel prize winner with Arrow in 1972, who first wrote about value in 1934 and later suggested a simple theory of money. He was one of the first to introduce risk in his quantitative analysis. He noted "The risk-factor comes into our problem in two ways: First, as affecting the expected period of investment, and second as affecting the expected net yield of the investment". He described the effect of risk on yield: "an increase in the risk of investment will act like a fall in the expected rate of net yield; an increase in the uncertainty of future out payments will act like a shortening of the time which is expected to elapse before those out-payments; and all will ordinarily tend to increase the demand for money." Hicks did not mention the standard deviation as measure for risk, he called it 'some appropriate measure of dispersion'. He only classified the risk in two classes, namely time and yield. Hicks also described diversification. Sharpe who won the Nobel price together with Markowitz worked out the Sharpe ratio. The Sharpe ratio was created to answer the question "Given the same amount of risk, which investment provides me with the highest reward." To do this the Sharpe ratio balances the returns in excess of a risk free benchmark with the standard deviation of the return set. This provides a uniform risk platform which funds with different risk levels can benchmark against.

In his autobiography written for the Nobel prize institute Markowitz explains about his second inspiration source.

"The basic concepts of portfolio theory came to me one afternoon in the library while reading John Burr Williams's Theory of Investment Value. Williams proposed that the value of a stock should equal the present value of its future dividends. Since future dividends are uncertain, I interpreted Williams's proposal to be to value a stock by its expected future dividends. But if the investor were only interested in expected values of securities, he or she would only be interested in the expected value of the portfolio; and to maximize the expected value of a portfolio one need invest only in a single security. This, I knew, was not the way investors did or should act. Investors diversify because they are concerned with risk as well as return. Variance came to mind as a measure of risk. The fact that portfolio variance depended on security covariances added to the plausibility of the approach. Since there were two criteria, risk and return, it was natural to assume that investors selected from the set of Pareto optimal risk-return combinations."

The Dividend Discount Model of Williams (1937) Ph.D. dissertation implied only that investors should focus on the highest expected return. Markowitz's writes in 1999 that his idea of standard deviation or variance came from the work of Uspensky in 1937, who wrote an introduction to mathematical probability. The concept of covariance itself was developed before 1851 since notion was made by mathematician Sylvester, but it took 50 years before it was introduced in finance. Probably Markowitz did not now that Fischer already put some of the same thoughts on paper long before him in 1906. Bernstein (1992) notes that only a few scholars had mentioned the risk/return trade-off in portfolio selection before Markowitz: Fisher, Hicks and the Cowles Commission's Dickson Leavens (1945). Dimand wrote about this part of the history of financial economics in 2007 and described that Fisher's (1906) went much further than only a mere mention of the risk/return relationship. The work of Fisher gives us a good sense of the state of knowledge 100 years ago and we can still learn from his view on the subject matter.

"The Fisher diagram(1907) finds the optimum where a linear consumption possibility frontier is tangent to an indifference curve that is convex to the origin, implying consumption smoothing over time. If the axes represented states of the world

instead of time periods, the convexity of the indifference curves to the origin would represent risk aversion and lead to smoothing of consumption across states of the world."
The diagram shows some coincidental familiarities with the theory of Markowitz. Smoothing of consumptions over time can be related to diversification over time and its associated risks and rewards related to time. Fisher emphasized the present discounted value of expected income flows, treating the time pattern of income and spending (rather than a stock of capital or wealth) as fundamental knowledge. Fisher developed a good notion of risk and abstracted from risk in the early sections of his first book.
"The rate of interest acts as a link between income-value and capital-value, and by means of this link it is possible to derive from any given income-value, its capital-value, i.e. to 'capitalize' income. To do this we assume that the expected income is foreknown with certainty, and that the rate of interest (in the sense of an annual premium) is foreknown, and also that it is constant during successive years" (Fisher, 1906).
Fisher was very conscious and knew that future income flows and rates of returns are expectations, not known with certainty:
"Up to this point we have ignored the element of chance, by assuming that the entire future income-stream, or at any rate, such portions of it as needed to influence present choice, are foreknown and mapped out in advance . . . This assumption, like the assumption that bodies fall in vacuo, in the ordinary presentation of the theory of gravitation, has enabled us to complete our formal statement of the theory more easily, although at the expenses of exact conformity to actual historical fact; for, in the concrete world, the most conspicuous characteristic of the future is its uncertainty" (Fisher, 1907).
Three elements were important in his analysis:

1.) Expected return, measured as the mean of the distribution of percentage dividend paid on a stock;
2.) Risk, measured by the standard deviation of a subjective probability distribution over possible outcomes;
3.) Individual attitude towards risk, measured by Fisher's 'coefficient of caution'.

In his book he included the appendix on "Variability about a Mean, as measured by the 'Standard Deviation"', which showed that he was advanced in quantifying the risk/reward relationship, though he did not make notion of diversification effects or covariance.

Fisher's knowledgeable view on risk may be clear from his writings. He describes that some market participants will estimate more highly than other the risks taken and that from this fact it might be seem that there is a distinction between the actual risk incurred and the estimate which individuals put on it. *"…chance is always an estimate. Chance is subjective. Although one man's estimate may be better than another's through superior knowledge, intuition, or experience, the best estimate is still only an estimate, not a certainty. In the actual world of events there is no uncertainty. Aside from human opinion, there is no such thing as chance. To an omniscient being, all things are certain. It must be admitted that this view of chance is not familiar to ordinary man, nor is it universally accepted by the professed students of chance."*
And further he continues with making his statements
"It is only as the conditions vary slightly from time to time in their unknown elements that there is a change of winner; and the instant the unknowness of these elements is introduced into the problem, the observer unconsciously shifts his ground from the long run to the true theory of chance. Chance is, then, an affair of human knowledge or ignorance. According to this the ignorance theory, chance is not objective, but subjective. … Chance exists only so far as ignorance exists; varies with different persons according to their comparative ignorance of the matter under consideration; and is in fact a measure of ignorance."
A logical conclusion from his statements is that Fisher sees chance as a subjective estimate of risk ,which is in fact a measure of ignorance and that this can be measured by the standard deviation.

The notion of risk aversion and the coefficient of caution is also explained in Fisher (1906) by referring to the casino of Monte Carlo.
"a person who will pay more than the mathematical value of chance. At Monte Carlo, the "bank" makes its profit in this way, although its victims know full well that they are paying more than the mathematical value of their chances. The consequence, of course, is ruin to most of them. Fortunately, persons who deliberately gamble are in most communities in the minority. The ordinary man is unwilling to pay even the full mathematical value of the chance. He is reluctant to assume any risks, and is, on the contrary, willing to make sacrifices to rid himself of them."
Dimand noted that Fisher credited Norton with the insight that individuals become less cautious with increasing capital and named Fisher as pioneer in proposing a measure of risk aversion. Risk aversion and non-satiation are insights which in modern terms is often referred to in behavioral finance literature when referred to prospect theory of Tversky and Kahneman(1974) or to the theory of Nobel prize winners

Arrow and Pratt, which in essence shows that relative risk aversion is constant, the degree of absolute risk aversion decreases as wealth increases. Levy and Markowitz(1959) related the concept of expected utility to the mean-variance representation, they presented empirical results and showed that the function of the mean and the variance is equal to expected utility plus half the second derivative of utility times the variance.

$$f(E,V) = U(E) + .5U''(E)V$$

This second derivative is characterized by three economists between 1862 and 1873, Jevons, Menger and Walras. In economics this is known as the effect of scarcity on the price, induced by human action, and is resting on psychology and the law of satiation of wants. This equation may be thought of as a rule by which, if you know the mean and variance of a distribution, you can guess it's expected utility.

1.2.2 Modern Portfolio Theory

Modern portfolio theory (MPT), as advocated by Markowitz (1952), attempts to find a combination of assets which maximizes the expected return of a portfolio for a given level of risk, or similarly minimizes the variance of a portfolio for a given amount of expected return. The rationale behind this theory is that investors will only choose a riskier portfolio if they will get compensated by a higher expected return. The main assumptions made here are that investors are risk-averse and that risk can be described by the variance of returns. The article assumed that "beliefs" or projections about securities follow the same probability rules that random variables obey. This assumption leads to the following mathematical rules

1.) The expected return on the portfolio is a weighted average of the expected returns on individual securities
2.) The variance of return on the portfolio is a function of the variances of, and the covariance's between, securities and their weights in the portfolio.

When the covariance is normalized, one obtains the correlation matrix. Portfolio risk can be reduced by adding assets to the portfolio which are characterized by non-perfect correlation with the respective portfolio. Hence, including low-correlation assets in a portfolio can offer significant diversification benefits to investors.

Markowitz made a difference between efficient and inefficient portfolios and proposed an optimization framework by geometrical analysis. When we optimize every single portfolio by means of the MPT and plot the results in a risk-return space, we will obtain a combination of optimal portfolios which will form a hyperbola. The upper part of this hyperbola is dubbed the efficient frontier. Without going short on the risk free rate, one cannot improve the set of portfolios on this frontier in terms of expected return without increasing the risk. Therefore the efficient frontier provides us with the optimal portfolios for given amounts of risk and return.

Markowitz explained that during the time his research was published there was another scientist, Roy (1952), who proposed a striking similar framework, which was different in the assumptions made. Roy did not continue his work in finance compared to Markowitz.

1.) Roy's allowed the amount invested in any security to be positive or negative
2.) Markowitz proposed allowing the investor to choose a desired portfolio from the efficient frontier, where Roy recommended choice of a specific portfolio which maximizes the equation return divided by standard deviation.

From the literature on one-period mean-variance optimization models, building on the work of Markowitz (1952) it is known that assumptions on risk and return have impact on the optimal solutions obtained. Small changes in expected values, volatilities and covariance's of future asset returns change the risk and return tradeoff between the various asset classes and thereby influence the optimal investment portfolio.

Theory of estimating variance, covariance and mean has progressed during the years that passed. First we will elaborate variance, risk and hedging and after that we focus on the mean and the risk premium.

1.2.3 Variance

As mentioned before Fisher would be one of the first to apply probability theory in finance by using the measure of 'variability about a mean' as measured as the standard deviation. The standard deviation is often squared for appropriate averaging or calculation, called variance. We can distinguish three approaches for determining the standard deviation.

1.) Historical averaging or averaging with a moving window (MA)
2.) Exponential Weighted Moving Averaging(EWMA)
3.) AutoRegressive Moving Average(ARMA) and in particular Generalized AutoRegressive Conditional Heteroskedasticity (GARCH)

The research into risk metrics at investment bank J.P. Morgan has spurred the development of an exponential weighted standard deviation as a widely accepted risk measure worldwide during the 90's. In 1992 they launched the RiskMetrics methodology to the marketplace and when the RiskMetrics Group outgrew the firm's internal risk management resources it was spun off from J.P. Morgan. A well known and widely used decay factor was 0.94 resulting from the empirical research of this group. Here I would like to emphasize the difference between risk measurements and metrics. Were the standard deviation is a measure of the 'uncertainty' or 'chance' definition of risk. 'Value at Risk' and 'Semi variance' are metrics of the asymmetric characterization within the definition of risk, in which on tries to only capture the negative variations from the mean.
The best known form of an ARMA model is the Generalized AutoRegressive Conditional Hetereoskedasticity, a definition that follows from the work of Engle and Bollerslev, who developed this third approach of determining the standard deviation.

1.2.4 Covariance

Following the success of the GARCH model in describing the time-varying variances of economic data in the single-asset or univariate case, many researchers have extended the single-asset or time-series univariate GARCH model to the multivariate domain in which multiple assets can be handled. Bollerslev, Engle and Wooldridge (1988) developed the basic framework for a multivariate GARCH model (MGARCH). They extended the GARCH representation in the univariate case to a conditional-variance matrix, which is vectorized and known as VECH or diagonal representation. MGARCH is often difficult to use for empirical applications since it involves a large number of parameters and it is difficult to assure that the matrix is positive-definite (Tse & Tsui, 1999). Empirical applications demand further simplifications and restrictions. To verify the condition that the conditional-variance matrix of an estimated MGARCH model is positive definite, Bollerslev (1990) suggested the constant-correlation MGARCH (CCC-MGARCH) model that can overcome these difficulties.
In statistics, dependence refers to any statistical relationship between two random variables or two sets of data. Correlation refers to any of a broad class of statistical relationships involving dependence. Due to its computational simplicity, the CCC-MGARCH model is very popular among empirical researchers. However, while the constant-correlation assumption provides a convenient MGARCH model for estimation, some studies found that this assumption is not supported by some financial data. Researchers found that the returns across different markets exhibit time-varying correlations. Thus, there is a need to extend the MGARCH models to incorporate time-varying correlations and keep the feature of satisfying the positive-definite condition during the optimization. Engle and Kroner (1995) proposed a class of MGARCH model called the BEKK (named after Baba, Engle, Kraft and Kroner) model (Tse & Tsui, 1999). However, it is a great problem that it becomes difficult to estimate the conditional covariance's as the sample size increased (Yilmaz, 2010).

The DCC-GARCH model (Engle, 2002), is a generalized version of the CCC model (Bollerslev, 1990). The challenging problem of constant correlation is solved by the dynamic conditional correlation GARCH (DCC-GARCH), proposed by Engle (2001). An important assumption of the CCC model is that the time-varying of conditional covariance's is caused by the time-varying of the conditional variance of the individual return series. The DCC model differs in allowing the conditional correlation matrix, R_t to be time-varying. Mathematical framework of this model, developed by Engle and Sheppard (2001), has main two steps algorithm to have time varying covariance matrix. First step is to find conditional standard deviations through the univariate GARCH and second step is to model the time varying correlations relying on lagged values of residuals and covariance matrices. After that, the conditional covariance matrix could be found by using conditional standard deviations and dynamic correlations (Yilmaz, 2010).

Both the Modern Portfolio Theory and Generalized AutoRegressive Heteroskedasticity are abstracts from reality and quantitative approaches to analyze financial data to gather information. Both of these methods did not infer with estimating the direction of the mean return, which is the toughest issues of the investment subject since the mean is subject to uncertainty. When someone is unwillingly exposed to a particular risk and is not determined to spend time to circumvent this uncertainty one often had the possibility to hedge against these risks.

1.2.5 Hedging of risk

There are many types of risk present in financial markets, the earliest product were forwards or standardized future contracts used by the farmers to sell short and reach an agreement with the buyer. This enables them to hedge against a bad harvest, which could be caused by flood or draught, other weather conditions or natural factors like pests and diseases. By selling short one agrees on a price before delivering the good, this good can even be 'in production'. These kind of contracts are called derivatives since they derive their value from the underlying. In the financial market these agreements are often sold at margin, which means that the buyer has to deposit or preserve a part of his capital as collateral. These derivative markets are often highly leveraged up till 50 times the initial deposit, so one also understands that many fortunes are made and lost in these markets.

Nowadays hedging is possible for many risk factors beside commodity risk, the largest market is the interest rate market and second largest is the currency market. In the interest rate market the concepts of swaps is often used. In finance, a swap contract is a derivative in which counterparties exchange certain benefits of one party's financial instrument for those of the counterparty's financial instrument. The benefits in question depend on the type of financial instruments involved. For example, in the case of a swap involving two bonds, the benefits in question can be the periodic interest (or coupon) payments associated with the bonds. A pension fund could use them for instance to hedge for their long term obligations or a bank could sell them for his asset transformation

The second important market is the currency future market with which we have to deal for instance if we would like to invest abroad. One must know that certainty often is possible at the cost of a premium implied in the price or as fee for intermediation.

2. Methodology

In this section the methodology used will be discussed and we shall elaborate on the econometric issues which were encountered during the analysis. The methodology used can be characterized in six sections. The first section Portfolio theory, in which we define the methodology of Modern Portfolio Theory, the second section is about the variance and Generalized AutoRegressive Conditional Heteroskedasticity to model the variance. The third section is about the correlation coefficient. The fourth section is about the mean and methodology like the probit model and the Markov Regime Switching model to model and estimate the mean. Then we elaborate shortly on dynamic estimation and the last section is about the test methodology put to practice.

2.1 Modern Portfolio Theory

The mean-variance rule states that the investor would (or should) want to select one of those portfolios which give rise to the mean-variance combinations indicated as efficient. These are those with minimum variance σ_p^2 for a given mean return π_p or maximum π_p for given σ_p^2. First lets answer the elementary question as proposed in the book 'Asset Pricing' of Scholes (2005) "When does the mean-variance frontier exist?" This is when the set of portfolio means and variances is less than the whole expected return-risk space in which all combinations are possible. This is when we rule out a specific case of two perfect correlated returns, but with different means. In this case one could short one, buy the other, and achieve infinite expected returns with no risk assuming that we have unlimited supply of these assets. It would be a violation of the economic law of one price. Scholes stated this in a theorem.
Theorem: *So long as the variance-covariance matrix of returns is non-singular, there is a mean-variance frontier.*
This is a requirement which we will also need when we develop time-variant correlation coefficients.

The standard computation of the mean-variance frontier uses a brute-force approach. There are techniques by which we can compute the set of efficient portfolios and efficient mean-variance combinations associated with given $\underline{\mu}$ and $\underline{V}$. The goal of the proposed model is to determine the optimal minimum variance portfolio for a given level of risk subjecting to the following constraints written out by Balvers (2001) and Renström (2002).

$$\underline{\boldsymbol{w}}^T \underline{\boldsymbol{\mu}} = r$$
$$\underline{\boldsymbol{w}}^T \underline{\boldsymbol{l}} = 1$$

where r denotes the expected portfolio return, $\underline{\boldsymbol{l}}$ represents a $1 \times n$ column vector of 1's and the sum of the weights is equal to 1, or the whole portfolio.

The minimization problem can be solved using the Lagrange multiplier method with multipliers λ_1 and λ_2 for the constraints. Set

$$L = \underline{\boldsymbol{w}}^T \underline{\boldsymbol{V}\boldsymbol{w}} - \lambda_1 \left(\underline{\boldsymbol{w}}^T \underline{\boldsymbol{\mu}} - r \right) - \lambda_2 \left(\underline{\boldsymbol{w}}^T \underline{\boldsymbol{l}} - 1 \right)$$

The first-order condition with respect to $\underline{w}$ is

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = 2\underline{\boldsymbol{w}\boldsymbol{V}} - \lambda_1 \underline{\boldsymbol{\mu}} - \lambda_2 \underline{\boldsymbol{l}}$$

By pre multiplying the Lagrangian by ½ we get more convenient expressions and

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = \underline{\boldsymbol{w}\boldsymbol{V}} - \lambda_1 \underline{\boldsymbol{\mu}} - \lambda_2 \underline{\boldsymbol{l}}$$

Solving this derivative for the optimal portfolio weights is done by applying the following steps
The result is a vector and at the optimum this vector must be equal to the null vector, so we have the necessary first-order condition for an optimum.

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = \underline{\boldsymbol{wV}} - \lambda_1 \underline{\boldsymbol{\mu}} - \lambda_2 \underline{\boldsymbol{l}} = 0$$

The solution is then obtained by pre multiplying the necessary condition by the inverse to the variance-covariance matrix

$$\underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{wV}} = \lambda_1 \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} - \lambda_2 \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$$

$$\Leftrightarrow$$

$$\underline{\boldsymbol{w}}^{*} = \lambda_1 \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$$

Based on the fact that $\underline{\boldsymbol{V}}$ is positive definite we can conclude that $\underline{\boldsymbol{w}}^{*}$ minimizes the variance and that the solution obtained is unique.

Pre multiplying the first-order condition by $\underline{\boldsymbol{w}}^{*}$ and using constraints gives

$$\sigma_p^2 = \lambda_1 r + \lambda_2$$

To replace the Lagrange multipliers we pre multiply vector $\underline{\boldsymbol{w}}^{*}$ by the vector of expected returns, $\underline{\boldsymbol{\mu}}$, and separately by the unit vector, $\underline{\boldsymbol{l}}$, to obtain

$$\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{w}}^{*} = \lambda_1 \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$$

$$\underline{\boldsymbol{l}}^T \underline{\boldsymbol{w}}^{*} = \lambda_1 \underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$$

$\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}}$, $\underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$ and $\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$ are scalars and we can label them A, B, and C.
$\underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} = \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}} = \mathrm{C}$ and $D = AB - C^2$

Then

$$r = \lambda_1 A + \lambda_2 C$$

$$1 = \lambda_1 C + \lambda_2 B$$

Solve these to obtain

$$\lambda_1 = \frac{Br - C}{D}, \lambda_2 = \frac{A - Cr}{D}$$

Finally substitute λ_1 and λ_2 into the function for the optimal weights and we have the optimal weights as function of the vector of expected returns, the unity vector and the variance-covariance matrix.

$$\underline{\boldsymbol{w}}^{*} = \frac{Br - C}{D} \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{\mu}} + \frac{A - Cr}{D} \underline{\boldsymbol{V}}^{-1} \underline{\boldsymbol{l}}$$

$$= \frac{A\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}} - C\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}}}{D} + \frac{B\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} - C\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}}{D} r$$

And substitution of λ_1 and λ_2 in σ_p^2 yields an explicit expression of the portfolio frontier:

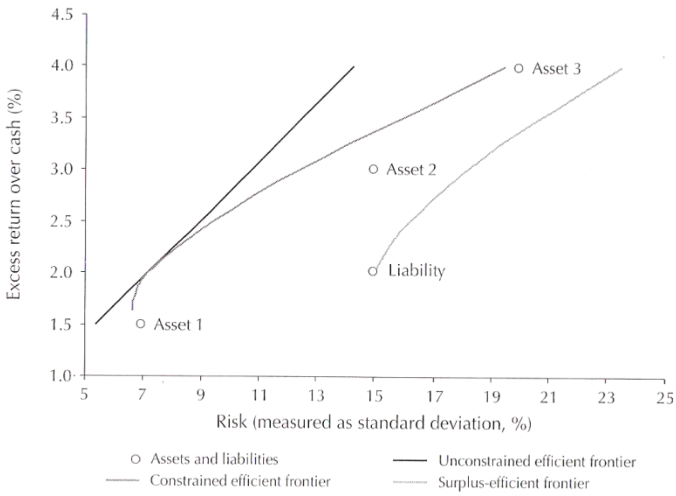
$$\sigma_p^2 = (Br^2 - 2Cr + A)/D$$

Plot the μ_p against σ_p^2 to draw the efficient frontier. The intersection between the line with origin at the risk-free rate and the efficient frontier is known as the tangency or optimal portfolio. For a more extensive elaboration on the theory please refer to Ingersoll (1987).

2.1.1 Adding leverage

To be realistic the constraint the weights of the assets had to sum up to 1 in $\underline{\boldsymbol{w}}^T\underline{\boldsymbol{l}}$. To be more realistic we could release this assumption. In essence this means that we would allow for leveraged of the portfolio. When the leverage constraint is released one is able to select any combination of the tangency portfolio and the risk-free asset on the unconstrained efficient frontier of figure 1.1. During this research we did not add any leverage to our portfolios.

Figure 1.1: Efficient frontiers (Scherer, 2002)



2.1.2 Adding liabilities

Scherer (2002) gives an excellent compilation of theory for portfolio construction and describes the method to add liabilities to the framework. Since assets cannot isolate from liabilities for firms such as banks, insurance companies and pension funds we need to consider them in our portfolio analysis for determining an optimal asset allocation. Even investors need to think about liabilities as a real consumption stream after retirement. Asset-Liability Management (ALM) becomes a requirement for every investor who seeks to define his or her potential liabilities carefully. It focuses on managing the difference between assets and liabilities, also called "surplus". The change in surplus depends on the returns of the asset portfolio, r_p^2, as well as the liability returns (percentage changes in the value of outstanding liabilities), r_l^2:

$$\Delta Surplus = Assets \times r_p - Liabilities_t \times r_l$$

the surplus returns are expressed as the change in surplus relative to assets

$$\frac{\Delta Surplus}{Assets} = r_p - \frac{Liabilities}{Assets} r_l$$
$$= r_p - f r_l$$

In which f is the ratio of liabilities to assets. If we set $f = 1$ and $r_l = c$, we are back in a world without liabilities (alternative is to think of liabilities as cash). Surplus volatility $\sigma^2_{surplus}$, can now be incorporated into the framework established in section 2.1 by including a short position in liabilities:

$$\sigma^2_{surplus} = \begin{bmatrix} w_1 \\ \vdots \\ w_k \\ -f \end{bmatrix}' \begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,k} & \sigma_{1,l} \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{k,1} & \cdots & \sigma_{k,k} & \sigma_{k,l} \\ \sigma_{l,1} & \cdots & \sigma_{l,k} & \sigma_{l,l} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_k \\ -f \end{bmatrix}$$
$$= var(\pi_{surplus}) = \underline{\boldsymbol{w}}^T \underline{\boldsymbol{V}}_{surplus}\, \underline{\boldsymbol{w}}$$

Let's assume that liabilities can be summarized as one single asset, l, whereas, for example, $\sigma_{k,l}$ summarizes the covariance of the k^{th} asset with our liabilities. Scherer emphasizes that one of the difficulties in asset-liability management arises from the non-existence of a liability mimicking asset that, if bought, would hedge out all the liability risk completely. This is often caused by inflation-indexed liabilities or final wage-related schemes, which create unhedgeable equity-linked liabilities. Pension funds will always have to accept some liability noise. Since our goals is to arrive at an surplus-efficient frontier we can transform the ALM problem into the portfolio optimization framework. This is done by expressing the covariance matrix in terms of surplus risk. Each $\underline{\boldsymbol{V}}_{surplus}$ can be expressed as $cov(r_i - f r_l, r_j - f r_l)$. During this research we did not add liabilities to our portfolios, this is an important section for the practical application for insurance companies and pension funds, which deal with considerable liabilities.

2.1.3 Stability and out of sample characteristics of the efficient frontier

The next and practically important question addressed in this section is the stability of efficient portfolios. Ziemba and Mulvey (1998) explained that it makes little sense to optimize market shares and currency hedges, if the covariance matrix and expected (or average) returns are so unstable over time. Unstability causes that the ex-ante efficient portfolio proves to be inefficient ex post. Here the problem is partly solved by relying on better estimates for the future and accepting that the level of the return is for the largest share unforeseeable. Since we use time varying structures for the risk (variance) as well as for the relation (correlation) parameters it is possible to fix the portfolio at a certain risk level with its accompanying 'reliable' historical risk premium (return) parameters. The covariance matrix of assets and liabilities is transformed via a matrix of long-short positions into the covariance matrix of surplus returns.

$$\underline{\boldsymbol{V}}_{surplus} = \begin{bmatrix} 1 & 0 & \cdots & 0 & -f \\ 0 & 1 & . & . & -f \\ \vdots & . & \ddots & . & \vdots \\ 0 & 0 & . & 1 & -f \end{bmatrix} \begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,k} & \sigma_{1,l} \\ \vdots & \ddots & \vdots & \vdots \\ \sigma_{k,1} & \cdots & \sigma_{k,k} & \sigma_{k,l} \\ \sigma_{l,1} & \cdots & \sigma_{l,k} & \sigma_{l,l} \end{bmatrix} \begin{bmatrix} 1 & 0 & \cdots & 0 & -f \\ 0 & 1 & . & . & -f \\ \vdots & . & \ddots & . & \vdots \\ 0 & 0 & . & 1 & -f \end{bmatrix}'$$

While cash is a very conservative asset in a asset-only framework, in an asset-liability framework it becomes one of the most risky assets. This since it has no covariation with liabilities (often similar to long bonds). Cash therefore cannot serve as a liability-hedging asset. Now we manipulate the expected returns to reflect the relative return of assets versus liabilities:

$$\boldsymbol{\mu}_{surplus} = \begin{bmatrix} \mu_1 - f\mu_l \\ \vdots \\ \mu_k - f\mu_l \end{bmatrix} + c(1-f)$$

Then we can optimize the Modern Portfolio Theory framework with inputs $\underline{\boldsymbol{V}}_{surplus}$ and $\boldsymbol{\mu}_{surplus}$ and solve for the optimal weights. Following Scherer (2002), this is the method with which we can trace out the surplus-efficient frontier. The unconstrained (asset-only) efficient frontier and surplus-efficient frontier coincide if:

- Liabilities are cash so when assets have no covariation with liabilities
- All assets have the same covariation with liabilities
- There exists an asset which behaves like the liabilities and which lies on the efficient frontier

Summarized we can give the following convenient formula to calculate the surplus portfolio variance.

$$\sigma_{surplus}^2 = \begin{bmatrix} w_1 \\ \vdots \\ w_k \\ -f \end{bmatrix}' \begin{bmatrix} \underline{V} & \underline{\Gamma} \\ \underline{\Gamma}' & \sigma_{l,l} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_k \\ -f \end{bmatrix}$$

Where $\underline{\Gamma}$ is the vector which expresses the covariance between asset and liability returns the surplus-efficient frontier that arises is shown in figure 1.1.

2.2 The Variance

2.2.1. Variance computation

Portfolio variance in relation with portfolio return is often used to determine the optimal asset allocation. The measure most often used to calculate the variance of a portfolio is a time-invariant measure of variance . It is dubbed naïve variance since it puts equal weights on past variance data, which is the simplest manner to deal with risk. Hull (2009) defines σ_t as the volatility of a market variable on day t, as estimated at the end of day t-1. The square of the volatility is the variance rate. The univariate unbiased estimate of the variance rate of asset i, σ_i^2 as a function of returns π_t and mean return $\bar{\pi}$, using the most recent *m* observations is:

$$\sigma_{i,t}{}^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\pi_{t-1} - \bar{\pi})^2$$

To calculate portfolio variance, a measure of the strength of the correlation between two assets is used, which is the covariance.

$$\begin{aligned} \sigma_{1,2} &= \mathrm{Cov}(\pi_1, \pi_2) = \mathrm{E}(\pi_1 - \mu_{\pi_1})(\pi_2 - \mu_{\pi_2}) \\ &= E(\pi_1 \pi_2) - \mu_{\pi_1} \mu_{\pi_2} \\ &= \sum_{i=1}^{N} \frac{(\pi_{1_t} - \overline{\pi_1})(\pi_{2_t} - \overline{\pi_2})}{N} \end{aligned}$$

in which N is the sample size, μ_{π_i} is the sample mean and the expected return and *t* denotes the observation of π.

Statistical correlation as function of covariance

$$\rho_{1,2} = \frac{\sigma_{1,2}}{\sigma_1 \sigma_2}$$

$$\sigma_{1,2} = \rho_{1,2} \sigma_1 \sigma_2$$

The variance of a two-asset portfolio is

$$\begin{aligned} \sigma_p^2 &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \rho_{1,2} \sigma_1 \sigma_2 \\ &= w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_{1,2} \end{aligned}$$

w_1 and w_2 are portfolio weights which satisfy unity $w_1 + w_2 = 1$.
σ_1 and σ_2 are respectively the standard deviations of asset 1 and asset 2.

We can calculate this for n assets.

$$\sigma_p^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} w_i w_j \sigma_{ij}$$

If we use the fact that the variance of R_i is σ_{ii} then

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}$$

For a multiple assets it is easier to use matrix notation, and set

$\underline{\boldsymbol{\pi}} = (\pi_1, \pi_2, \dots, \pi_n)^T$, asset returns
$\underline{\boldsymbol{w}} = (w_1, w_2, \dots, w_n)^T$, weights of the assets
$\underline{\boldsymbol{\mu}} = (\mu_1, \mu_2, \dots, \mu_n)^T$, expected returns
$\underline{\boldsymbol{V}} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \cdots & \sigma_{1,n} \\ \sigma_{1,2} & \sigma_2^2 & \cdots & \sigma_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{1,n} & \sigma_{2,n} & \cdots & \sigma_n^2 \end{bmatrix}$, the $n \times n$ variance-covariance matrix of returns

The covariance and correlation matrices are symmetric because the covariance and the correlation between π_i and π_j is the same as the correlation between π_j and π_i.
The variances of the individual assets are on the diagonal of this matrix.

We would like to know the variance of the portfolio return

$$\pi_p = w_1\pi_1 + w_1\pi_1 + \dots + w_n\pi_n = \underline{\boldsymbol{w}}^T \underline{\boldsymbol{\pi}}$$

by choice of $\underline{w}$ we obtain the portfolio variance

$$\sigma_p^2 = var(\pi_p) = \underline{\boldsymbol{w}}^T \underline{\boldsymbol{Vw}}$$

We could also take the square root of the result and obtain the portfolio standard deviation or volatility. Here I would like to refer to an excellent paper of Reider (2009), who gave us generalized equations.

2.2.2 AutoRegressive Conditional Heteroskedasticity Model

The AutoRegressive(AR) part comes from the fact that the model is an autoregressive model in squared returns. The conditionality comes from the fact that the model's next period's volatility is conditional on information this period. Heteroskedasticity means non constant volatility. In a linear regression where $y_t = \alpha + \beta x_i + \epsilon_t$, when the variance of the residuals, ϵ_t is constant, we name that homoskedastic and use Ordinary Least Squares (OLS) to estimate α and β. When the variance of the residuals is not constant, we call that heteroskedastistic and we can use Weighted Least Squares(WLS) to estimate the regression coefficients. Let's say that that the return on an asset is

$$r_t = \mu + \sigma_t \epsilon_t$$

in which ϵ_t is a sequence of N(0,1) i.i.d. random variables. We will define the residual return at time t, $r_t - \mu$, as

$$\alpha_t = \sigma_t \epsilon_t$$

The ARCH(1) model developed by Engle (1982) gives us the equation.

$$\sigma_t^2 = \alpha_0 + \alpha_1 \alpha_{t-1}^2$$

In which $\alpha_0 > 0$ and $\alpha_1 \geq 0$ to ensure positive variance and $\alpha_1 < 1$ for stationary. In an ARCH(1) model, when the residual return, α_t is large in magnitude, the forecast for the next period's conditional volatility, σ_{t+1} will be large.

2.2.3 Generalized AutoRegressive Conditional Heteroskedasticity

The second measure is a time variant variance called General Autoregressive Conditional Heteroscedasticity (GARCH) variant of variance. This variant is well known since it allows to take the auto correlative as well as the mean-revering characteristics of variance into account.

For high order ARCH processes volatility can be modelled by GARCH(p,q). Generalized ARCH was developed by Bollerslev (1986), compared to ARCH(q), dependencies are permitted on q lags of past ϵ_t^2 in addition to p lags of past h_t^2 as shown below. In a GARCH(p,q) model the volatility depends on last periods volatility and the return residuals:

$$r_t = \sigma_t \varepsilon_t \qquad h_t^2 = \omega_0 + \sum_{i=1}^{q} a_i r_{t-i}^2 + \sum_{i=1}^{p} \beta_j \, h_{t-j}^2 \qquad 0 < a + \beta < 1$$

The rate of decay of the autocorrelation is measured by $\alpha + \beta,$ the closer to 1, the slower the decay. Following Cont (2005) estimations on returns often yield $\alpha + \beta$ close to 1. Volatility forecast from GARCH(1, 1) can be made by repeated substitutions of the obtained parameters, the last value of the variance and the .

By the use of programming one can estimate the GARCH parameters and calculate the variance series for all asset classes.

2.3 The correlation coefficients

2.3.1 Correlation computation

For calculation of the correlation coefficients a series of n measurements of X and Y written as x_i and y_i where $i = 1, 2, \ldots, n,$ then the sample correlation coefficient can be used to estimate the population Pearson correlation r or often written as ρ between X and Y. The sample correlation coefficient is written

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of X and Y, and s_x and s_y are the sample standard deviations of X and Y.

This can also be written as:

$$r_{xy} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{(n-1)s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum {x_i}^2 - (\sum x_i)^2}\sqrt{n\sum {y_i}^2 - (\sum y_i)^2}}$$

2.3.2 Dynamic Conditional Correlation

The multivariate GARCH model assumes that returns from k assets are conditionally multivariate normal with zero expected value and covariance matrix $\boldsymbol{H}_t$. The returns can be either mean zero or the residuals from a filtered time series. The proposed dynamic correlation structure by Engle and Sheppard (2001) is constructed as follows

$$r_t | F_{t-1} \sim N(0, \boldsymbol{H}_t)$$

and

$$\boldsymbol{H}_t = \boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t$$

H_t is the conditional covariance matrix and as a function of
D_t is the $k\ x\ k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the i^{th} diagonal and h_{it} denotes the conditional variance of the *i*-th return.
R_t is the time varying correlation matrix.

The log-likelihood of this estimator can be written as

$$\begin{aligned} L &= -\frac{1}{2}\sum_{t=1}^{T}(\text{k}\log(2\pi) + \log(|\boldsymbol{H}_t|) + r_t{}'\boldsymbol{H}_t{}^{-1}r_t) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(\text{k}\log(2\pi) + \log(|\boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t|) + r_t{}'\boldsymbol{D}_t{}^{-1}\boldsymbol{R}_t{}^{-1}\boldsymbol{D}_t{}^{-1}r_t) \\ &= -\frac{1}{2}\sum_{t=1}^{T}(\text{k}\log(2\pi) + 2\log(|\boldsymbol{D}_t|) + \log(|\boldsymbol{R}_t|) + \epsilon_t{}^{T}\boldsymbol{R}_t{}^{-1}\epsilon_t) \end{aligned}$$

where $\epsilon_t \sim N(0, R_t)$ are the residuals standardized by their conditional standard deviation. We propose to write the elements of D_t as univariate GARCH models, so that

$$h_{it} = \omega_i + \sum_{p=1}^{P_i}\alpha_{ip}r_{it-p}^2 + \sum_{q=1}^{Q_i}\beta_{iq}h_{it-q}$$

for $i = 1,2, \dots, k$ with the usual GARCH restrictions for non-negativity and stationarity being imposed, such as non-negativity of variances and $\sum_{p=1}^{P_i}\alpha_{ip} + \sum_{q=1}^{Q_i}\beta_{iq} < 1$. The subscripts are present on the individual P and Q for each series to indicate that the lag lengths chosen need not be the same. The specification of the univariate GARCH models is not limited to the standard GARCH (p,q), but can include any GARCH process with normally distributed errors that satisfies appropriate stationary conditions and non-negativity constraints.

$$\boldsymbol{Q}_t = \left(1 - \sum_{m=1}^{M}\alpha_m - \sum_{n=1}^{N}\beta_n\right)\bar{\boldsymbol{Q}} + \sum_{m=1}^{M}\alpha_m\,(\boldsymbol{\epsilon}_{t-m}\boldsymbol{\epsilon}^{T}{}_{t-m}) + \sum_{n=1}^{N}\beta_n\,\boldsymbol{Q}_{t-n}$$

$$\boldsymbol{R}_t = \boldsymbol{Q}_t^{*-1}\boldsymbol{Q}_t\boldsymbol{Q}_t^{*-1}$$

where $\bar{Q}$ is the unconditional covariance of the standardized residuals, $\boldsymbol{\epsilon}_t$ resulting from the first stage estimation, and

$$\boldsymbol{Q}_t^* = \begin{bmatrix} \sqrt{q_{1,1}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{q_{2,2}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{q_{k,k}} \end{bmatrix}$$

so that $\boldsymbol{Q}_t^*$ is a diagonal matrix composed of the square root of the diagonal elements of $\boldsymbol{Q}_t$. The typical element of $\boldsymbol{R}_t$ will be of the form

$$\rho_{i,j,t} = \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}$$

The following useful result from linear algebra simplifies finding the necessary conditions for R_t to be positive definite and hence a correlation matrix.

2.4 The mean

2.4.1 Mean computation

The benefit of using the means or the returns, versus prices, is normalization. Expectations about the mean can be formed by many different methods, the most intuitive is to use the historical average. One of the problems encountered using historical averages is that real returns can differ substantially from this estimated mean. First of all we need to differentiate between arithmetic and geometric mean. The most commonly familiar mean is the arithmetic mean. We are familiar with how to average grades. In mathematics it is defined as

$$Arithmatic\ mean = \frac{\alpha_1 + \alpha_2 + \cdots + \alpha_n}{n}$$

The geometric mean is defined for all α's being real and positive numbers as

$$Geometric\ mean = \sqrt[n]{\alpha_1\alpha_2 \cdots \alpha_n}$$

To determine if one should use the arithmetic or geometric mean determine if returns are dependent or independent. The return of different assets are independent, but returns over time are dependent, since total return over a period depends on which return occurs first. The amount n relates to the amount of compounded returns within the period and is relevant in discrete time and is also known as the True Time-Weighted Rate of Return when cash flows are present or when there is a sequence of returns during the period. When r_c is the return on a continuous compounding basis this means that the number of periods is infinite.

$$\bar{r}_{geometric} = -1 + \sqrt[n]{\prod_{i=1}^{n}(1 + r_{arit\,h,i})}$$

When means are continuous and continuously compounded it is convenient to use return which is time additive and mathematically convenient to use.

$$r_{log} = \ln(\frac{V_t}{V_{t-1}})$$

Now we showed the difference between arithmetic mean and geometric mean we can go one step further and discuss the difference between geometric and logarithmic mean. The logarithmic mean of two numbers is smaller than the arithmetic mean but larger than the geometric mean (unless the numbers are the same, in which case all three means are equal to the numbers).
In the formula above V_t is the asset or index value at time t. The logarithmic mean is calculated as the area under an exponential curve or integral. The log function can be in any base, e.g. natural log (ln), as long as consistent bases are used all throughout calculation. The interest rate expressed as a continuously compounded rate is called the force of interest.

2.5 Dynamic estimation

2.5.1 Rolling Windows

For dynamic estimation of the model parameters and the return, variance and correlation estimations, a method based on rolling windows is used. Let say that one has a long time series of data, of length **T**, available for estimation, where **T** is much larger than t, the number of observations used to form the estimations in constructing efficient portfolios. We consider the general case where we use T monthly return observations to form portfolios and hold them for one month before rebalancing. When the holding period is one month, t = 120, the number of months in 10 years for the long term estimation and t = 24 for the short term estimation. The expected portfolio allocation that needs to be rebalanced will be relatively smaller when using the long term estimation window. We use the rolling window method described below as the benchmark, that provides consistent estimators of the 1-month out-of-sample parameters and variables. To test the GARCH-DCC model against the GARCH model a 10 year period is used. During the estimation procedure, the window shifts 1 month for each time step.

2.6 Test methodology

When the model is developed in MATLAB we can estimate means, variances and covariance's and test which setting is delivers the best estimates and forecasts. Secondly we can calculate optimal portfolios. The performance of time varying GARCH(P,Q) variances and DC Correlations is tested against a model which uses historical or so called naïve variances and correlations instead.

It is possible to compare the expected return and standard deviations with the realized returns and standard deviations by means the Diebold-Mariano test statistics. To compare forecast errors from two different models the Diebold-Mariano test statistic is used. The statistic compares the deviations of the forecasts from the realizations, corrects for serially correlated forecasting errors and determines if this is statistically significant.

We would like to know how we would be invested when we would have used certain methodology for estimation of parameters. This can be done by comparing realized Sharpe ratios of the tangency portfolios of the separate models using a 1-month rolling windows.

2.6.1 Root Mean Squared Error

Mean squared error (MSE) of an estimator is a method to quantify the difference between values implied by a forecasted estimator and the realized values of the estimated quantity. MSE is a risk function for the expected value of the squared error loss or quadratic loss. The MSE of estimator $\hat{\theta}$ compared to the estimated parameter θ is formulized as

$$MSE(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

which can be separated in the sum of the variance and the squared bias of the estimater

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + (Bais(\hat{\theta}, \theta))^2$$

An MSE of zero, meaning that the estimator $\hat{\theta}$ predicts observations of the parameter θ.

In this paper the MSE will be used as out-of-sample mean squared error, which refers to the mean value of the squared deviations of the predictions from the true values over an out-of-sample period. The MSE of different models, including the random walk model which simply implies that the forecast equal to the most recent value, can be compared. The MSE can be used as relative measure as proposed by Thompson (1990) to compare the performance of the forecasted means and variances against the random walk model or the naïve forecasts.

In academic literature and MATLAB the RMSE value is a basic measure of how closely a model fits some data, which measures the average mismatch between each realized data point and the model.

We should look at the RMSE values as a first tool to inspect the quality of the fit, high RMSE values can indicate problems. The smaller the RMSE, the closer our model follows the data; if a model goes through each data point exactly or when the errors are on average zero, then the RMSE is zero. RMSE is simply the root of the MSE.

2.6.2 Diebold-Mariano

For statistical comparison of forecasting accuracy between models we can us the Diebold-Mariano (DM) statistics (Diebold and Mariano, 1995) for practical application see Chueng, Chinn and Pascual, (2005). This statistic is used for testing whether the performance of the forecast series is significantly different from that of the random walk forecast and expresses this in a comparable figure. The most important feature of the DM test statistic is that it corrects for economic loss resulting from the volatility of the tested variable. By comparing the DM test statistics we will be able to assess which model predicts the returns better. The statistics are calculated using the squared error and absolute error loss functions.

To determine if one model predicts better than another we may test null hypotheses

$$H_0\colon E\left[L\left(\varepsilon^1_{t+h|t}\right)\right] = E\left[L\left(\varepsilon^2_{t+h|t}\right)\right]$$
$$H_1\colon E\left[L\left(\varepsilon^1_{t+h|t}\right)\right] \neq E\left[L\left(\varepsilon^2_{t+h|t}\right)\right]$$

We can calculate the loss differentials as follows

$$d_{sq,t} = (\varepsilon_t^{m2})^2 - (\varepsilon_t^{m1})^2$$

$$d_{abs,t} = |\varepsilon_t^{m2}| - |\varepsilon_t^{m1}|$$
$$\overline{d} = \frac{1}{T_0}\sum_{t=t_0}^{T} d_t$$

in which ε_t^{m1} and ε_t^{m1} respectively denote the forecasting errors of model 1 and 2. These forecasting errors can be calculated using rolling 1-step ahead forecasts from model 1 and 2. When the loss differentials are positive this indicates that model 2 produces a larger forecast error than model 1. The DM statistic

$$DM = \frac{\overline{d}}{SE(\overline{d})}$$

$$DM = \frac{\overline{d}}{\sqrt{(LRV_{\overline{d}}/T)}}$$

$$LRV_{\overline{d}} = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j\,, \gamma_j = cov(d_t d_{t-j})$$

in which $SE(\overline{d})$ is the standard error of the loss differential, which is calculated as the asymptotic long-run variance(LRV) of $\sqrt{T}\overline{d}$, $\sqrt{(LRV_{\overline{d}}/T)}$. The long-run variance is used, because the sample of loss differentials is serially correlated for $h > 1$ (Zivot, 2004). So the loss differentials are regressed on a constant using the Newey-West correction to the standard error. This to test the null hypothesis, that the models have equally forecasting accuracy. When the t-statistic is positive we can conclude that model 1 is more accurate than model 2.

2.6.3 Sharpe ratio

One way to assess the performance of the Naïve, the GARCH and the DCC-GARCH model is by using the Sharpe ratio. The Sharpe ratio of the portfolio is calculated by the function

$$SR = \frac{R_p^{tot} - R_F^{tot}}{\sigma_p^{tot}}$$

In which the total portfolio return R_p^{tot} and σ_p^{tot} are calculated during the whole period of holding the portfolio.
The aggregate return over N days

$$R^{tot} = \prod_{t=1}^{N}(1 + R_t) - 1$$

where R_t is the monthly return.
The Sharpe ratio is calculated using the naïve approximation for risk, since we used GARCH estimates of the variance to forecast 1-step or month ahead, the squared realized return is used as proxy of the variance.

$$\sigma_p^{tot} \approx \sum_{t=1}^{N} w_a R_a^{m\,2}$$

in which *a* denotes the separate asset classes.

We are interested in comparing the expected Sharpe ratio with the realized Sharpe ratio.

1. The Sharpe ratio is calculated by first estimating the optimal portfolio weights and registering the expected return and variances.
2. We calculate one expected Sharpe ratio as if we would rebalance our portfolio every month towards an optimal asset allocation. This is done by using the mean and variance of the optimal portfolios.
3. We use the weight and the realized returns and estimated covariance's matrix to determine our realized Sharpe ratio.
4. If the Sharpe ratio is different we can determine which one is higher and conclude which asset allocation approach is outperforming the other.

2.6.4 Maximum drawdown

The maximum cumulative loss from a market peak to trough, often called the maximum drawdown (MDD), is a measure of how sustained one's losses are. The measure is known well in the asset management industry since large drawdowns usually lead to fund redemptions, and so the MDD is the risk measure of choice for many money management professionals. A reasonably low MDD is critical to the success of any fund. While long-term investors are ultimately rewarded with strong absolute returns, short-term losses are often sharp enough to make even the most rational investors question their allocations. (Altegris, 2011). The maximum drawdown is measured by the cumalative sum of consequent losses.

When the return at time t, is smaller or equal to 0 and when the last Maximum DrawDown observation is smaller than 0 then sum the return with the last MDD observation and else the value is the minimum of the return and zero.

$$\sum \mathrm{r_t} + \mathrm{MDD_{t-1}}$$

when

$$\mathrm{r_t} \leq 0$$
$$\mathrm{MDD_{t-1}} \leq 0$$

else

$$Min(\mathrm{r_t}, 0)$$

3 Assumptions

3.1 Modern Portfolio Theory

During this research the same assumptions as Hull (2007) are made which are all true for some market participants.

1.) There are no transaction costs for trading

2.) Each is subject to the same tax rate on all net trading profits

3.) Participants can lend and borrow money at the same risk-free rate

4.) Market participants will make markets efficient by taking advantage of arbitrage opportunities

Versijp (2011) adds the following assumptions for modern portfolio theory to our list.

5.) Agents prefer more over less (no satiation)

6.) Agents dislike risk (risk-aversion)

7.) Traders maximize utility, and do so for 1 period

8.) Utility is a function of expected return and variance and nothing else

9.) There is no distortion from inflation

10.) All information is available at no costs

11.) All investments are infinitely divisible

And last one which should be on our assumption list for proper analysis

12.) The unit of measurement contains a constant Purchasing Power

Of course this list is not the best representation of reality, but allows to do valuable analysis.

4 Data and software

4.1 Data

4.1.1 Asset classes

Investors have many investment possibilities and following Modern Portfolio Theory they should search for the best risk/reward ratio in combination with diversification. Since assets within particular asset classes tend to be influenced by the same risk factors, one will find difficulties to diversify. Great opportunity for diversification can be found in investing in other asset classes which will bring different risk factors and premium for bearing them. From a report about commodity trading of TheCityUK (2011) we learn from their rough estimates that many investors recently discovered commodity futures as an investment. This since the world wide asset class commodities, value of contracts under management doubled in three years from 2008 till 2010 from $200bn. to about $380bn. The U.S. investor has excellent access to these markets by trading for instance on the commodity exchange (COMEX) or by buying Exchange Traded Products (ETP), which almost tripled in these three years from $113bn. to $310bn. at the end of 2010. SIFMA (2011) registers the total sizes of U.S. credit markets and the Bank of International Settlements(BIS) (2011) registers sizes of international credit as well as equity markets. The sizes of the traditional asset classes, government securities, corporate securities and traded equities were respectively $9,123bn., $7,598bn and $251bn..

In this research standard liquid asset classes like fixed income, commodities and equity are used. A sub divisions of fixed income bonds is made, in government bonds and (corporate) investment-grade bonds. Data necessary was obtained by use of Datastream, Bloomberg and other databases of the Erasmus University Rotterdam(EUR) and pension organization PGGM. I will use monthly total return data of indices of publically traded bonds for government bonds like S&P 30 Year US Treasury Bond Futures Index available from 1988, JP Morgan Global Aggregate (Investment-grade) Bond Index available from 1973 onwards. For commodities I will use GSCI indices available from 1973.

The following data series were used as being representative as price index for the asset classes were U.S. investors are able to choose from to determine their portfolios:

1.) Standard & Poor Goldman Sachs Commodity Index
2.) J.P. Morgan Aggregate Bond Index
3.) Standard & Poor 500 Index
4.) U.S. 30-year Treasury Bill

These series were mainly chosen because there is considerable data available to make statistically correct inferences. Of each index we used 279 data points. Each data point represented the return over the period of a month. To have comparable values we only used the first part of the data for estimation of the parameters to develop expectations for the month ahead.

4.1.2 Time frame

I will focus on the period after the end of Bretton woods in August 1971 and especially on the period with available data for all asset classes from 1988, this since data for both the bond indexes is non-existent before 1988.

The current dataset contains the following 4 indices

1	S&P GSCI - Spot index USD	SPGSCI Index
2	JPM Global Aggregate Bond - Total Return Index USD	JGAGGUSD Index
3	S&P 500 – Total Return Index USD	SPX Index
4	S&P 30 Year US Treasury Bond Futures - Total Return Index USD	SPGSCI Index

4.2 Software

4.2.1 MATLAB

MATLAB is a powerful program, which offers more flexibility and standard functions than the language and interface of Visual Basic present in versions of Excel. During my research I made extensively use of the program and would like to thank the organizations Erasmus University of Rotterdam as well as PGGM for empowering me with their licensed versions. The power of computing enabled me to conduct procedures which would not have been possible to conduct by pen and paper. During this research I have learned the language of MATLAB programming, extensive hours of debugging has enabled me to fully understand every facet of each code. Thanks to the writers of the language as well as the codes we save time and are enabled to compute beyond the limits of human capability.

To calculate the optima and minima and to estimate our parameters, MATLAB codes in combination with computing power have been used extensively. Here you find a summary of packages and codes used together with the adjustments made to them. For convenience a large share is included in the appendices.

For the estimation of the time-variant variance parameters the UCSD_GARCH toolbox of Sheppard was found to be useful, functions are specified in the next section. Levin (2004) wrote the largest share of the framework I used for estimating the Dynamic Conditional Correlation, which was also supported by the UCSD_GARCH toolbox of Sheppard. For all three estimation procedures, codes to conduct rolling windows, to loop estimation procedures, were programmed. The codes that are used to conduct the procedures of Modern Portfolio Theory makes use of the formulas which are standard present in MATLAB, just like the codes used to conduct the test procedures.

5 Model specifications and assumptions

In this section the procedures are not always described in formulas, but often in codes. When % is used in the beginning of the sentence it is a comment in MATLAB, without it is working code and it could be directly copied to an Editor or Compiler. It is possible to use the procedures in other programs, but one has to rewrite the codes used. To outline the model specifications and assumptions the following structure is used. At first the programming code is considered shortly, secondly we will test the reliability of the differences among the inputs and third we shall elaborate on the comparison between rolling Sharpe ratios. Then in the fourth section we will elaborate on the distributions used in determining an optimal allocation. During this research we limited the weight of the commodities and equity part of the portfolio to 25%, this is to cap the maximum amount of risk in the portfolio resulting from relatively higher volatility, during extreme and/or unexpected market conditions, of these asset classes.

5.1 GARCH-DCC framework

The model in MATLAB makes use of the following stepwise procedure to determine the covariance's and correlations matrices, the covariance matrix and the resulting data can be used in the Modern Portfolio Theory framework to construct the portfolio frontier.

GARCH is a time variant variance called General Autoregressive Conditional Heteroscedasticity (GARCH) variant of variance. This variant is well known since will allow to take the auto correlative as well as the mean-revering characteristics of variance into account. For the estimation procedure of variances and covariance matrix fat-tailed GARCH is used. The model is specified as follows and the maximum likelihood is carried as described in the methodology, the name and source of the package is described in section 4. For high order ARCH processes volatility can be modelled by GARCH(p, q). Generalized ARCH was developed by Bollerslev (1986), compared to ARCH(q), dependencies are permitted on q lags of past ϵ_t^2 in addition to p lags of past h_t^2 as shown below. In a GARCH(p, q) model the volatility depends on last periods volatility and the return residuals:

$$r_t = \sigma_t \varepsilon_t \qquad h_t^2 = \omega_0 + \sum_{i=1}^{q} a_i r_{t-i}^2 + \sum_{i=1}^{p} \beta_j\, h_{t-j}^2 \qquad 0 < a + \beta < 1$$

The rate of decay of the autocorrelation is measured by $\alpha + \beta$, the closer to 1, the slower the decay. Following Cont (2005) estimations on returns often yield $\alpha + \beta$ close to 1. Volatility forecast from GARCH(1, 1) can be made by repeated substitutions of the obtained parameters, the last value of the variance and the .

By the use of programming one can estimate the GARCH parameters and calculate the variance series for all asset classes.

For the Dynamic Conditional Correlation estimation procedure of the covariance matrices the DCC multivariate GARCH with quasi maximum full likelihood is used. The name and source of the package is described in section 4.

The multivariate GARCH model assumes that returns from k assets are conditionally multivariate normal with zero expected value and covariance matrix $\boldsymbol{H}_t$. The returns can be either mean zero or the residuals from a filtered time series. The proposed dynamic correlation structure by Engle and Sheppard (2001) is constructed as follows

$$r_t | F_{t-1} \sim N(0, \boldsymbol{H}_t)$$

and

$$\boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R}_t \boldsymbol{D}_t$$

H_t is the conditional covariance matrix and as a function of
D_t is the $k\ x\ k$ diagonal matrix of time varying standard deviations from univariate GARCH models with $\sqrt{h_{it}}$ on the i^{th} diagonal and h_{it} denotes the conditional variance of the *i*-th return.
R_t is the time varying correlation matrix.

The log-likelihood of this estimator can be written as

$$
\begin{aligned}
L = & -\frac{1}{2}\sum_{t=1}^{T}(\mathrm{k}\log(2\pi) + \log(|\boldsymbol{H}_t|) + \ r_t{}'\boldsymbol{H}_t{}^{-1}r_t) \\
= & -\frac{1}{2}\sum_{t=1}^{T}(\mathrm{k}\log(2\pi) + \log(|\boldsymbol{D}_t\boldsymbol{R}_t\boldsymbol{D}_t|) + \ r_t{}'\boldsymbol{D}_t{}^{-1}\boldsymbol{R}_t{}^{-1}\boldsymbol{D}_t{}^{-1}r_t) \\
= & -\frac{1}{2}\sum_{t=1}^{T}(\mathrm{k}\log(2\pi) + 2\log(|\boldsymbol{D}_t|) + \log(|\boldsymbol{R}_t|) + \epsilon_t{}^T\boldsymbol{R}_t{}^{-1}\epsilon_t)
\end{aligned}
$$

where $\epsilon_t \sim N(0, R_t)$ are the residuals standardized by their conditional standard deviation. We propose to write the elements of D_t as univariate GARCH models, so that

$$h_{it} = \omega_i + \sum_{p=1}^{P_i}\alpha_{ip}r_{it-p}^2 + \sum_{q=1}^{Q_i}\beta_{iq}h_{it-q}$$

for $i = 1,2, \dots, k$ with the usual GARCH restrictions for non-negativity and stationarity being imposed, such as non-negativity of variances and $\sum_{p=1}^{P_i}\alpha_{ip} + \sum_{q=1}^{Q_i}\beta_{iq} < 1$. The subscripts are present on the individual P and Q for each series to indicate that the lag lengths chosen need not be the same. The specification of the univariate GARCH models is not limited to the standard GARCH (p, q), but can include any GARCH process with normally distributed errors that satisfies appropriate stationary conditions and non-negativity constraints.

$$\boldsymbol{Q}_t = \left(1 - \sum_{m=1}^{M}\alpha_m - \sum_{n=1}^{N}\beta_n\right)\bar{\boldsymbol{Q}} + \sum_{m=1}^{M}\alpha_m\,(\boldsymbol{\epsilon}_{t-m}\boldsymbol{\epsilon}^T{}_{t-m}) + \sum_{n=1}^{N}\beta_n\,\boldsymbol{Q}_{t-n}$$

$$\boldsymbol{R}_t = \boldsymbol{Q}_t^{*-1}\boldsymbol{Q}_t\boldsymbol{Q}_t^{*-1}$$

where $\bar{Q}$ is the unconditional covariance of the standardized residuals, $\boldsymbol{\epsilon}_t$ resulting from the first stage estimation, and

$$\boldsymbol{Q}_t^* = \begin{bmatrix} \sqrt{q_{1,1}} & 0 & 0 & \dots & 0 \\ 0 & \sqrt{q_{2,2}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \sqrt{q_{k,k}} \end{bmatrix}$$

so that $\boldsymbol{Q}_t^*$ is a diagonal matrix composed of the square root of the diagonal elements of $\boldsymbol{Q}_t$. The typical element of $\boldsymbol{R}_t$ will be of the form

$$\rho_{i,j,t} = \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}$$

The following useful result from linear algebra simplifies finding the necessary conditions for R_t to be positive definite and hence a correlation matrix.

5.2 Modern Portfolio Theory framework

The preprogrammed codes which are prominent in a standard MATLAB installation we don't have to make many adjustments. Important is the adjustments of the constraints. During this research we limited the weight of the commodities and equity part of the portfolio to 25% and the whole portfolio to 100% so that we do not use leverage. We also limited the use of the risk-free asset, since the risk-free asset has zero risk it does not has an effect on our Sharpe ratio and can only be used to borrow or lend. Dealing with the adjustment of the Sharpe ratio, when including the risk-free asset, is not part of this research.

The standard computation of the mean-variance frontier uses a brute-force approach. There are techniques by which we can compute the set of efficient portfolios and efficient mean-variance combinations associated with given $\underline{\mu}$ and$\underline{V}$. The goal of the proposed model is to determine the optimal minimum variance portfolio for a given level of risk subjecting to the following constraints written out by Balvers (2001) and Renström (2002).

$$\underline{\boldsymbol{w}}^T\underline{\boldsymbol{\mu}} = r$$
$$\underline{\boldsymbol{w}}^T\underline{\boldsymbol{l}} = 1$$

Where r denotes the expected portfolio return, $\underline{\boldsymbol{l}}$ represents a $1 \times n$ column vector of 1's and the sum of the weights is equal to 1, or the whole portfolio.

The minimization problem can be solved using the Lagrange multiplier method with multipliers λ_1 and λ_2 for the constraints. Set

$$L = \underline{\boldsymbol{w}}^T\underline{\boldsymbol{V}\boldsymbol{w}} - \lambda_1\left(\underline{\boldsymbol{w}}^T\underline{\boldsymbol{\mu}} - r\right) - \lambda_2\left(\underline{\boldsymbol{w}}^T\underline{\boldsymbol{l}} - 1\right)$$

The first-order condition with respect to $\underline{w}$ is

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = 2\underline{\boldsymbol{w}\boldsymbol{V}} - \lambda_1\underline{\boldsymbol{\mu}} - \lambda_2\underline{\boldsymbol{l}}$$

By pre multiplying the Lagrangian by ½ we get more convenient expressions and

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = \underline{\boldsymbol{w}\boldsymbol{V}} - \lambda_1\underline{\boldsymbol{\mu}} - \lambda_2\underline{\boldsymbol{l}}$$

Solving this derivative for the optimal portfolio weights is done by applying the following steps:
The result is a vector and at the optimum this vector must be equal to the null vector, so we have the necessary first-order condition for an optimum.

$$\frac{\partial L}{\partial \underline{\boldsymbol{w}}} = \underline{\boldsymbol{w}\boldsymbol{V}} - \lambda_1\underline{\boldsymbol{\mu}} - \lambda_2\underline{\boldsymbol{l}} = 0$$

The solution is then obtained by pre multiplying the necessary condition by the inverse to the variance-covariance matrix

$$\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{w}\boldsymbol{V}} = \lambda_1\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} - \lambda_2\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$$

$$\Leftrightarrow$$

$$\underline{\boldsymbol{w}}^* = \lambda_1 \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$$

Based on the fact that $\underline{\boldsymbol{V}}$ is positive definite we can conclude that $\underline{\boldsymbol{w}}^*$ minimizes the variance and that the solution obtained is unique.

Pre multiplying the first-order condition by $\underline{\boldsymbol{w}}^*$ and using constraints gives

$$\sigma_p^2 = \lambda_1 r + \lambda_2$$

To replace the Lagrange multipliers we pre multiply vector $\underline{\boldsymbol{w}}^*$ by the vector of expected returns, $\underline{\boldsymbol{\mu}}$, and separately by the unit vector, $\underline{\boldsymbol{l}}$, to obtain

$$\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{w}}^* = \lambda_1 \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$$

$$\underline{\boldsymbol{l}}^T \underline{\boldsymbol{w}}^* = \lambda_1 \underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} + \lambda_2 \underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$$

$\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}}$, $\underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$ and $\underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$ are scalars and we can label them A, B, and C.
$\underline{\boldsymbol{l}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} = \underline{\boldsymbol{\mu}}^T \underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}} = \mathrm{C}$ And $D = AB - C^2$

Then

$$r = \lambda_1 A + \lambda_2 C$$

$$1 = \lambda_1 C + \lambda_2 B$$

Solve these to obtain

$$\lambda_1 = \frac{Br - C}{D}, \lambda_2 = \frac{A - Cr}{D}$$

Finally substitute λ_1 and λ_2 into the function for the optimal weights and we have the optimal weights as function of the vector of expected returns, the unity vector and the variance-covariance matrix.

$$\underline{\boldsymbol{w}}^* = \frac{Br - C}{D}\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} + \frac{A - Cr}{D}\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}$$

$$= \frac{A\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}} - C\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}}}{D} + \frac{B\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{\mu}} - C\underline{\boldsymbol{V}}^{-1}\underline{\boldsymbol{l}}}{D} r$$

And substitution of λ_1 and λ_2 in σ_p^2 yields an explicit expression of the portfolio frontier:

$$\sigma_p^2 = (Br^2 - 2Cr + A)/D$$

Plot the μ_p against σ_p^2 to draw the efficient frontier. The intersection between the line with origin at the risk-free rate and the efficient frontier is known as the tangency or optimal portfolio. For a more extensive elaboration on the theory please refer to Ingersoll (1987).

6 Results

6.1 Returns

When compare the mean of the returns on basis of the Root Mean Squared Error statistic, we can determine which forecast produced the lowest errors from the realized values. From this statistic we cannot make inferences about the direction of change which caused the error, since differences in returns are squared. During the research we used a method in which we estimate statistics for a certain time window over a time period from January 1988 up till April 2011 or equal to returns of 279 months. A shorter time period for calculating averages was used for the first time periods to deal with the missing information. In appendix 9.1.1 table 1 also the statistics for other common used time windows are shown.

From table in table 1, we can conclude that each asset class has its own individual setting to create the optimal forecasts. For commodities it is optimal to use the last month's return to forecast the return one month ahead, for long term U.S. government bonds the period is 3 months. For corporate bonds use the one year averaged return and considering equity the input for the mean-variance framework should be the 10 years average return. Here I would remind you that the knowledge obtained over time could increase your mean and in reality you could differ from the proposed mean, when this process is registered you could compare the mean in similar fashion. Though we used the Root Mean Squared Error to compare return models we should proceed by correcting for the economic loss resulting from the volatility of our return variable. The statistic is used for testing whether the performance of the forecast series is significantly different from that of the random walk is the Diebold-Mariano statistic. The most important feature of the DM test statistic is that it corrects for economic loss resulting from the volatility of the tested variable. By comparing the DM test statistics we will be able to assess which model predicts the returns better. The statistics are calculated using the squared error and absolute error loss functions. From tables in appendix 9.3 we find or commodities that the random walk followed by the historical sample average for investment grade bonds and treasury bonds we should use the historical sample average. Comparing the models with the DM statistic for equity we should use the historical average over 10 years.

6.2 Variances

We must note that the RMSE statistics for GARCH are distorted since our GARCH forecast is autoregressive for this statistic we need a Diebold-Mariano statistic. This statistic corrects for autoregressive elements.
When compared to the squared realized return, we see that we should use the historical sample variance for each asset except for the government bonds were we should use the historical 10 year average deviation. We see that GARCH underperforms, one much ask the question if both measures are comparable to the squared returns. Since there is no real observable variance it goes

6.3 Correlations

For the correlation estimates we used two methods one is to assume that the historical average is relevant and one in which we make use of the Dynamic Conditional Correlation. Comparison of the correlation coefficients, per pair requires considerable computing power. We compared the MPT with and without GARCH-DCC for the 2 year and 10 year rolling window. This will be discussed in the following section. Resulting from the Realized Rolling Sharpe ratios we can conclude that the GARCH DCC framework adds considerably to the realized Sharpe ratio.

6.4 Model comparisons

For comparison of models, we should compare the efficient frontier and especially the location of the tangency portfolio of the portfolios that uses the different forecasted returns, variances and correlations and compare this to a situation with full knowledge, in which we insert the realized returns and realized covariance matrix.

Now that we build the model in MATLAB we can put the model to practice. This is done by using historical data up to the start of 1988 and by calculating the Sharpe ratios over time by the use of rolling windows. By using different settings and comparing these we will configure the model so that the likelihood of finding the best combination of the assets is maximal, which will bring us the optimal Sharpe ratio for each month ahead. Since reliable methodology decision for asset managers should be based on fast amounts of data and depends on the investment horizon, we will show the Realized Rolling Sharpe ratios over a measurement period of 2 years and 10 year. The second measurement we use to compare the portfolios is the Maximum Draw Down measure which measures the cumulative consequent losses. In which the 2 year measure is more vulnerable for dynamic environments which could turn out to be positive or negative depending on the macro-economic business environment. These results are fully based on the tangency portfolios.

The settings which we will test are the following:
The MPT-framework with and without
- Historical Covariance matrix,
- GARCH-DCC Covariance matrix
The MPT-framework with data estimation and performance measurement period of
- 2 year
- 10 year
With and without full knowledge of the returns 1-month ahead.

The forecasts of the MPT-framework versus the realizations of the actual investments in the weights defined by the framework are measured. The first setting is so called historical setting, an unweighted historical average over the estimation period, is used to estimate the most likely returns, variances and correlations. Next we will consider the Realized Rolling Sharpe statistics which are plotted in the figures. The Sharpe ratio was created to answer the question "Given the same amount of risk, which investment provides me with the highest reward". To do this the Sharpe ratio balances the returns in excess of a risk free benchmark with the standard deviation of the return set. This provides a uniform risk platform which funds with different risk levels can benchmark against. We used Realized Rolling Sharpe ratios which imply that we computed the risk and return that would have resulted from using the investments weights recommended by the model.

Considering the settings under full knowledge of the returns and we are able test whether the historical or the GARCH-DCC framework performs better. Comparing the Rolling Sharpe ratios we find mixed results. When we look at a 10 year horizon the GARCH-DCC framework performs slightly better up to 2006 considered the performance measured over 10 years. When looking at the Sharpe ratio of the 2 year horizon this GARCH-DCC framework also slightly out performs the framework using the historical covariance matrix, this is shown in figures 2 and 4 in appendix 9.2.

7 Concluding Remarks

During the last few years many inventions have been made in the area of financial economics, during the time of Fisher, 100 years ago, there were no appropriate quantitative methods available to estimate dynamic conditional variances and correlations. These methods have been developed in the last 20 years. In this research paper we combined some of the important advances. Many of the firms active in the financial sector still make use of the MPT framework as developed by Markowitz, which imply that they all make use of linear optimization techniques. These firms differ in the method in which they estimate or gather these inputs. Goal of this research was to develop a model and especially one which delivers an optimal asset allocation. This optimal asset allocation is defined in portfolio theory as the portfolio which delivers the best return for each unit of risk or has the lowest amount of risk relative to return. This portfolio is known as the tangency portfolio. This model improves the average Sharpe ratio over time up to 50% on a 2 year time frame, but the outperformance seems time-invariant. Our results with the 2 year and 10 year GARCH-DCC MPT framework are convenient, but it is questionable if the benefits out weights the costs of complexity. We must also see the fact that a 1-month switch of 50 percent of the portfolio from corporate bonds to other assets classes requires an extreme amount of flexibility. And let's not forget that we did not take into account the costs of switching, we would incur up till 0.5% in costs and possibly costs resulting from illiquidity. The results are too weak to claim that the GARCH-DCC framework is considerably stronger than using historical covariance matrix. Improvement of the return expectations could do the trick. This thesis shows a framework which is simple, easy to adjust and which can be tailored to your preferences.

For further research I have six recommendations, one should test the framework for other inputs, like risk factor based assets Secondly one could adjust the framework for the inclusion of liabilities and calculations for the surplus of the coverage ratio of pension funds and insurance companies. Another interesting subject to work out in this framework is risk budgeting in which one aims to keep the risk of the portfolio below a target risk level. Fourth recommendation is to work out the framework on the daily and intraday level for management of portfolios and trading books. The fifth recommendation is to improve the expectation of the returns and accompanying volatility by using macro-economic figure density, Hamilton's Regime-switching, Kalman filters. Last but not least I would recommend to look further than macro-economic variables to forecast risk and return, naming natural factors such as the influence of the solar system on human behavior, behavioral factors acting on the humans who deal with the financial markets and which determine trends and cycles and other cycles such as the Kondratieff cycles, Armstrong cycles and political cycles. For this I would like to refer to experts like Gann (1909), Elliot, and more recently Armstrong and Steenhouwer (2005),

8 References

Altegris Academy, Research Series, (2011), Managed Futures: Staying the Course
Short-term Drawdowns, Long-term Focus, may 2011
Balvers, R.J., (2001), Foundations of asset pricing, Chapter II Mean-Variance portfolio choice
Bank of International Settlements, (2011), *BIS Quarterly Review March 2011*
Bollerslev, T., (1986), Generalized Autoregressive Conditional Heteroscedasticity
Bollerslev, T., (1990), Modeling the coherence in short-run nominal exchange rates, a multivariate generalized arch model
Bollerslev, T., Engle, R.F. & Wooldridge, J.M., (1988), A Capital Asset Pricing Model with time-varying covariances
Diebold, F.X. and Mariano, R., (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, 253-265.
Dimand, R.W., (2007), Irving Fisher and financial economics: the equity premium puzzle, the predictability of stock prices, and intertemporal allocation under risk
Chueng, Y-W., Chinn, M.D., and Pascual, A.G., (2005), Empirical exchange rate models of the Nineties
Cont, R., (2005), Volatility clustering in financial markets: Emperical Facts and Agent-Based Models
Cowan, J., (2008), article: "Brain waves patterns themselves after rhythms in nature", The university of Chicago news office, February 15, 2008 (http://www-news.uchicago.edu/releases/08/080215.cowan.shtml)
Diebold, F.X. and Mariano, R.S., (1995), Comparing predictive accuracy, *Journal of Business and Economic Statistics*, 13, 253-263.
Elton, E., and Gruber, M., (1973), Estimating the dependence structure of share prices-implications for portfolio selection, *Journal of Finance*, 28:1203-32
Engle, R.F., and Sheppard, (2001), Theoretical and Empirical properties of Dynamic Conditional Correlation Multivariate GARCH
Engle, R.F., Shephard, N. and Sheppard, K. (2008), Fitting vast dimensional time-varying covariance models, Oxford Financial Research Centre, *Financial Economics Working Paper* n. 30.
Engle, R.F. and Kroner, K.F., (1995), Multivariate simultaneous generalized ARCH, Econometric Theory, 11, 122-150.
Fisher, I, (1906), The Nature of Capital and Income, New York: Macmillan, reprinted in: Fisher (1997), Vol. 2.
Fisher, I, (1907), The Rate of Interest, New York: Macmillan, reprinted in: Fisher (1997), Vol. 3
Levin, F., (2004), Looking for independence: reducing the risk of groups of long-short strategies during market swings through correlation forecasts, LSE (get perm)
Gann, W.D., (1909), Ticker and Investment Digest, Interview by R. D. Wyckoff, Volume 5, 2, 54
Hicks, J.R., (1935), A Suggestion for Simplifying the Theory of Money, *Economica*, 1-19
Hull, J., 2009. Options, Futures and Other derivatives
Ingersoll, J.E., Jr., (1987), Theory of Financial Decision Making, Rowman & Littlefield, 82-87.
Leavens, D.H., (1945), Diversification of Planning, *Trusts and Estates*, 80, 469–73.
Levin, F., (2004), Looking for independence: reducing the risk of groups of long-short strategies during market swings through correlation forecasts
Levy, H. and Markowitz, H.M., (1979), Approximating expected utility by a function of mean and variance, *American Economic Review*, June.
Lucas, R.E., Jr., (1987), Models of Business Cycles, Oxford: Basil Blackwell, London and New York, ISBN 0-631-147918.
Meese, R. and Rogoff, K., (1983), Empirical exchange rate models of the seventies, do they fit out of sample
Maes, I., (1991), On the Origins of Portfolio Theory, *Kyklos*, 44, 3–18
Markowitz, H.M., (1952), Portfolio selection, *The Journal of Finance*, 7(1): 77-91
Markowitz, H.M., (1990), Foundations of portfolio theory, Nobel lecture
Markowitz, H.M., (1999), The early history of portfolio theory: 1600-1960, *Financial Analyst Journal*, 55, 4, 5-16
MISFA, (2011), Credit market statistics

Plato, (368 B.C.), Theatetus, (http://plato.stanford.edu/entries/plato-theaetetus/)
Reider, R., (2009), Volatility forecasting I: GARCH models
Renström, T., (2002), Financial markets: Theory & Evidence, (http://www.econ.rochester.edu/Wallis/Renstrom/Eco217.html), lecture 4 and 5
Robbins. L., (1932), An Essay on the Nature and Significance of Economic Science, Macmillan, London
Scherer, B., (2002), Portfolio construction and risk budgeting
Scholes, M.S. under pseudonym Cochrane, J.H., (2005), Asset Pricing, Princeton University Press, ISBN 978-0-691-12137-0.
Sharpe, W.F., (1990), Asset allocation, management style and performance measurement
Sheppard, K., (2010), Financial Econometrics MFE MATLAB introduction
Steenhouwer, H., (2005), Macroeconomic scenarios and reality: a frequency domain approach for analyzing historical time series and generating scenarios for the future, Optima Grafische Communicatie, Rotterdam, ISBN 90-8559-091-4.
The Bank of International Settlements, (2011), BIS Quarterly Review June 2011
TheCityUK, (2011), Financial markets series: commodities trading
Thastrom, P., (2008), Dynamic Conditional Correlation models in a multiple financial asset portfolio
Thompson, P.A., (1990), An MSE statistic for comparing forecast accuracy across series, *International Journal of Forecasting*, Vol. 6, pp219-227
Tse, Y.K. and Tsui, K.C., (1999), A multivariate GARCH model with time-varying correlations
Tversky, A. And Kahneman, D., (1974), Judgment under uncertainty: Heuristics and Biases
Versijp, P., (2011), Advanced Investments, lecture notes, Erasmus University of Rotterdam Course
Yilmaz, T., (2010), Improving Portfolio Optimization by DCC and DECO GARCH, Evidence from Istanbul Stock Exchange
Zhang, K. and Chan, L., (2010), Efficient factor GARCH models and factor-DCC mode
Ziemba W.T. and Mulvey J.M., (1998), Worldwide asset and liability modeling, Cambridge University Press, ISBN 0-521-57187
Zivot, E., (2004), Notes on Forecasting

9 Appendices

9.1 Result tables: Root Mean Squared Error and Diebold-Mariano statistics

9.1.1 Root Mean Squared Error test statistics

Table 1: The mean: Root Mean Squared Error statistics

Measurement	Commodities	Government bonds	Corporate bonds	Equity
Random walk (0 mean)	0,004538	0,005095	0,011912	0,031408
Historical sample	0,004139	0,002877	0,007139	0,038371
Historical average 10 years	0,003471	0,002444	0,007067	0,031061
Historical average 1 year	0,004111	0,002382	0,006458	0,040390
Historical average 6 months	0,004015	0,002326	0,006488	0,042827
Historical average 3 months	0,003193	0,002098	0,007075	0,034517
Historical average 1 month	0,002925	0,002413	0,007001	0,043703

The smallest Root Mean Squared Errors are highlighted, for information about the statistic please refer to the methodology in section 2.6.1 and 2.6.2.

9.1.2 Diebold-Mariano test statistics

Table 2: Diebold-Mariano return

Commodities

Test statistic	Random walk (0 mean)	Historical sample average	Historical average 10 years	Historical average 1 year	Historical average 6 months	Historical average 3 months	Historical average 1 month
Random walk (0 mean)	1.0000	1.2607	1.4624	2.2294*	2.2611*	3.0151**	4.7068**
Historical sample average	-1.2607	1.0000	0.4527	1.9323	1.9085	2.9552**	4.7364**
Historical average 10 years	-1.4624	-0.4527	1.0000	1.8500	1.9115	2.8512**	6.2793**
Historical average 1 year	-2.2294*	-1.9323	-1.8500	1.0000	0.5227	2.8512**	6.1084**
Historical average 6 months	-2.2611*	-1.9085	-1.9115	-0.5227	1.0000	2.4017*	4.5568**
Historical average 3 months	-3.0151**	-2.9552**	-2.8512**	-2.8512**	-2.4017*	1.0000	4.4800**
Historical average 1 month	-4.7068**	-4.7364**	-6.2793**	-6.1084**	-4.5568**	4.4800**	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 3: Diebold-Mariano return

JPM Global Aggregate Bond Index

Test statistic	Random walk (0 mean)	Historical sample average	Historical average 10 years	Historical average 1 year	Historical average 6 months	Historical average 3 months	Historical average 1 month
Random walk (0 mean)	1.0000	-0.3395	0.3564	0.7254	2.0205*	3.7024**	5.9473**
Historical sample average	0.3395	1.0000	0.0675	0.7912	2.0305*	3.7595**	6.0949**

Historical average 10 years	-0.3564	-0.0675	1.0000	0.8935	2.2572*	4.0388**	6.2301**
Historical average 1 year	-0.7254	-0.7912	-0.8935	1.0000	2.2938*	3.3514**	5.6157**
Historical average 6 months	-2.0205*	-2.0305*	-2.2572*	-2.2938*	1.0000	2.4805*	5.6248**
Historical average 3 months	-3.7024**	-3.7595**	-4.0388**	-3.3514**	-2.4805*	1.0000	5.2322**
Historical average 1 month	-5.9473**	-6.0949**	-6.2301**	-5.6157**	-5.6248**	-5.2322**	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 4: Diebold-Mariano return

S&P 500

Test statistic	**Random walk (0 mean)**	**Historical sample average**	**Historical average 10 years**	**Historical average 1 year**	**Historical average 6 months**	**Historical average 3 months**	**Historical average 1 month**
Random walk (0 mean)	1.0000	-1.3172	-1.8149	1.3921	2.5503*	3.5285**	4.2393**
Historical sample average	1.3172	1.0000	-1.3486	4.1395**	4.0929**	4.3123**	4.5540**
Historical average 10 years	1.8149	1.3486	1.0000	4.7063**	4.4930**	4.5444**	4.6272**
Historical average 1 year	-1.3921	-4.1395**	-4.7063**	1.0000	2.4814*	3.5416**	4.3080**
Historical average 6 months	-2.5503*	-4.0929**	-4.4930**	-2.4814*	1.0000	2.8132**	4.3117**
Historical average 3 months	-3.5285**	-4.3123**	-4.5444**	-3.5416**	-2.8132**	1.0000	3.9568**
Historical average 1 month	-4.2393**	-4.5540**	-4.6272**	-4.3080**	-4.3117**	-3.9568**	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 5: Diebold-Mariano return

US Treasury bonds 30y

Test statistic	**Random walk (0 mean)**	**Historical sample average**	**Historical average 10 years**	**Historical average 1 year**	**Historical average 6 months**	**Historical average 3 months**	**Historical average 1 month**
Random walk (0 mean)	1.0000	-0.6210	0.5443	1.9182	3.5484**	5.3648**	5.2836**
Historical sample average	0.6210	1.0000	0.6910	2.8567**	4.2859**	5.5511**	5.3166**
Historical average 10 years	-0.5443	-0.6910	1.0000	2.8259**	4.3314**	5.5706**	5.3211**
Historical average 1 year	-1.9182	-2.8567**	-2.8259**	1.0000	2.854**	5.2443**	5.0298**
Historical average 6 months	-3.5484**	-4.2859**	-4.3314**	-2.8540**	1.0000	3.9244**	4.6711**
Historical average 3 months	-5.3648**	-5.5511**	-5.5706**	-5.2443**	-3.9244**	1.0000	3.7516**
Historical average 1 month	-5.2836**	-5.3166**	-5.3211**	-5.0298**	-4.6711**	-3.7516**	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 6: Diebold-Mariano variance

Commodities

Test statistic	GARCH	Historical sample	Historical 10 years
GARCH	1.0000	-4.4704**	-1.6139
Historical sample	4.4704**	1.0000	-2.9423
Historical 10 years	1.6139	2.9423	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 7: Diebold-Mariano variance

JPM Global Aggregate Bond Index

Test statistic	GARCH	Historical sample	Historical 10 years
GARCH	1.0000	-1.4850	-0.4731
Historical sample	1.4850	1.0000	1.9031
Historical 10 years	0.4731	-1.9031	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 8: Diebold-Mariano variance

S&P 500

Test statistic	GARCH	Historical sample	Historical 10 years
GARCH	1.0000	-2.1248*	-1.9658*
Historical sample	2.1248*	1.0000	1.2617
Historical 10 years	1.9658*	-1.2617	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

Table 9: Diebold-Mariano variance

US Treasury bonds 30y

Test statistic	GARCH	Historical sample	Historical 10 years
GARCH	1.0000	0.5395	0.8571
Historical sample	-0.5395	1.0000	1.0181
Historical 10 years	-0.8571	-1.0181	1.0000

** = Significant at the 1% level; * = Significant at the 5% level; The negative D-M statistics indicate that the model in the first column is worse, a positive values indicate that the model in the first row is worse, for information about the statistic please refer to the methodology in section 2.6.3.

9.2 Result figures: Sharpe ratio and Maximum Draw Down statistics

9.2.1 10 year time frame

Figure 1: 10 year Sharpe ratios of Asset classes

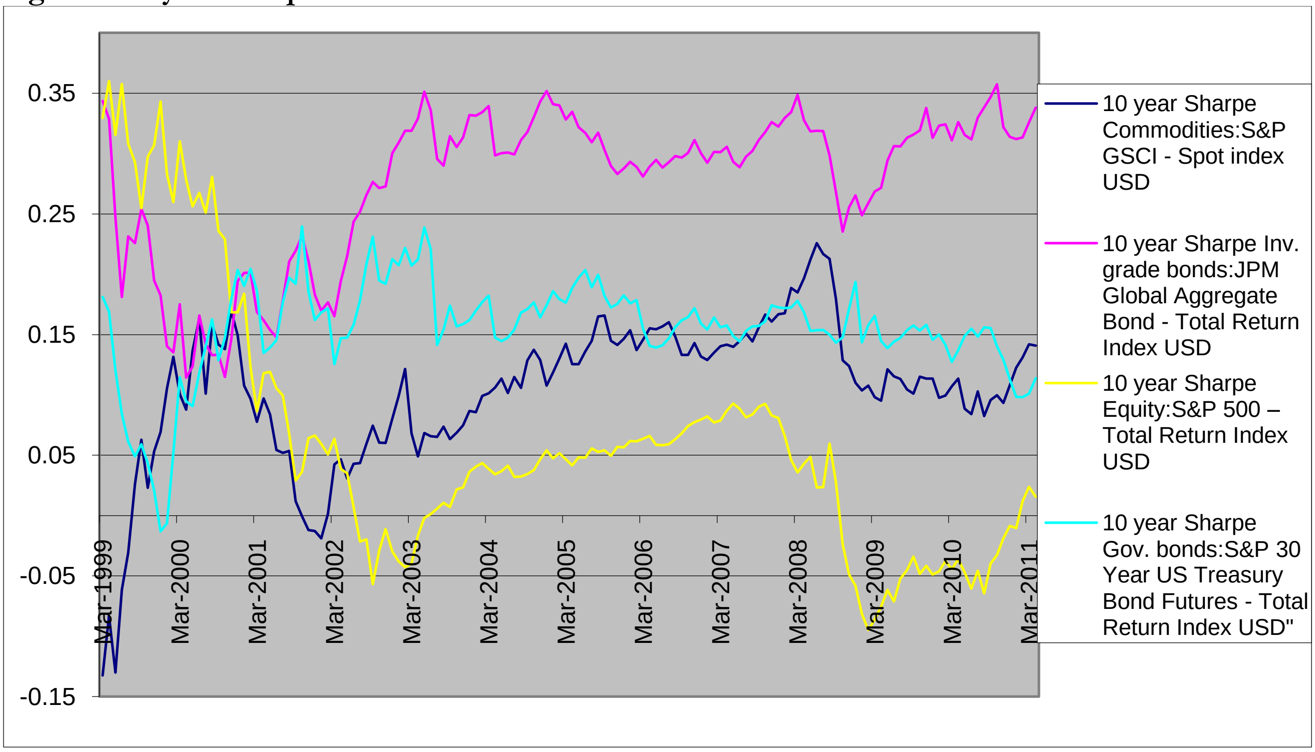
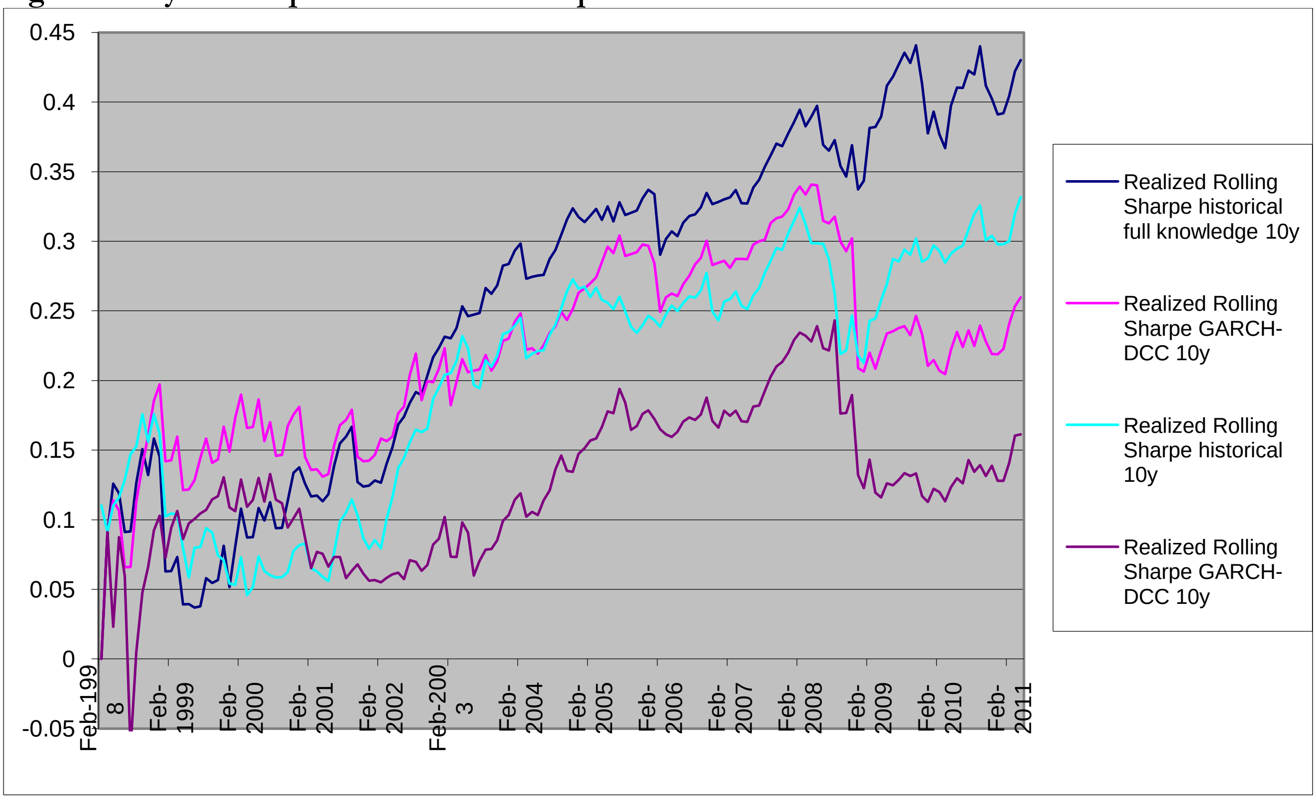


Figure 2: 10 year Sharpe ratios of allocated portfolios



9.2.2 2 year time frame

Figure 3: 2 year Sharpe ratios of Asset classes

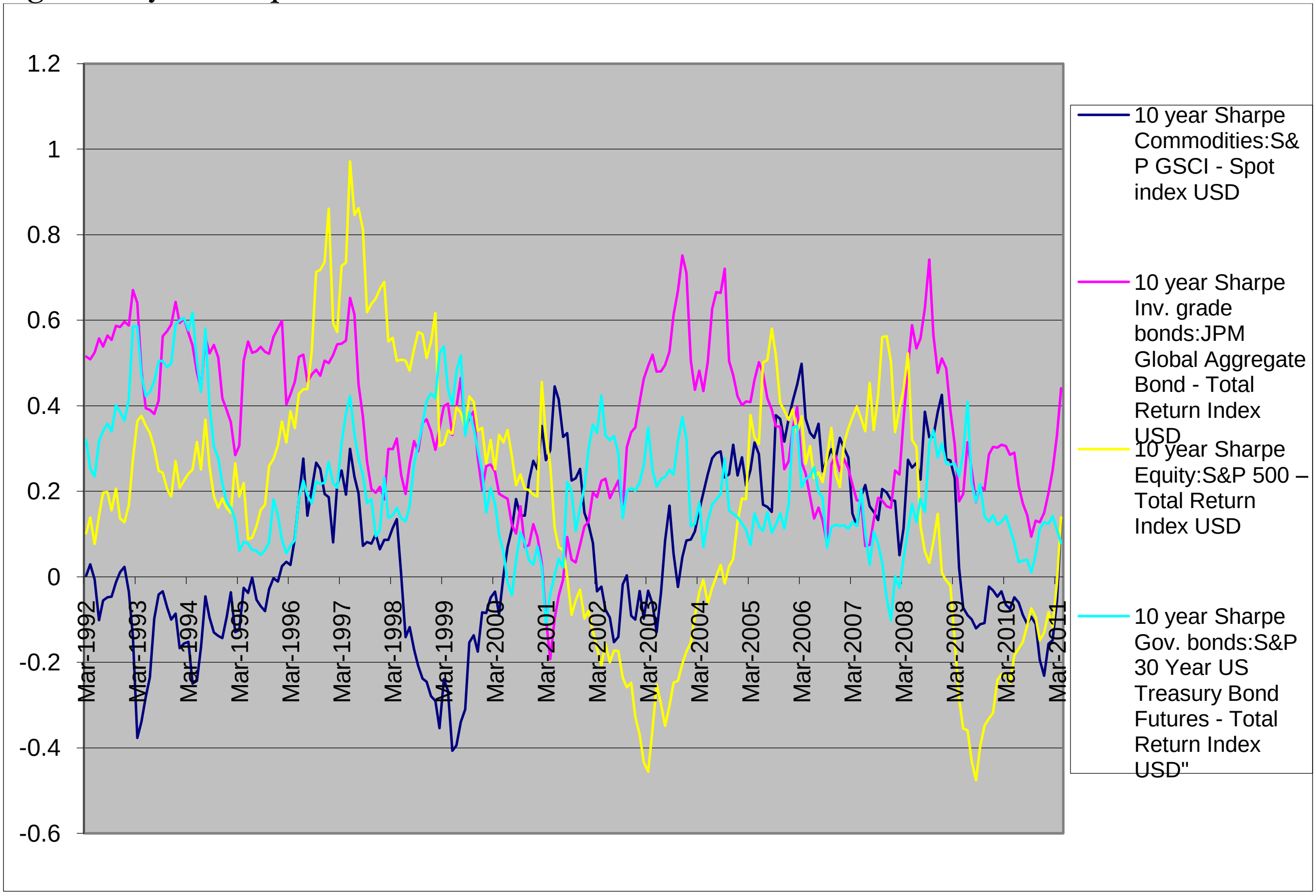


Figure 4: 2 year Sharpe ratios of allocated portfolios

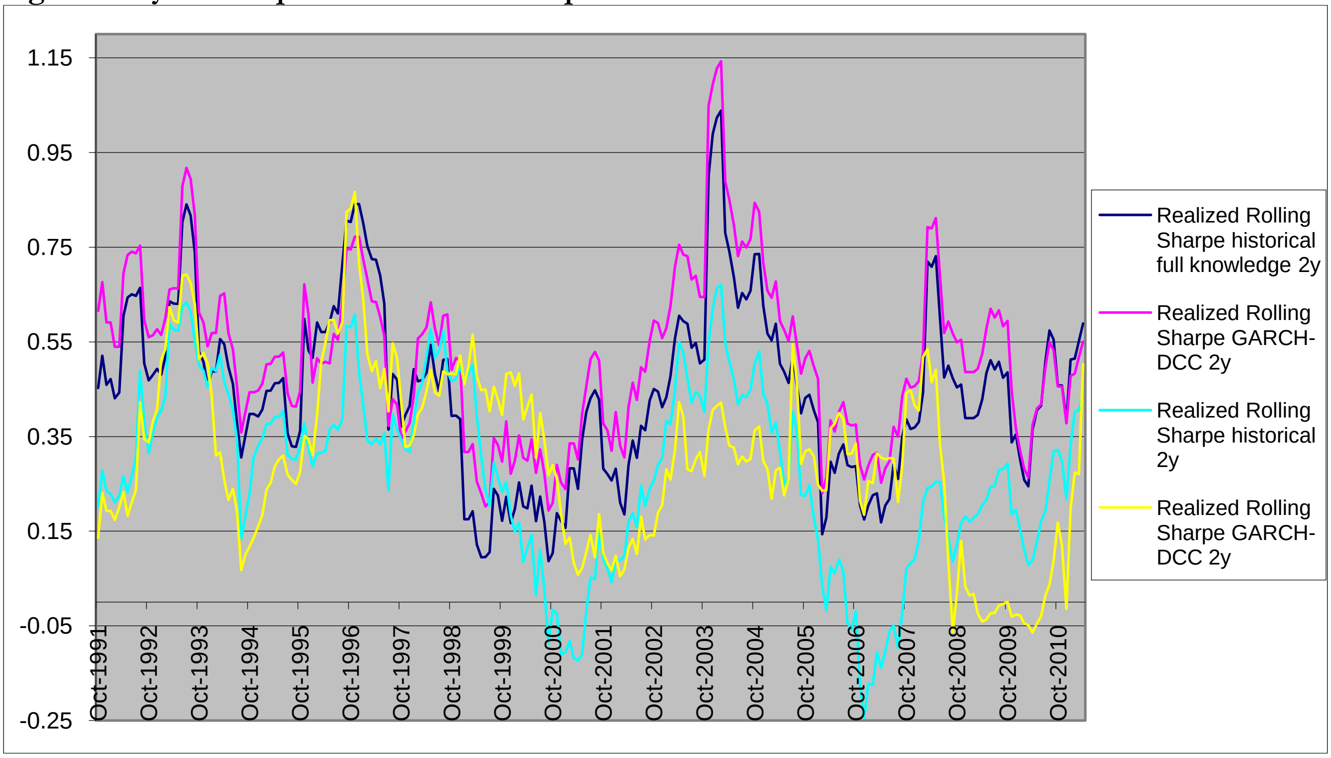
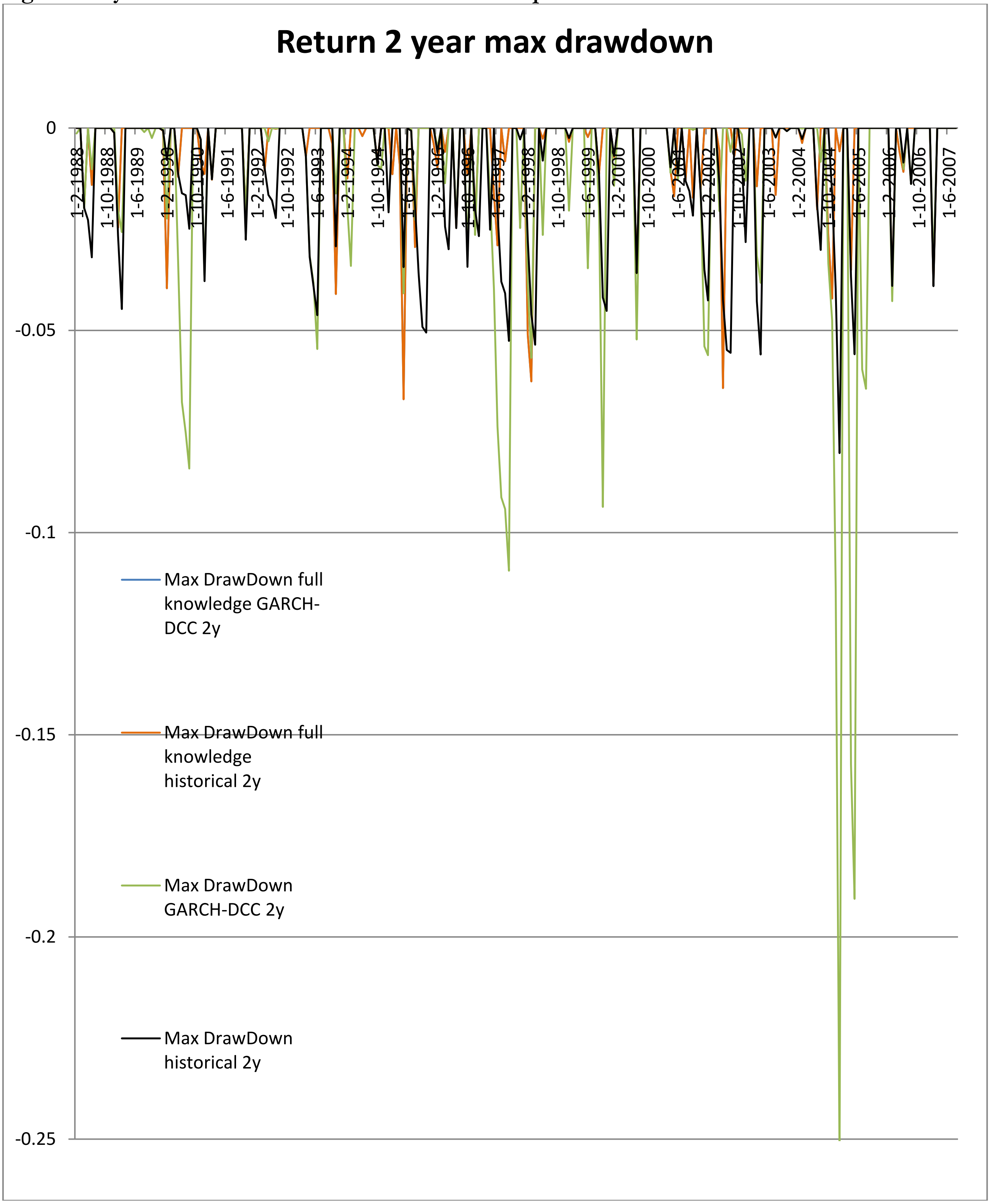


Figure 5: 2 year Maximum Draw Downs of allocated portfolios



9.3 MATLAB Codes

9.3.1 GARCH-DCC framework

```
clear all;
clc; % clear command startwindow
tic % starts a stopwatch timer, toc to print elapsed time


% create a file with asset returns
% adjust datafile names of %portfolio %load
% change number of assets in portfolio function and constraints %portfolio,
%efficient
% adjust riskless rate, borrowrate and riskaversion level %optimal
% adjust current settings %current

% GARCH-DCC predictions, out of sample, sliding (jumping) startwindow
(F.L.) %

%load, LOADING DATA SET
load 'assets4.csv' -ascii % define datafile and format
load 'data1mrate.csv' -ascii % short term 1m interest rate
date = assets4(1:279,1); % define dates column
data = assets4(1:279,[2 3 4 5]); %[2:S&P GSCI spot 3:JPM Global Aggregate
Bond 4:S&P 500 5:30 year Treasury yield]); % define data columns 2:5
rate = data1mrate(1:279,2);
clear assets4;
[T,k] = size(data); % the first "dimension" is the number of rows, and the
second "dimension" k is the number of columns

%inputs, INPUTS AND RESTRICTIONS
ar = 10; % 5 10* 15 AR terms 1~10 %*chosen in F. Levin paper* and text
ma = 0;  % 0* no MA terms %*chosen in F. Levin paper* and text
p = 1; % 1* 1 2 3x 5x % chosen in F. Levin paper
q = 1; % 1* 3 2 3x 5x % chosen in F. Levin paper
dccP = 1;
dccQ = 1;
archP = 1;
garchQ = 1;  % settings carried over from ARMA GARCH

DCC_garchQ=ones(1,k)*garchQ; % k = data series
DCC_archP=ones(1,k)*archP; % k = data series

offset = 1;
months = 120; %% months or datapoints, otherwise change to 12 (months)

Est_GARCH = ones(T,k) * 0;
F_pred_ARMA = ones(T,k) * 0;
F_pred_GARCH = ones(T,k) * 0;
F_pred_CORR = ones(k,k,T) * 0;
F_param1_DCC = ones(T+1,14) * 0;
F_pred_COV = ones(k,k,T) * 0;
F_portfolio = ones(T,k) * 0;

startwindow = 0; % set to datawindow+2 below the end of the dataset for a
quick test
Tend = 120; % length datawindow
lag = 0;
```

```
while (Tend < T)

    col = 1;

    while (col <= k)

        % TRAIN/ ESTIMATION OF PARAMETERS

        Tstart = max(1, startwindow + 1);
        Tend = startwindow + months;


            % GARCH

            [GARCH_parameters, GARCH_likelihood, GARCH_stderrors,
GARCH_robustSE, GARCH_ht, GARCH_scores] =
fattailed_garch(data(Tstart:Tend,col:col) , p , q , 'NORMAL');
            GARCH_likelihood;
            GARCH_stderrors;
            GARCH_robustSE;
            clear GARCH_scores;


        % PREDICT

        startwindow = startwindow + 1;
        Tend2 = min(T+1, startwindow + months);


            % GARCH

            predict_GARCH = GARCH_parameters(1,1) +
GARCH_parameters(2,1)*data(Tend2-1,col)^2 +
GARCH_parameters(3,1)*F_pred_GARCH(Tend2-1:Tend2-1,col:col); % Forecast t+1
%GARCH_parameters(2,1)*data(Tend2-1,col)^2


        % STORE
            F_param1_GARCH(Tend:Tend,col:col) = GARCH_parameters(1:1,:)';
            F_param2_GARCH(Tend:Tend,col:col) = GARCH_parameters(2:2,:)';
            F_param3_GARCH(Tend:Tend,col:col) = GARCH_parameters(3:3,:)';

            Est_GARCH(Tstart:Tend,col:col) = GARCH_ht(:,:);
            F_pred_GARCH(Tend+1:Tend+1,col:col) = predict_GARCH(1,:); %
Forecast t+1

        col = col + 1;
        startwindow = startwindow - 1;

    end


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
    %dcc, DCC estimation
```

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
        Tstart = max(1, startwindow);
        Tend = startwindow + months;

    % DCC TRAIN

        [DCC_parameters, DCC_loglikelihood,
DCC_Ht]=dcc_mvgarch(data(Tstart:Tend,:),dccP,dccQ,archP,garchQ); %
estimates parameters DCC
        clear DCC_loglikelihood;

        [pred_DCC_logL, pred_DCC_Rt,
pred_DCC_likelihoods]=dcc_mvgarch_full_likelihood(DCC_parameters,
data(Tstart:Tend,:), DCC_archP,DCC_garchQ,dccP,dccQ); % input parameters
output Rt + forecast
        clear pred_DCC_logL;
        clear pred_DCC_likelihoods;

    % DCC PREDICT % estimate

        startwindow = startwindow + 1;

        Tstart2 = max(1, startwindow + 1);
        Tend2 = min(T+1, startwindow + months);

        % from dcc_mvgarch

        if isempty(archP)
           archP=ones(1,k);
        elseif length(archP)==1
            archP=ones(1,k)*archP;
        end

        if isempty(garchQ)
            garchQ=ones(1,k);
        elseif length(garchQ)==1
           garchQ=ones(1,k)*garchQ;
        end
        index=1;
        [t,k]=size(data);
        parameters=zeros(t+1,14);
        F_param1_DCC(1,:) = zeros(1,14); %3xk+2
        F_param1_DCC(Tend:Tend,:) = DCC_parameters(1:14,1); %3xk+2
        parameters = F_param1_DCC(:,:);

        % from dcc_mvgarch_full_likelihood ADJUSTED CODE FOR FORECAST

            index=(k+sum(archP)+sum(garchQ));
            H1 = zeros(t,k);
            H1 = Est_GARCH(1:t,:);
            stdresid1(startwindow:Tend,:)=
data(startwindow:Tend,:)./sqrt(H1(startwindow:Tend,:)); %uses H1 uses
Return estimates of variance forecast with: H1 is pred_GARCH
            parameters(T+1,1:14) = DCC_parameters(1:14,1)'; %3xk+2
            a=parameters(Tend,index:index+dccP-1); % last garch index
parameter
            b=parameters(Tend,index+dccP:index+dccP+dccQ-1);
            sumA=sum(a);
```

```
            sumB=sum(b);
            F_parama_DCC(Tend:Tend,:) = a;
            F_paramb_DCC(Tend:Tend,:) = b;
            g = data(Tstart:Tend,1:k); % remember, still 1 window with same
Tstart and Tend as start.
            [tt,k]=size(stdresid1);


            %First compute Qbar, the unconditional Correlation Matrix
            Qbar=cov(g); % Calculate covariance matrix over datawindow


            % Next compute Qt
            m=max(dccP,dccQ);
            Rt=zeros(k,k,t+m);
            H=[zeros(m,k);H1];
            P=dccP;
            Q=dccQ;
            stdresid=[zeros(m,k);stdresid1];
            for j=Tend2 % forecast Qt, the unconditional covariance of the
standardized residuals stresid e(t) resulting from first stage estimation
                Qt(:,:,j)=Qbar*(1-sumA-sumB);%-sumA
                    for i=1:P
                    Qt(:,:,j)=Qt(:,:,j)+sumA*(stdresid1(j-
i,:)'*stdresid1(j-i,:));
                    end;
                    for i=1:Q
                    Ft(:,:,j)=Qt(:,:,j)+sumB*Qt(:,:,j-i);

Rt(:,:,j)=Ft(:,:,j)./(sqrt(diag(Ft(:,:,j)))*sqrt(diag(Ft(:,:,j)))'); %
calculate Rt, the time varying correlatio matrix
                    end;
                    F_pred_Rt(:,:,Tend2:Tend2) = Rt(:,:,Tend2:Tend2);
            end;

Rt(:,:,j)=Ft(:,:,j)./(sqrt(diag(Ft(:,:,j)))*sqrt(diag(Ft(:,:,j)))'); %
calculate Rt, the time varying correlatio matrix


    % STORE

        F_pred_CORR(:,:,Tend2 :Tend2) = pred_DCC_Rt(:,:,120:120);
        F_pred_COV(:,:,Tend2 :Tend2) = Qbar(:,:,:);



end
toc % time

%dcccovariance, DCC Returned a CORRELATION matrix, now to mix with the
GARCH and get a COVARIANCE matrix
F_pred_COVe = ones(k,k,T) * 0;
F_pred_GARCH_STDEV=F_pred_GARCH.^(0.5);
for i=1:T

F_pred_COVe(:,:,i)=diag(F_pred_GARCH(i,:))*F_pred_CORR(:,:,i)*diag(F_pred_G
ARCH(i,:)');
end
% The non synchronous shift
if (lag == 1)
```

```
    F_pred_COVe(:,:,T) = F_pred_COVe(:,:,T-1);
    months = months + 1;
end
```

9.3.2 Modern Portfolio Theory framework

```
startwindow = 2; % 2 more than variable estimations/forecasts
Tend = 120; % 119 more than startwindow
io(Tend2,:)=1;


while (Tend-2 < T)

Tstart = max(1, startwindow + 1);
Tend = startwindow + months;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%mpt, MODERN PORTFOLIO THEORY/ MEAN-VARIANCE OPTIMIZATION
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Portfolio = data(max(startwindow,121):Tend,:);
[m n] = size(Portfolio); % the first "dimension" is the number of columns,
and the second "dimension" is the number of rows
[AvrRet]= [0 0 0 0]; %mean(Portfolio); % calculates the average return of
the Portfolio (return) vector
[ExpRetVec]= (1+ AvrRet).^12-1; % expected return is (1 + average return)
annualized - 1
[CovMtx]= F_pred_COV(:,:,Tend); % calculates the covariance of the
Portfolio returns


%efficient, COMPUTE THE EFFICIENT FRONTIER (SHORT NOT SELLING ALLOWED)
NumPorts = n; % defines the variable n, number of portfolios
AssetMin = [ 0 0 0 0]; % sets minimum constraint per asset (6 assets)
AssetMax = [ 0.5 0.5 0.5 0.5]; % sets maximum constraint per asset (6
assets)
pval = 1; % sets constraint total portfolio adds up to 1
Constraint = portcons('PortValue',pval, NumPorts, 'AssetLims', AssetMin,
AssetMax, NumPorts); % sets the portfolio constraints
[PortRisk,PortRet,PortWts] = portopt(ExpRetVec, CovMtx, 50, [],
Constraint); % calculate the expected return vector and covariance matrix
of 50 portfolios
[OptRisk,OptRet,OptWts]=portopt(ExpRetVec, CovMtx, NumPorts, [],
Constraint); % calculate the expected return vector and covariance matrix
of n portfolios


%optimal, OPTIMAL ASSET ALLOCATION
RisklessRate = rate(Tend); % define the riskless rate
BorrowRate = rate(Tend); % define the borrowing rate
RiskAversion = 3; % define the level of risk aversion, Coefficient of
investor's degree of risk aversion. Higher numbers indicate greater risk
aversion. Typical coefficients range between 2.0 and 4.0 (Default = 3).
[RiskyRisk, RiskyReturn, RiskyWts, RiskFraction, OverallRisk,
OverallReturn] = portalloc (PortRisk, PortRet, PortWts, RisklessRate,
BorrowRate, RiskAversion); % calculate the optimal portfolio allocation


%%plot, PLOT THE RESULT
```

```
plot(PortRisk, PortRet, 'm-', OptRisk, OptRet, 'x', ...
     0, RisklessRate, 'k:square',  RiskyRisk, RiskyReturn, 'k:diamond',...
     [0; RiskyRisk], [RisklessRate; RiskyReturn],'r--') % plot variables
xlabel('Portfolio Risk'); % label x
ylabel('Portfolio Return '); % label y
title('Efficient Frontier'); % titel graph
legend('Efficient Frontier', 'Optimal Portfolio', 'Risk free asset', 'Risky
Portfolio', 'Asset Allocation Point')
grid on


%%STORE

F_PortRiskWts(Tend,1) = RiskyWts(1,1);
F_PortRiskWts(Tend,2) = RiskyWts(1,2);
F_PortRiskWts(Tend,3) = RiskyWts(1,3);
F_PortRiskWts(Tend,4) = RiskyWts(1,4);
F_PortRisk(Tend,1) = RiskyRisk(1,1);
F_PortReturn(Tend,1) = RiskyReturn(1,1);

startwindow = startwindow + 1;

end
toc % time
```